

기/획/특/집

한국의 집단유전학 : 엘리코박터에서 한국인 집단까지 (Korean Population Genetics : from Helicobacter pylori to Korean Population)



삼성생명과학연구소 유전체연구센터
책임연구원
이정호
cyi3639@samsung.co.kr

집단유전학(population genetics)은 한 생물종의 집단의 수준에서 그 유전적 다양성(genetic diversity)을 규명하고 집단내(內) 및 집단간(間)의 변이의 양을 추정하며 집단의 유전적 변동과 진화적 힘들(evolutionary forces)간의 관계를 규명하는 유전학 분야이다. 통계학적 이론과 계산이 응용된 통계유전학(statistical genetics)의 상당부분이 이러한 집단유전학의 발전과 주제에 뒷바침이 된다. 그러나 이론과 계산만이 집단유전학은 아니며 집단수준의 표본뜨기(sampling)와 생물종에 대한 유전학적 통찰, 그리고 집단규모의 대규모 실험이 동반된다. 또한 생성할 데이터가 가져올 해석적 응용을 만들 수 있는 지식기반이 있어야 한다. 따라서 이러한 학문적 배경이 세워져 있는 전문가의 연구설계(study design)는 가장 중요하면서도 어렵다. 광의의 집단유전학은 '학문적으로' 다시 협의의 집단유전학, 양적 형질을 다루는 양적유전학(quantitative genetics), 그리고 생물종의 진화에 대한 유전적 배경을 깊이 다루는 진화유전학(evolutionary genetics)의 세가지로 분류할 수 있다[1].

1. 엘리코박터 필로리의 집단유전학

한국인에게 많은 위암과 위장 질환의 발병에 엘리코박터 필로리(Helicobacter pylori)가 중요한 관련성을 가진다는 것은 잘 알려져 있다. 한국에서 미생물의 집단유전학은 많이 연구되지 않은 분야들중의 하나이다. 집단유

전학 기초데이터와 연구결과의 응용성에 대한 인식도 낮은 편이다. 따라서 이러한 분야의 지식기반이 거의 없고 새로이 창출할 의지도 거의 없었다. 그런데 전세계의 인류의 거의 반에 감염하고 있는 엘리코박터의 집단유전학은 흥미로울뿐 아니라 임상적으로 중요한 기초자료를 가져올 가능성이 높다. 그리고 여러 임상적 반응의 개인적 차이 혹은 변이에 대한 근거나 문맥(context)를 형성하게 되어 있다.

한국에서 잘 알려지지 않고 전문가도 희박한 집단유전학에 다시 엘리코박터라는 세균의 집단유전학을 이야기하면 그러한 분야도 있느냐고 반문하는 사람이 많을 것으로 예상한다. 그러나 이 세균종의 스트레인(strain)을 규명하고 그 스트레인의 다양성에 근거한 임상 적용에 관한 연구는 거의 집단유전학적 배경이 없이 수행된 것도 사실이다. 유전적(유전체적) 변이와 병독력(virulence)과의 관계나 약제저항성과 유전체 돌연변이와의 관계는 살펴보는 연구자들이 많을 것으로 추정된다.

집단유전학에서 단백질(효소) 표지와 같은 분자적 표지(molecular marker)의 사용은 1960년 후반 혹은 1970년 초부터 거의 유행처럼 이용되었는데 세균의 집단유전학에도 많이 이루어 졌다.

엘리코박터를 대상으로 하는 집단유전학적 연구에도 효소 유전표지를 사용하여 이 생물종의 유전적 다양도를 추정한 것으로 대표적인 것이 미국 텍사스 휴스턴에서 낸 고(Go)등의 논문이다[2].

논문의 제목을 한국어로 번역하면 "복수좌위 효소 전기 영동법에 의한 엘리코박터 필로리의 집단유전학적 분석 : 광범위한 대립인자적 다양성 및 재조합적 집단 구조". 전분 젤 전기영동에 의한 6개 유전자 (각각 대사 효소를 암호화하는 유전자들)의 대립인자(allele)를 지정하여 각 좌위(locus)의 대립인자의 수를 관찰하였고, 이에 따라 유전적 다양도의 한 지수(H)를 계산하여 다른 세균종의 값들과 비교하였다 (Table 1과 Table 2 참조).

기/획/특/집

Table 1. Allelic frequencies and genetic diversity at six enzyme locus in 73 ETs of *H. pylori* (extracted from Go MF et al., 1996).

Enzyme Locus	No. of alleles	Genetic diversity(h)
GLD	12	0.849
IPO	6	0.549
G6P	12	0.866
IDH	9	0.521
NSP	12	0.710
ADK	16	0.913
Mean	11.2	0.735

a. GLD: glutamate dehydrogenase ; IPO : indophenol oxidase ; G6P : glucose-6-phosphate dehydrogenase; IDH : isocitrate dehydrogenase ; NSP : nucleotide phosphorylase ; ADK : adenylate kinase.

아주 흥미로운 것은 다른 세균종들에 비해서 유전적 다양도가 굉장히 높은 것으로 나타났다는 것이다(Table 2 참조).

Table 2. Genetic diversity among ETs in representative species of pathogen bacteria (extracted from Go MF et al., 1996).

Taxon	No. of ETs	Average no. of alleles per locus	Mean diversity (H)
Bordetella spp.	14	2.4	0.284
Borrelia spp.	35	5.9	0.673
Erysipelothrix spp	50	3.7	0.314
Haemophilus influenzae	280	6.4	0.467
Helicobacter pylori	73	11.2	0.735
Legionella pneumophila	50	3.2	0.312
Listeria monocytogenes	45	3.6	0.424
Mycoplasma spp.	33	NA	0.580
Neisseria meningitidis			
8 primary serogroup	331	7.2	0.547
Asymptomatic carriers	78	6.0	0.538
Neisseria gonorrhoeae	89	2.5	0.410

Salmonella spp	80	NA	0.627
Staphilococcus aureus	252	6.7	0.289
Streptococcus pneumoniae	53	3.9	0.413
Streptococcus pyogenes	33	4.4	0.420

이는 다음의 몇가지 흥미로운 과학적 의문을 가지게 한다.

- 1) 전세계의 헬리코박터 집단들의 비교를 통해서 지역내(內) 및 지역간(間)의 차이가 나타날 수 있을까?
- 2) 이러한 차이 혹은 변이들의 존재는 인간집단내 및 집단간의 변이와 어떠한 상관성이 있을까?
- 3) 헬리코박터와 상호작용하는 인간집단내의 개인들의 유전적 프로파일(genetic profile) 혹은 유전적 체질은 다른 것인가?
- 4) 헬리코박터, 숙주 인간 그리고 환경 사이의 상호 관련성은 어떠한 양상일까?

한국의 헬리코박터 집단의 유전적 구조는 어떻게 되어 있을까? 그리고 그 구조는 전세계적 변이 양상과 어떠한 관계와 위상을 가지는가? 그리고 숙주 한국인들의 위장 질환의 발병과의 관계는 어떠할까? 이러한 한국적인 질문들에 대한 대답에는 집단유전학적 연구가 필히 기본적인 데이터와 지식을 가져다 주어야 한다.

한국에서도 헬리코박터에 대한 연구는 임상적인 측면과 그 일부 유전체적 측면에 대한 연구는 오랫동안 진행되어 왔고, 경상대학교 의과대학 BK21 헬리코박터 필로리 연구센타(센터장: 이광호 교수)에서 축적한 지식이 많다. 한국인 숙주에서만 나타나는 듯한 스트레인 51의 물리적 지도화에 의한 규명이 대표적인 것이다[3]. 잊달아 스트레인 51의 전체 유전체의 서열이 결정되었다[4].

최근에 전세계의 여러 지역에서 분리한 균주들을 대상으로 한 헬리코박터의 집단유전학적 연구에서 헬리코박터의 변이가 지역별 인류집단의 지역적 차이와 대동소이하게 나타나며 현재로서는 가장 강력한 유전학적 증거를 갖추고 있는 인류의 아프리카 기원설(Out of Africa hypothesis)을 지지하는 증거를 가져다 주었다 [5].

기/의/특/집

2. 인간집단유전학

한국인 집단의 유전적 구조(조성)에 대한 연구가 가져올 기초데이타는 헬리코박터와 위장 질병과의 공진화 (coevolution) 연구와 같은 양상을 포함하여 한국인의 복합성 질병연구에 다양한 응용성을 예고하고 있다. 복합성 질병 혹은 빈발성 질병들의 유전자를 찾는 연구에도 한국인의 집단유전학적 데이터들과 그에 기반한 기초적 유전지식들이 필요할 것은 당연하다. 효과적인 임상적 응용연구가 가능하게 하기 위해서도 한국인 집단에 대한 기초적 유전지식과 데이터는 필수적이다.

1970년대에 이미 인간집단유전학은 실험적으로는 단백질(효소) 표지를 이용한 방법론을 성숙시켰고 [6] (Harris and Hopkins, 1976), 그 이론적 발전과 함께 전세계적 규모의 연구를 통해서 거시적인 데이터와 지식이 축적된 상황이다 [7] (Cavalli-Sforza et al., 1994). 이제는 유전자 의 화학적 조성에 직접연결된 DNA변이를 유전표지 (genetic marker)로 이용하는 인간집단유전학이 꽃을 피우고 있다. 그리고 그 규모의 광대성으로 인해서 엄두도 내지 못하던 대용량의 유전자형결정실험(genotyping experiment)이 유전체학실험의 기계화 및 전산화의 도래에 발맞추어 폭발적으로 늘어났다. 더욱 대용량이면서 더욱 정확도가 높은 실험장비의 출현이 예상되고 있기 때문에 집단유전체학의 발전은 현재에 상상하는 규모이상으로 빠른 시간안에 성취될 것으로 보인다. 이른바 집단유전체학(population genomics)의 명실상부화(名實相符化)가 진행되고 있는 것이다. 이론적인 측면에 있어서도 그동안 한 두 유전자 혹은 좌위의 집단적 변동에 대한 생물수학적 배경만이 있었던 것 [8]을 생물학전용 슈퍼컴퓨팅을 통한 전산화를 통해서 수백, 수천, 수만의 유전자 혹은 좌위를 대상으로 실제로 계산해 볼 수도 있게 될 것 같고, 실제의 인간집단의 표본을 이용하여 그 실상을 검증해 볼 수도 있을 것으로 예상된다. 이론과 데이터처리 그리고 실험이 같이 발달할 것이다.

한국인 집단에 대한 연구는 1980~90년대에 한국유전학회를 중심으로 상당수의 단백질(효소) 유전표지에 근

거하여 이루어 졌고, 한국인 집단의 유전적 다양성에 대한 추정이 이루어져 왔으나 종합되지 않았다. 이들에 대한 메타분석(meta-analysis)이 가능할지 확신할 수 없으나 가능성이 없는 것은 아니다. 그리고 인간유전자들이 대부분 그 염기서열이 다 알려진 이상 이러한 효소 유전표지들에 근거한 데이터와 DNA변이에 근거한 집단데이타는 비교가능하리라고 본다.

한국인의 기원에 관한 연구의 일부로 부계와 모계의 유전양식을 가지는 Y염색체와 미토콘드리아 DNA의 변이를 유전표지로 이용한 집단유전학이 가능하다. 그중에 Y염색체의 변이를 이용하여 이러한 분야의 결과를 내고 있는 경우가 단국대학교 생물학과 김욱 교수의 연구팀이다 [9, 10].

최근에 심심치 않게 몽골리안 유전체연구사업이니 아시아유전체연구사업이니 하는 말이 회자되고 있다. 한국인 집단에 대한 유전적 다양성에 대한 연구가 민족적 기원이나 아시아전체의 다양성이라는 문맥안에서 그리고 전세계적인 연구추세 – 예를 들어 '인간유전체다양성연구사업(Human Genome Diversity Project)'이나 '중국유전체다양성 연구사업(Chinese Genome Diversity Project)'등 – 와 어깨를 나란히 해야 할 차원의 과업이라고 할 수 있다. 다른 연구팀이나 컨소시엄에서 나오는 데이터나 재료를 옮바로 비교하고 한국인 집단의 데이터가 가지는 전세계적 의미나 위상을 제대로 찾을 수 있는 기반위에서 수행되어야 할 과업이다. 현재는 초기이지만 한민족 유전체연구사업으로 부르기를 좋아하는 미국 컬럼비아대학 유전체센타 조우 터윌리거(Joseph Terwilliger)박사의 "한국인 이산집단연구(Korean Diaspora Project)"가 미국립보건원의 100억원가량의 연구비 규모로 진행되고 있다. 주로 중국 조선족, 우즈베크 및 구소련연방의 고려인들, 유럽의 입양아 집단들 같은 국외 한국인 이산집단의 집단과 가계를 중심으로 유전형과 환경의 상호작용등에 대해서도 연구할 예정으로 있다[11].

현재의 집단유전학 혹은 집단유전체학 연구에는 DNA변이를 이용한 유전표지들이 압도적으로 이용되고 있다.

기/획/특/집

유전적 지도 작성, 유전적 다양성 연구, 질병유전자 찾기 연구에서는 공히 미위성체(microsatellite) 혹은 단지반복다형(short tandem repeat polymorphism, STRP)으로 불리우는 유전표지들이 많이 사용된다. 최근에 약 400개 가량의 미위성체 유전표지를 이용한 전세계적 규모의 '인간 집단들의 유전적 구조(genetic structure of human populations)'에 관한 연구 결과가 발표되었다[12]. 인간 유전체다양성연구사업 컨소시엄에서 모은 전세계적 DNA표본 패널을 이용한 미래의 수많은 후속 연구의 디딤돌(stepping stone)을 형성하는 결과이다. 한국에서도 이러한 유형의 DNA변이에 대한 연구가 소규모로 수행되었다. 단지반복다형은 수사과학(forensic science)에도 사용되는 유전자지문으로도 사용되는데 한국에도 국립과학수사연구소나 친자확인생명공학사들에서 사용하기도 한다[13]. 요즈음 상당한 불이 일고 있는 단일염기다형(single nucleotide polymorphism, SNP)도 대립인자적 다양도(allolelic diversity)에서는 엄청 떨어지지만 유전체전체에서 가장 수가 많고 그 반수형(haplotype)을 연계하여 이용할 가능성이 충분하기 때문에 질병유전자 찾기 연구에 많이 이용되고 있다. 국제적으로 반수형지도(HapMap)작성 컨소시엄이 차세대 유전체연구사업으로 시작되었는데 가능성만을 믿고 이끌어 가는 개척 연구(pioneering research)의 전형적인 예다.

3. 집단유전학의 지식기반의 형성

대규모, 대용량, 기계화, 전산화로 대표되는 생물학적 거대연구 흐름인 유전체학의 발달이 한국에서는 자칫 잘못하여 선진국이 잘 닦아놓은 기반위에서 별반 연구투자나 학문적 발전을 일구지도 않으면서 소위 '선택과 집중'을 통해서 온전히 발달시킬 수 있다고 보는 사람이 많은 것 같다. 이러한 국가적 전략이면 생명공학의 산업적 기술과 경제를 발전시킬 수 있다는 확신에 차 있었지 않았는가 생각된다. 한국의 생명과학계에서 문제가 되는 것은 연구자나 연구책임자급의 수준에서 스스로의 전문분야에 대한 '선택과 집중'이 이루어져야 하는데 그간의 연구투자의 열악성으로 인해서 평생을 자신의 전문분야에

서 깊이 있게 연구하는 분위기가 없었다. 생명공학전반에 관한 문제는 선택과 집중의 원칙으로는 별로 승산이 없을 것 같고 개별 연구자 레벨의 연구분야에서의 선택과 집중은 국제경쟁력의 차원에서 아주 중요한 것 같다. 깊이 있게 연구한 전력(track record)과 능력(potential)을 갖춘 연구자층이 다양화되어야 전반적인 국제경쟁력이 생길것으로 보인다.

생물학적 혹은 임상적 응용을 위해서 꼭 필요한 기초데이터도 없이 진단이나 제약의 개발과 산업성장을 희망한다는 것은 어쩌면 국제적으로는 선진국이라는 황소뿔에 올라탄 악착빠른 개도국 생쥐를 연상케 하는 빌상인 것 같다. 아무런 정당한 노력을 기울이지도 않고 어부지리로 일화천금을 얻으려던 빌상이 아니었나 반성해야 하지 않을까? 이러한 주장에 잘 동원되는 것이 한국은 선진국 만큼의 경제력, 인프라, 인력이 없기 때문이라는 궁색한 배경 설명이다. 그러나 상대비교를 통해서 살펴보는 바대로 경제규모에 걸맞는 규모의 투자와 인력양성이 이루 어지지 않은 것이라는 외국 분석가들의 지적이 더 타당할지도 모른다. 그리고 집단유전학과 같이 어려운 분야에 인력이 양성되도록 배려하는 것 보다는 외국의 전문가들의 분석에 의존하려는 생각을 가질 수가 있는데, 이는 '미래의 고부가치를 창출하는 젊출인 유전적 데이터 관리 및 분석과 이론의 발전을 저해하는 근시안적인 시견이다.

다른 국가 특히 아이슬란드와 영국의 사례는 집단유전학적 연구가 튼튼히 자리잡은 배경위에 임상적 응용을 지향하는 나름대로 균형진 연구전략을 보여주고 있다. 아이슬란드의 경우는 국민 전체 인구의 족보/가계와 임상적 데이터들이 모두 국가관리체제의 데이터베이스에 들어가 있고 이와 병행하여 집단유전학적, 유전역학적 방법에 근거한 질병유전자 찾기 연구를 생명공학회사인 '디코우드유전학사(deCode Genetics)'가 수행하고 있다[14]. 영국의 경우는 국가연구관리기구중의 하나인 의학연구회(MRC, Medical Research Council), 민간의 기초의학연구자선기금인 웰컴트러스트(Wellcome Trust) – 이

기/의/특/집

산하에 수개의 기초의학연구센타와 연구소(biomedical centers and institutes)가 있고 인간유전체연구사업으로 유명한 생거연구소(Sanger Institute)도 이 기금이 운영한다 – 와 영국정부의 보건부가 연합하여 '영국바이오뱅크(UK BioBank)' 연구사업을 컨소시엄형태로 출범시켰다 [15]. 영국 시민들중에서 중년 50만의 자원대상자를 중심으로 생활방식과 질병, 유전적 체질과 질병, 그리고 환경과 유전과 질병의 삼자의 관계에 대한 연구를 하도록 하는 거대 장기(long-term) 연구사업이다. 이러한 국가적 거대사업의 근간에 필요한 집단유전학 및 유전역학(genetic epidemiology)에 대한 인력이 충분히 마련되어 있고 그들의 개척 연구 혹은 탐색적 연구(exploratory research)와 전문가 양성을 체계적으로 지원하고 활성화시키고 있다.

학문적, 임상응용적, 산업적 측면으로 보아도 집단유전학의 발전에 중심주가 되는 지식기반의 형성이 한국의 유전체연구사업에서 시급한 필요성으로 대두되고 있다. 유전체연구에서 주축적 지식기반을 형성하는 집단유전학 없이 그동안 한국에서 진행되던 대로의 추종연구를 주축으로 하는 상태로서는 연구비대비 연구생산성의 개선을 기대할 수 있을지도 의문이다. 사상누각을 지어가는 지름길이 될 개연성도 충분하다. 집단유전학적 지식기반에는 이론, 계산, 소프트웨어, 기초데이터의 관리, 분석, 해석등의 작업과 그에 필요한 인력이 가장 중요한 부분을 차지한다. 기반과 기초가 튼튼하면 지어진 건물이 단단하고 오래간다. '한국인 집단을 주요 대상으로 하는 한국인의 집단유전학'이 올바로 서야 학문적, 임상응용적, 산업적 발전이 따라온다.

인용문헌

- Crow JF (1986) Basic Concepts in Population, Quantitative and Evolutionary Genetics, W. H. Freeman, New York
- Go MF, Kapur V, Graham DY, Musser JM (1996) Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. *J. Bacteriol.* 178:3934–3938.
- Lee WK, Choi SH, Park SG, Choi YJ, Choe JY, Park JW, Jung SA, Byun EY, Song JY, Jung TS, Lee BS, Baik SC, Cho MJ, Youn HS, Ko GH, Kim YS, Park JH, Lee DS, Yoo HS, Ghim SY, Rhee KH (1999) Genomic diversity of *Helicobacter pylori*. *J. Korean Soc. Microbiol.* 34:519–532.
- Lee WK, Baik, Song JY, Park JU, Cho MJ, Youn HS, Ko GH, Choi SH, Park HS, Kang SS, Jeong JE, Hahn YH, Kim S, Kim YS, Yoo HS, Kim JJ, Rhee KH (2002) Genome sequence of a Koran isolate (strain 51) of *Helicobacter pylori*. p.62, Abstract A-17, Proceedings of the 11th Korea Genome Conference, KOGO(Korean Genome Organization), August 22~23, 2002.
- Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, Yamaoka Y, Megraud F, Otto K, Reichard U, Katzowitsch E, Wang X, Achtman M, Suerbaum S (2003) Traces of human migrations in *Helicobacter pylori* populations. *Science* 299: 1582–1585.
- Harris H, Hopkinson DA (1976) Handbook of Enzyme Electrophoresis in Human Genetics. North – Holland Publishing, the Netherlands.
- Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The History and Geography of Human Genes. Abridged Paperback Edition. Princeton University Press.
- Balding DJ, Bishop M, Cannings C (eds.) (2001) Handbook of Statistical Genetics. Wiley. Chichester, UK.
- Kim W, Shin DJ, Harihara S, Kim YJ (2000) Y chromosome DNA variation in East Asian populations and its potential for inferring the peopling of Korea. *J. Hum. Genet.* 45:76–83.
- Kim W, Shin DJ, You SA, Kim YJ (1998) Y-specific

기/획/특/집

DNA polymorphisms of the YAP element and the locus DYS19 in the Korean population. *J. Hum. Genet.* 43:195–198.

11. Terwilliger JD, Goring HHH, Magnusson PKE, Lee JH (2002) Study design for genetic epidemiology and gene mapping for the Korean Diaspora Project. *Life Science Research(China)* 6:95–115.
12. Rosenberg NA, Pritchard JK, Weber JL, Cann MH, Kidd KK, Zhivotovsky LA, Feldman MW (2002) Genetic structure of human populations. *Science* 298:2381–2386.
13. Han MS, Hong SB, Choi SK, Cho YH, Jin HJ, Kwak KD, Kim W (2002) Population genetic data on the thirteen CODIS short tandem repeat loci in Koreans. *Kor. J. Genet.* 24:83–87.
14. deCode Genetics (<http://www.decode.com>)
15. UK BioBank (<http://www.wellcome.ac.uk/en/1/biovenpop.html>)



존경하는 한국유전체학회 회원여러분
안녕하십니까

유전체연구회로 출발하여 지난 2000년 4월 15일 창립한 한국유전체학회는 날로 발전하는 유전체학과 생물정보학의 국내 연구를 보다 활성화하는데 기여하고자, 2003년 2월 한국유전체학회 공식학술지인 "Genomics & Informatics"를 발간하게 되었습니다. 회원 여러분의 노력의 결실이 되는 논문을 우리 학술지에 출판함으로써, 국내 유전체연구와 생물정보학 연구자들간의 교류가 활발해지고, 나아가 이 Genomics & Informatics 학술지가 국제적으로 인지되는 학술지로 발전할 수 있기를 소망합니다. 이러한 꿈을 이루기 위해서 회원 여러분의 적극적인 논문 투고가 무엇보다도 절실합니다. Genomics & Informatics는 genetics, genomics, proteomics, metabolomics, physiomics, microarray, chip technologies, high-throughput biology, systems biology, bioinformatics 등에 관련된 모든 논문을 환영하며, 논문의 종류로는 original article은 물론 rapid communication, application note 등도 접수하고 있습니다. Mini-review는 편집위원회에서 초청하는 형식으로 진행되며, 회원 여러분의 주천도 환영합니다. 논문투고 양식은 한국유전체학회 웹사이트인 www.kogo.or.kr에서 보실 수 있습니다. 저희 한국유전체학회 편집위원회를 비롯한 운영위원회 위원들은 Genomics & Informatics를 국내 학술진흥위원회 학술지로 등록하는 것은 물론 나아가 SCI Journal로 발전시키는 것을 목표로 최선을 다할 것을 약속드립니다. 회원 여러분의 적극적인 협조를 부탁드립니다. 감사합니다.

Genomics & Informatics 편집위원장 서정선
사단법인 한국유전체학회 회장 김병동