한국에서 발견된 나균의 구조-이웃 군집법에 의한 분자 역학 조사

김종필
한국한센복지협회 연구원

Molecular Epidemiology of *Mycobacterium leprae* as Determined by Structure-Neighbor Clustering in Korea found cases

Jong-Pill Kim
Institute for Leprosy Research, Korean Hansen Welfare Association

Background : It has proven challenging to investigate the molecular epidemiology of *Mycobacterium leprae*, the causative agent of leprosy, due to difficulties with culturing of the organism and a lack of genetic heterogeneity between strains. Recently, A panel of variable-number tandem-repeat (VNTR) markers and an alternative method, structure-neighbor clustering, which assigns isolates with the most similar genotypes to the same groups and, subsequently, subgroups, without inferring how the strains descended from a common ancestor have been developed.

Methods : A total of 29 samples from Korea found cases were studied by 14 VTRN typing and an alternative method, structure-neighbor clustering with 13 and 14 VNTRs by Structure Program(k=10).

Results : Only 286 cases of 522 total cases(including database of Bellingham Research Institute) showed p>0.8(in 13 and 14 VNTRs). Almost Korea found cases(18 cases) were included in group 3(13 VNTRs), in group 9(14 VNTRs)(by Structure Program, k=10).

Conclusions : The structure-neighbor clustering by Structure Program with panels of VNTR is a useful approach for investigating the molecular epidemiology of *Mycobacterium leprae.*

※ Key Words : *Mycobacterium leprae*, VNTR, structure-neighbor clustering

서 론

한센병은 나균(*Mycobacterium leprae*)에 의해 야기되는 만성 질환으로 피부, 말초 신경 및 호흡기 점막에 침범하는 특징이 있다. 약제 치료의 성공에도 불구하고 한센병은 여러 개발 도상국에서 널리 퍼져 있으며 일부 지역에서는 새로운 환자의 발생률이 상대적으로 변하지 않고 있다.[1] 최근 효과적인 치료제의 보급에도 불구하고, 아직도 매년 수십만 명의 신환자가 확인되고 있는데 이는 항생제에 기반을 둔 한센병 퇴치전략[2] 환경 내

에서도 지속적인 전파가 일어나고 있다는 것을 의미한다. 이에 대한 대책으로 성공적인 감염을 일으키는 데 필요한 미생물의 수와 감염과 증상의 첫 발병 사이의 시간을 포함하는 나균의 병인에 대한 기본적인 이해 또한 필요하다. 또한 한센병 전파 역동을 이해하는 것은 한센병 박멸을 목적으로 하는 새로운 전략 수립에 필수적이다.

나균 분리주에서 16개의 VNTR의 loci를 이용하여 분류할 수 있는 방법이 비교적 비싸지 않은 분자생물학적 분리주 유전자형 분류 방법으로 소개되었고[3], 전파에 관한 이해를 위한 나균 분리주 간의 거의 유사한 집단이나 소집단의 개체들로부터 신뢰성 있게 구분할 수 있는 새로운 2단계 접근법인 구조-이웃 군집법(structure-beighor clustering)이 개발되었다. 이는 접근성의 정확도를 평가하고 분리 균주가 다른 다양한 조건에 90% 이상의 정확도를 가지고 적절한 무리나 소무리로 나열되었는가를 찾아내기 위한 시뮬레이션 연구의 자료를 이용하고 있다.

Bellingham Research Institute의 웹사이트(http://web.me.com/barrygahall/Leprosy/Database/Database.html)에 저장되어 데이터베이스 내의 분리주 간의 75% 정도의 군집화가 가능하다. 새로운 군집화 알고리즘이 개발되어야 90% 이상의 군집화가 가능하다. 역학적 연구를 위해 연구자가 데이터베이스 내의 분리주와 새로운 분리주와의 관계만을 빠르게 알기 위해서는 NearestNeighbor 프로그램을 이용한다.[4] 이를 통하여 나균의 전세계적인 집단 구조와 국가 간의 나균의 전파에 대해서도 전체적으로 평가할 수 있게 되었다.

이에 저자는 2001년 이후 한국에서 진단된 환자들에서 얻는 나균 시료를 이용하여 집적 계층적 클러스터링과 함께 구조-이웃 군집법을 이용하여 한국에서 발견된 나균의 분자역학 조사를 실시하고 이를 보고한다.

※ 교신저자 : 김종필
전자우편 : dr_jpkim@hotmail.com
주 소 : 경기도 의왕시 원골로 59
한국한센복지협회(031-452-7094)

연구방법

1. 대상

2001년부터 2016년 사이에 한국한센복지협회 연구원을 내원한 한센병 환자(신환 및 재발환자 포함)에서 분리된 나균 시료 29개를 이용하였다.(Table 1) 29개 중 2개의 시료(x100, x101)는 착오에 의해 기초자료를 확인할 수 없었다.

2. 나균 분리 및 나균DNA 추출

나균 시료를 QIAGEN DNA mini-prep kits®을 이용하여 제조사가 지정한 방법을 이용하여 나균 DNA를 얻었다.

3. 중합효소연쇄반응(PCR)에 기초한 증폭 검사 및 시켄싱

나균의 RLEP을 대상으로 실시간 PCR을 위한 Primer-BLAST를 이용하여 나균 RLEP4 (X17152.1)에 대한 시발체 (forward:ACCATTTCTGCCGCTGGTAT, reverse:ATCTGCGCTAGAAGGTTGCC)를 제작하여 QuantiTect SYBR® Green PCR Kits(Qiagen, USA)를 사용하여 제조사의 지침에 따라 실시하여 실시간 PCR 중합효소연쇄반응을 시행하여 CT값을 계산하여 다음 과정으로의 진행 여부를

Table 1. Korea found Cases

ID	age	sex	WHO classification	Disease status	Year_found	Birth Country
n011	59	male	MB	new case	2001	Korea
n031	67	male	MB	new case	2003	Korea
n051	61	male	MB	new case	2005	Korea
n052	63	female	MB	new case	2005	Korea
n061	68	male	MB	new case	2006	Korea
n071	69	male	MB	new case	2007	Korea
n072	63	male	MB	new case	2007	Korea
n073	70	male	MB	new case	2007	Korea
n091	64	female	MB	new case	2009	Korea
n092	68	female	MB	new case	2009	Korea
n101	71	male	MB	new case	2010	Korea
n121	81	male	MB	new case	2012	Korea
n122	57	male	MB	new case	2012	Korea
n131	64	male	MB	new case	2013	Korea
n141	73	female	MB	new case	2014	Korea
n161	71	female	MB	new case	2016	Korea
r031	63	female	MB	relapse case	2003	Korea
r032	65	male	MB	relapse case	2003	Korea
r051	66	male	MB	relapse case	2005	Korea
r052	70	male	MB	relapse case	2005	Korea
r071	55	male	MB	relapse case	2007	Korea
r081	55	female	MB	relapse case	2008	Korea
r091	58	male	MB	relapse case	2009	Korea
r101	55	female	MB	relapse case	2010	Korea
f081	31	male	MB	new case	2008	Sri Lanka
f082	26	male	MB	new case	2008	Thailand
f101	30	male	MB	new case	2010	Indonesia
x100	?	?	?	unknown	?	?
x101	?	?	?	unknown	?	?

판단하였다. 그 결과에 따라 Gillis 등[3]에 의한 방법에 따라 14종의 VNTR loci에 대한 중합효소연쇄반응을 실시한다. 14종의 VNTR loci에 대한 시발체에 이용하여 Qiagen사의 multiplex-PCR kit를 사용하여 10ul PCR Mastermix(2X Qiagen multiplex-PCR, Qiagen, USA), 2.0ul Q solution, 시발체 각각 1ul (stock concentration 25mM), 시료 2ul을 첨가하여 최종량 20ul로 하여 동일한 과정의 중합효소연쇄반응하여 최종산물을 전기영동에 의해 확인한 후 외부 회사에 시켄싱을 의뢰하여 결과를 확인한다. 14종의 VNTR loci에 대해 사용한 시발체는 Table 2와 같다.

Table 2. Short tandem repeat loci and the primer pairs(forward and reverse) required for PCR amplification and sequencing[3]

Locus	Forward Primer			Reverse Primer			Total length
	No.	Sequence	Tm	No.	Sequence	Tm	
AC8a	F1	GTGTTACGCGGAACCAGGCA	65.5	R1	CCATCTGTTGGTACTACTGA	53.5	124
AC8b	F2	GATGCGACTATCACTCGCACGCAGTT	68.0	R3	GCTGGTTTCCTTCTAGTCCC	60.1	140
AC9	F2	GCCTGGTGCCCGGACAATGC	69.2	R2	ACATCACACTGATCTCGCCGGCGCT	71.4	145
TA10	F1	TAGATTCAAACGACCATGCA	57.2	R1	TGATAATCACGTGTTTCCGC	58.4	185
AT17	F2	GACACACTCGATCTCAGTAG	56.8	R1	TTAGCAGGACGATTGTACAG	57.0	180
GTA9	F4	CGCAGATGCAACGATCAC	59.3	R4	AATATGCATGCCGGTGGT	60.0	122
GAA21	F1	CTACAGGGGGCACTTAGCTC	62.1	R2	GGACCTAAACCATCCCGTTTT	60.4	201
GGT5	F2	TCACCATCGACGCTCCGGGT	68.4	R1	TCGGCCTGGTTGTCTGCCTT	66.8	161
6-7	F2	CTACTTGCGCGCCACCGCCA	70.6	R2	GCCGTCGCCAGGTTTTGCAG	67.2	191
12-5	F1	CTGGTCCACTTGCGGTACGAC	65.1	R1	GGAGAAGGAGGCCGAATACA	61.4	289
18-8	F1	GCCCGTCTATCCGCATCAA	62.5	R1	GCAAAGATCAGCACGCCAAT	61.8	348
21-3	F2	TGTTGAAATTTGGCGGCCAT	61.5	R3	TGCAAGGAGTGCTCAGCTAT	61.5	180
23-3	F3	CAGTCGCCCGGATACTGTTA	61.5	R2	TAAATCCGCTCCCAAATCTT	57.1	190
27-5	F2	GTGCTGTGCCTGCAGCCGTT	68.8	R2	TCCCCAAAGCCGCCGAATCC	67.9	270

4. 자료 처리

1) 데이터베이스화

14종의 VNTR loci에 대한 중합효소연쇄반응의 결과물을 시켄싱을 의뢰하여 결과를 확인한 후, 데이터베이스화를 실시하는데, 데이터베이스는 Microsoft Excel spreadsheet에 나균의 VNTR 특성은 각 분리주의 이름(ID), 근원 국가, 14개의 VNTR 반복수 등으로 구성되도록 하였다.

우리나라 증례와 비교하기 위해 함께 Bellingham Research Institute의 웹사이트(http://web.me.com/barrygahall/Leprosy/ Database/Database.html)에 저장되어 데이터베이스를 함께 포함하여 Structure-neighbor clustering, 집적계층적 클러스터링 및 Nearest Neighbor를 실시하기 위해 Microsoft Excel spreadsheet의 자료를 일반 text파일로 변환하여 준비한다.

2) 구조-이웃 군집법 (Structure-neighbor clustering)

연결되지 않은 표지자로 구성된 유전자형 데이터를 사용하여 모집단 구조를 추론하기 위한 모델 기반 클러스터링 방법을 구현하기 위해 structure(구조) 프로그램(http://pritch.bsd.uchicago.edu/structure.html)을 사용하였다. Hall 등[4]의 제안에 따라 클러스터(집단) 수(K)는 10으로 하였으며, VNTR 수는 13개(18-8 제외), 14개 두 가지로 분석하였다. 그 결과에서 각 개별에 대해 구조 프로그램은 K그룹 각각에 속하는 사후 확률을 추정하여, p가 0.8 이상일 때만 가장 가능성이 높은 집단에 할당하고 그렇지 않으면 다음 분석에서 제외하였다.

3) 집적 계층적 클러스터링 (Hierarchical Clustering)

Dendrogram이라고 하는 트리 다이어그램으로 표시되는 클러스터 계층 구조를 구축하기 위해 NCSS11(https://www.ncss.com/software/ncss/)을 이용하였다. VNTR의 수는 13개(18-8 제외), 14개 두 가지에 대해 한국에서 발견된 증례의 시료 전체와 이가 포함되어 있는 클러스터(집단)에 대해 각각 분석하였다.(NCSS11 option: Clustering Method Group ; Average (Unweighted Pair-Group), Distance Type; Euclidean, Scale Type; Standard Deviation)

4) NearestNeighbor

역학적 연구를 위해 연구자가 데이터베이스 내의 분리주와 새로운 분리주와의 관계를 빠르게 알기 위해서는 NearestNeighbor 프로그램을 이용하였다. 이 프로그램은 Bellingham Research Institute의 웹사이트(http://web.me.com/barrygahall/Software/Software.html)에서 다운받아 실시하였다. 관계가 확인된 후에 GraphViz(version 2.24)를 이용하여 도포화하였다. VNTR의 수는 13개(18-8 제외), 14개 두 가지에 대해 한국에서 발견된 증례의 시료 전체와 이가 포함되어 있는 집단에 대해 각각 분석하였다. 외국 사례는 한국에서 발견된 사례와의 관계가 확인된 경우만 도포화하였다.

결 과

1. 구조-이웃 군집법

클러스터 수(K)는 10으로 하였으며, VNTR 수는 13개(18-8 제외), 14개 두 가지로 분석하여, p가 0.8 이상일 때만 가장 가능성이 높은 집단에 할당하고 그렇지 않으면 다음 분석에서 제외하였다. VNTR 수 13개 및 14개 모두에서 286예만이 p가 0.8 이상이었다. 한국에서 발견 증례는 VNTR 수 13개에서는 주로 집단 (3)에, 14개에서는 집단 (9)에 포함되어 있었다.

2. VNTR 13(18-8 제외)

1) 한국에서 발견된 모든 증례

NCSS11의 dendrogram에서 미분류 외에 5개의 소집단으로 분류되었고, NearestNeighbor 프로그램상 6개의 소집단으로 분류되었다. (Table 2, Fig. 1, Fig. 2)

2) 집단(1)

대다수가 필리핀 증례이고, 한국에서 발견된 n141이 포함되어 있었으며, NCSS11의 dendrogram에서 미분류 외에 10개의 소집단으로 분류되었고, 외국 사례는 한국에서 발견된 사례와의 관계가 확인된 경우만 포함하여, NearestNeighbor 프로그램상 1개의 소집단으로 분류되었다. (Table 3, Fig. 3, Fig. 4)

3) 집단(3)

대부분의 한국 발견 증례에 포함되어 있었으며, NCSS11의 dendrogram에서 미분류 외에 4개의 소집단으로 분류되었고, 외국 사례는 한국에서 발견된 사례와의 관계가 확인된 경우만 포함하여, NearestNeighbor 프로그램상 4개의 소집단으로 분류되었다.(Table 4, Fig. 5, Fig. 6)

4) 집단(5)

대다수가 중국 증례이고, 한국에서 발견된 n121, n131, n161, f081, f101, x100 등이 포함되어 있었으며, NCSS11의 dendrogram에서 미분류 외에 4개의 소집단으로 분류되었고, 외국 사례는 한국에서 발견된 사례와의 관계가 확인된 경우만 포함하여, Nearest Neighbor 프로그램상 2개의 소집단으로 분류되었다. (Table 5, Fig. 7, Fig. 8)

3. VNTR 14(18-8 포함)

1) 한국에서 발견된 모든 증례

NCSS11의 dendrogram에서 미분류 외에 7개의 소집단으로 분류되었고, NearestNeighbor 프로그램상 6개의 소집단으로 분류되었다. (Table 6, Fig. 9, Fig. 10)

2) 집단(1)

대다수가 인도 증례이고, 한국에서 발견된 f101, x100이 포함되어 있었으며, NCSS11의 dendrogram에서 미분류 외에 10개의 소집단으로 분류되었고, 외국 사례는 한국에서 발견된 사례와의 관계가 확인된 경우만 포함하여, NearestNeighbor 프로그램상 1개의 소집단으로 분류되었다.(Table 7, Fig. 11, Fig. 12)

3) 집단(2)

대다수가 필리핀 증례이고, 한국에서 발견된 n141이 포함되어 있었으며, NCSS11의 dendrogram에서 미분류 외에 7개의 소집단으로 분류되었고, 외국 사례는 한국에서 발견된 사례와의 관계가 확인된 경우만 포함하여, Nearest Neighbor 프로그램상 1개의 소집단으로 분류되었다.(Table 8, Fig. 13, Fig. 14)

4) 집단(6)

대다수가 중국 증례이고, 한국에서 발견된 n121, n131, n161, f081 등이 포함되어 있었으며, NCSS11의 dendrogram에서 미분류 외에 2개의 소집단으로 분류되었고, 외국 사례는 한국에서 발견된 사례와의 관계가 확인된 경우만 포함하여, NearestNeighbor 프로그램상 1개의 소집단으로 분류되었다. (Table 9, Fig. 15, Fig. 16)

5) 집단(9)

대부분의 한국 발견 증례 포함되어 있었으며, NCSS11의 dendrogram에서 미분류 외에 4개의 소집단으로 분류되었고, 외국 사례는 한국에서 발견된 사례와의 관계가 확인된 경우만 포함하여, NearestNeighbor 프로그램상 4개의 소집단으로 분류되었다.(Table 10, Fig. 17, Fig. 18)

4. 착오에 의해 기초자료를 확인할 수 없는 2개의 시료(x100, x101) 분석

VNTR 13개의 경우 x100은 한국에서 발견된 모든 증례에서는 dendrogram에서는 f101과 같은 소집단으로 분류되었고, x101은 미분류(n031, f081, n122) 소집단으로 분류되었다. NearestNeighbor 프로그램에서는 x101은 동일하나, x100은 f081보다 n121과 더 가깝게 확인되었다. VNTR 14개의 경우 x100은 한국에서 발견된 모든 증례에서는 dendrogram에서는 f101과 같은 소집단으로 분류되었고,

x101은 n141 및 n122와 같은 소집단으로 분류되었다. NearestNeighbor 프로그램에서는 x101은 동일하나, x100은 f081보다 n121과 더 가깝게 확인되었다.

Table 2. Sub-clusters of samples(total case found in Korea, used 13 VNTR data)

sub-cluster	samples
1	x100_n,
2	n141_n,
3	n011, n092, r091
4	n051, n052, n061, n071, n072, n073, n091, n101, r031, r032, r051, r052, r071, r101, r081
5	n131_n, n161_n, n121_n
unclassified	n031, x101_n, f081_n, n122_n

Table 3. Sub-clusters of samples(cluster 1 by Structure Program(K=10, 13 VNTRs))

sub-cluster	samples
1	ph63, ph64, ph114
2	ph70, ph107, ph109, ph121, ph122
3	ph11, ph25, ph49
4	ph6, ph53, ph60, ph72
5	ph38, ph94
6	ph26, ph78, ph89
7	in82, ph1, ph9, ph30, ph43, ph54, ph55, ph56, ph65, ph68, ph76, ph79, ph84, ph86, ph87, ph90, ph95, ph101, ph103, ph105, ph116, ph117, ph118, ph125, ph126, ph127, ph128
8	ph61, ph88
9	ph10, ph13, ph17, ph23, ph24, ph32, ph34, ph35, ph52, ph58, ph66, ph91, ph108, ph115, ph123, ph124
10	ph75, ph93
unclassified	**n141**, ph15, ph20, ph33, ph40, ph57

Table 4. Sub-clusters of samples(cluster 3 by Structure Program(K=10, 14 VNTRs))

sub-cluster	samples
1	n072, r031
2	n011, n092, r091
3	r051, r052
4	n051, n061, n071, n073, n091, n101, r032, r101, r081
unclassified	n031, n052, r071, ml1

Table 5. Sub-clusters of samples(cluster 5 by Structure Program(K=10, 13 VNTRs))

sub-cluster	samples
1	**x100**, **f101**
2	**n131**, **n161**, **n121**
3	ch1, ch2, ch3, ch5, ch9, ch10
4	ch16, ch19
unclassified	**f081**, ch13

Table 6. Sub-clusters of samples(total case found in Korea, used 14 VNTR data)

sub-cluster	samples
1	x100_n, f101_n
2	f081_n, f082_n
3	x101_n, n141_n, n122_n
4	n131_n, n161_n, n121_n
5	n011, n031, n051, n071, n072, n073, n101, r031, r032, r051, r052, r101, r081
6	n052, n061, n091, n092, r091
unclassified	r071

Table 7. Sub-clusters of samples(cluster 1 by Structure Program(K=10, 14 VNTRs))

sub-cluster	samples
1	**x100**, **f101**
unclassified	in4

Table 8. Sub-clusters of samples(cluster 2 by Structure Program(K=10, 14 VNTRs))

sub-cluster	samples
1	ph63, ph64, ph114
2	ph6, ph53, ph60, ph72
3	ph54, ph105
4	ph75, ph93
5	in1, in82, ph1, ph9, ph20, ph26, ph30, ph43, ph55, ph56, ph65, ph68, ph76, ph78, ph79, ph84, ph86, ph87, ph89, ph95, ph101, ph103, ph116, ph117, ph118, ph125, ph126, ph127, ph128
6	ph10, ph11, ph13, ph17, ph23, ph24, ph25, ph32, ph34, ph35, ph38, ph40, ph58, ph66, ph67, ph70, ph91, ph94, ph107, ph108, ph109, ph110, ph112, ph115, ph121, ph122, ph123, ph124
7	ph48, ph61, ph88
unclassified	**n141**, in81, ph33, ph57, ph90

Table 9. Sub-clusters of samples(cluster 6 by Structure Program(K=10, 14 VNTRs))

sub-cluster	samples
1	**n131**, **n161**, **n121**
2	ch1, ch2, ch3, ch4, ch5, ch6, ch11, ch14
unclassified	**f081**, ch13

Table 10. Sub-clusters of samples(cluster 9 by Structure Program(K=10, 14 VNTRs))

sub-cluster	samples
1	n072, r031
2	n011, n092, r091
3	r051, r052
4	n051, n061, n071, n073, n091, n101, r032, r101, r081
unclassified	n031, n052, r071, ml1

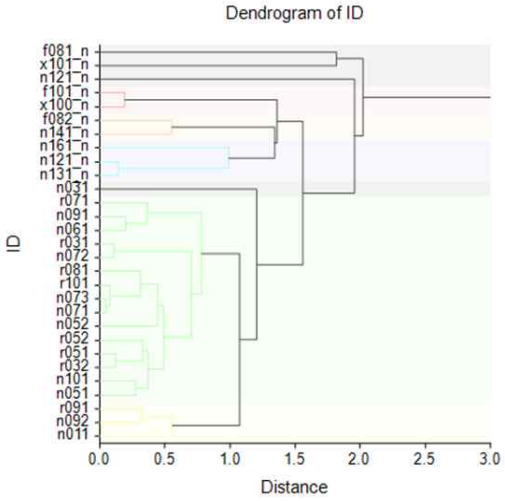

Fig. 1. Dendrogram of samples(total case found in Korea, used 13 VNTR data)

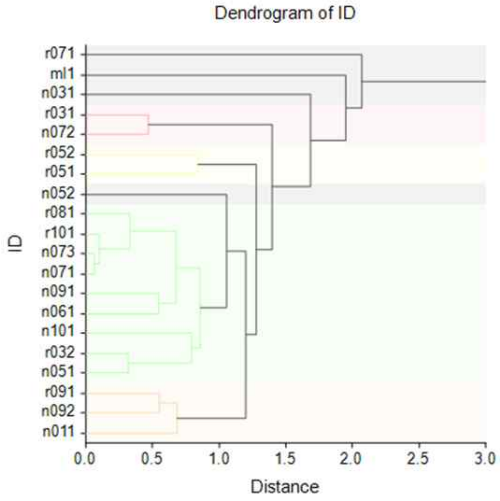

Fig. 5. Dendrogram of samples(cluster 3 by Structure Program(K=10, 13 VNTRs))

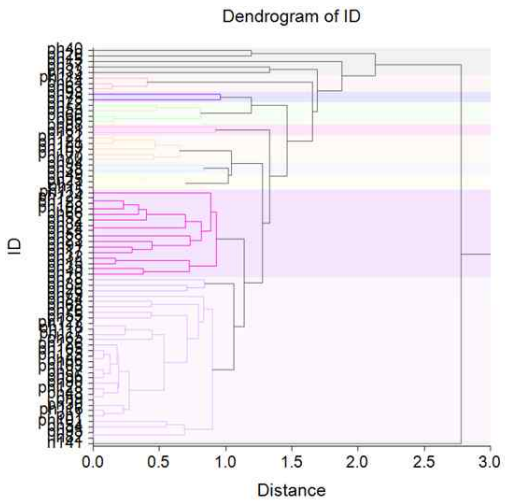

Fig. 3. Dendrogram of samples(cluster 1 by Structure Program(K=10, 13 VNTRs))

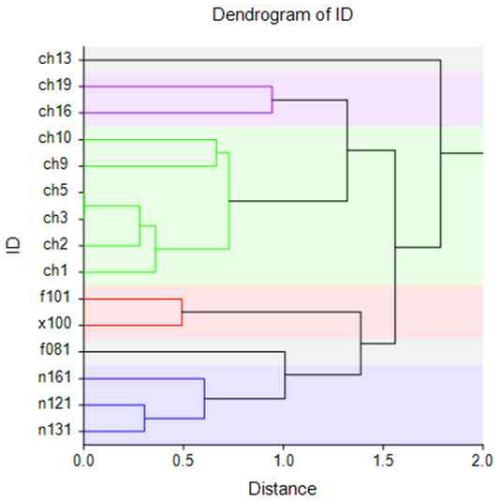

Fig. 7. Dendrogram of samples(cluster 5 by Structure Program(K=10, 13 VNTRs))

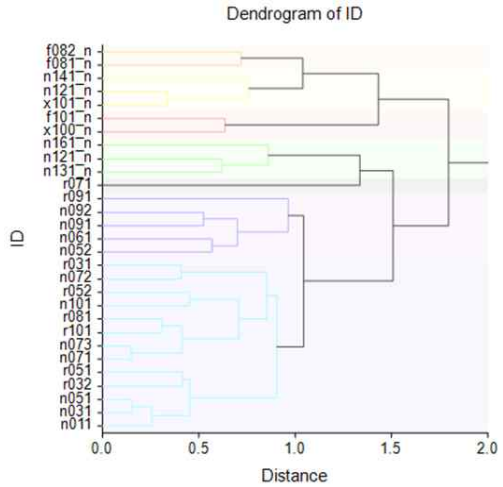

Fig. 9. Dendrogram of samples(total case found in Korea, used 14 VNTR data)

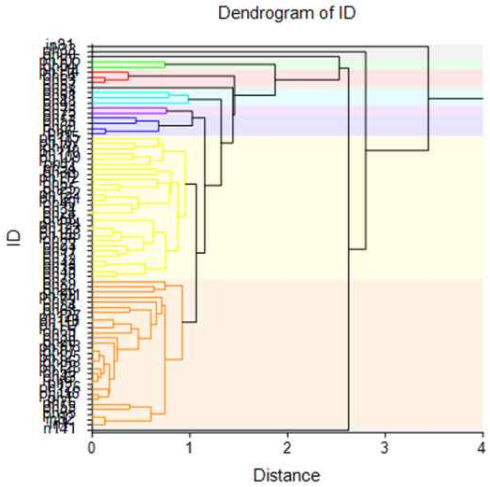

Fig. 13. Dendrogram of samples(cluster 2 by Structure Program(K=10, 14 VNTRs))

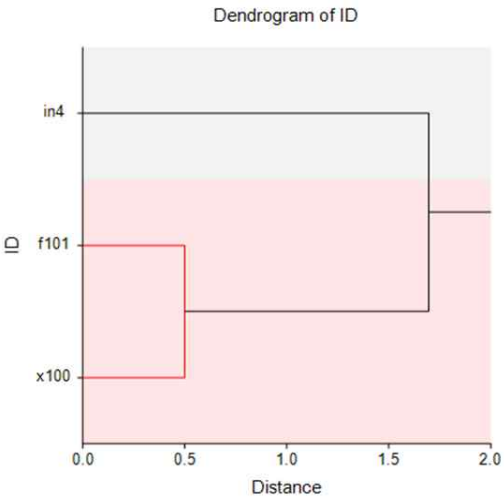

Fig. 11. Dendrogram of samples(cluster 1 by Structure Program(K=10, 14 VNTRs))

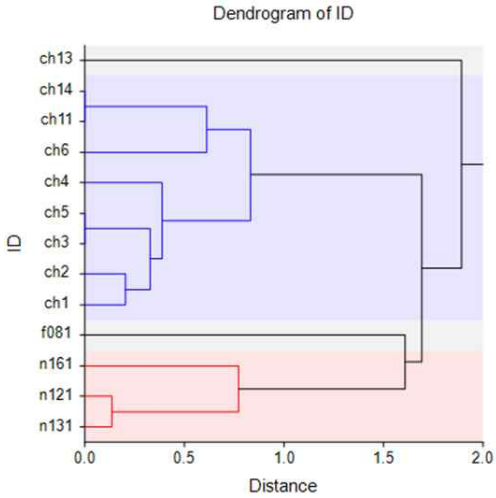

Fig. 15. Dendrogram of samples(cluster 6 by Structure Program(K=10, 14 VNTRs))

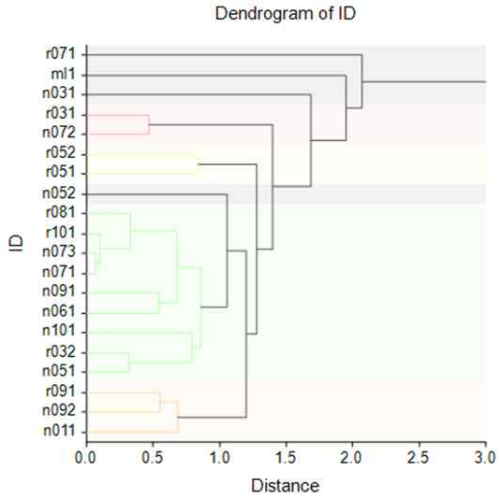

Fig. 17. Dendrogram of samples(cluster 6 by Structure Program(K=10, 14 VNTRs))

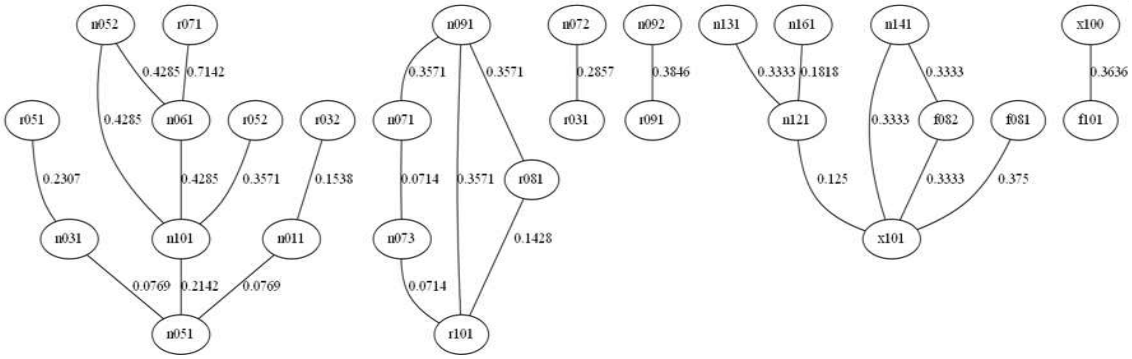

Fig. 2. Figure of NearestNeighbor(total case found in Korea, used 13 VNTR data)

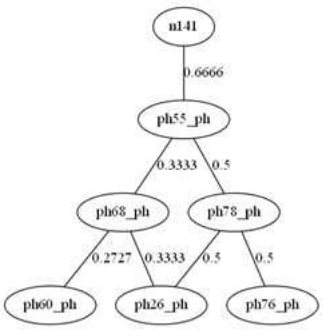

Fig. 4. Figure of NearestNeighbor(samples : cluster 1 by Structure Program (K=10, 13 VNTRs), only relative with Korea found cases)

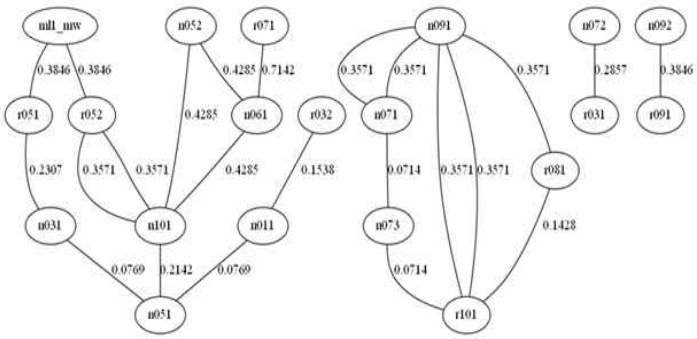

Fig. 6. Figure of NearestNeighbor(samples : cluster 3 by Structure Program(K=10, 13 VNTRs), only relative with Korea found cases)

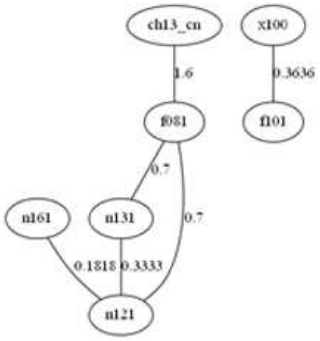

Fig. 8. Figure of NearestNeighbor(samples : cluster 5 by Structure Program(K=10, 13 VNTRs), only relative with Korea found cases)

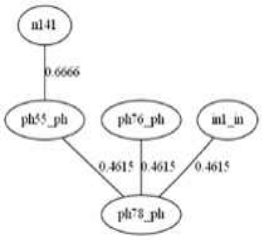

Fig. 14. Figure of NearestNeighbor(samples : cluster 2 by Structure Program (K=10, 14 VNTRs), only relative with Korea found cases)

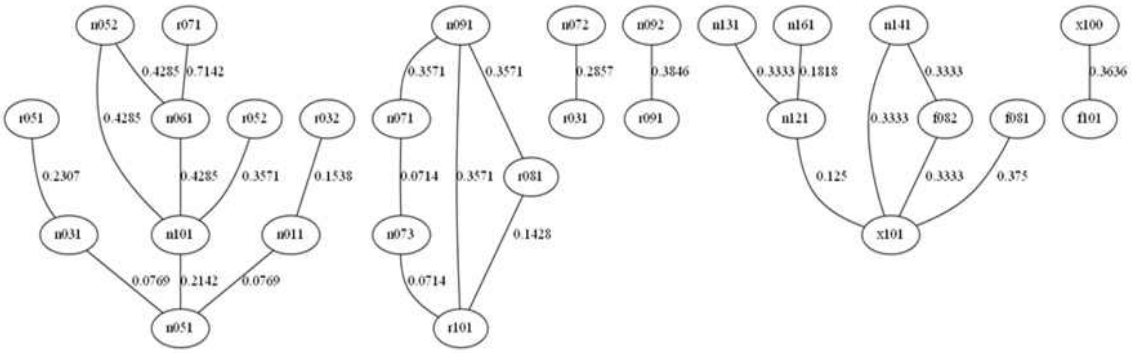

Fig. 10. Figure of NearestNeighbor(total case found in Korea, used 14 VNTR data)

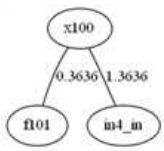

Fig. 12. Figure of NearestNeighbor(samples : cluster 1 by Structure Program (K=10, 14 VNTRs), only relative with Korea found cases)

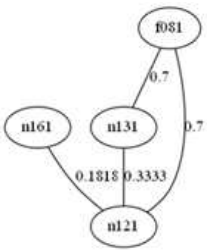

Fig. 16. Figure of NearestNeighbor(samples : cluster 6 by Structure Program (K=10, 14 VNTRs), only relative with Korea found cases)

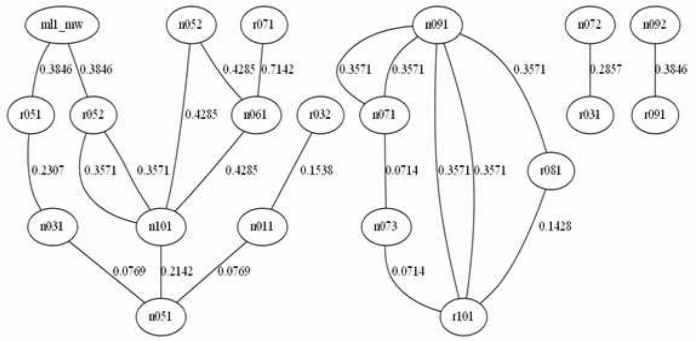

Fig. 18. Figure of NearestNeighbor(samples : cluster 9 by Structure Program(K=10, 14 VNTRs), only relative with Korea found cases)

고 찰

복합요법을 이용한 치료의 성공에도 불구하고, 한센병은 여러 개발도상국에 만연한 채로 있고, 일부 지역에서는 새로운 질병의 발생이 거의 변하지 않은 채로 남아 있다.[1] 세계보건기구는 현재 20여만 건 이상이 국제적으로 발병했을 것으로 추산하고 있지만, 병의 피해 및 예방적 측면에서 근거한 가능한 손실을 포함하면 연간 24만 건 이상의 새로운 발병이 추정되고 있다. 이런 계속되는 숫자는 효과적인 복합요법 치료가 한센병의 전염 고리를 방해하는데 실패했다는 것을 말해준다. 병원체의 관점에서 정확한 전염 방법과 사람과 사람 사이의 접촉에 대한 잠재적 중요성을 포함하는 나균의 역학적 관점에 대한 기본적인 정보는 알려지지 않은 채 남아 있다.[3] 오랫동안 유행률과 신환자 발견율 사이의 차이는 한센병의 전파 및 잠복기에 대한 지식 부족에 의한 것으로 평가된다. 한센병의 역학, 감염원, 전파의 정확한 방식, 접족 방식의 중요성 등에 대한 기본 지식의 부족에 이러한 문제를 해결하는 방법을 지연시키고 있다. 한센병의 전파 양상과 감염원에 대한 이해의 잠재능력으로 다양한 유행지역의 나균 지역적 분포를 확인하는 나균의 유전자형 분류 도구를 적용하는 것이다. 그 데이터는 독립적으로 보고되었기 때문에 결과에 대한 포괄적인 분석은 현재까지는 없었고, 최근 기존의 계통발생학적 분석이 이러한 균주 간의 관계를 평가하는데 부족함이 있다.[4] 세계보건기구는 현재 한센병 전 세계 유병률이 20만 건 이상인 것으로 추정하고 있지만 질병 수축 및 예방의 기본 측면에 대한 상당한 이해 부족을 강조하면서 매년 그 이상의 새로운 사례가 발견된다고 추정하고 있다.[3] *M. leprae*의 역학에 대한 기본 정보는 잠재적 감염원, 개체의 정확한 전달 방식, 사람과 사람의 접촉의 잠재적 중요성을 포함하여 알려지지 않았다.

한센병 감염의 가장 기본적인 역학을 이해하는 것은 합리적인 수준의 신뢰와 정확성으로 격리 균주 간의 관계를 추정할 수 있는 충분한 유전 변이를 밝힐 수 있는 값싼 분자 스트레인 타이핑 방법이 없기 때문에 크게 방해 받았다. *Mycobacterium leprae* 균주를 16종의 VNTR 유전자좌를 사용하고 10개의 세포를 출발 물질로 사용하여 최근에 개발되고 검증된 방법[3]에 사용되는 방법은 저렴한 타이핑방법의 필요성을 크게 충족시킨다.

Structure 프로그램은 연결되지 않은 표지자로 구성된 유전자형 데이터를 사용하여 모집단 구조를 추론하기 위한 모델 기반 클러스터링 방법을 구현한다. 이 방법은 Pritchard 등[5]의 논문에서 소개되었으며 Falush 등의[6,7] 후속 논문에서 확장되었다. 이 방법의 응용은 인구 구조의 존재를 증명하는 것, 독특한 유전자 집단을 확인하는 것, 개체를 개체군에 할당하는 것, 이민자와 혼혈된 개체를 확인하는 것을 포함한다. 간단히 말해서, 우리는 K개체군(K는 알려지지 않음)이 있는 모델을 가정하며, 각각은 각 유전자좌의 대립 유전자 세트에 의해 특징지어진다. 표본의 개체는 유전자형이 혼합된 것으로 나타나면 모집단에(확률적으로) 할당되거나 두 개 이상의 모집단에 공동으로 할당된다. 인구 집단 내에서 유전자좌는 하디-와인버그(Hardy-Weinberg) 평형과 연결 평형에 있다고 가정하고, 개인은 이것을 성취할 수 있는 방법으로 모집단에 배정된다. 이 모델은 특정 돌연변이 과정을 가정하지 않으며, 마이크로 위성, SNP 및 RFLP를 포함하여 일반적으로 사용되는

대부분의 유전자 표지자에 적용될 수 있다. 모델은 부분 집단 내에서 표지자가 연쇄 불균형이 아니라고 가정하므로 극단적으로 가까운 표지자를 처리할 수 없다. Hall 등은[4] 유전적으로 유사한 균주를 그룹과 하위그룹으로 클러스터링 할 수 있는 좀더 일반화된 접근법을 개발했다. 이 방법을 구조-이웃 클러스터링이라고 한다. 이 접근법은 두 단계를 필요로 한다.

첫 번째 단계에서, 구조는 개체군 구조추론 프로그램[5]을 사용하여 유전적으로 유사한 계통의 큰 개체군으로 격리되고, 두 번째 단계에서는 각 개체군의 개체가 하위 군집 가장 가까운 이웃 네트워크를 기반으로 한다. 진정한 관계를 알 수 있는 시뮬레이션 된 데이터를 사용하여 클러스터에 신뢰할 수 있게 할당할 수 있는 개인의 비율은 이들 사이의 유전적 거리와 그보다는 적지만 누락된 데이터의 양에 따라 결정된다.

NCSS11 프로그램 모듈[8]에서 사용 가능한 집적계층적 클러스터링 알고리즘은 일반적으로 dendrogram이라고 하는 트리 다이어그램으로 표시되는 클러스터 계층 구조를 구축하며, 개별 클러스터의 각 개체로 시작한다. 각 단계에서 가장 유사한 두 클러스터가 단일 새 클러스터로 결합되고, 융합되면 객체가 분리되지 않는다. Dendrogram의 가로축은 클러스터 간의 거리 또는 비 유사성을 나타내고, 세로축은 객체와 클러스터를 나타낸다. 유사성과 클러스터링이 구현되는데, 두 클러스터의 각 결합(융합)은 수평선을 두 개의 수평선으로 분할하여 그래프에 표시되고, 짧은 수직막대로 표시된 스플릿의 수평 위치는 두 클러스터 사이의 거리(비 유사)를 제공한다.

Hall 등은[4] 어떤 VNTR 유전자좌가 신뢰할 수 없고 분석을 혼란시킬 가능성이 있는지를 결정하는 전략을 기술했으며 이용 가능한 VNTR 마커 중 13가지가 분자 역학에서의 사용에 적합하다는 것을 발견했다. 구조-이웃 클러스터링의 품질은 분석에 사용되는 기본 데이터의 품질에 따라 다르다는 것을 분명히 알 수 있다. 또한 모든 나균 조사자가 적절한 품질 검사를 사용하여 식별한 14개의 신뢰할 수 있는 VNTR 유전자좌 중 locus 18-8은 신뢰할 만하지만, 격리된 전역 집합의 클러스터링에는 거의 영향을 미치지 않는 것처럼 보인다. 단일 국가 또는 한 국가 내의 단일 지역에 국한된 것과 같이 더 많은 지역 비교에서 변형 해결에 긍정적으로 기여할 수 있습니다. 향후 연구에서 이 문제를 구체적으로 다루고 데이터 세트에 포함하기 위해 최적의 유전자좌를 선택하는 방법을 더 고려할 것이다.

구조-이웃 군집법에서 Hall 등이[4] 제시한 클러스터 수(K)는 10으로 하였을 때, VNTR 수 13개(18-8 제외), 14개 모두에서 한국에서 발견 증례는 주로 집단(3)에, 14개에서는 집단(9)에 포함되어 있어, locus 18-8은 신뢰할 만하지만, 격리된 전역 집합의 클러스터링에는 거의 영향을 미치지 않는 것처럼 보인다는 보고와 일치하였다. 이는 실용성 면에서 13개의 VNTR을 사용하여도 무방하다고 고려할 수 있겠다.

한국에서 발견된 모든 증례에서 13개(18-8 제외) 및 14개의 VNTR을 사용한 경우 dendrogram에서 후자가 6개의 소분류가 되어 미분류 소집단 내 포함 증례가 적어 분류상 유용한 것으로 사료되고, NearestNeighbor 프로그램에서는 양자간의 상이점이 없었다. 단지 외국인 증례 모두가 착오로 인한 기초자료가 없는 증례와의 연관이 있어 이에 대

한 향후 평가가 필요할 것으로 사료된다. 말라이 예가 포함되어 있고 대부분의 한국 발견 증례인 13개의 VNTR 의 집단(3)과 14개의 VNTR의 집단(9)에는 모두 미분류 외에 4개의 소집단으로 분류되었고 미분류 수도 유의한 차이가 없었다. 말라위 예와는 r051, r052과의 관련성이 확인되어 이에 대한 역학적 평가가 향후 필요할 것으로 사료된다. NearestNeighbor 프로그램에서는 외국 증례의 국내 증례와의 관련성에서 양자간의 상이점이 없었다. 대다수가 필리핀 증례인 13개의 VNTR의 집단(1)과 14개의 VNTR의 집단(2)에는 한국에서 발견된 n141이 dendrogram에서 미분류 외에 10개 및 7개의 소집단으로 분류되었는데 후자가 미분류 소집단 내 포함 증례가 적어 분류상 유용한 것으로 사료된다. NearestNeighbor 프로그램에서는 외국 증례의 국내 증례와의 관련성에서 양자간의 상이점이 없었다. 대다수가 중국 증례인 13개의 VNTR의 집단(5)와 14개의 VNTR의 집단(6)에는 각각 한국에서 발견된 n121, n131, n161, f081, f101, x100 및 n121, n131, n161, f081 등이 dendrogram에서 미분류 외에 4개 및 2개의 소집단으로 분류되었는데, 한국 증례가 전자에서는 2개 소집단으로 후자는 1개 소집단으로 포함되어 이에 대한 유용성의 평가 필요할 것으로 사료된다. NearestNeighbor 프로그램에서는 외국 증례의 국내 증례와의 관련성에서 양자간의 상이점이 없었다. Dendrogram 및 Nearest Neighbor 간의 비교 유용성에 대한 평가에 대한 향후 연구가 필요하기는 하지만, dendrogram 및 NearestNeighbor에서 14개의 VNTR 집단수가 13개의 VNTR 집단수보다 세분되어 자료의 소집단 분석을 위해서는 가능하면 14개의 VNTR에 대한 자료를 얻어 이에 대한 13개 및 14개를 대상으로 통계처리 후 이에 대한 분석을 하는 것이 유용할 것으로 사료된다.

분자생물학적 유형 체계는 빠른 기술의 진보를 겪고 있다. 아종의 단계에서 미생물의 생물학적 다양성의 생물학적 기초를 이해하는 것에 대한 진보는 질병전파의 적절한 역학적 이해에 필요한 개념적 기초를 진보시킬 것이다. 집적 계층적 클러스터링 및 NearestNeighbor 도구와 병행한 나균의 구조-이웃 군집법에 의한 분자 역학조사의 활용은 한센병 전염의 역학에 공헌하고 결국 한센병에서 자유로운 세계에 도달하기 위한 효과적인 통제와 예방 전략을 가능하게 할 것이라고 사료된다.

참고문헌

1. Cardona-Castro N, Beltran-Alzate JC, Romero IM, Montoya E. Me lendez F. Torres RM, et al. Identification and comparison of Mycobacterium leprae genotypes in two geographical regions of Colombia. Lepr. Rev. 2009;80:316-321.
2. Young SK, Taylor GM, Jain S, Suneetha LM, Suneetha S, Lockwood DN, Young DB. Microsatellite mapping of Mycobacterium leprae populations in infected humans. J ClinMicrobiol, 2004;42:4931- 4936.
3. Gillis T, Vissa V, Matsuoka M, Richardus JH, Truman R, Hall BH, et al. Characterisation of short tandem repeats for genotyping Mycobacterium leprae. Lepr.Rev. 2009; 80:250-260.
4. Hall BG, Salipante SJ. Molecular epidemiology of Mycobacterium leprae as determined

by structure-neighbor clustering.J Clin Microbiol. 2010;48(6):1997-2008.

5. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data.Genetics. 2000;155:945-959.
6. Falush D, Stephens M, Pritchard JK. Inference of population structure: Extensions to linked loci and correlated allele frequencies. Genetics. 2003;164: 1567-1587.
7. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes. 2007;7:574-578.
8. https://www.ncss.com/software/ncss/