

미생물의 역학적 분석을 위한 검사 과정의 기본 개념

장 숙 진

조선대학교 의과대학 진단검사의학교실

Basic Concept of Laboratory Procedures for the Epidemiological Analysis of Microorganism

Sook-Jin Jang

Department of Laboratory Medicine, Chosun University College of Medicine, Gwangju, Korea

서 론

병원 감염의 역학적 연구에 있어서 기본적인 전제는 역학적 군집에서 얻은 분리주는 클론성으로 연관되어 있어 공통된 특징들을 가지는 반면 역학적으로 무관한 분리주는 동일 균종이라 하더라도 여러 가지 특징이 다르다는 것이다. 이러한 전제에 의거하여 형별 검사를 하여 양자를 구별하려는 연구들이 역학적 분석을 위해 지속적으로 수행되어져 왔다. 형별 검사는 미생물에 의해 표현되는 특징들을 검사하는 표현형적 기법과 염색체나 염색체 이외의 유전적 요소들을 DNA에 근거한 방법으로 직접 검사하는 유전형적 기법으로 분류된다[1]. 역학적 목적으로 사용된 최초의 형별 검사는 표현형적 기법이었지만 표현형적 특성에 주안을 둔 기법들은 같은 균종이나 아종(subspecies)의 분리주들간의 유전적 관계를 측정하는데 필요한 분석능을 제공해 주는 경우가 드물다. 이러한 분석능은 다양하게 제기되는 역학적 의문을 해결하는데 필수적이다. 병원 감염의 근원이나 각 개인들간의 질환 전파, 새로운 고도 독성 또는 약제-내성 균주의 출현, 상재균 또는 감염 균주의 미세진화(microevolution), 병원균의 일반적인 집단 구조 등을 정확하게 동정하기 위해서는 형별 분석의 범위가 균종 동정으로부터 균주 및 아균주(substrain) 형별검사로 옮겨가야만 한다. 많은 경우들에서 분리주들을 군집 분석을 통해 군별로 분별하기 위

해 유전적 관련성을 정확히 측정할 필요가 있다. 그러한 정보를 제공할 수 있는 방법들이 개발되어 왔으며 연구자들은 관련성 여부를 규명하기 위해 그들이 선택한 방법이 적절한지를 확인해야 한다. 연구할 때 대답해야 할 질문을 분명하고 정확하게 서술하고, 도달해야 할 유전적 관련성의 수준을 결정한 다음, 그 질문에 답하는데 필요한 분석능을 가진 유전적 지문 식별법을 선택하고, 선택한 방법의 효율을 입증하는 것은 필수적이다.

1. 생물형 형별 검사 (biotyping) 및 유전적 지문 확인 기법들

유전적 연관성을 평가하기 위해 DNA에 근거를 둔 기법들이 발달되기 전에는 유전적 차이보다 표현형적 차이를 측정하는 생물형 형별 검사에 의존하였다. 표현형이 유전형을 반영한다는 논리 하에 다수의 표현형적 매개 변수들을 채택하여 분석하면 유전적 관련성을 측정할 수 있다는 접근법이다. 세균과 진균 양자의 균종과 속명(genus)들을 서로 감별하기 위해 생물형 형별 검사는 여전히 빠르고 믿을만한 진단 방법들을 제공하고 있다. 동화 (assimilation)의 양상(생화학적 양상)들과 향체에 근거를 둔 검사법들이 세균과 진균 감별에, phage 형별 검사는 세균 감별에 지속적으로 사용되고 있다. 그러나 이런 방법들은 일반적으로 역학적 질문들에 답변하는데 필요한 정도의 유전적 감별력을 제공하지 못한다. 그 이유는 다음과 같이 두 가지로 설명할 수 있다. 첫째, 대부분의 이러한 검사들은 유전형을 잘 반영하기 위해 필요한 서로 무관한 매개 변수들을 충분히 제공하지 못한다. 그들은 대개 혈청형 검사와 같이 제한된 수의 군들만 감별해낸다. 둘째, 많은 유전

접수번호 : CM 7-1-11

교신저자 : 장숙진

(501-717) 광주광역시 동구 서석동 588

조선대학병원 진단검사의학과

TEL: 062)220 - 3259, 3272, 011)614 - 8853 FAX: 062)232 - 2063

E-mail: sjbjang@chosun.ac.kr

자들의 표현은 환경의 변화나 발육 계획(developmental programs)들과 함께 진균내에서 고빈도로 보이는 표현형 교체와 같은 가역적인 표현형적 변화들에 의해 영향을 받기 때문이다. 또한 phage와 플라스미드는 수평적으로 전파된다.

대부분의 생물형 형별 검사 방법들은 유전적 지문 식별법보다 기대에 미치지 못하나 그 예외로 다중 유전자 자리 효소 전기영동 (multilocus enzyme electrophoresis, MLEE)이 있다. MLEE는 대부분의 효율적인 DNA 지문 식별법과 동등한 수행능을 나타내는 강력한 식별법이다.

유전적 물질의 차이를 평가하는 DNA 지문 식별법은 가장 정확한 유전적 지문식별법으로 가정되고 있으나 항상 그런 것은 아니다. 일부는 MLEE처럼 효과적이나 다른 것은 그보다 못한 성적을 보인다. DNA 지문 식별법으로 일반적으로 흔히 쓰이는 방법에는 제한 분절 길이 다형성법(restriction fragment length polymorphism, RFLP)과 탐색자를 사용하는 RFLP, 다형성 DNA의 무작위 증폭법(random amplification of polymorphic DNA, RAPD), 파형 전기장 겔 전기영동(pulsed field gel electrophoresis, PFGE), 염기순서분석법 등이 있다.

2. 효율적인 유전적 지문 식별법의 일반적 필요 조건

지문 식별법에 필요한 일반적인 요건들을 열거할 수는 있으나 서로 다른 역학적 문제에 대해 서로 다른 수준의 해상도나 엄중성이 필요하다는 것을 인식해야 한다.

(1) 제기된 문제에 대답하는데 필요한 수준에서 유전적 거리를 반영할 수 있는 자료를 제공하는 방법을 선택해야 한다.

우리가 고려해야 할 관련성의 수준은 무엇인가를 염두에 두어야 한다. 왜냐 하면 유전적 관련성을 측정할 수 있는 참된 측정법은 거의 없고 어떤 방법은 다른 것보다 한 수준을 분별하는데 기능이 더 낫다는 정도이고 진화의 시간(evolutionary time)은 추정될 수 있을 뿐이기 때문에 우리는 그것을 연속된 것으로 보기 보다는 다음의 4가지 수준들로 카테고리를 나눈다. 즉 "동일함"과 "고도로 연관되었으나 동일하지 않음", "중등도로 연관되었음", "관련되지 않았음"의 4가지 카테고리로서 이러한 카테고리를 사용하면 유용하다.

(2) 유전적 지문 식별법은 환경의 간섭에 의한 변화와 고빈도의 유전체 재조직에 대해 영향을 받지 않아야 한다.

MLEE에서와 마찬가지로 DNA 지문 식별법이 분석하려고 하는 대상 연쇄는 주의깊게 선택해야만 한다. 예를 들어 플라스미드 DNA와 소형 염색체(minichromosomal) DNA는 일부 경우들에서 좋지 않은 선택이 될 것이다. 왜냐 하면 성장속도와 세포의 표현형이 유전적 표지자의 유지와 개편속도에 영향을 줄 수 있기

때문이다.

(3) 대부분의 경우에서 유전적 지문 식별법은 빠르고 편리하고 비용이 감당할 만하며 컴퓨터로 분석하기 쉬운 방법이어야 한다.

다수의 분리주를 대상으로 하거나 연관성을 복합적으로 측정하거나 미래에 그 자료를 후향적으로 이용할 계획이 있으면 컴퓨터를 이용하여 자동적으로 분석하고 저장할 수 있는 방법을 선택해야 한다. 보유한 기술력으로 검사할 수 있고 비용과 소요시간이 알맞으며 제기된 문제에 답을 줄 수 있는 방법을 선택해야 한다.

3. 흔히 사용되는 유전적 지문 식별법

원핵생물과 진핵세포생물 병원균들에 대해 반복적으로 사용되면서 다양한 역학적 질문들에 답변할 수 있는 몇 가지 유전적 지문 식별법중 가장 흔히 사용되는 방법에 대해 간단히 기술하면 다음과 같다.

(1) 다중 유전자 자리 효소 전기영동(multilocus enzyme electrophoresis, MLEE)

MLEE는 효소의 전기영동에 근거를 둔 방법이므로 DNA 분석법이 아니라 표현형적 다형성을 측정하는 방법이다. 여러 가지 효소의 단백질내 아미노산의 변이로 인해 하전량(the net charge of the protein)이나 구조(protein conformation)가 변화함에 따라 전기영동의 이동도가 달라지는 것에 근거를 두어 분석하며[2, 3], 모든 수준의 분석능에 도달하는 우수한 방법이다. 이들 다형성은 단백질 구조에 뿌리를 두고 있는데 이는 대개 그들의 유전 정보를 지정(coding) 하고 있는 구조적 유전자의 연쇄를 반영한다. 전기 영동적 이동도는 단백질의 최종적인(net) 전하에 주로 의존하는데 이는 단백질의 일차 구조에 의해 주로 결정되지만 단백질의 이차나 삼차, 사차 구조에 의해서도 영향을 받는다. 이러한 후자 수준의 구조는 하전된 아미노산 잔기들을 은폐하거나 노출시키는데, 이들이 단백질의 전체 하전량에 기여하고 전기영동 이동도에도 영향을 준다. 또한 MLEE는 유전자가 유전 정보를 지정하고 있는 영역 안에 있는 다형성만 검출한다. 일반적으로 보통 단백질의 아미노산의 약 15%만 변화해도 MLEE법으로 분별될 수 있다고 알려져 있다[4]. 이 수준의 민감도라면 유전적 관계의 서로 다른 수준을 평가하는데 충분하다.

MLEE 법의 일반적인 실험 방식의 주요 요소들은 (1) 분리주를 복제(cloning)하고 세포를 키우는 등 세포를 준비하여; (2) 세포를 분쇄하고 원침하여 상층액을 수집하여 단백질 extract를 준비하는 과정과; (3) 전기영동하고; (4) 동종효소(isozyme) 염색하여; (5) 결과를 분석하는 과정으로 구성되어 있다.

(2) 탐색자를 적용하지 않는 제한 분절 길이 다형성

법 (restriction fragment length polymorphism, RFLP)

RFLP는 진핵세포생물 분석에는 심한 제한점이 있으나 세균에 적용할 때는 훨씬 유용하여 주로 세균 분석에 사용된다. 세균을 지문 식별하는 주된 방법중 하나는 유전체를 한정된 횡수만큼 잘라주는 제한 효소로 RFLP 양상을 생성한 후 PFGE로 분석하는 방법이다. 전자는 큰 DNA 분절을 생성하고 후자는 이 크기가 큰 분절들을 agarose gel에서 분리한다. 세균의 유전체는 대개 단일한 DNA 분자(즉, 단일한 염색체)로 구성되어 있다 이 유전체는 크기와 복잡성에 있어 현저한 차이가 있을 수는 있으나 거의 모든 경우에서 세균의 유전체는 저등한 진핵세포생물 유전체보다 훨씬 덜 복잡하다. 원핵생물 유전체에서 생성된 RFLP 양상은 분별능이 훨씬 좋을 것 같으나 실제로는 그렇지 않다. 자주 잘라주는 제한효소를 사용하면 ethidium bromide 염색된 양상은 여전히 번잡(crowded)하기 때문에 탐색자를 쓰지 않는 표준 RFLP 분석은 세균 분석에 인기 있는 역학적 도구가 아니다. 대신 세균학적 역학자들은 DNA 분절을 소수로 만들기 위해 드물게 잘라주는 제한효소를 사용해 왔다. RFLP 법의 일반적인 실험 방식의 주요 요소들은 (1) 분리주를 복제하고 세포를 키우는 등 세포를 준비하고; (2) 세포를 분쇄하여 DNA를 준비한 다음; (3) DNA 의 제한 효소 처리로 소화하고; (4) PFGE 전기영동하여; (5) ethidium bromide 염색한 후; (6) 결과를 분석하는 과정으로 구성되어 있다[4,5].

(3) 부합화 (hybridization) 탐색자를 사용하는 RFLP

표준적 진핵세포생물 RFLP 양상에서 ethidium bromide는 모든 DNA 분절을 염색한다. 이것이 그렇게 많은 band가 나오고 양상이 복잡하여 band들 사이에 감별이 불가능하거나 어렵게 되는 이유이다. 부합화 탐색자는 RFLP 양상중 일부 분절들만 선택적으로 염색되게 하므로 양상의 분별능이 높아진다.

제한효소-소화된 DNA를 지문 식별용 탐색자로 Southern blot hybridization시키는 것이 원핵생물과 진핵세포생물 양자의 병원균을 분석하는데 사용되어져 왔다. 이 방법의 역량과 효율은 선택된 부합화 탐색자의 우수성에 달려있다[2]. 탐색자를 사용하는 RFLP 법의 일반적인 실험 방식의 주요 요소들은 (1) 분리주를 복제하고 세포를 키우는 등 세포를 준비하여; (2) 세포를 분쇄하고; (3) DNA를 준비한 다음; (4) DNA를 제한 효소로 소화하여; (5) 전기영동한 후; (6) ethidium bromide 염색하여; (7) DNA를 막에 옮긴 다음; (8) 탐색자로 부합화하여; (9) 양상을 가시화한 후; (10) 결과를 분석하는 과정으로 구성되어 있다.

(4) 다형성 DNA의 무작위 증폭법 (random amplification of polymorphic DNA, RAPD)

RAPD는 진핵세포생물 병원균의 DNA 지문 식별법

에 가장 흔히 사용되는 방법이다[6,7]. 정확히 개발되면 이 방법은 분별능의 모든 수준에서 연관성을 평가할 수 있는 고도로 효율적인 방법이다. 그러나 RFLP와 탐색자를 사용하는 RFLP 및 PFGE법과 대조적으로 RAPD법의 문제점은 검사실간 재현성에 문제가 있다는 점이다. 약 10 bases 정도의 무작위 시동체를 사용하여 유전체 전반에 걸쳐 증폭대상을 PCR로 증폭한다. 증폭산물을 agarose gel에서 분리하여 ethidium bromide 염색하여 가시화한다. 시동체 부합화 장소들 간에 간격이 변화하거나 시동체 결합장소가 나타나거나 사라지고 삽입, 결실, 재조합에 의해 위치가 변화할 때 다형성이 나타난다. 대개 10-bp의 시동체 각자마다 한 개 내지 몇 개의 주 band와 소수의 부 band가 생성된다. 어떤 부 band 들은 동일한 실험실에서 반복 실험을 할 때 재현성이 높지 않을 수도 있으므로 주 band들만 대개 분석에 사용한다. 이러한 복잡성이 역학적 연구를 수행하기에는 너무 낮기 때문에 유전적 관련성을 평가하는데 필요한 복잡성을 얻기 위해 여러 가지 시동체를 사용하여 자료를 축적해야 한다. 일반적으로 최소한 2개의 재현성 있는 강한 band를 생성하는 시동체 약 8개를 사용하여 최소한 16가지의 다형성 bands가 나오도록 하는 것이 일반적 지침으로 추천된다. 다형성의 정도가 분리주간의 유전적 변이를 분해하는데 충분할 만큼 커야 한다. 다형성의 정도가 너무 낮아도 연구를 편향시키기 때문에 10-90 %의 분리주에 나타나는 다형성 bands를 선택할 것을 추천한다[8-10].

RAPD법의 일반적인 실험 방식의 주요 요소들은 (1) 분리주를 복제하고 세포를 키우는 등 세포를 준비하여; (2) DNA를 준비한 후; (3) PCR 증폭하여; (4) 전기영동하고; (5) ethidium bromide 염색하여 band를 가시화한 후; (6) 결과를 분석하는 과정으로 구성되어 있다.

(5) 전기영동에 의한 핵형분석 (karyotyping)

하급 진핵세포생물들이 가지고 있는 다수의 염색체는 PFGE법으로 크기가 분리될 수 있는 범위 안에 있기 때문에 전기영동적 핵형분석이 진핵세포생물 병원균들의 지문 식별법으로 사용되어져 왔다[11]. 세포를 세포벽을 제거하는 효소와 혼합함으로써 형성된 세포벽 불완전 제거균(spheroplasts)를 agarose plug 안에 포매한다. 막을 제거하기 위해 계면활성제 (detergent)와 proteinase를 첨가하여 단백질을 소화하고 핵산을 유리한다. agarose 바탕질 (matrix)은 전단력(shear forces)을 줄여서 큰 염색체 DNA 가 분절되는 것을 막는다. 그 후 agarose plug를 agarose slab gel의 well 안에 넣어 PFGE 전기영동한 후 염색체 DNA를 ethidium bromide로 염색한다. 염색체는 염색체-특이 탐색자들로 Southern blot hybridization하여 동정할 수 있다. 감염성 효모 분석에 적용했을 때 양상들이 재현성이 있고 준비 방법에 별 영향을 받지 않는다는 보고도 있으나[12] 시약, 검체 준비, 전기 영동 조건에 따라 양상이 변화

한다는 연구들도 있으며 균들이 형질형 전환(switching)을 보일 수도 있다[13]. 실제로 전기 영동적 핵형 분석상 변화의 빈도가 고빈도의 형질형 전환(phenotypic switching)에 의해 급격하게 영향을 받을 수도 있다는 연구 등이 보고되어 PFGE법으로 동정되는 변화율(rate of change)이 항상 유전적 거리(distance)를 반영하는 것은 아닐 수도 있다는 문제가 제기되어 있다[4].

(6) 염기순서분석

염기순서분석이 균종내의 파생 양상(divergence)과 연관성을 평가하는데 가장 정확한 자료를 제공함에도 불구하고 그 주요한 영향은 균종간의 수준에서 계통발생학적 분지도를 생성하는데 한정되어진 데에는 이유가 있다. 균종간의 수준에서 단일 연쇄내 변화는 계통발생을 해석하는데 충분하다. 그러나 아종 수준에서는 대부분의 경우들에서 단일 장소내 연쇄의 변화 속도가 고도로 연관되었으나 동일하지는 않은 분리주들을 서로 구분할 만큼 충분히 빠르지 않아서 군집 분석에 충분한 정보를 제공하지 못한다. 그런 정보를 얻으려면 다수의 장소에서 염기순서분석을 하여 자료를 모아야 하므로 자료 수집에 필요한 기간이 길어진다. 이에 필요한 시간과 비용 문제가 극복되면 염기순서분석이 유망한 방법으로 사용될 수 있을 것이다. 유전적 연관성의 서로 다른 수준에서 감별에 필요한 수준의 변화를 보이는 유전자를 목표로 하는 주의 깊게 선택한 시동체 set에 근거를 둔 다중 유전자 자리 염기순서 형별 분석(multilocus sequence typing : MLST) 방법이 세균 분석에 적용되고 있다[14-16].

4. 유전적 지문 식별법의 효율 확인

제기된 문제에 대한 정확한 수준의 유전적 해상도를 그 지문 식별법이 제공하는가를 확인해야 한다. 그러므로 이미 검증된 방법을 채택하거나, 사용한 방법을 검증하기 위한 검사들을 수행해야 한다. 가장 직접적인 검사는 관련성이 나타났거나 알려져 있는 검사 분리주들의 set를 사용하여 서로 연관되지 않는 방법과 비교를 하는 것이다. 검사되어야 할 수준들을 분류한다면 첫째, 동일한 수준; 둘째, 고도로 유사하나 동일하지 않은 수준; 셋째, 중등도로 관련된 수준; 넷째, 무관한 수준으로 나눌 수 있다.

이러한 관련성의 각 수준들에 따라 다음의 4 가지 군으로 구성된 시험 수집균을 확보해야 한다. 첫째, 동일해야만 하는 분리주들(즉 신선하게 배양된 한 분리주에서 즉각 파생된 다수의 clones); 둘째, 고도로 유사하나 아마도 동일하지는 않은 분리주들(즉, 아마도 동일 기원으로 추정되는, 서로 다른 개인들에서나 몸의 위치들에서 짧은 간격을 두고 회수된 분리주들); 셋째, 중등도로 관련된 분리주들(즉 같은 지리학적 장소에서 나온 독립적인 분리주들의 집합에 포함된 균들); 넷째,

무관한 분리주들(즉, 지리학적으로 떨어져 있는 서로 다른 장소에서 나온 균들)이다. 시험 수집균을 각자의 방법으로 지문 식별하여 각 쌍의 분리주들에 대해 동일한 분석 방법으로 유사성 계수(similarity coefficient)를 계산한다. 각자 독립적인 set의 자료들에 근거한 계통수를 생성한 다음 양자를 서로 비교한다.

만약 유전적으로 무관한 방법들이 둘 다 동일한 분리주들을 동일하게, 고도로 연관된 분리주들을 비슷하지만 동일하지 않은 것으로, 무관한 분리주들을 서로 다른 것으로 동정하고, 또한 서로 무관한 두 방법들이 분리주들을 동일한 집단으로 균을 나눈다면 그 둘은 연관성의 모든 수준에서 서로의 효율을 함께 확증해 준다고 본다. 만약 그들이 동일하다고 추정되는 분리주들을 동일한 것으로 동정하고, 고도로 관련되었으나 동일하지 않은 분리주들을 비슷한 균들로 동정하였으나, 중등도로 관련된 균들을 무관한 분리주들로 동정함으로써 비슷한 풍조(fashion)로 군집을 분류하지 않았다면 그 방법 중 하나나 둘 다가 낮은 수준의 연관성을 평가하는데 비효율적일 수도 있다고 본다. 만약 두 방법이 분리주들을 일반적으로 동일한 균들로 분류하나 그중 한 방법만 고도로 연관되었으나 동일하지 않은 분리주들을 감별할 수 있다면 그 방법들은 서로 교차 확인을 하지만 둘 중 하나만이 클론성으로 집락형성한 한 집단(a clonal colonizing population) 내의 미세진화를 감별하는데 효과적일 것이다.

이 일반적인 교차 확인법은 세균과 진균, 기생충 병원균을 위한 다양한 지문 식별법의 효율을 평가하는데 사용되어져 왔다[3,10,17]. 유전적 지문 식별법은 그 체계가 얼마나 확실하게 보이는지에 상관없이, 교차 확인을 하지 않고서는 서로 다른 수준의 연관성을 규명하는 감별능력에 대해 항상 의심받을 수도 있다는 것을 강조해야만 한다. 그러나 그러한 교차 확인은 클론성의 집단 구조를 가진 세균들에 대해서만 가능하다는 것을 인식해야 한다[3,10,18,19].

5. 자료 분석

(1) 유사성 계수의 계산

유전적 지문 식별법의 자료들은 서로 다른 형태로 나타난다. MLEE는 다수의 점 양상으로 나오고 RAPD와 RFLP, 탐색자를 사용하는 RFLP 및 PFGE는 단독이나 다수의 band 양상으로 나온다. 자료의 형태에 무관하게 유전적 지문 식별법의 목표는 계통수(dendrogram)나 계통나무(tree)를 생성하기 위해 분리주들의 각 쌍간에 유사성 계수(S_{ABS})를 계산하는데 있다. S_{ABS} 를 계산하기 위한 공식과 S_{ABS} 로부터 계통나무를 생성하기 위한 공식은 주의 깊게 선택해야 한다[20].

1) band의 위치에 대한 단독 자료에 근거를 둔 방법 band의 양상들만 만드는 방법에서는 band의 유무를

1과 0의 이진수 값 (binary value)으로 각각 기술한다. n_{AB} 는 A와 B 두 양상에 공통으로 있는 band (1,1 로 표기) 수이다. a는 B에는 없으나 A에는 있는 band 의 수(1,0 으로 표기)이고 b는 A에는 없으나 B에는 있는 band 의 수(0,1 로 표기)이며 c는 A와 B에 없는 band 의 수(0,0 으로 표기)이다. 검체 크기인 x는 band의 총 수를 나타내고 수식으로는 $x = n_{AB} + a + b + c$ 이다. 공통 (match)되는 수 m은 $m = n_{AB} + c$ 이고 불일치(mismatch)의 수인 u는 $u = a + b$ 이다. 이런 기본적인 논리로 band의 위치에만 근거를 두어 생성된 다음과 같은 여러 가지 S_{ABS} 의 공식을 개발할 수 있다. 이중 가장 주목할 만한 것은 Jaccard(S_j)의 공식이다.

$$S_j = \frac{n_{AB}}{n_{AB} + a + b}$$

the coefficient of Dice (S_D) :

$$S_D = \frac{2n_{AB}}{2n_{AB} + a + b}$$

the sample matching coefficient (S_m) :

$$S_m = \frac{n_{AB} + c}{n_{AB} + a + b + c}$$

the Pearson coefficient (S_ϕ) :

$$S_\phi = \frac{(n_{AB}c - ab)}{\sqrt{[(n_{AB} + a)(n_{AB} + b)(b + c)(a + c)]}}$$

MLEE나 일련의 시동체들로 이항식의(binomial) 자료를 제공하는 RAPD와 같은 방법으로 얻어진 자료들은 banding 자료에서 얻어진 것과 유사한 바탕질(matrix)을 생성하기 위해 이진수 값으로 변환할 수 있다. 그런 다음 그것은 분리주들의 각 쌍간의 S_{ABS} 를 계산하기 위해 사용될 수 있다. S_{ABS} 를 산출하기 위해서 동일한 공식을 고려할 수 있다.

2) band 위치와 강도 양자에 근거를 둔 방법

많은 경우 band 강도에서 얻어진 정보를 band 위치에 대한 정보에 첨가한다. 이러한 공식들에서 X_{iA} 와 X_{iB} 는 A와 B 양상에서의 band 강도를 각각 나타낸다. X_{iA} 와 X_{iB} 는 A와 B 각각에서 모든 강도들의 평균들을 나타낸다. band의 위치와 강도에 근거한 S_{ABS} 의 여러 공식들이 나와 있으며 가장 유명한 것은 Pearson의 product-moment 상관계수 (S_r)이다.

$$S_r = \frac{\sum_{i=1}^n [(X_{iA} - \bar{X}_{iA})(X_{iB} - \bar{X}_{iB})]}{\sqrt{\sum_{i=1}^n (X_{iA} - \bar{X}_{iA})^2 \sum_{i=1}^n (X_{iB} - \bar{X}_{iB})^2}}$$

absolute difference/total area (S_{AB}) :

$$S_{AB} = 1.0 - \frac{\sum_{i=1}^n |X_{iA} - X_{iB}|}{\sum_{i=1}^n |X_{iA} + X_{iB}|}$$

S_{AB} 값을 계산하는 방법을 선택할 때 몇 가지 주의점을 명심해야 한다. 첫째, band 강도의 재현성이 좋지 않으면 band 위치에만 근거한 S_{AB} 가 분명히 더 정확하다는 것이다. 둘째, 사용한 방법이 band의 유무를 대립유전자 (allele)의 유무와 같이 취급한다면 band 위치에만 근거한 S_{ABS} 가 사용된다. 그러나 사용되는 DNA 지문 식별법의 대부분에서 대립(allelism)이 항상 확증되는 것은 아니라는 것을 인식해야 한다. 셋째, 반복연쇄에 근거를 둔 방법들에서 band 강도는 최소한 band 위치에서 얻는 것만큼은 연관성에 대한 정보를 제공할 수도 있으므로 양자에 근거를 둔 방법이 선호될 수도 있다[4].

서로 다른 유사성 계수들이 나왔다고 주어진 유전자 자리(loci)에서 대립 유전자를 꼭 동정해야 하는 것은 아니다. 이런 이유로 교차-확인법이 적용될 때 그들이 선호된다. 왜냐 하면 교차-확인법에서는 한 개 이상의 방법들이 우세한 표지자들에 의지하여 분석하기 때문이다. 또한 유전적 유사성은 상호우성(codominant) 표지자의 빈도에 근거를 둔 유전적 거리로부터 평가할 수도 있다[20].

(2) 계통수의 생성

S_{ABS} 를 계산한 후 역학 연구에서 분리주 각 쌍간에 생성된 바탕질(matrix) 값을 얻게 된다. 이 값들은 계통수를 생성하는 근거가 된다. 분리주들을 계통 나무(tree)로 연결하는 가장 흔히 쓰이는 방법은 Rohlf가 처음 사용한 unweighted pair group method using arithmetic averages (UPGMA) 이다. 이 방법은 뿌리 있는 계통 나무(rooted tree)를 생성한다[21]. 진화 시계를 나타내는 Fitch- Margoliash법도 뿌리 있는 계통수(rooted dendrogram)를 생성한다. 그에 추가되는 특징으로 유전적 거리들과 가지의 길이들 사이에 계산되는 제곱(squares)의 합계를 최소화하기 위해 집합점(node) 들을 재배열함으로써 계통 나무에 대해 선택 가능한 대안적 모양(alternative topology)을 검사하기 때문에 수행하는데 시간이 더 걸린다. 모든 균주들에 대해 공통적인 분자적 시계를 가정하지 않기 때문에 뿌리 없는 계통 나무(unrooted tree)를 생성하는 방법들도 갖추어져 있다. neighbor-joining법과 Fitch-Margoliash 법 등과 같이 진화 시계를 나타내지 않는 방법들은 뿌리 없는 계통 나무들을 생성하며 사실 더 정확한 분포도들(topographies)을 생성한다[21,22].

neighbor-joining법과 특히 UPGMA는 다른 방법들보다 신속하여 큰 실험적 검체들에 대해서는 더 선호되

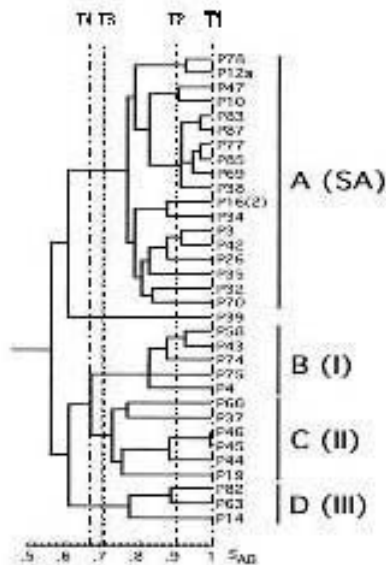


Fig. 1. Example of a dendrogram. The similarity coefficients (S_{AB} s) were generated from the patterns of 33 isolates of *C. albicans* fingerprinted with the complex Ca3 probe. The value T1 represents identicalness, at an S_{AB} of 1.00; threshold T2 demarcates highly related isolates at an S_{AB} of 0.90; threshold T3 demarcates clades at an 0.70; and the value T4 represents the average S_{AB} at 0.66. Using an arbitrary S_{AB} threshold of 0.70, four groups (A through D) were readily distinguishable in a dendrogram of these isolates. Only 1 of the 33 isolates, P39, did not fall into one of the four groups, which is an outlier (i.e., members of no group). In this example, a South African-specific clade of *C. albicans* (SA) was identified in addition to the three clades (I, II, and III) present in U.S. collections. Reprinted from reference 23.

지만 각 방법들간에 서로 장단점들이 있다. 이러한 방법들이 연관성의 아종 수준이라기 보다는 균종 수준에서 계통 나무를 만들도록 개발되었으나 아종과 균주 수준에서 해석될 수 있는 계통 나무를 생성한다는 것을 인식해야 한다. 자료들이 강건하다면 어떤 방법이 적용되었든지 상관없이 비슷한 계통수들이 얻어져야 한다. 균종내 분석을 하여 얻어진 계통수들은 전적으로 혈통을 계통발생학적으로 대변하는 자료로 여겨서는 안된다. 그보다는 균주들간의 유전적 유사성과 분기를 가시적으로 서술하고 군들과 공통의 조상으로부터 진화한 생물군인 단계통군을 동정하는데 유용할 수도 있는 실용적인 도구로 보아야 한다. 그렇게 얻어진 계통수는 각 균주가 순수하게(즉, 유전적 교환 없이) 클론성의 혈통을 통해 파생되고 homoplasy가 무시할 수 있을 만큼 미약할 때에만 진정한 계통발생을 나타낸다. 이러한 조건들은 역학적 연구에 사용되는 표지자들에 대해서는 드물게 증명된다.

UPGMA가 어떻게 수행되는가를 고려해 보면 다음과 같다. 가장 유사한 분리주들을 찾기 위해 S_{AB} 바

탕질(matrix)을 정밀 조사(scan)한다. 만약 한 군보다 더 많은 군(둘 이상)들이 동정되면 처음 나온 군을 임의로 1군(group 1)으로 삼는다. S_{AB} 축을 따라 적절한 위치에서 분리주들이 인접하게 된다. matrix를 다시 한번 정밀 조사하여 그 다음 순으로 가장 유사한 분리주나 분리주들의 군을 찾은 다음 그것을 S_{AB} matrix를 따라 첫번째 군에 연결시킨다. 이런 기능이 모든 분리주들이 계통 나무에 편입될 때까지 계속 반복된다. *C. albicans* 분리주들로 구성된 표본 계통수가 Fig. 1에 제시되었다[23]. 이 계통수에서 역치들과 경계표지들이 표시되었다. T1는 S_{AB} 가 1.00으로서 동일한 균주들을 나타내고 T2는 S_{AB} 0.90에서 고도로 연관된 분리주들의 경계선을 그어 주고 T3는 중등도로 연관된 군들을 구획해 주는데 이 예에서는 S_{AB} 가 0.70인 지점에서 구분해 주었으며 T4는 계통수에 있는 모든 균주에 대한 평균 S_{AB} 를 나타내며 이 예에서는 S_{AB} 가 0.66이다. T3 역치는 단계통군(clades)을 가장 잘 구획해 주는 역치이다. 가끔 계통수에서 역치들과 경계표지들의 선택은 다소 임의적이지만[2] 계통 나무를 해석하는데 극히 중요하다. 분리주들의 기원들과 이들 기원들에 근거한 연관성의 가정들에 대해 익숙해지게 되면 이들 값들에 대해 확신이 서게 된다.

(3) 군집 분류의 완전성(integrity)에 대한 검사 및 통계적 분석들

군집들을 정의하는 역치의 선택이 다소 임의적일 수도 있음에도 불구하고 군집의 내용이 안정되고 군집의 동정이 사실 믿을 만하다는 것을 보여줄 수 있는 다음과 같은 방법이 있다.

첫째, 계통수를 생성하는데 있어서 분리주들이 선택되는 순서가 임의 추출(randomize)될 수 있다. 이것은 UPGMA를 사용할 때 중요한 관리방법(control)이다. 그러므로 이 방법은 더 높은 순위의 군집(higher-order clusters)에서 실수할 수 있는 경향이 있다. 그러한 경우 10가지 무작위 출발(random start)을 적용하여 군집이 그대로 남아있는지 평가할 수 있다. 둘째, 모든 두 쌍간(pairwise) 비교에 대해 S_{ABS} 값의 백분율의 set를 첨가하거나 제거함으로써 "잡음(noise)"을 도입할 수 있다. 잡음은 두 번째-계층의 군집 배치(second-tier groupings)의 안정성에 영향을 줄 수 있고 그것을 평가하기 위해 사용될 수 있다[4]. 셋째, 군집의 안정성을 평가하기 위해 두 가지 독립적인 유전적 지문 식별법에 의한 군집분석의 결과를 비교할 수 있다[3,10,17].

군집 분류의 완전성(integrity)을 검사하기 위해 진화적 생물학자에 의해 가장 흔히 사용되는 방법은 boots-trapping이다[24]. 이 과정에서 임의 표본 추출(random sampling)에 의해 삭제와 복제(duplication) 과정들이 일어나는데 이 과정은 정상적으로 1,000 번씩 수행되

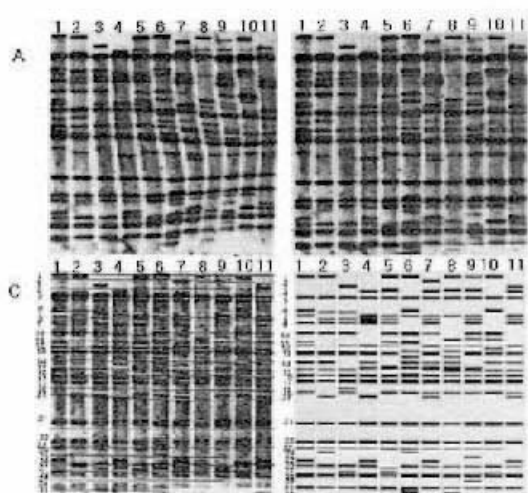


Fig. 2. Computer-assisted processing and analysis of DNA fingerprint patterns generated with the complex Ca3 probe of 11 isolates of *C. albicans*. (A) Original digitized image with inherent distortions; (B) straightened patterns; (C) correlations of bands with universal standards; (D) model generated from data. Adapted from reference 13.

어 모든 계통수에 조화되는 계통 나무(consensus tree)를 생성한다. "majority rule consensus" 연산법이 각 집합점(node)의 발생의 백분율을 계산하는데 사용된다. 백분율이 80 % 이상이면 안정된 집합점을 나타낸다.

두 번째로 흔한 방법인 jackknifing에서는 수집군의 부분집합(subsets)을 약 100 번쯤 무작위로 선택하여 공감적인 계통수를 생성하고 각 집합점에서 발생의 백분율을 계산한다[25]. 그러나 이러한 방법들은 동정된 유전자 자리의 대립유전자들을 나타내는 자료들을 위한 DNA 지문 식별법을 위해 개발되었지 RFLP나 RAPD에 의해 생성되는 양상과 같은 자료들을 위해 개발된 것이 아니다. 그래서 이들은 일반적인 역학적 연구들에서는 흔히 사용되지 않는다. 그러나 각 시동체를 독립된 유전자로 여긴다면 이들은 RAPD 자료에 사용될 수 있다.

상이한 분리주 집단들의 상관성을 비교하는데 있어서 서로 다른 집단들의 평균 S_{ABS} 를 비교하는 것이 유익할 때도 가끔 있지만 전 집단의 평균 SAB에서 얻는 정보가 많지 않을 때가 흔하다. 평균 SAB가 유의하게 다른지 여부를 Student t test로 평가할 수 있는데 이렇게 Student t test를 적용할 때 대개 확률 값(probability value)이 0.05 (5 %) 이하이면 유의성을 나타낸다고 본다.

6. 유전적 지문 식별법에서 컴퓨터의 역할

더 많은 자료를 취급하거나 정량적 비교가 필요할 때는 컴퓨터 프로그램을 동원해야 한다. 군집분석의

거의 모든 단계를 보조하기 위한 컴퓨터 프로그램이 나와 있다. 복잡한 양상을 생성하는 지문 식별법에 대해서는 시스템이 자동적으로 원래의 상을 획득하여 사용자의 도움을 받아 뒤틀린 부분(distortion)을 제거하고 lane을 동정하며 보편적인 표준에 대해 표준화(normalize)하여 band를 동정하고 band의 강도를 측정한다. 유사도 계수(similarity coefficients)를 계산하고 계통수를 생성하고 군집을 동정한 후 통계적 검사를 시행한다[2]. 컴퓨터 시스템은 후향적 연구와 미래에 새로 수집된 자료들과 비교하기 위한 모든 수준의 자료를 보관한다. 자동화된 유전적 지문 식별법은 지문 식별 결과의 측면(즉 특정한 band 길이)이나 환자의 특징(즉 HIV 양성) 또는 감염균의 특징(약제 감수성) 중 어느 것이나 상관없이 모든 면에 근거를 두어 탐색을 수행할 수 있다. 일반적인 software를 프로그램이라고 부른다.

(1) 양상을 숫자로 표시하기(digitizing the pattern)

올바른 형식(bitmap or tiff)으로 상을 저장하는 gray-scale scanning을 지원하는 software를 가진 scanner를 사용하여 자가방사성도나 그에 상응하는 것들을 컴퓨터 hard drive에 숫자로 표시할 수 있다(Fig. 2A). scanner는 합리적인 범위의 grayscales (256 grayscales)에 민감해야만 한다.

(2) 국부적(local) 표준

자동화된 분석을 하고 서로 다른 gels에서 얻어진 양상들을 비교하기 위해서는 양상들을 반듯하게 정돈(straighten)해야 한다. 그러기 위해 필요한 경계 표지(landmark) band로서 표준을 test 검체들 양쪽의 옆 통로(lanes)에 set로 세우거나, 검사하는 검체들에 추가하거나, 단형성(monomorphic) bands를 검사하는 양상의 구성 요소들로 넣을 수 있다. 세 경우 모두에서 표준들은 검사 양상들에 있는 전체 분자량의 범위에 미치는 bands를 포함시켜야 한다.

(3) 전체적인(global) 표준

여러 가지 gel의 양상들을 비교하는 연구들에서는 각 검사 양상은 자료의 전체적인 표준에 대해 표준화(normalize)해야만 한다. 전체적인 표준은 국부적 표준의 band들을 공유해야 한다.

(4) 최초의 상 처리 과정

양상을 왜곡시키는 허상(artifact)을 피할 수 있는 처리 체계(processing stems)의 필요성을 살펴보면 다음 두 가지를 들 수 있다. 첫째, 불균일한 중합체 형성(polymerization)이나 전기영동 때문에 전체 gel 양상에 웃는 모양이나 찌그러짐 경사진 모양 및 다른 선상 또는 비선상의 왜곡들이 포함될 수 있다. 양상을 직선화시키는 과정을 '탈왜곡(unwarping)'이라고 하는데 이 과

정에는 국부적 표준이 꼭 필요하다. 수평적으로 왜곡된 공통된 band들 사이에 수평선이 그어지고 수직적 왜곡들의 공간에 수직선이 그어진다. 한번 그려지고 수정되면 프로그램이 처리된 양상을 재구성한다(Fig. 2B). 둘째, gel이나 검체가 불균일하게 채워졌다면 단일 lane에 있는 양상은 강화되거나 탈강화될 수 있다.

(5) 자동화된 lane 및 band 검출

프로그램이 정밀 조사(scan)하여 자동적으로 lane들을 검출하고 그들 사이에 간격을 주어 band 인지를 쉽게 한다. 이 시점에서 lane이 밀리거나 당겨지거나 압축되었는지(sliding, stretching 및 compressing) 여부를 포함한 최종적인 band 정렬(alignment)에 대해 프로그램이 살펴보고 편집할 수 있도록 해야 한다. 만약 lane들에 gel이나 검체가 불균일하게 채워졌다면 같은 분자량을 가진 band들도 서로 약간씩 다른 속도로 이동될 수 있기 때문에 이 과정이 필요할 수 있다. 그 후 프로그램은 배경(background)의 강도를 제거하면서 lane을 따라 grayscales의 음영계측(densitometric) 분석을 수행해야 한다. 최고점(peaks)에서의 적분(integration)이나 강도 역치(threshold)를 band 동정을 위해 사용할 수 있다. band 강도는 0-5까지 category를 나누거나 직접 측정하거나 표준화 곡선(normalization curve)에 따라 변환할 수 있다. 그 후 gel의 모형을 band의 위치와 강도로 부터 생성할 수 있다(Fig. 2D).

(6) 전체적인 표준에 대해 눈금을 조정(calibrating)하고 연관시키는(linking) 과정

국부적 표준은 전체적인 표준에 맞춰 눈금이 조정되며, 표준 band의 분자량을 제공해 준다(Fig. 2C). 그런 다음 프로그램은 검체 lane들의 band들을 국부적 표준의 bands에 대해 자동적으로 연결시킨다. 사용자는 같은 크기의 band들을 고려하여 관용의 정도를 지정해야만 한다. 표준에 상응하는 것이 없는 bands도 연관시킬 수 있다. 만약 필요하다면 전체적인 표준에 새로운 band를 추가하는 선택기능(option)이 있어야만 한다.

(7) 기본적 자료 파일의 생성

이전의 자료들이 band 수와 분자량 또는 각 band의 화소(pixel) 거리 등을 표지된 각 분리주의 수직축을 따라서 내림차순으로 나열하는 원문 지도(text map)를 제공한다. 각각의 위치에서 위치 단독으로 또는 band 강도에 대해 0 또는 1의 이진수 값의 자료(binary datum)들이 나열된다. 만약 MLEE나 RAPD, 또는 다른 다중 유전자 자리(multilocus) 분석법을 사용하여 이진수 값의 자료가 수집되었다면 이 시점에서 컴퓨터 프로그램에 자료 file을 수동적으로 생성할 수 있거나 software가 자료를 자동적으로 합병할 수 있다.

(8) 유사성 계수(similarity coefficient)를 계산하고 계통수를 생성하는 과정

프로그램은 사용자가 S_{ABS} 를 계산할 공식을 선택할 수 있도록 해주어야 한다. 특정한 공식이 선택되면 프로그램은 분리주의 모든 쌍간에 S_{ABS} 의 matrix를 생성한다. 프로그램은 S_{AB} matrix로부터 계통수를 생성할 여러 종류의 연산법과 함께 역치에 의해 군들을 분리하고 군집의 안정성을 검사하며 통계적 분석을 할 software를 제공해야만 한다.

(9) 다른 유용한 기능들

프로그램은 접근 가능한 저장능과 비교능을 제공하여 유전적으로 지문 식별된 새 set의 분리주들은 어느 것이나 이전에 지문 식별된 자료내의 분리주들과 비교할 수 있어야 한다. 환자와 분리주 특성들과 같은 다른 자료 요소들에 근거를 둔 검색(search) 들을 허용해야 한다. 전자는 환자의 질병 상태와 치료, 지리적 위치, 분리주의 해부학적 근원과 나이, 성별 체중과 선행 조건들을 포함한다. 후자는 분리주의 성장 특성들과 당 동화 양상, 항원성, 약제 감수성, 변환되는 종목(switching repertoire)들과 특히 독성과 병원성에 관련된 다른 표현형적 특성들을 포함한다. 또한 특정 MLEE 대립 형질(allele)이나 RFLP 분절의 유무와 같은 특정한 자료의 특징들에 대한 유전적 지문 식별의 자료들을 축적하고 개발할 수 있어야 한다[4].

7. 수집

유전적 지문 식별 검사를 수행하는데 시간이 오래 걸리기 때문에 처음에 균을 수집할 때 여러 가지를 주의 깊게 고려해야만 한다. 그러나 역학적 연구의 이 첫 단계가 가장 빈약하게 생각되는 경우가 흔하다. 이 문제의 일부는 의심할 여지없이 병원균들의 집락 형성(colonizing) 집단들은 유전적으로 균일하다는 가정에서 기인한다. 흥미롭게도 많은 경우들에서, 특히 면역능이 손상된 환자에서는 감염 집단들은 다수의 균종이거나 같은 균종이지만, 다수의 균주들 또는 상이한 아균주(diverging substrains)들로 구성될 수도 있다는 것이 점차 드러나고 있다. 재발성 감염들이나 편리공생균(commensal)들에서 균종들의 불균일성이 관찰되어져 왔다[26,27]. 각 경우들에서 균을 처음 수집할 때 주의를 기울였음에도 불구하고 이질성이 나타났다.

대부분의 표준적인 수집 방법에서 원래의 검체들은 흡인물, 조직, 혈액, 다른 체액이나 면봉검체들이다. 이런 물질들은 혐기성이나 호기성 배양을 위해 미생물 검사실에 운반되었을 수도 있다. 어떤 경우들에서는 미생물은 물질들로부터 유리되어야만 한다. 혈액 세포들은 용해시키고 조직들은 잘게 썰어줄 수도 있으며 액체들은 여과되거나 원침될 수도 있다. 문제는 클론형성능(clonality)이다. 만약 채집된 집락 형성(colonizing) 집단이 균일하다면(즉, 하나의 유전적으로 균일

Table 1. 분자적 형별 검사시 고려하고 주의해야 할 점

1. 균이 있어야 검사가 가능하다. 보관한 분리주들을 버리기는 쉬우나 그것을 쓰레기통에서 회수할 수는 없다.
2. 쉽게 이용할 수 있고 비용이 저렴한 기법을 먼저 이용한다. 초기의 결과와 역학적 자료상 좀더 연구가 필요하다고 판단될 때에만 더 복잡한 검사들을 하는 단계로 진행해 간다.
3. 형별 검사 결과는 대개 균주들이 동일한지 또는 연관되었는지에 대한 판단을 나타내주는 것이지 양성이나 음성 결과를 내주는 것이 아니다.
4. 여러 가지 형별 검사법들은 서로 다른 특성들을 평가하기 때문에 상이한 형별 검사법으로 얻어진 결과가 항상 완전하게 연관되는 것은 아니다. 그것을 나쁘다고 보기 보다는 그것이 삶 자체의 모습이라고 생각하면 된다.
5. 역학적으로 연관되지 않았음에도 불구하고 모든 균주들이 동일한 결과를 보이거나, 그들이 모두 상이하다는 결과가 나오면 그 형별 분석법 결과의 신뢰성을 의심해 본다.
6. 한 가지 이상의 형별 분석법에 의해 상이한 것으로 나타난 분리주들은 그들이 역학적으로 연관된 것처럼 보이더라도 동일한 균주를 나타내는 것이 아닐 가능성이 높다.
7. 균들은 불안정한 유전적 속성들(즉 서로 다른 유전자들이나 plasmid 내용의 표현 등)을 가질 수도 있다. 형별 분석법을 감정할 때는 평가하려는 대상 특성이 안정한가 여부를 숙고해 본다.
8. 어떤 균주들(예: methicillin-resistant *S. aureus*)은 유전적 다양성이 제한되어 있다. 그래서 역학적으로 무관한 균주들이 형별 기법에 상관없이 동일하게 나타나는 경향이 있어서 돌발감염을 나타내는 것으로 잘못 해석될 수도 있다.
9. 일반적으로 한 균종(예: *E. coli*)에 유효한 방법은 연관된 다른 대상균(예: 다른 *Enterobacteriaceae*)에 대해서도 유효하게 작용할 가능성이 많다.
10. 한 방법이 항상 최선책인 것은 아니다. 어떤 것이 최상의 방법인지는 검사한 균주와 조사하는 역학적 문제나 연구 계획에 따라 다르다.
11. 어떤 실험실에서는 검사료를 받고 분자적 형별검사를 해준다. 또한 어떤 연구자들은 그들이 논문의 공동저자가 될 수 있다면 균주들을 검사해줄 의향이 있다.

Adapted from reference 31, 32.

한 균주들로 구성되어 있으면) 문제가 없다. 그러나 그 집단이 불균일하다면(다수의 균종이나 균주들 또는 아균주들로 구성되어 있다면) 몇 가지 문제되는 시나리오가 일어난다. 첫째, 유전적 지문 식별 검사를 할 시기에 배양이 유전적으로 불균일하면 혼합된 지문이 나타날 것이다. 둘째, 만약 일차 배양이 혼합된 균주들과 아균주(substrain)들을 함유하고 있고 그들이 함께 배양되었다면 사용된 배양 조건에서 더 빨리 자랄 수 있는 균주들이나 아균주들이 더 증균될 것이다. 마지막으로, 혼합된 균주들과 아균주를 함유한 일차 배양이 배양되어 오직 한 클론만 분석에 사용되었다면 문제가 일어난다. 단일한 클론은 감염을 일으킨 균집단을 대표하지 않을 수 있다. 이 문제들에 대한 해결책은 명백하다. 첫째, 일차적 검체를 클론성으로 배지에 심어 그 농도를 평가하거나 연속 희석을 수행하거나 대략적인 밀도를 아는데 사용할 수 있게 한다. 둘째, 지시계 배지(indicator agar)를 균종의 초기 검색을 위해 사용할 수 있다. 셋째, 한 개 이상의 집락을 유전적 지문 식별법을 위해 딸 수 있다. 채취되고 분석된 집락의 수는 수집균의 크기와 연구 목적에 따라 달라질 것이다.

검체의 균일성에 덧붙여 일부의 경우들에서는 공간과 시간에 관련된 두 번째 문제가 일어난다. 편리공생

(commensalism)에 체위 특이성(body location specificity)과 균주들의 지리적 특이성(geographical specificity)과 급속한 미세진화(microevolution)가 있다는 증거가 축적되고 있다[2, 28, 29]. *C. albicans* 같은 병원균은 여러 장소에 집락화할 수 있으므로 몸의 어떤 위치에서 검체가 채취되었는지를 고려하는 것이 현명하다[30]. 만약 어떤 사람이 특정한 질환에 대한 균주의 특이성을 검사하려고 한다면 질환에 따른 차이보다 지리적 차이가 더 가치가 있을 수 있다는 점을 인식하는 것도 중요하다. 예를 들어 특정한 질환을 가진 환자들 신체의 한 장소에서 나온 분리주들은 유전적으로 비슷하지만, 동일 질환이 있는 환자들 신체의 서로 다른 장소에서 수집한 분리주들을 비교할 때 그들이 유전적으로 서로 같지 않다는 것을 발견할 수도 있다[29]. 그러므로 어떤 장소에서 균을 수집할지 그 지리적 경계들을 제한하거나 독특(distinct)한 다수의 집단(population)들을 선택할 것을 고려해야 한다. 끝으로, 미세진화의 속도와 검체 채취시기를 고려해야만 한다. 아균주 특수화(substrain specialization) 과정을 분석하고 있으면 미세진화의 직선적 속도(linear rate)가 서로 다른 질환 상태들과 항생제 내성에 관련된 아균주의 차이들보다 더 중요할 수 있다는 것을 볼 수 있다. 그러므로 수집의 시간 간격(window)을 제한하는 것을 고려해야 한다.

결 론

분자적 형별법은 돌발 감염과 토착성 병원 감염을 조사하는데 있어서 기본적인 역학 및 미생물학적 방법이 할 수 있는 것보다 더 신속하고 철저한 조사를 할 수 있게 해준다. 분자적 형별법이 병원 역학과 감염 관리를 위한 귀중한 도구이기 때문에 한 개 이상의 확립된 분자적 형별법을 익숙하게 사용할 수 있다면 바람직할 것이다. 또한 분자적 형별 검사시 고려하고 주의해야 할 점에 대해 전문가가 조언한 내용에 귀를 기울이는 것도 도움이 될 것이다(Table 1)[31,32].

미생물의 지문 식별법에 사용된 방법들에 대한 선행 연구들이 모든 병원균의 범주들(세균, 진균, 기생충)이나 심지어는 한 범주 내의 모든 균종들에 대해서도 우세한(dominant) 방법이 없다는 것을 나타냈다. 예를 들어 세균에서는 *Streptococcus pneumoniae*와 *Staphylococcus aureus*에 대해 RFLP- PFGE가 가장 흔히 쓰이는 방법이다. *Mycobacterium tuberculosis*의 경우에는 탐색자를 사용하는 RFLP가 가장 흔히 쓰이는 방법이다. 비슷한 다양성(variability)이 진균과 기생충 균종에 대해서도 적용된다. 이러한 다양성의 이유는 두 가지이다. 첫째, 서로 다른 균종들의 유전체의 차이가 일부 경우들에서 서로 다른 방법들을 정당화한다. 둘째, 서로 다른 수준의 연관성을 측정하는데 서로 다른 방법들이 효과적이다. 셋째, 과학적이지는 않을지 모르나 그 영역에서 과학자들의 공감에 의해 그리고 처음 적용된 방법들의 이력(history) 때문에 특정 균종들에 대해 특정 방법이 지속적으로 사용될 수도 있다. 넷째, 어떤 경우들에서는 편의성에 따라 선택할 수도 있다.

끝으로 지문 식별법을 사용하려고 할 때 유념해야 할 사항으로는 제기된 질문에 답할 수 있는 지문 식별법을 선택해야 한다는 점과 필요한 유전적 관련성의 수준에서 효율이 입증된 방법인지를 확인하라는 점, 컴퓨터-지원된 체계를 사용하고 자료 저장시 후향적 분석을 할 수 있으면서 새 자료와 비교할 수 있는 형식으로 저장하라는 점과 검체를 수집하는데 사용하는 방법을 주의 깊게 고려하라는 점을 강조할 수 있다.

참 고 문 헌

- Arbeit RD. Laboratory procedures for the epidemiologic analysis of microorganisms. In : Murray PR, Baron EJ, et al. eds. Manual of Clinical Microbiology. 7th ed. Washington, D.C. :American society for microbiology, 1999: 116-7.
- Soll DR. The ins and outs of DNA fingerprinting the infectious fungi. Clin Microbiol Rev 2000 ;13: 332-70.
- Tibayrenc M, Neubauer K, Barnabe C, Guerrini F, Skarecky D, Ayala FJ. Genetic characterization of six parasitic protozoa: parity between random-primer DNA typing and multilocus enzyme electrophoresis. Proc Natl Acad Sci USA 1993;90:1335-9.
- Soll DR, Lockhart SR, et al. Laboratory procedures for the epidemiologic analysis of microorganisms. In : Murray PR, Baron EJ, et al. eds. Manual of Clinical Microbiology. 8th ed. Washington, D.C. : American society for microbiology, 2003;139-61.
- Sambrook J, Russell DW, et al. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, 2001:5.55-5.90.
- Welsch, J, McClelland. M, Fingerprinting genomes using PCR with arbitray primers. Nucleic Acids Res 1990;18:7213-24.
- Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 1990;18:6531-5.
- Enger L, Joly S, Pujol C, Simonson P, Pfaller M, Soll DR. Cloning and characterization of a complex DNA fingerprinting probe for *Candida parapsilosis*. J Clin Microbiol 2001;39:658-69.
- Lockhart SR, Pujol C, Joly S, Soll DR. Development and use of complex probes for DNA fingerprinting the infectious fungi. J Med Mycol 2001; 39:1-8.
- Pujol C, Joly S, Lockhart SR, Noel S, Tibayrenc M, Soll DR. Parity among the randomly amplified polymorphic DNA method, multilocus enzyme electrophoresis, and Southern blot hybridization with the moderately repetitive DNA probe Ca3 for fingerprinting *Candida albicans*. J Clin Microbiol 1997; 35:2348-58.
- Carle GF, Olson MV. An electrophoretic karyotype for yeast. Proc Natl Acad Sci USA 1985;82:3756-60
- Sangeorzan JA, Zervos MJ, Donabedian S, Kauffman CA. Validity of contour-clamped homogeneous electric field electrophoresis as a typing system for *Candida albicans*. Mycoses 1995;38:29-36.
- Soll DR. High-frequency switching in *Candida albicans*. Clin Microbiol Rev 1992;5:183-203.
- Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci USA 1998;95:3140-5.
- Dingle KE, Colles FM, Wareing DR, Ure R, Fox AJ, Bolton FE, et al. Multilocus sequence typing system for *Campylobacter jejuni*. J Clin Microbiol

- 2001;39:14-23.
16. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J Clin Microbiol 2000;38:1008-15.
 17. Kremer K, van Soolingen D, Frothingham R, Haas WH, Hermans PW, Martin C, et al. Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. J Clin Microbiol 1999;37:2607-18.
 18. Tibayrenc M. Population genetics and strain typing of microorganisms: how to detect departures from panmixia without individualizing alleles and loci. C R Acad Sci III 1995;318:135-9.
 19. Tibayrenc M. Towards a unified evolutionary genetics of microorganisms. Annu Rev Microbiol 1996;50:401-29.
 20. Felsenstein J. Distance methods for inferring phylogenies; a justification. Evolution 1984;38:16-24.
 21. Fitch, WM, Margoliash. E. Construction of phylogenetic trees. Science. 1967;155:279-84.
 22. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406-25.
 23. Blignaut E, Pujol C, Lockhart S, Joly S, Soll DR. Ca3 fingerprinting of *Candida albicans* isolates from human immunodeficiency virus-positive and healthy individuals reveals a new clade in South Africa. J Clin Microbiol 2002;40:826-36.
 24. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985;39:783-91.
 25. Wu, CFJ. Jankknife, bootstrap and other resampling plants in regression analysis. Ann. Stat 1986;14:1261-95.
 26. Lockhart SR, Joly S, Vargas K, Swails-Wenger J, Enger L, Soll DR. Natural defenses against *Candida* colonization breakdown in the oral cavities of the elderly. J Dent Res. 1999;78:857-68.
 27. Vargas K, Messer SA, Pfaller M, Lockhart SR, Stapleton JT, Hellstein J, et al. Elevated phenotypic switching and drug resistance of *Candida albicans* from human immunodeficiency virus-positive individuals prior to first thrush episode. J Clin Microbiol 2000;38:3595-607.
 28. Lockhart SR, Fritch JJ, Meier AS, Schroppel K, Srikantha T, Galask R, et al. Colonizing populations of *Candida albicans* are clonal in origin but undergo microevolution through C1 fragment reorganization as demonstrated by DNA fingerprinting and C1 sequencing. J Clin Microbiol 1995;33:1501-9.
 29. Pfaller MA, Lockhart SR, Pujol C, Swails-Wenger JA, Messer SA, Edmond MB, et al. Hospital specificity, region specificity, and fluconazole resistance of *Candida albicans* bloodstream isolates. J Clin Microbiol 1998;36:1518-29.
 30. Soll DR, Galask R, Schmid J, Hanna C, Mac K, Morrow B. Genetic dissimilarity of commensal strains of *Candida* spp. carried in different anatomical locations of the same healthy women. J Clin Microbiol 1991;29:1702-10.
 31. Weber S, Pfaller MA, Herwaldt LA. Role of molecular epidemiology in infection control. Infect Dis Clin North Am 1997;11:257-78.
 32. Maslow J, Mulligan ME. Epidemiologic typing systems. Infect Control Hosp Epidemiol 1996;17:595-604.