

균주간 유전체 지문 비교분석에서 유전형질 일치성의 확률적 한계 분석

조영근

경성대학교 생물학과 및 기초과학연구소

Analysis of Probabilistic Limits of Trait Identity in Inter-Strain Comparison of Genomic Fingerprints of Bacteria

Young-Gun Zo

Department of Biology, Kyungsung University, Busan 608-736, Republic of Korea

(Received June 20, 2011 / Accepted July 18, 2011)

Genomic fingerprinting methods are useful in determining relatedness among bacterial strains. However, random coincidences in sizes of two DNA fragments in two different fingerprints may occur, resulting in erroneous interpretation of relatedness between two bacterial genomes. In this study, I estimated the probability of occurrence of DNA bands of identical size in fingerprints of two unrelated genomes, so that the significance of fingerprint-based estimation of genome relatedness could be analyzed. The probability could be estimated as outputs of a function formulated with the three parameters: the numbers of observed fragments, all possible sizes of fragments and observed fragments common in a given pair of fingerprints. The parameter most instrumental to significance of relatedness estimation was the number of all possible sizes of fragments. To keep the number of coincidentally-common size of fragments below 10, about 200 fragments should be distinguishable in the fingerprints.

Keywords: gel resolution, genomic fingerprinting, genome relatedness, probability of coincidence, uniform distribution

세균 유전체 지문분석법은, 균주의 일치성(clonal identity 또는 clonality)을 판정하거나, 균주 유전체간 유전형질 분포의 유사성에 근거하여 유전체간 친연성(relatedness)을 측정하는데 유용한 방법이다(6). 다양한 유전체 지문분석법들 중, 핵산 중합효소 연쇄반응(polymerase chain reaction; PCR) 및 DNA 제한효소를 사용하여 세균 유전체의 일부 정보를 추출하여 다양한 크기의 DNA 단편으로 구성된 지문을 생성하는 방법은, 비교적 단순한 절차에 따라 신속하게 재현성 있는 결과를 제공하는 저비용의 방법으로, 널리 이용되고 있다(1, 2, 6, 7, 9). 유전체 전체의 제한효소의 절편을 생성하는 방법으로, pulsed-field gel electrophoresis (PFGE)법이 표준적으로 사용되며(3), 균주간 변이도가 높은 유전체 상의 일부 영역을 PCR로 선별 증폭하고 제한효소 절편의 다양성을 비교하는 방법도 제한적으로 사용된다. 반대로, 제한효소 절편을 선택적으로 증폭하여 전개하는 amplified fragment length polymorphism (AFLP)법

또한 표준적으로 사용된다. 한편, PCR 과정만을 수행하는 방법은 절차가 간단하고 최소한의 장비만을 사용하여 수행할 수 있는 방법으로, random amplified polymorphic DNA (RAPD), arbitrarily primed (AP)-PCR, repetitive (rep)-PCR 등이 있다. 이와 같이 지문분석법에는 다양한 형식이 존재하지만, DNA 단편들을 각 단편들의 크기에 따라 분류하여 지문이 구성되며, 지문간 크기가 일치하는 단편의 수량에 따라, 유전체 상 유전형질의 일치성 또는 지문간 유사성이 계산하는 과정이 공통된 결과 해석 방법이다.

그러나, 지문분석법은 두 지문간 비교에서 단편의 크기의 일치성만을 근거로 자료를 해석함에 있어 기인한 친연성 과대평가의 문제점을 안고 있다. 크기가 일치하는 단편들이 동일한 유전형질에 관여하는 단편에 의해 만들어진 것인지, 아니면 다른 유전형질들을 만드는 단편들이 우연히 같은 크기를 갖게 되었는지 구분되지 않아, 형질의 일치성을 여부를 판정하는 단계에서 우연한 단편 크기의 일치를 형질 일치도로 평가하는 오류가 존재하는 것이다. 이는 실제 인과관계가 없는 두 정보

* For correspondence. E-mail: zoyful@gmail.com; Tel.: +82-51-663-4643; Fax: +82-51-627-4645

를 인과관계가 있는 허구적 상관(suprious correlation)으로 해석하는 오류의 한 전형으로 볼 수 있다(8). 특히, 친연성이 낮은 두 균주의 지문 비교에서, 이 오류는 지문의 비교해석에 큰 영향을 줄 수 있다. 친연성이 매우 높은 균주들은 대부분의 유전형질들이 비교되는 균주간에 공통되어, 같은 크기의 단편들이 같은 형질에서 만들어질 가능성이 높다. 하지만, 친연성이 낮은 두 균주를 비교하는 경우, 균주간 공통된 유전형질의 수가 적어, 같은 크기의 단편들이 다른 형질에서 유래하며 우연히 같은 크기를 보여, 허구적인 유전형질 일치성 오류가 친연성 판정에 기여하는 정도가 높아지기 때문이다. 따라서, DNA 단편의 크기만으로 유전형질의 상동성을 판정하고, 균주 유전체간 친연성을 평가하는 분석법을 사용함에 있어, 허구적 연관성에 의한 오류의 정도를 산정하고, 이를 반영하는 해석의 기법이 필요하다. 본 연구는 우연의 결과로 동일 크기의 단편밴드가 만들어 지는 확률을 정량하고, 허구적 상관의 영향을 최소화하며 균주 유전체간 친연성을 해석할 수 있는 통계적 유의성의 범위를 판정하였다.

DNA지문 자료의 특성과 매개변수

DNA 지문 자료는 아가로스 젤(agarose gel) 등의 핵산 이동 매질 안에서 이동하면서, 그 크기에 따라 분리된 DNA 단편들의 존재 유무로 구성된다. 매질 상의 이동거리가 직접적인 측정값이고, 크기가 알려진 표준크기의 단편들(size marker)의 이동거리와 알려진 단편 크기의 비교를 통해 만들어진 표준곡선에 따라, 측정된 이동거리는 DNA 단편의 크기(bp)로 변환된다. 아가로스 젤 등의 매질에서는 형광발색으로 시각화 또는 광량을 측정하여 존재유무를 표지하며, 특정 크기의 DNA 단편 밴드(band)의 유무가 지문의 자료의 구성요소가 된다. 따라서, 한 균주의 유전체 지문은 사용한 매질에서 사람의 시각 또는 광학적 기구에 의해 시각화하여 얻을 수 있는 DNA 단편 밴드들의 크기의 집합으로 표현할 수 있다.

사용한 매질에서 얻을 수 있는 DNA 단편밴드들의 모든 관찰 가능한 크기의 집합을 유전체 지문을 구성하는 DNA 단편 밴드들의 크기들의 전체집합(U)으로 생각할 수 있다. n 번째 크기의 단편밴드를 β_n 으로 표현하기로 하고, 가능한 전체 단편의 수가 i 개일 때, $U = \{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \dots, \beta_i\}$ 라는 집합으로 표현된다. 이때, 분석 과정을 단순화하기 위해, 우선 가장 단순한 밴드 분포 모델을 설정하고, 복잡한 유형의 밴드 분포 유형의 영향을 단계적으로 고려해 나가기로 한다. 가장 단순한 단편밴드의 분포 유형은, 모든 β_n 의 유전체상 존재 빈도가 모두 동일한, 균질분포(uniform distribution)를 갖는 경우이다. 대부분의 유전형질은 세균 유전체에 하나의 유전자 좌(genetic locus)를 가지며, 적용하는 지문 분석법이 각 형질에 무작위적으로 결정되는 크기의 단편밴드를 만들어 낸다면, 이 가정은 성립되며, 대부분의 지문분석방법의 경우 성립될 수 있을 것이다.

한편, 각 균주 별로 관찰되는 유전체 지문은 U 의 원소 중 특정한 개수를 택한, U 의 한 부분 집합으로 표현될 수 있다. A와 B로 명명된 한 쌍의 균주를 대상으로 유전자 지문을 비

교한다면, 단편의 수가 i 개인 균주 A의 지문자료는 U 를 구성하는 t 개의 단편밴드 중, i 개를 선별한 부분집합(F_A)이며, 단편밴드의 수가 j 개인 균주 B의 지문자료 역시, t 개의 단편밴드 중, j 개를 별도로 선별한 부분집합(F_B)이다. 이때, 임의의 두 균주는 별도의 유전체를 가지므로, 지문자료인 i 개의 단편밴드의 선별과 j 개의 단편밴드를 선별하는 사건은 서로 독립적으로 일어난다. 따라서, t 개의 가능한 단편 중, 임의의 두 균주의 유전체 지문자료인 F_A 와 F_B 가 만들어지면서 나타날 수 있는 모든 가능한 β_n 의 조합의 수(N)는 F_A 와 F_B 각각의 지문을 만들기 위한 β_n 의 조합(combination)의 곱으로 계산된다(수식 1).

$$N = \binom{t}{i} \binom{t}{j} \tag{수식 1}$$

지문 단편 일치성과 유전형질의 일치성

친연성이 있는 균주간에는 두 유전체에 공통으로 존재하는 유전형질이 존재하고, 유전체 지문에 그 공통형질이 포함된다면, 크기가 일치하는 DNA 단편밴드를 형성하게 된다. 이런 형질의 일치성에서 비롯된 밴드의 일치성은 균주간의 친연성을 실질적으로 반영하는 경우로, 실제 밴드일치(true band-matching; TBM)라 지칭하기로 한다. 한편, 친연성이 없는 두 균주들의 유전체 지문에서 서로 다른 두 유전형질이 우연히 같은 크기의 단편밴드가 형성된다면 허구 밴드일치(false band-matching; FBM)이라 지칭하고, 균주간 친연성을 과대평가하게 되는 원인으로 규정할 수 있다. FBM은 우연의 산물로서, 지문 F_A 와 F_B 가 각각 U 의 부분집합을 형성할 때, 우연히 동일한 크기의 β_n 들을 원소로 갖게 된 경우이다.

친연성 없는 균주간 지문 일치성이 나타날 확률

친연성이 없는 두 균주의 유전체가, 각각 i 개와 j 개의 단편밴드를 원소로 하는 지문 F_A 와 F_B 를 구성할 때, 두 지문은 0에서 i 개 사이의 공통 밴드(FBM)를, 우연히, 가질 수 있다. 균주간 친연성이 없다는 점은 곧 지문으로 반영되는 유전형질들이 실질적으로 공유되지 않음을 의미하므로, F_A 와 F_B 가 각각 독립적으로 임의의 i 개와 j 개의 단편밴드를 U 에서 취해서 조합을 구성하는 경우와 같다. 우연히 만들어지는 공통 단편밴드(FBM)의 수를 k 라 한다면, k 개의 공통밴드를 갖는 F_A 와 F_B 가 임의로 형성되는 빈도(C)는, 수식 2와 같이 세 조합의 곱으로,

$$C = \binom{t}{k} \binom{t-k}{i-k} \binom{t-i}{j-k} \tag{수식 2}$$

계산된다. 이때, 첫 번째 조합은 U 에서 k 개의 임의로 선정하는 조합이고, 두 번째는 F_A 에만 있는 단편밴드를 선정하는 조합이며, 세 번째는 F_B 에만 나타나는 밴드를 선정하는 조합이다. 따라서, 친연성이 없는 두 균주의 유전체가, 형질의 공통성이 아닌, 우연에 의해 k 개의 공통밴드(FBM)를 가진 두 지문을 형성하게 될 확률(P)은 수식 3과 같은 비율로 계산된다.

$$P = \frac{C}{N} \quad \text{수식 3}$$

P 값은 0에서 1까지 범위의 값으로, 적용한 지문분석 방법의 특성인 t , 균주의 특성인 i 와 j , 그리고 두 균주의 조합에 의한 특성인 k 에 따라 결정되는 확률값으로 볼 수 있다. P 값이 0인 경우, 지문분석에서 나타난 i, j, t 의 조건에서, 친연성이 없는 두 균주가 k 개의 FBM을 만들어낼 확률이 전혀 없는 것으로 해석된다. 따라서, 이 조건의 관측된 두 유전체지문이 $k > 0$ 인 k 개의 공통밴드를 보인다면, TBM이 k 개 형성됨을 의미하여, 균주간에 친연성이 없다는 가설은 기각되고, 오히려 두 균주간에는 친연성이 있다는 결론을 내릴 수 있다. 나아가, Type I 검정 오류를 5%로 정한다면, $P < 0.05$ 인 경우, 두 지문간 k 개 또는 그 이상의 공통밴드가 관찰된다면, 95% 유의수준에서 두 균주 간에는 친연성이 있다고 결론을 내릴 수 있다. 한편 $P \geq 0.05$ 인 경우 k 개 또는 그 이하의 공통밴드가 관찰된다면, 95% 유의 수준에서 그것들이 모두 FBM인 것으로 해석되어, 친연성 부재의 가설을 수용하는 검정이 가능하다.

장편 rep-PCR법의 지문자료에 대한 적용

위 수식 3에 의한 P 값은 i, j, t 라는 세 매개변수에 대한 의존성이 있다. t 는 지문분석방법에 따라 다르며, 같은 분석법을 사용하더라도, 매 분석 시기마다 어떤 범위 안에서 다소 변화한다. 따라서, P 의 특성은 각 지문분석방법 별로 고찰해야 한다. 이를 위해, 고정된 t 값에 대하여, 각각 i 와 j 개의 밴드 수를 갖는 두 지문의 공통밴드 수 k 가 가정될 때 P 값을 계산하여, k 가 의미하는 바를 해석하였다.

Chokesajjawatee *et al.* (1)은 *Vibrio cholerae* 세균의 유전체를 대상으로, 장편(long-range) rep-PCR을 이용하여, 이전의 방법들 보다 상대적으로 크기가 큰 DNA 단편들을 포함한 유전체 지문을 생성하는 분석법을 확립하였다. 이 방법은 1% 아가로즈 젤에서 DNA 단편들을 분리하였는데, 단편밴드들은, 균주에 따라, 12-35개의 범위를 보였으며, 190-6,470 bp 크기의 밴드들을 지문자료로 사용하도록 제시되었다. 본 연구에서, 이 지문자료들을 다시 분석한 결과, 아가로즈 젤의 해상도는 젤에 따라 다소 변동을 보여, 190 bp의 밴드를 대상으로, 7-15 bp 차이가 있는 단편밴드들이 밴드들이 중첩되지 않고 확연히 구별될 수 있는 정도의 해상도를 보였다. 이는 3.7-7.9%의 해상도에 해당하며, t 값은 각 해상도의 차이를 보이는 모든 밴드 위치의 개수로 결정할 수 있다. 특정 해상도(r)만큼의 차이를 보이는 밴드크기의 총집합은 $(1+r)$ 의 공비를 갖는 등비수열로 표현가능하며, 초항을 190 bp로 하고, 6,470 bp를 넘지 않는 수열의 항의 수를 계산하여, t 값을 결정할 수 있다. 계산결과, 3.7%의 해상도를 갖는 젤은 $t=98$, 7.9%의 해상도에서는 $t=47$ 로 결정되었다. 자료 제시의 편의상 $i=j$ 라 설정하고, P 값의 계산결과, 특정 i 또는 j 에 대하여 $P \geq 0.05$ 인 하나의 k 범위(가설 기각 범위; Figs. 1A and B의 회색범위)와 $P < 0.05$ 인 두 개의 분리된 k 범위(가설수용 범위; Figs. 1A and B의 백색범위)가 관찰되었다. 가설기각 범위의 경우, k 개의 공통밴드 모두

FBM으로 해석되지만, 가설기각 범위가 두 개의 범위로 나타난 점은 각 범위에 대하여 다른 논리로 해석되어야 할 부분이다. k 가 가설수용 범위보다 큰 경우는, 공통밴드의 수가 임의적 일치성의 경우 보다 많은 상황으로, 이 범위의 k 값은 FBM에 TBM이 더해진 수를 의미한다. 반면, Figs. 1A and B에서, k 가 가설수용 범위보다 작은 가설기각의 영역이 또한 광범위하게 관찰되었는데, 이 영역은 t 개로 제한된 전체 밴드 위치 중 상당수가 i 개 또는 j 개의 밴드에 사용되어, FBM 발생 빈도가 높아, 친연성이 없는 지문 사이에도 적은 수의 공통밴드가 반드시 나타나는 경우이다. 결과적으로, 가설수용 영역의 최대 k 값 보다 많은 공통밴드를 보여야, 친연성에 의한 밴드의 일치성이 방증된다고 볼 수 있다. $t=47$ 인 경우, 가설수용 영역의 최대 k , 즉 FBM의 수는 t 의 41-80%에 달해, 1 중 40% 이상이 FBM이며, 허구적 친연성이 강하게 나타난다. 따라서, $t=47$ 인 해상도는 매우 친연관계가 높은 균주들의 친연성 여부를 판가름하는데만 사용 가능한 것으로 사료된다. $t=98$ 인 경우, 가설수용 영역의 최대 k 는 t 의 25-46%에 달해, i 개의 밴드들 중 약 절반 이상이 TBM일 정도로 친연성이 강한 균주들의 지문 해석에 사용될 수 있는 것으로 사료된다. 위의 두 해상도의 지문분석법을 허구적 상관의 오류의 큰 영향없이 사용하려면, 여러 쌍의 균주들 간의 친연성을 각 쌍에 대하여 상대적으로 비교할 때에 만 사용하여야 할 것이다.

위의 결과는, 해상도가 낮은 젤에서 친연성 없이 형성되는 공통밴드의 수가 상당한 수준으로 존재하며, 매우 높은 친연성이 있는 균주들 간에만 유의미한 해석이 가능함을 시사한다. 보다 높은 해상도의 아가로즈 젤로는, 약 2%의 해상도를 보이는 Metaphor 아가로즈 젤을 고려할 수 있다(10). 이 경우, $t=179$ 로 계산되며, 지문에서 관찰되는 밴드의 2-10개의 밴드들이 임의적 지문들 간에 공통 밴드로 나타날 수 있는 것으로 계산되었다(Fig. 1C). 가장 높은 해상도의 DNA 단편 분리는 1 bp의 단편을 구별할 수 있는 해상도로, 일반적으로 염기서열 결정을 위한 모세관전기영동(capillary electrophoresis)의 경우에 보이는 해상도이다. 이 방법을 적용할 시에는, 50 bp에서 600-900 bp 크기의 단편들만 서로 분리가 가능한 단점이 있기는 하지만, 단편의 길이가 1 bp 단위로 명백히 구분되므로, t 가 600-900의 범위를 가진다고 볼 수 있다. P 값의 계산결과, 모세관전기영동 적용시, 우연한 공통밴드의 수는 1-4개로 낮게 나타나며, k 값의 가설기각 영역을 하나만 보였다(Fig. 1D).

공통밴드에 의한 유전체지문 해석의 한계

이상의 결과에 따르면, 수식 3의 확률값은 우연히 유전체 지문 밴드가 일치하는 경우와 균주간 친연성에 근거한 형질의 일치성 때문에 지문 밴드가 일치하는 경우를 구별하여, FBM의 비중을 친연성 판정의 유의성의 척도로 사용할 수 있도록 하였다. 또한, 특정 지문분석법이 균주간 친연성을 측정하는데 사용되기 위하여 가장 중요한 방법론의 특성은 단편분리 매질의 이동 중 분리 해상도인 총 밴드의 수(t)가 200개 이상으로 많아야, FBM의 수가 10개 이하로 유지되며, TBM 존재

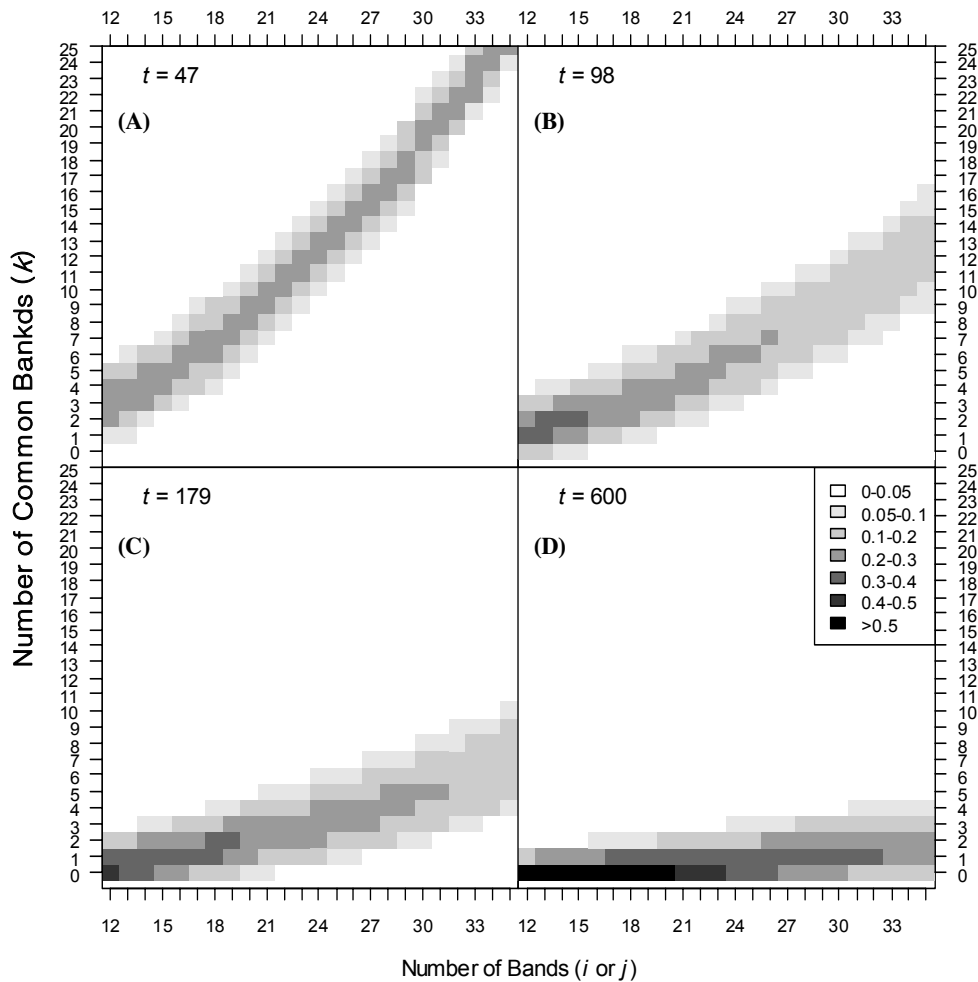


Fig. 1. Probability of occurrence of k number of bands common to two fingerprints with i and j numbers of bands when t number of bands have potential to appear in the fingerprints (inset: gray levels correspond to probability values written next to the gray-colored rectangles)

의 유의성을 높여, 친연관계의 유의성을 용이하게 판정할 수 있게 된다는 점이었다. 일반적인 아가로즈 젤을 사용하는 지문분석법의 경우 밴드의 크기를 결정할 때, 시각적 방법보다, 매우 정밀한 영상분석법에 의한 교정(calibration)을 통하여, 젤에 나타나는 밴드들을 해상도 2% 또는 200개 이상 크기단위로 분리하여 측정할 수 있도록 하여야 할 것이다. 현재 일반적으로 사용되는 유전체 지문분석법들 중 이 조건을 가장 잘 충족 시키는 방법은 AFLP 법으로 사료된다. 이는 대장균을 대상으로 한 방법론 비교 실험에서 AFLP 방법이 다른 방법에 비하여 높은 균주 변별력을 지닌 것으로 나타난 결과와 일치하는 해석이다(5). 그러나 제한효소를 사용하는 방법들이 상대적으로 복잡한 과정과 비용을 요한다는 단점이 있으므로(4), 한 단계의 PCR만을 사용하는 보다 간단한 지문분석법들을, 총 단편밴드 수가 충분히 나타날 수 있도록, 단편분리 과정 또는 자료 해석과정을 개선하는 것이 지문분석법의 유용성을 극대화하는 방안일 것이다.

더불어, 본 연구에서 사용된 밴드크기의 균질 분포를 가정한 단순한 확률 모델의 한계가 지적되어야 한다. 임의의 중심값을 갖는 불균질 분포를 갖는 두 지문에서 우연히 밴드들이 일치할 확률은, 대부분의 경우 균질분포에 비하여 낮다. 따라서 가정된 균질분포가 P 값을 과소평가하지 않으므로, 본 연구에서 사용된 확률모델과 친연성의 유효도를 평가하기 위한 공통밴드의 수는 보수적으로 결정된 것이다. 실제 유전체지문들에서는 균질분포 뿐만 아니라 다양한 불균질 분포가 예상되므로, 본 연구에서 제시된 균주간 친연성을 추론하기 위한 최소 공통밴드의 수는, 추론의 유의성을 평가하기 위한 충분조건은 아니지만, 필요조건 중의 하나로 고려되어야 한다.

적요

유전체 지문 분석법은 세균 균주간의 친연성을 판정하는데 유용하다. 그러나 친연성이 낮은 두 균주의 지문 사이에서

우연히 발생하는 DNA 단편 크기의 일치성은 유전형질의 일치성의 해석에 오차를 유발한다. 본 연구는 임의의 두 유전체 지문에서 우연히 DNA 단편의 크기가 일치할 확률을 정량하여, 유전체 지문에 근거한 친연성 해석의 유의성을 고찰하였다. 유전형질 일치성 없이 단편 크기가 일치할 확률은 관찰되는 단편의 수, 관찰 가능한 전체 단편의 수와 크기가 일치하는 단편의 수로부터 계산될 수 있는 함수로 분석되었다. 유의성에 가장 큰 영향을 미치는 독립 매개변수는 전체 단편의 수였으며, 우연한 공통 단편의 수를 10개 미만으로 유지하기 위해서는 약 200개 이상의 단편이 지문에서 관찰될 수 있어야 하는 것으로 계산되었다.

감사의 말

이 논문은 2011학년도 경성대학교 신입교수정착연구비에 의하여 연구되었음.

참고문헌

- Chokesajjawatee, N., Y.G. Zo, and R.R. Colwell. 2008. Determination of clonality and relatedness of *Vibrio cholerae* isolates by genomic fingerprinting, using long-range repetitive element sequence-based PCR. *Appl. Environ. Microbiol.* 74, 5392-5401.
- Foley, S.L., A.M. Lynne, and R. Nayak. 2009. Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. *Infect. Genet. Evol.* 9, 430-440.
- Gerner-Smidt, P., K. Hise, J. Kincaid, S. Hunter, S. Rolando, E. Hyytiä-Trees, E.M. Ribot, and B. Swaminathan. 2006. PulseNet USA: A Five-Year Update. *Foodborne Pathog. Dis.* 3, 9-19.
- Hahm, B.K., Y. Maldonado, E. Schreiber, A.K. Bhunia, and C.H. Nakatsu. 2003. Subtyping of foodborne and environmental isolates of *Escherichia coli* by multiplex-PCR, rep-PCR, PFGE, ribotyping and AFLP. *J. Microbiol. Methods* 53, 387-399.
- Jonas, D., B. Spitzmüller, K. Weist, H. Rüdén, and F.D. Daschner. 2003. Comparison of PCR-based methods for typing *Escherichia coli*. *Clin. Microbiol. Infect.* 9, 823-831.
- Li, W., D. Raoult, and P.E. Fournier. 2009. Bacterial strain typing in the genomic era. *FEMS Microbiol. Rev.* 33, 892-916.
- Loy, A. and L. Bodrossy. 2006. Highly parallel microbial diagnostics using oligonucleotide microarrays. *Clin. Chim. Acta.* 363, 106-119.
- Sapsford, R. and V. Jupp. 2006. Data collection and analysis, p. Pages. SAGE, London.
- Terletski, V., G.B. Michael, and S. Schwarz. 2004. Subtracted restriction fingerprinting—a new typing technique using magnetic capture of tagged restriction fragments. *FEMS Immunol. Med. Microbiol.* 41, 1-8.
- Zo, Y.G., I.N. Rivera, E. Russek-Cohen, M.S. Islam, A.K. Siddique, M. Yunus, R.B. Sack, A. Huq, and R.R. Colwell. 2002. Genomic profiles of clinical and environmental isolates of *Vibrio cholerae* O1 in cholera-endemic areas of Bangladesh. *Proc. Natl. Acad. Sci. USA* 99, 12409-12414.