

# 게놈연구 동향

박종훈 · 김명희

유전공학연구소 유전자원센터 게놈사업

게놈이란 생물의 유전형질을 결정짓는 유전정보의 총칭으로 이는 각 생물을 구성하고 있는 세포들의 생존을 위하여 요구되는 특정 구조나 활동능력을 위한 설계도라고 할 수 있다. 게놈연구는 인간, 동물, 식물, 미생물 및 곤충류를 포함하는 모든 생물을 대상으로 수행되고 있으며, 이들에 대한 유전정보를 해독하기 위한 유전자지도 작성 및 유전자 배열을 분석하려는 노력들이 이루어지고있다. 이들 연구대상을 인체로 하면 인체게놈연구, 미생물로 하면 유용미생물 게놈연구라고 할 수 있다. 그러나, 이러한 게놈연구는 *E. coli*와 같은 작은 크기의 게놈은 물론 큰 크기의 게놈(Table 1) 연구에 많은 노력과 막대한 연구비가 필요하지만 그 이상의 연구결과와 파급효과를 기대할 수 있기 때문에 대부분의 선진국에서는 게놈연구를 적극적으로 지원하고 있다. 미국은 국립보건원 (NIH), 에너지성 (DOE) 및 대학과 기업 연구소를 중심으로 유전자질환 치료와 관련된 인체 게놈을 체계적으로 연구하고 있으며 가장 많은 연구비를 투자하고있다. 독일의 경우는 국립연구소를 중심으로 산업미생물 균주, 효모, *Arabidopsis* 및 *Drosophila* 게놈을 연구하고 있으며 기술개발 및 교육에도 힘을 기울이는 실정이며, 일본의 경우는 이화학연구소, 동경대학, 오사카대학, 코오베대학 및 농업생물자원연구소 등에서 각기관별로 인체게놈, 인체 21번 염색체 집중해석, 벼(rice genome project), 고초균 게놈 등을 분담하여 집중투자 연구에 전념하고 있다. 프랑스의 경우는 CEPH 및 Human genetic center를 중심으로 사람염색체 9, 10, 11번과 MHC complex 등을 연구하고 있다. 영국 및 이태리에서도 MRC, SERC, 국립연구소 등에서 인체게놈, 밀, 미생물, 인체성염색체 및 마우스 게놈에 관심을 기울이고 있다. 최근에 미국의 국립보건원과 에너지성에서는 1990년부터 합동으로 수행

하고있는 게놈연구 5개년 계획(1991-1995)을 보완하여 새로운 5개년(1994-1998) 계획을 마련하였다. 이 계획에는 인체게놈의 지도작성 및 염기서열결정, 물리지도에 알려진 유전자들의 위치결정, 게놈연구에 필요한 기술, 모델 organism, 정보의 축적, 윤리, 법 및 사회적인 문제, 게놈연구에 관련된 교육훈련, 기술이전, 축적된 기술 및 데이터의 상호교환 등에 관한 구체적인 계획들의 변경을 포함하고있다. 본문에서는 인체게놈연구 수행을 위하여 필요로하는 몇몇 게놈연구 관련 기술과 현재 이용되고 있는 유전자지도 작성방법 등에 대하여 알아보고 향후의 게놈연구 방향에 관하여 기술하려고 한다.

## 인체게놈연구와 관련된 기술

인체게놈연구에는 많은 기술적인 어려움이 있지만 거대 염색체를 순수하게 분리하고 이에 대한 도서관을 작성하는 것이 매우 중요하다. 여기서는 게놈 연구를 위하여 가장 기본적인면서도 필수적인 몇개의 기술, 즉, 거대 DNA의 분리 및 YAC클로닝 그리고 자동 염기서열 결정기술에관하여 간략히 서술하고자 한다.

### 1. Pulsed field gel electrophoresis

게놈에 대한 물리적 분석을 하려면 전기영동을 통한 큰 크기의 DNA를 분리할 수 있어야 하며, 기존의 전기영동방법으로는 불가능하다. 이를 극복하기 위하여 개발된 pulsed field gel electrophoresis (PFGE)는 큰 크기의 DNA 분자를 agarose 겔에 통과시킬 때 주기적으로 전장의 방향을 바꾸어 주면 효모 염색체와 같은 큰 크기 (200 kbp-3 Mb)의 분자가 분리되게 고안한 실험방법으로서 정확한 메카니즘은 아직 잘 알려져 있지 않다(1,2). PFGE를 통한 거대 DNA 분자를 분리할 때 가장 중요한 점은

**Table 1.** The subjects of genome projects and their DNA contents (in base pairs)

Organism	Classification	Base pairs
<i>E. coli</i>	Bacterium	4,000,000
<i>Saccharomyces</i>	Yeast	4,000,000
<i>Arabidopsis</i>	Plant	100,000,000
<i>Caenorhabditis elegans</i>	Nematode	100,000,000
<i>Drosophila melanogaster</i>	Insect	165,000,000
<i>Mus musculus</i>	Mammal	3,000,000,000
<i>Homo sapiens</i>	Mammal	3,500,000,000

온전한 염색체를 분리하는 것이며, 이를 위하여 배양된 세포를 저 용점 agarose 겔과 혼합하여 불력을 만들고 적당한 효소를 처리한다. 분리한 염색체를 *NotI*과 *SfiI*과 같은 rare-cut 제한효소로 절단하여 PFGE 방법으로 전기영동하면 큰 크기의 절단된 DNA 단편들이 따로 나타난다. 경우에 따라서 메가 단위의 DNA 단편들은 수일 동안 전기영동하여야 분리가 가능하다(3).

일반적으로 게놈의 지도작성을 말할 때는 마커 사이의 거리를 recombination 빈도수로 측정하고 있는 유전적인 지도작성을 말하지만, 게놈의 여러 부위에서 recombination 빈도수가 다르기 때문에 염색체 상에 물리적인 거리가 다를 수가 있다. 실제적인 물리지도 작성을 위해서는 PFGE를 이용하여 마커 사이의 거리를 DNA 제한효소단편들의 크기로 결정하고 있다.

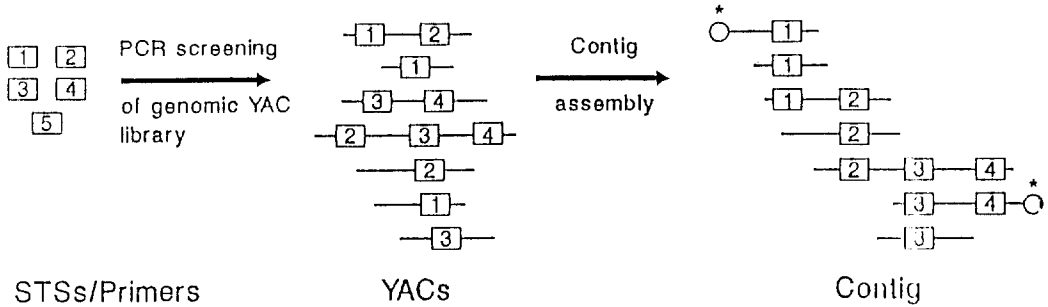
## 2. YAC(yeast artificial chromosome) 벡터의 도입

Cosmids 벡터를 이용한 유전자 지도작성 및 분석방법은 매우 노동집약적이고 복잡하기 때문에 새로운 개념의 벡터개발의 필요성이 대두되었으며, 이에 수 백 kbp를 클론할 수 있는 YAC(yeast artificial chromosome), BAC(bacterial artificial chromosome) 및 PAC(phage artificial chromosome) 벡터가 도입되었다(4, 5). 이 중에서 YAC 벡터의 경우에 가장 큰 유전자(200 kbp 이상)를 클로닝할 수가 있으며, yeast DNA를 초기화하는 염기서열인 yeast centromere, 선택 마커인 두 쌍의 telomere 및 클로닝 site가 존재한다. 이러한 YAC 벡터에 클로닝하기 위해서는 genomic DNA를 부분절단시키고 작은 크기

를 제거시킨 다음, 분리된 두 개의 YAC arm에 연결시킨 후 yeast에 transfection시킨다. 이들은 정상적인 yeast chromosome과 같이 진행되며, 만들어진 수 많은 클론으로부터 PCR screening 방법을 이용하거나 고 밀도 filter에 plating시켜서 hybridization 방법으로 특정 유전자를 얻거나 STS-contig 지도작성(Fig. 1)을 할 수 있게 된다(6, 7, 8). 확인된 YAC 클론의 외부 DNA는 위에서 설명한 PFGE 전기영동으로 쉽게 분리가 가능하며 probe로서 전체 인체 DNA 또는 특정유전자를 이용하여 Southern blotting해 보면 전기영동상에 yeast 염색체와 구분하여 특정유전자에 대한 띠를 확인할 수가 있다. 현재 YAC 클로닝에 대한 문제점은, 일반적으로 구축된 YAC 클론들의 약 10%가 1개 이상의 YAC을 포함하고 있고, 일부 library의 경우 약 60% 정도가 cocloning이 되어 있다는 점이다. 이런 점들은 긴 폭의 물리 지도작성에 많은 어려움을 유발시키므로 새로운 균주나 벡터의 개발이 진행 중에 있다.

## 3. 자동 염기서열 결정 기술

게놈연구에서 가장 중요하고 더욱 개발되어야 할 기술로는 DNA 염기서열을 결정하는 일로서 여러 가지 방법들이 수행되고 있다. 게놈전체의 염기서열을 결정할 때 속도와 비용편을 고려하여야 하는데, Multiplex 염기서열 결정법(9)과 자동염기서열 결정법(10, 11)이 이러한 점을 고려하여 나온 방법이다. Multiplex 방법은 여러 개의 시료를 혼합하여 동시에 Maxam-Gilbert 염기서열 결정 방법을 진행시킨 다음에, 전기영동 및 filter에 DNA 단편들을 transfer시킨 다음, 각 시료를 확인할 수 있는 특이한 probe를 차례로 사용하여 hybridization시켜서 염기서열을 결정하는 것이다. 또한 자동염기서열 결정 방법은 manual 방법에서의 단점인 DNA 단편들의 detection과 염기서열정보의 분석능력 부족을 보완하여 고안한 기계를 사용한 것으로서 fluorochrome을 표식자로 사용한다. 사용된 네 개의 dye는 레이저에 의하여 흥분되고 이들은 서로다른 빛을 발산하며 photomultiplier에 의해 detection된다. Dye 표식 방법에는 두가지가 있는데 하나는 일반 염기서열 primer, 또 하나는 네 개의 dideoxy chain terminator에 표식하는 것이다. 실제로 자동염기서열 결정법은 *C. elegans* 게놈의 염기서열 결정에 적용하여 160 kbp를 완성하였고 계속해서 나머지 부분을



**Fig. 1.** General strategy for constructing STS-content maps. (\*circled ends are potential candidates for the next cycle of STS development and chromosome walking).

진행하고 있다.

게놈 크기가 거대한 경우에 전체 염기서열을 결정하기 위해서는 위에서 알아본 방법으로는 한계가 있으며, 새로운 개념의 방법들이 개발되어야만 속도와 비용의 문제를 극복할 수가 있다. 예를들면, 기존의 slab젤 대신에 모세관 튜브를 이용하면 매우 높은 전장 형성이 가능하며 DNA 단편 피의 높은 해상력을 기대할 수가 있으며(12), 기존의 젤을 10  $\mu\text{m}$ 까지 얇게 만들어서 사용하는 HUGE(horizontal ultrathin gel electrophoresis) 방법, STM(scanning tunneling microscope)을 이용하여 직접 DNA 가닥의 뉴클레오타이드의 순서를 정하는 방법 등이 개발 중에 있다(13,14,15).

#### 4. Understanding of DNA sequence

게놈의 염기서열이 많이 축적되면 이들을 유용한 정보로 만드는 작업이 필요하게 된다. 따라서 염기서열이 결정되면, 이들을 모으고 체계화하고 분석하는 일들이 수행되어야 하며 databank로부터 염기서열의 유사성을 조사하는 일들이 선행되어야 한다(16). DNA 염기서열에대한 주요 databank에는 GeneBank(Los Alamos National Laboratory)와 EMBL Data library(European Molecular Biology Laboratory, Heidelberg)가 있으며 1991년 말까지 이들 기관에는 인체 염기서열의 20%에 해당하는 65,000,000 염기들이 축적되어 있다. 한편, 새로운 DNA 염기서열의 기능을 부여하는 일은 간단한 일이 아니므로, 새로운 알고리즘의 개발 및 컴퓨터 개발이 수반되어야 가능하리라고 본다.

### 인체게놈의 염색체 지도 작성과 염기서열 결정

인체게놈 연구의 1차 목표는 일련의 염색체 지도를 점진적으로 자세하게 작성하는 것이다. 이것을 위하여 일차로 염색체를 분리하여(17) 작은 단편으로 나누고, 이 작은 조각들을 클로닝하여 조작과 분석이 용이하도록 하는 일이 필요하고, 그 다음엔 이들 조각들을 원래의 염색체 위에 차례대로 정확히 표시하여 지도를 작성해야 한다.(Fig. 2) 염색체 지도 작성이 끝나면 위치가 결정된 DNA 단편들의 염기서열을 결정하게 된다.

염색체 지도란 각각의 염색체 상에서 각 유전자들의 순서와 이들 상호간의 거리 등을 나타내 주는 지표인데, 해상능력이 떨어진 genetic linkage map (DNA 지표, 즉 유전자 혹은 몇몇 염기서열이 알려진 DNA 단편들이 후대에 유전이되는 양상에 의하여 염색체 상에서의 상대적 위치가 결정된 지도)과 DNA 분자 자체의 화학적 특성에 의하여 작성된 보다 정교한 유전자 물리지도(physical map)가 있다. 현재까지 유전학자들에 의하여 염색체상에서 수천 개의 유전자 위치가 결정되었으며 최근들어서는 고 해상능을 갖는 유전자 지도 작성을 위하여 지금까지와는 다른 체계적인 새로운 연구가 필요하게 되었고 아래에서 유전자 지도 작성과 염기서열 결정을 위해 현재 사용하고 있는 몇가지 방법에 대하여 서술하고자 한다.

### 유전자 지도 작성 방법들

#### 1. Genetic Linkage Maps

Genetic Linkage Map이란 염색체상에서 특정 유전자 혹은 DNA 지표의 상대적인 위치를 표시해주는 것인데, 물리적 혹은 화학적 특성이 개인간

차이를 나타내고 유전이 되며 또 실험실에서 수월하게 그 차이점의 분석이 가능하다면 어떤 것이든 잠정적인 유전자 지표(genetic marker)가 될 수 있다. 이때 유전자 지표가 구조 유전자로서 발현이 되어 외견상 관찰이 가능할 수도 있고 혹은 단백질을 코드하고 있지는 않으나 유전되는 형태의 추적이 가능한 것 일 수도 있다. DNA염기배열에 차이가 있으면 이 차이점을 염기서열상으로 정확히 분석할 수 있으므로 매우 유용한 지표가 될 수 있다.

지도 작성에 유용한 지표는 다형성(polymorphism)을 나타내어야만 한다: 다형성이란 DNA 염기서열에 있어서의 차이를 말하는데 평균 약 300~500 염기쌍마다 1염기 꼴로 다형성이 나타나고 있다. 이때 다형성이 단백질을 코드하고 있는 엑손(exon)에서 나타나면 눈의 색깔이나 혈액형의 차이 혹은 질병에의 민감성 등과 같이 쉽게 관찰이 가능하나, 반면 단백질을 코드하고 있지 않은 인트론(intron) 내에서 일어나서 개체의 모양이나 기능에는 거의 영향을 미치지 않으나 DNA상에서 다형성의 분석이 가능하다면 지표로서의 사용이 가능하다. 예를들어 제한효소 절편 길이 다양성(RFLP: restriction fragment length polymorphism)은 DNA 절단 효소인 제한효소에 의한 절단 여부에 의하여 염기서열의 차이점 분석이 가능하고 tandem repeat sequences의 중복되는 단위염기의 수에 따라 길이의 다양성 등으로 분석이 가능하다. 인체에서의 유전자 linkage map은 2개의 지표가 어떤 빈도로 함께 유전이 되는가를 관찰함으로써 만들어 졌다.

두개의 지표가 같은 염색체상에서 근접하게 위치해 있다면 부모로부터 자식에게로 늘 함께 유전이 될 것이다. 난자와 정자가 만들어 질 때 DNA가 잘려서 다른 상동 염색체와 재조합(meiotic recombination)을 하게 되는 경우가 있는데, 이때 두 지표가 근접해 있으면 있을수록 두 지표 사이에서 재조합이 일어날 확률은 줄어들게 되고 이런 경우를 두 지표가 밀접히 연결(tightly linked)되어 있다고 한다. 그러므로 재조합이 일어나는 빈도는 두 지표 간의 거리를 측정할 수 있게 하여준다.

유전자 지도에서 지표들간의 거리는 미국인 유전학자 Thomas Hunt Morgan의 이름을 딴 centimorgan(cM)으로 나타내는데 1cM은 정자나 난자 형성에 1%의 확률로 재조합을 일으키는 거리를 말한

다. 유전적 거리인 1cM은 대략 백만 염기쌍(1 Mega base pair)에 해당하고 현재의 인체 유전자 지도의 해상 능력은 약 천만 염기쌍(10 Mb; 10 Mega base pair)정도이다.

## 2. 물리적 지도

### 1) 염색체 지도

염색체 지도에서는 유전자 즉, DNA 단편들의 염색체 상에서의 거리를 염기쌍(base pair)으로 표시한다. 유전자 지표들은 세포유전학적 염색법으로 관찰하면 염색체상의 특정 밴드에 위치하고 있음을 알 수 있는데 이는 in situ hybridization으로 확인할 수 있다. Genetic linkage map에서와 마찬가지로 완전한 개체 단위에서만 관찰할 수 있는 유전자 지표도 지표들간의 거리를 염기쌍으로 표시 한다면 그것은 염색체 물리지도라고 한다. 현재까지 가장 잘 만들어진 염색체 지도가 약 10 Mb 정도의 해상 능력 밖에 갖지 못하였으나 최근 새로운 방법(FISH; fluorescence in situ hybridization)의 개발로 약 2~5 Mb까지 해상이 가능하게 되었고, 이 방법을 조금 변형시켜 덜 꼬여있는 interphase시의 염색체를 사용할 경우 약 100,000 bp까지 해상이 가능하다.

### 2) cDNA 지도

cDNA 지도란 게놈 중 단백질을 코드하고 있는 엑손(exon)의 위치를 염색체위에 표시를 해 놓은 지도를 말하는데, cDNA는 실험실에서 분리된 mRNA를 주형(template)으로 하여 만들 수 있고 이 DNA는 전체 게놈상에서 표시가 가능하다. cDNA 지도의 유용성은 유전자가 전사되어 발현이 되는 부위만을 표시하여 주므로 생물학적으로나 의학적으로 특히 중요하다고 생각되어지는 부위 만을 특이하게 표시해 준다고도 할 수 있다. 예를들어 어떤 질병 관련 유전자가 앞에서 언급한 genetic linkage 방법으로 염색체 상에서의 위치가 대략 확인 되었다면 조금 더 정교한 cDNA 지도를 통하여 그 부위에서 발현되고 있는 관련 유전자들을 추적하므로서 질병관련 유전자를 수월하게 분석할 수 있다.

### 3) Macrorestriction map

현재 고도의 해상능을 갖는 염색체 지도를 작성하기 위하여 사용하는 방법 중의 하나로 top-down 지도작성법에 의하여 만들어진 지도이다. 여기서는 염색체를 하나씩 분리하여 각각의 염색체를 제한 효소로 절단하여 몇개의 큰 조각으로 만들고 이것을

순서대로 나열하여 저 해상능의 지도를 작성하고, 각각의 조각을 다시 다른 제한효소로 절단하여 작은 조각을 얻어 점점 더 세분화된 지도를 작성하게 된다. 이 지도에서는 제한효소 인식부위와 그 부위들 간의 거리 등을 알 수 있다. 다음 장에서 설명 할 contig map보다 불연속점(gap)이 드물고 연속적인 지도를 작성할 수 있는 장점이 있으나 해상능이 상대적으로 떨어지고 또 유전자를 찾기 위한 방법으로는 그리 유용하지 않다는 단점이 있다.

최근 PFGE(pulsed field gel electrophoresis)방법의 개발은 이전의 전기 영동방법으로는 40 kb 정도까지 밖에 분리할 수 없었던 점에 비추어 PFGE 방법으로는 10 Mb까지 분리가 가능하다는 등 거대 DNA의 클로닝과 지도작성을 용이하게 하였다.

#### 4) Contig map

역시 고도의 해상능을 갖는 염색체 지도를 작성하기 위하여 사용하는 방법 중의 하나인데, 여기서는 분리된 염색체를 처음부터 작은 조각으로 자르고 각각을 클로닝한 후에 조각들을 차례차례 순서를 결정해 가며 지도를 작성해 가는 방법이다. 차례대로 나열되어 연결된 절편들이 합해져서 하나의 긴 DNA 블록을 형성하게 되는데, 이것을 콘티그(contig)라고 부른다. 이 방법의 장점은 연구자들 상호간에 안정된 클론들의 공유가 가능하다는 점이다. 또 FISH 방법으로 염색체 위에서 특정 콘티그의 위치를 확인할 수도 있다.

콘티그 지도는 서로 조금씩 중복되는 DNA 서열을 갖는 클론들의 연결된 유전자 도서관(linked library)으로 이루어진 전염색체 지도라고 할 수 있다. 이 방법은 보통 2 Mb 이하되는 유전자를 찾는 데는 유용하나, 어떤 DNA는 특이하게도 클론이 잘 안되는 경우가 있으므로 염색체 전체를 커버하는 지도를 작성 하기에는 어려운 단점이 있다.

지난 몇년간 효모의 인공염색체(YAC: Yeast artificial chromosome) 벡터에 1 Mb 정도되는 거대 유전자를 클로닝 할 수 있는 방법이 개발되어 콘티그 지도를 작성하기 위해 필요한 클론의 수를 현저히 줄였다. 좀 더 자세한 지도를 작성하려면 효모인공염색체에 들어가 있는 인체유전자 절편을 잘라내어 다시 더 작은 조각으로 자르고 이들 절편들을 클론하기에 적합한 벡터(이때 박테리오파지 람다나 코스미드가 쓰임)에 넣어 싹클론하고 이들의 콘티그

지도를 작성하면 된다.

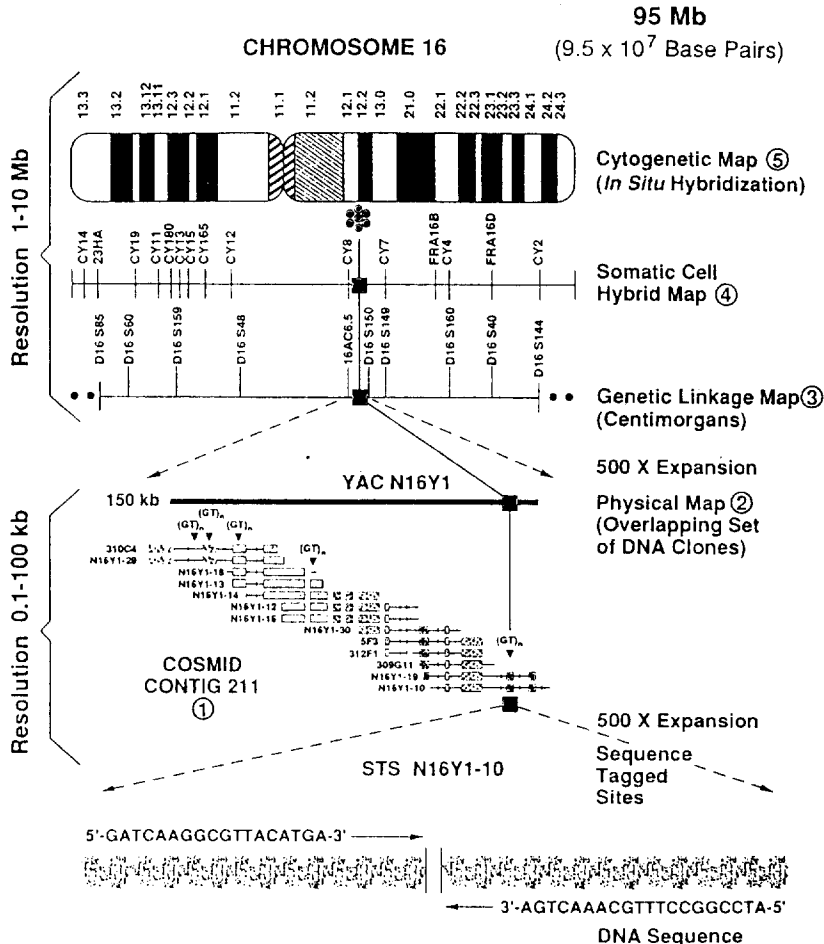
최근들어 효모인공염색체의 효모 내에서의 불안정성을 개선한 박테리아 인공염색체가 개발이 되어 몇몇 실험실에서 사용되고 있다.

## 염기서열결정

인체게놈연구의 최종 목표인 인체 게놈 전체의 염기 서열을 밝히기 위해서는 많은 부분에서의 개선--즉, 염기서열 결정의 속도 및 효율, 현 염기서열결정 방법의 신뢰성 등--이 필요하다. 고효율의 DNA 분리 정제의 자동화연구와 기존의 방법과는 완전히 다른 몇개의 단계를 없앤 새로운 염기서열결정법의 개발 연구가 진행되고 있으나, 현재 많이 사용되고 있는 방법은 코스미드와 박테리오파지 람다에 클론되어 있는 인체 DNA 단편을 잘라서 분리하고, 이것을 더 작은 조각으로 만들어 염기서열결정 벡터에 싹클론한 후 염기서열을 결정하는 방법이다.

### 1. 염기서열 결정방법

현재 분자 생물학자들이 주로 사용하고 있는 염기서열 결정 방법은 Maxam-Gilbert 방법과 썩거(Sanger)방법 두가지이다. 두 방법 모두 겔 전기영동을 하여 염기의 서열을 분석하는 방법이고 지금은 거의 모든 단계를 자동으로 할 수 있는 시스템이 개발되었다. Maxam-Gilbert 방법은 화학 물질을 이용하여 DNA의 특정 염기를 자르고 결과적으로 길이가 서로 다른 단편들을 전기영동하여 겔 상에서 길이에 따라 분리된 양상을 해석하게 된다. 썩거의 염기 서열 결정 방법(chain termination 혹은 di-deoxy method라고 함)은 효소를 이용하여 단사인 DNA주형으로부터 새로운 염기를 하나씩 붙여가며 DNA를 합성하고 특정 염기를 만나면 더 이상 합성을 차단하여 반응을 끝내도록 되어 있으며 이 합성된 DNA 단편을 전기영동하여 겔 상에서의 분리 양상을 해석하는 것이다. 이 두 방법을 제 1세대 염기서열 결정법이라고 하는데, 현재 소규모적인 게놈 해석에 사용되고 있다. 인체 염색체중 가장 작은 Y 염색체만 해도 그 크기가 50 Mb 또 가장 큰 제 1 염색체는 약 250 Mb 정도로 추정되는데 현재까지 염기서열이 결정된 것 중 가장 긴 것이 약 350,000bp 밖에 되지 않으며 현재의 기술로는



**Fig. 2.** This diagram illustrates how a cloned DNA fragment can be related to the cytogenetic map and how sequence tagged site (STS) markers (a short segment of sequenced DNA) can aid in integrating the genetic and physical maps. The location of an STS (STS N16Y1-10) is shown from the bottom up in (1) an ordered set of cosmid clones (cosmid contig 211); (2) a 150-kb YAC insert (YAC N16Y1); (3) a genetic linkage map of chromosome 16 between two genetic markers (16AC6.5 and D16S150); (4) a somatic cell hybrid map between two markers (CY8 and (5) a cytogenetic map between 16q12.1 and 16q12.2, as determined by in situ hybridization (from LANL; Los Alamos National Lab.).

최적의 조건에서 1년에 50,000~100,000 bp 정도의 염기서열 결정이 가능하고 한 염기서열을 결정 하는데 약 1~2 달러 정도가 소요된다. 이 수준으로는 인체게놈을 완전히 해석 하는데 30,000년간 30억 달러가 필요하게 된다. 그러므로 인체게놈 연구에 있어서 중점적으로 다루어야 할 부분 중 하나가 하루에 100,000 염기 이상을 정확하게 또 한 염기 해석당 0.5달러 이하의 경비로 염기서열을 결정할 수 있는 염기서열 자동결정 방법의 개발이다.

제 2세대(interim) 염기서열 결정법은 적은 경비로

약 10배 이상의 속도와 정확성을 갖는 방법이 될 것이며 제 3세대의 겔이 필요없는 염기서열 결정법은 수 백, 수 천배의 효율을 갖는 염기서열 결정방법이 될 것이며 이것은 대부분의 인체 게놈 해석에 사용될 것으로 예상된다.

## 2. 부분염기서열 결정: 염색체 지도 작성과 유전자 확인에의 유용성

서로 다른 실험실에서 만들어진 염색체 지도는 DNA 단편을 만드는 방법이나 이들을 분리하여 지도를 작성하는 방법의 차이로 인하여 연계성에 문

제가 있어왔다. 그리하여 공동의 인용 지표가 필요하게 되었고 부분적으로 염기서열이 결정된 특정부위(약 200~300 염기쌍)가 클론이나 콘티그 그리고 긴 DNA 절편 등을 분석하는데 인용 지표로서 매우 유용하다. 이를 STS(sequence tagged sites)라고 부르며 이 짧은 DNA가 물리지도 작성에 있어서 기준 지표가 되고 있다.

전체 게놈 중 발현이 되는 부위는 몇 퍼센트에 지나지 않으나 잠정적으로 매우 유용한 정보를 갖는 곳이라고 판단하여 과학자들은 무작위의 DNA보다 cDNA의 부분 염기서열 결정을 시도하고 있다. 이들 발현이 되는, 또 부분적인 염기 서열이 결정된 작은 DNA 단편들(ESTs; expressed sequence tags)은 유전자 지표로 사용될 뿐만 아니라 동시에 RNA로 전사되는 유전자들을 나타내 주고 있다. 이 방법으로 대부분의 인체 유전자를 빠른 속도로 발견하게 될 것이다. ESTs의 또 다른 장점은 염색체 상에서 유전자들만의 위치를 알려주고 특히 게놈에서의 코딩 부위만을 특이하게 표시하여 준다는 점이다.

게놈연구는 지금까지 알아본 기술 및 방법 뿐만 아니라 새로운 기술개발 및 막대한 연구비를 투자해야만 하는 일이지만, 심지어 중진국 및 개도국에서도 자국실정에 맞는 대상을 선정하여 집중적인 투자를 하고 있다. 중국의 경우 중국과학원 식물과 학연구소에서 식물게놈, 상하이대학을 중심으로 인체게놈 및 초파리연구 등에 몰두하고 있다. 더욱이 선진국에서는 표면적으로는 국제공동노력을 내세우고 있지만, 자국의 기술보호를 앞세우고 있으며 축적된 정보 및 기술을 특허화하고 있다. 이에 대응하여 국내에서도 생명과학자들간에 게놈 연구에 대한 학술 발표와 정보 교환, 세계 동향 파악 등에 취지를 둔 학술단체인 한국 인체 유전자 연구회가 1989년 결성 되었으며, 1992년에는 유전공학 연구소 유전자원센터 내에 게놈사업 담당을 개설하였고, 1993년 인체게놈연구회가 연구계 중심의 산·학·연 협동 연구회로 선정되어 과기처로부터 지원을 받고 있으며 기존의 한국인체 유전자 연구회와 통합 운영되어 게놈 연구에 대한 학술활동, 정보교환 및 국내 게놈연구에 대한 국가 기획사업 등을 추진하고 있다.

인체게놈연구의 최종 목적인 인체 게놈 전체의 염기 서열이 밝혀지면 모든 유전자들의 염색체 상

에서의 위치가 밝혀질 것이고, 그 다음은 각 유전자들의 기능(gene function)이 밝혀질 것이며, 더 나아가 이것은 생물학자들에게는 인체 생물학 연구의 Rosetta stone으로서 21세기의 생물학 연구를 주도하게 될 것이다.

## 참고문헌

1. Lai, E., B. W. Birren, S. M. Clark, M. I. Simon, and L. Hood, 1989. Pulsed field gel electrophoresis. *Biotechniques*. 7: 34-42.
2. Schwartz, D. C. and C. R. Cantor, 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*. 37: 67-75.
3. Smith, S. B. S. Gurrieri and C. Bustamente, 1990. Fluorescence microscopy and computer simulations of DNA molecules in conventional and pulsed-field gel electrophoresis. *Electrophoresis of large DNA molecules: Theory and applications*. 55-79.
4. Schlessinger, D. 1990. Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. *Trends Gen.* 6: 248-258.
5. Burke, D. T., G. F. Carle and M. V. Olson, 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science*. 236: 806-812.
6. Green, E. D., and M. V. Olson, 1990. Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA*. 87: 1213-1217.
7. Coulson, A., R. Waterson, J. Kiff, J. Sulstun, and Y. Kohara. 1988. Genome linking with yeast artificial chromosomes. *Nature*. 335: 184-186.
8. Green, E. D. and M. V. Olson. 1990. Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: a model for human genome mapping. *Science*. 250: 94-98.
9. Church, G. M., and S. Kieffer-Higgins. 1988. Multiplex DNA sequencing. *Science*. 240: 185-188.
10. Traino, G. L. 1990. DNA sequencing, automation, and the human genome. *Anal. Biochem.* 62: 418-426.

11. Wilson, R. K., C. Chen, N. Avdalovic, J Burns, and L. Hood. 1990. Development of an automated procedure for fluorescent DNA sequencing. *Genomics*. **6**: 626-634.
12. Swerdlow, H., and R. Gesteland. 1990. Capillary gel electrophoresis for rapid, high resolution DNA sequencing. *Nuc. Acids. Res.* **18**: 1415-1419.
13. Davis, L. M., F. R. Fairfield, C. A. Harger, J. H. Jett, R. A. Keller, J. H. Hahn, L. A. Krakowski, B. L. Marrone, J. C. Martin, H. L. Nutter, R. L. Ratliff, E. B. Shera, D. J. Simpson, and S. A. Soper. 1991. Rapid DNA sequencing based upon single molecule detection. *GATA*. **8**: 1-7.
14. Edstrom, R. D., X. R. Yang, G. Lee, D. F. Evans. 1990. Viewing molecules with scanning tunneling microscopy and atomic force microscopy. *FASEB*. **4**: 3144-3151.
15. Beebe, T. P., T. E. Wilson, D. F Ogletree, J. E. Katz, R. Balhorn, M. B. Salmeron, and W. J. Siekhaus. 1989 Direct observation of native DNA structures with scanning tunneling microscope. *Science*. **243**: 370-372.
16. Doolittle, R. F., ed. 1990. Molecular evolution: computer analysis of protein and nucleic acid sequences. *Meth. Enzymol.* **183**. 1-736.
17. Gray, J. W., P. N. Dean, J. C. Fuscoe, D. C. Peters, B. J. Trask, G. J. van den Engh, and M. A. van Dilla. 1987. High-speed chromosome sorting. *Science*. **238**: 323-329.