

## DNA배렬결정기술의 발전과 그 응용

김순의, 허명식

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《현시대는 과학과 기술의 시대이며 과학과 기술이 류레없이 빠른 속도로 발전하는것은 현대과학기술발전의 중요한 특징입니다.》(《김정일선집》 증보판 제15권 485페이지)

DNA배렬결정기술은 DNA재조합기술을 안받침하여 생명과학의 발전을 힘있게 추동하여온 생물공학의 기본기술의 하나로서 끊임없이 발전하여 생명과학의 발전과 사람들의 생활에 커다란 영향을 미치고있다.

1945년에 《형질전환인자는 DNA이다.》라는 에이버리의 충격적인 선언이 발표되고 1953년 왈손-크릭의 DNA두오리라선모형구조가 발표되어 분자생물학이 시작되었지만 근 20여년동안 DNA는 생물학적으로 가장 중요한 첫째가는 물질이면서도 생화학적으로 가장 다루기 힘든 물질이었다. 실제로 1953년 인슐린의 1차구조가 결정되면서 단백질의 1차구조결정법이 개발되어 1975년전까지는 단백질의 1차구조가 직접 결정되었지만 DNA의 1차구조결정은 꿈에 지나지 않았다. 그러나 1972년 재조합DNA기술과 1975년 막쌈-길버트법과 썬거법이 개발되면서 DNA는 생화학적으로 가장 다루기 쉬운 물질로 전환되어 그때부터 단백질의 1차구조는 특별한 경우를 제외하고 자기의 유전자의 1차구조로부터 알게 되었다. 이렇게 시작된 DNA배렬결정기술은 생명과학의 발전을 힘있게 추동하였을뿐아니라 그 발전과 더불어 급속히 발전하여 2003년에 인간게놈의 참고배렬을 13년동안에 수십억US\$의 자금으로 결정하였다면 현재는 한사람의 게놈배렬이 1 000US\$로 하루동안에 결정되게 되었으며 앞으로 더 발전하게 될것이다.

### 1. 제 1 세대 DNA배렬결정기술과 그 응용

1975년 길버트는 당시 대장균의 젓당오페론을 연구하면서 젓당오페론의 CAP-cAMP결합배렬에 젓당억제제가 결합하면 DNA를 화학적으로 처리할 때 그 부위가 절단되지 않는 현상으로부터 DNA를 특이적으로 절단할수 있다는 착상을 하게 되어 화학적방법에 기초한 막쌈-길버트법을 내놓았다. 당시 이 기술의 개발은 DNA의 연구에서 하나의 전환점으로 되었다.

이미 1950년대에 인슐린의 1차구조결정법을 내놓아 노벨상을 수여받은 썬거는 1976년에 DNA복제반응을 응용하여 특수한 물질인 디데옥시뉴클레오타이드를 교묘하게 리용한 또 다른 방법인 썬거법(혹은 디데옥시법)을 개발하여 DNA배렬결정기술의 발전에서 커다란 이정표를 마련하였으며 그후 이 방법에 기초하여 인간게놈배렬이 결정되었다. 이 방법에서는 DNA복제반응에서 반응액속에 보통의 데옥시뉴클레오타이드와 함께 적은 량의 디데옥시뉴클레오타이드를 넣어주어 반응과정에 이것이 들어가면 다음번 뉴클레오타이드가 결합되지 못하고 그러면 각이한 길이의 DNA단편이 생기는것을 리용하여 그것을 폴리아크릴아미드겔전기

영동법으로 분리하여 한번에 500~1 000bp의 배렬을 결정할수 있었다. 개발당시에는 반응을 A, T, G, C구로 나누어 진행하고 방사성동위원소를 리용하여 폴리아크릴아미드겔전기영동을 4개의 홈으로 나누어 진행하였다. 그후 막쌈-길버트법(화학적방법)과 썬거법(디데옥시법 혹은 효소적방법)을 내놓은 학자들이 다같이 노벨상을 수여받고 두 방법이 다같이 리용되었지만 썬거법이 자동화, 로봇트화에 더 유리한것으로 하여 썬거법만이 리용되게 되었다.

1980년대초 썬거법에서 4가지 디데옥시뉴클레오티드를 각각 서로 다른 형광색소로 표시할수 있게 되면서 폴리아크릴아미드겔전기영동을 1개의 홈에서 진행할수 있게 되고 시료량과 원가가 줄어들었다. 다음 폴리아크릴아미드겔전기영동을 평판식으로 하던것을 실판 전기영동법을 도입하면서 높은 전압하에서 영동시간을 단축하고 시료량이 줄어들어 원가가 대폭 줄어들었다. 1998년 《Applied BioSystems》회사가 1대의 장치에 96개의 실판을 설치한 《ABI PRISM 3700》형 배렬결정장치를 개발하면서 인간게놈해석에서는 커다란 전환이 일어났다. 이 장치가 개발됨으로써 《Celera》회사가 설립되고 인간게놈해석국제연합연구집단과 이 회사사이의 경쟁이 시작되어 인간게놈배렬결정이 가속화되었으며 결국 2000년 6월 26일 세계적으로 인간게놈배렬의 초판배렬이 발표되었다. 다음 이 방법이 보다 개량되어 인간게놈배렬의 완성판이 2003년 4월에 발표되었다. 그리고 이 방법으로 대장균과 효모, 애기장대, 초파리, 벼, 침팬지 등 수많은 생물의 게놈배렬이 결정되어 21세기 생명과학발전을 위한 튼튼한 기초가 마련되었다.

한편 그후 세계적으로 인간게놈배렬결정기술이 무어의 법칙에 따라 발전하리라는 예측이 나오고 그것을 자극하기 위하여 2012년까지 한사람의 게놈배렬을 1 000US\$로 결정할수 있게 한다는 인간게놈 1 000US\$프로젝트가 시작되었다.

썬거법이 이렇게 인간 및 수많은 생물의 게놈배렬결정에서 결정적인 역할을 수행하였지만 그후의 생명과학의 발전추세에 더는 따라갈수 없게 되었다. 20세기의 생명과학이 분자생물학의 시대였다면 21세기에는 게놈학의 시대로 되면서 모든 실험수법이 고속대량화되고 게놈배렬결정기술도 고속대량화의 요구를 충족시켜야 하였지만 더는 만족시킬수 없었다. 그것은 이 방법에서는 결국 폴리아크릴아미드겔전기영동을 진행하여야 하는데 아무리 실판전기영동법을 받아들여도 1대의 배렬결정장치에 실판을 500개이상을 장치하는것은 원리적으로 불가능하였다. 따라서 완전히 새로운 원리에 기초한 게놈배렬결정기술의 개발이 절실하게 제기되었다.

## 2. 제2세대 DNA배렬결정기술과 그 응용

썬거법으로 더는 고속대량화를 실현할수 없다는데로부터 썬거법의 원리에서 완전히 벗어난 새로운 배렬결정원리가 개발되었는데 그것이 바로 피로배렬결정법이었다. 이 방법은 역시 DNA복제반응에 기초하지만 4가지 뉴클레오티드를 차례로 한가지씩 넣어주면서 뉴클레오티드가 부가될 때마다 생기는 피로린산에 의하여 여러가지 효소반응으로 빛이 나오게 하고 어느 뉴클레오티드가 결합되었을 때 빛이 나오는가를 검측하여 DNA배렬을 결정하는 방법이다. 그러므로 이 방법에서는 차례로 뉴클레오티드를 넣어주고 결합되지 않은 뉴클레오티드는 세척하고 빛을 검출하는 조작을 반복하면 되는데 이것은 자동화, 로봇트화로 얼마든

지 할수 있는것이다. 그러므로 이것은 한번에 10만~수십억개의 반응을 동시에 진행해야 하는 고속대량화의 요구를 얼마든지 충족시킬수 있었다.

이 원리에 기초하여 2007년부터 454배열결정장치와 Illumina배열결정장치, SOLiD배열결정장치, 이온분출배열결정장치 등이 개발되어 리용되고있으며 이때부터 DNA배열결정기술이 무어의 법칙을 벗어나 보다 비약적으로 발전하게 되었다.(그림)

454배열결정장치에서는 약 100만개의 매 구멍에 차례로 뉴클레오티드를 넣어주고 어느 뉴클레오티드를 넣어줄 때 빛이 나오는가를 검출하여 배열을 결정한다.

Illumina배열결정장치에서는 유리지지체에 10억개의 DNA단편을 고정시키고 DNA복제반응을 진행시키면서 각각 서로 다른 형광색소로 표시한 4가지 뉴클레오티드를 차례로 넣어주며 매 뉴클레오티드를 넣어줄 때마다 수자식사진기로 사진을 찍고 그것을 중합분석하여 배열을 결정한다.

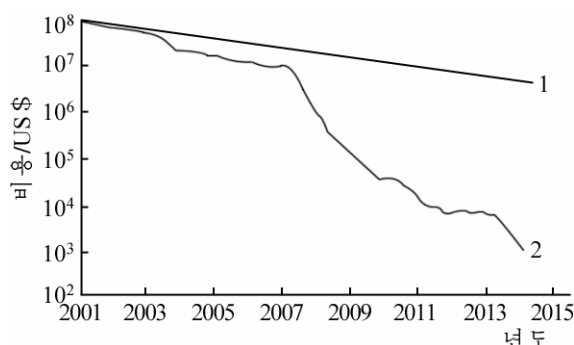


그림. 한 사람의 게놈을 결정하는데 드는 원가  
1—무어의 법칙, 2—원가변화

이온분출배열결정장치에서는 수십억개의 개별적인 홈에서 DNA복제반응이 진행되게 하고 뉴클레오티드가 부가될 때마다 나오는 수소이온을 검출하여 DNA배열을 결정한다.

이러한 장치들이 개발되어 2007년부터 DNA배열결정기술이 무어의 법칙에 따라 예측하였던것보다 훨씬 더 빠른 속도로 발전하고 이미 세계적으로 수만명의 사람들의 개인게놈이 결정되었다. 이렇게 2세대 배열결정장치로서 여러가지가 나오고 많은 성과가 이룩되었지만 한사람의 게놈을 1 000US\$로 결정한다는 목표에 도달하기에는 멀었다. 그것은 배열결정원가가 높고 모두 PCR로 목적하는 DNA분자를 수천개로 증폭한 다음 배열을 결정하기때문이다. 그러므로 세계적으로 단분자를 리용한 3세대 배열결정장치들이 개발되었다.

### 3. 제3세대 DNA배열결정기술과 그 응용

3세대 DNA배열결정원리와 방법에는 여러가지가 있지만 현재 상업화되어 널리 리용되고있는 장치로서는 PacBio배열결정장치와 Nanopore배열결정장치가 있다.

PacBio배열결정법에서는 1 000만개이상의 매 구멍의 바닥에 결합된 DNA폴리메라제가 한분자의 DNA분자를 복제하는 과정에 각각 서로 다른 형광색소로 표식한 뉴클레오티드가 부가되면서 생기는 형광발색의 전과정을 연속적으로 《동화상》으로 촬영하고 그것을 분석하여 DNA배열을 결정한다. 이 방법에서는 한분자의 DNA를 대상으로 하여 단계적으로가 아니라 연속적으로 모든 과정이 진행되므로 일명 단분자실시간(Single Molecule Real-Time: SMRT)법이라고도 한다. 이 방법의 우점은 1개의 배열결정단위의 길이가 매우 긴것이다. 1개의 배열결정단위의 평균길이가 10kb이상이고 총자료의 약 절반이 20kb이상이며 최대길이는 60kb이다. 이것은 1세대 및 2세대의 배열결정장치들에서는 생각조차 할수 없었던 길이(표)로서 매우 중요한 의미를 가진다.

실제로 인간과 진핵생물의 게놈에 많이 있는 긴반복배열이나 텔로메아, 공백 등의 배

렬을 결정하는데서 제기되던 문제들을 해결할수 있게 해준다.

이 방법의 결함은 장치의 배열결정생성량이 2세대 배열결정장치들보다 낮은것이다. 실제로 Illumina배열결정장치에서는 하루동안에 1 670억bp의 배열이 결정된다면 2015년 당시의 PacBio결정장치에서는 10억bp의 배열이 결정된다. 다른 하나의 결함은 배열결정의 오류가 11~15%로서 매우 높은것이다. 또한 PacBio배열결정법의 배열결정속도가 높지만 원가가 아직 높다는것이다.(표) 그러므로 현재 2세대 배열결정장치의 높은 정확도와 높은 배열결정생성량, 짧은 배열결정단위의 우점과 3세대 배열결정장치의 매우 긴 배열결정단위의 우점을 배합한 잡종배열결정법이 리용되고있다. 물론 PacBio배열결정장치자체가 계속 발전하여 2016년말에는 배열결정생성능이 하루에 약 4 500억bp로 늘어났고 원가도 떨어졌으므로 이 장치가 널리 리용되게 되었다.

Nanopore배열결정법에서는 나노기술을 리용하여 한오리사슬의 DNA분자를 nm크기의 매우 가는 구멍을 통과시킨다. DNA분자가 구멍을 통과하면서 뉴클레오티드배열에 따라 서로 다른 전류가 생겨난다. 다음 전류의 패턴을 리용하여 배열을 결정한다. 이러한 장치가 실제로 2014년에 두번째 3세대배열결정장치로서 상업화되었다. 앞으로 이 장치가 초당 1 000bp의 배열을 읽고 구멍을 500개 갖추고있다면 리론적으로 한 세균의 게놈은 1min내에, 한 사람의 게놈은 2h이내에 결정할수 있다. 이 장치는 PacBio장치와 비슷하게 배열결정단위의 길이가 길다는 우점을 가지고있다.(표) 그러나 배열결정의 오류가 초기에는 38%로서 더 높았으나 현재는 2%로 떨어졌다. 이 장치의 최대의 우점은 이동식으로서 매우 작고 USB장치이며 값이 매우 낮은것이다. 따라서 현재 이 장치는 생물의학실험실들에서 널리 리용되고있다.

그외에도 전자현미경이나 원자힘현미경을 리용하여 1개의 DNA분자를 가시화하고 DNA 모양의 약간한 차이를 리용하여 DNA분자를 주사하면서 뉴클레오티드배열을 읽는 장치 등이 연구되고있지만 아직 상업화되지는 못하였다. 그러나 앞으로 DNA배열기술은 보다 빠른 속도로 발전하여 세계적으로 수백만 혹은 수천만 사람들의 개인게놈이 보다 낮은 가격으로 결정되어 사람들이 자기의 게놈을 T카드속에 넣어가지고다니게 될것이며 그것은 의학의 발전에 커다란 기여를 하고 사람들이 자기 건강의 주인이 되어 병을 예방하고 개인 별의료법으로 병을 치료받게 될것이며 그러면 사람들의 앞으로의 생활에서는 거대한 전환이 일어나게 될것이다.

표. 여러 세대의 배열결정장치들의 성능비교

장치	세대	배열결정단위 길이/bp	오유률 /%	회전당 배열 결정단위수/건	회전당 시간	100만bp당 원가/US\$
Sanger ABI 3730x1	1세대	600~1 000	0.001	96	0.5~3h	500
Ion Torrent	2세대	200	1	$8.2 \cdot 10^7$	2~4h	0.1
454 (Roche) GS FLX+	"	700	1	$1 \cdot 10^6$	23h	8.57
Illumina HiSeq 2500 (High Output)	"	2 x 125	0.1	$8 \cdot 10^9$ (양쪽)	7~60h	0.03
Illumina HiSeq (Rapid Run)	"	2 x 250	0.1	$1.2 \cdot 10^9$ (양쪽)	1~6d	0.04
SOLiD 5500x1	"	2 x 60	5	$8 \cdot 10^8$	6d	0.11
PacBio RS II: P6-C4	3세대	평균 $(1.0 \sim 1.5) \cdot 10^4$	2~13	$(3.5 \sim 7.5) \cdot 10^4$	0.5~4h	0.04~0.08
Oxford Nanopore MinION	"	평균 $(2 \sim 5) \cdot 10^3$	2	$(1.1 \sim 4.7) \cdot 10^4$	50h	6.44~17.90

## 참 고 문 헌

- [1] A. Rhoads et al.; PacBio Sequencing and its Applications, Genomics Proteomics Bioinformatics doi.org/10.1016/j.gpb.2015.08.002 2015.
- [2] B. Alberts et al.; Molecular Biology of the Cell, Garland Science, 478~481, 2015.
- [3] J. D. Watson et al.; Molecular Biology of the Gene, Cold Spring Harbor Laboratory Press, 167~168, 2014.

주제107(2018)년 4월 5일 원고접수

## Development and Application of DNA Sequencing Technologies

*Kim Sun Ui, Ho Myong Sik*

DNA sequencing technologies, which have promoted the advances of bioscience together with recombinant DNA technology, provide one of major techniques in biotechnology, and have made great impacts to development of bioscience and life of people. Their more and more advancement suited for the era of personal genomics will lead to offer active assistance to basic research and application of bioscience in 21<sup>st</sup> century, medicine and agriculture. This review covers development, successes and perspectives of DNA sequencing technologies. Sanger's DNA sequencing, developed in 1976, has made a great contribution to completion of human genome project and beautiful decoration of 20<sup>th</sup> biology. Since then, DNA sequencing has been developed according to Moore's rule before the advent of the second genome sequencing in 2007, including Illumina, 454, SOLiD, and Ion Torrent, followed by the third genome sequencing such as PacBio sequencing and Oxford Nanopore Technologies, by which hundreds of thousand peoples' personal genomes have been determined and development of bioscience, medicine and agriculture will be driven in 21<sup>st</sup> century.

Key words: DNA, sequencing technology