

DNA 마커의 개발 현황과 이용 이 석 하 *

Development and Utilization of DNA Markers Suk-Ha Lee *

* 서울대학교 농업생명과학대학(College of Agriculture and Life Science, Seoul National University, Suwon 441-744, Korea)

유전자지도(genetic map)는 잡종 분리집단에서 대조되는 유전인자 (allele)의 분리 양상을 검정하여 만들어지며, 유전자간의 조환가로 표시되는 상대적인 유전거리를 이용하여 나타낸다. 유전자지도를 제작하기 위해서는 유전적 마커를 필요로하는데, 과거에는 형태적 특성의 변이에 따른 마커를 이용하였으나 최근 분자마커의 개발에 의하여 유전학적 연구가 급속도로 발전하게 되었다. 분자마커는 DNA 염기서열수준에서 나타나는 변이를 이용하는 것으로서, 대부분 neutral site이다. 이는 개체들간 DNA 변이에 의하여 표현형으로 나타나지 않는 것을 의미하며, 따라서 비기능성 DNA 부분의 변이 혹은 단순염기변이로서 거의 형태적으로 차이를 나타내지 않는다. 이와 같은 분자마커는 형태적 마커보다 비교할 수 없을 정도로 생물체에 많이 존재하고 생물체의 기능이나 생리에 장애를 주지 않아서 마커로 이용하는데 유리하다(Jones 등, 1997). 뿐만 아니라 주어진 분리집단에서 다형성을 보이는 마커가 충분하여, 한 개의 집단으로부터 유전자지도를 충분히 제작이 가능하다. 최근 생화학적 마커인 동위효소를 식물육종에 이용한 후 DNA 마커인 RFLP가 개발되고 계속하여 RAPD, SSR, AFLP, SNP 등 많은 종류의 마커가 개발되었다.

본 발표에서는 현재까지 개발된 마커들을 소개하고, 최근 개발된 마커의 미래 식물 육종상 이용 가능성에 대하여 논의할 것이다. 또한 마커들 간의 장점 및 단점을 비교 검토할 것이며, 또한 식물육종상 DNA 마커 이용에 있어서 가장 문제점이 되는 DNA 분석 자동화에 대하여 소개하고자 한다.

마커의 종류

1. 형태 마커

단일 유전자에 의하여 지배되고 다양한 환경조건에도 안정적으로 발현하게 되는 형태적 특성 (꽃색 등)이 마커로 이용될수 있다 (Staub 등, 1996). 그러나 상당수의 형태적 마커는 환경에 따라서 변화될 가능성이 있을 뿐만 아니라 (예: 오이의 신육형 유전자 *de*), 마커의 수가 극히 제한되어 주어진 한 개 집단에서의 유전자 지도 제작은 불가능하다.

2. Isozymes

전기영동을 통해 분리된 동위효소는 개체간 분자량이 다르기 때문에 마커로서 이용이 가능하다. 일반적으로 starch gel을 이용하여 분리된 동위효소는 기질과 반응하여 나타난 발색 반응으로 동위효소 유무를 탐색할 수 있으며, codominant 마커이다. 그러나 동위효소를 마커로 이용하는데 있어서 주어진 집단에 분리되는 마커가 한정되어 있을 뿐만 아니라, 때때로 post-translational modification에 의하여 정확한 유전자형을 식별하는데는 한계가 있다.

3. Restriction Fragment Length Polymorphism (RFLP)

RFLP는 최초의 DNA 마커로서 genomic DNA 상의 single base change, insertion, deletion 등 변이를 제한효소에 의하여 탐색하고 그 결과 나타나는 DNA 단편의 크기를 전기영동과 Southern blotting 기술에 의하여 다형성을 감지하는 기술로서 일반적으로 codominant 마커이다.

4. Randomly Amplified Polymorphic DNA (RAPD)

RAPD는 William 등 (1990)에 의하여 고안된 마커로서, 최초의 PCR-based 마커라고 할 수 있다. 이는 genomic DNA 상에서 약 10개의 base로 구성된 primer와 일치된 부위를 PCR에 의하여 증폭된 DNA 파편을 분석하는 마커로서, RFLP와는 달리 증폭된 DNA 양이 충분하여 radioisotope 혹은 기타 chemically-labelled 된 probe를 필요로 하지 않고, 분리된 gel 상에 ethidium bromide를 처리하여 hybridization 과정없이 DNA 변이를 분석할 수 있다. 그러나, RAPD는 재현성이 낮다는 문제점이 있다.

5. Amplified Fragment Length Polymorphism (AFLP)

AFLP는 PCR-based 기술로서 genome상의 고도로 선발된 부위(rare cutter와 frequent cutter이용)에 대한 DNA 변이를 나타내는 마커로서, RFLP 및 RAPD 마커 개발이후, 유전학자의 많은 관심을 가져왔다(Ridout & Donini, 1999). 실제로 각 작물의 유전자지도를 세밀화 시키는데 중요한 역할을 하였으나, genome상의 특정부위에 마커들이 주로 존재하는 것이 그 단점이다.

6. Simple Sequence Length Polymorphisms (SSLPs)

SSLP는 genome상 특정부위에서 simple sequence가 반복되는 수가 개체마다 다른 변이로서 이용될 수 있는 마커로서 simple sequence가 수십개의 염기로 구성되어 있을 경우 minisatellites (혹은 variable number of tandem repeats: VNTRs)이며, simple sequence가 2, 3, 혹은 4개의 base일 경우 microsatellites (simple tandem repeats: STRs)라 하며, 통상 SSR이라한다. Minisatellites는 chromosome의 끝부분에 주로 위치하나, microsatellites는 genome 전체에 널리 분포하여 마커로서 이용가치가 높고 loci마다 많은 alleles를 가지고 있어 PIC가 높아서 유전적 다양성을 예측하는데 이상적인 마커이다(Weising 등, 1998). 또한, microsatellites는 자동 염기서열 분석장치를 통한 SSR 분석이 가능하여 기존의 방법보다도 7-8배 이상의 속도를 가지고 DNA 변이를 분석할 수 있다(Cregan & Diwan, 1997).

7. Single Nucleotide Polymorphism (SNP)

생물체의 게놈구조가 구명됨에 따라서 염기의 개체변이에 대한 연구가 집중적으로 이루어져 왔는데 특히 단일 염기변이 (single nucleotide polymorphism: SNP)가 대표적인 변이로서 이를 point mutation이라고도 한다. SNP는 이미 전술한바 있는 single base change에 의한 RFLP 뿐만 아니라, 제한효소에 의하여 감지될수 없는 point mutation을 포함하기 때문에 생물체의 게놈에서 SNP의 출현율은 매우 방대하다. 따라서 SNP는 인체게놈연구(human genome project)에서 많은 발전을 가져 왔으며, 지난 2년동안에 약 45 백만 달러를 투자하여 300,000개의 SNP를 개발하였다 (Marshall, 1999).

이와 같은 SNP는 인간의 질병에 대한 유전적 진단을 가능하게 하고, 또한 진단된 결과에 대한 예방조치를 취함으로써 발병을 줄이게 할 수 있을 뿐만 아니라, 개개인에 대한

약물치료를 가장 적합한 형태로 이루어 질 수 있도록 하게 한다. 식물육종 분야에서도 다른 DNA 마커와 마찬가지로 QTL 탐색 및 유전적 마커로서 이용이 가능하여 차세대 마커로서 관심을 받고 있으며, 현재 일부 작물, 특히 콩에서 SNP 출현율에 대한 연구가 진행 중에 있다. Wang 등 (1998)에 의하여 조사된 인체 DNA의 1.98 megabase에서 1 kilobase 당 1.39 SNP가 발견되었다. 한편 콩에서는 미국 북부에서 주로 재배되고 있는 콩 18개 품종을 대상으로 18,000 base를 조사한 연구 결과에 의하면 (Cregan, 1999), 1 kilobase 당 3.4 SNP 출현율을 보여 인체 DNA 보다도, 콩의 SNP 출현율이 훨씬 높아서 식물 육종을 위한 마커로써 가능성을 나타내고 있다.

최근 분자생물학적 기술이 비약적으로 발전하게 됨에 따라 기계 전자공학의 기술을 접목하여 DNA chip 제조 기술이 발전하게 되었다. 유리와 같은 고상체에 막대한 종류의 DNA를 고정하여 제조된 DNA chip은 최소 500 bp 이상의 유전자가 붙여져 있는 cDNA chip과 15 내지 25개의 염기로 이루어진 oligonucleotide chip이 있는데, 이를 이용한 oligonucleotide microarray hybridization 기법을 이용하여 동시에 대량으로 수많은 SNP를 탐색을 가능하게 할 수 있어서 기존에 개발된 gel-based 시스템 마커인 RFLP, RAPD, SSR, AFLP 보다도 효율적으로 유전자 지도 제작을 할 수 있다. 게놈의 크기가 2500cM인 작물을 예로 할 경우 5cM의 간격의 유전자 지도를 제작하기 위해서는 5000 loci로 구성된 SNP map의 경우 주어진 집단에서 10% 만 다형성(polymorphism)을 보일 경우에도 가능하게 될 것이다. 따라서, SNP는 식물 육종을 위한 QTL 탐색, 유전적 다양성 측정, 혹은 유용유전자 보유 계통의 탐색에 적절하게 이용될 수 있다.

그러나 식물육종 과정 중 QTL 탐색은 SNP와 DNA chip의 결합기술에 의하여 효율적으로 저비용으로 가능하나, 일단 탐색된 QTL과 가깝게 연관되어 있는 SNP를 marker-assisted selection에 이용하기 위해서는 많은 loci를 대량으로 할 수 있는 SNP-DNA chip 기술과는 달리, 둘 내지 세 개의 loci를 가지고 수만개의 계통을 선발할 수 있는 다른 체계가 필요하다. 이는 gel-based 체계가 아닌 용액상태에서 목표 SNP 유전자의 유무를 탐색할 수 있는 oligonucleotide probe인 molecular beacon을 사용함으로써 (Tyagi and Kramer, 1996; Tyagi et al., 1998) 향후 SNP의 식물 육종상 간접선발을 가능하게 할 것이다.

마커들간의 비교

마커들을 작물육종에 이용하기 위해서는 1)마커 개발을 위한 비용과 2)개발된 마커들을 가지고 실제 유전자 지도를 작성하고, 유용유전자와 가깝게 연관된 마커 탐색 및 간접선발시 다량의 계통을 선발할 수 있는 체계를 가져야한다 (Rafalski & Tingey, 1993). 마커개발시 염기서열의 정보를 필요로 하는 SSR 및 SNP는 그 비용이 높다고 할 수 있다. 그러나, 일단 개발된 마커는 codominant marker로서 보다 많은 정보를 가져다 주고, genome상에 널리 존재하고, genotyping이 용이하여 육종가가 사용하기에는 바람직하다 (표 1).

DNA 변이 분석 자동화

식물육종과정상 마커를 이용하기 위해서는 DNA 분리, genotyping을 위한 PCR, 전기영동, hybridization 등 여러 과정에 있어서 자동화가 이루어져야 한다. 특히 DNA 분리의 자동화는 single-step protocol (Thomson & Henry, 1995)에 의하면 1인 하루 약 8백 여 계통으로 부터 DNA 분리가 가능하며, 따라서 월 평균 20,000 계통의 DNA 분리가 가능

한 것으로 알려져 있다. 따라서 마커에 의한 간접선발이 일반 실험실에도 충분히 가능하다. 또한 PCR을 용이하게 하기 위하여 Virk 등 (1999)에 의하면 애기장대로부터 DNA 추출없이 잎에 염을 처리하여 세포막과 핵막을 와해시켜 PCR을 직접하는 방법 등이 개발되었으나, 작물육종에 적용되기 위해서는 작물별 적합한 방법이 개발되어야 할 것이다.

한편 DNA 분석시 gel-based 체계는 전기영동에 걸리는 시간과 비용이 높아서, 전기영동을 하지 않고 DNA 변이를 분석하는 것이 효율적이다. 따라서, SNP는 전기영동과정 없이 hybridization을 고도로 정밀하게 하는 기술을 이용하여 육종시 선발효율을 높일 수 있는 차세대 마커로서 기대되어, 작물육종을 위한 SNP 탐색 및 그 이용에 대한 연구가 필요하다.

인용문헌

1. Cregan PB 1999. DNA markers, maps, and technologies. Proceedings of World Soybean Research Conference VI 46-58.
2. Cregan PB, Diwan N 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theor. Appl. Genet. 95:723-733.
3. Jones N, Ougham H, Thomas H 1997. Markers and mapping: we are all geneticists now. New Phytol. 137:165-177.
4. Marshall E 1999. Drug firms to create public database of genetic mutations. Science 284:406-407.
5. Rafalski JA, Tingey SV 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends in Genetics 9(8):275-280.
6. Ridout CJ, Donini P 1999. Use of AFLP in cereals research. Trends in Plant Sci. 4(2):76-79.
7. Staub JE, Serquens FC, Gupta M 1996. Genetic markers, map construction, and their application in plant breeding. HortScience 31(5):729-741.
8. Thomson D, Henry R 1995. Single-step protocol for preparation of plant tissue for analysis by PCR. BioTechniques 19(3):394-400.
9. Tyagi S, Kramer FR 1996. Molecular beacons: probes that fluoresce upon hybridization. Nature Biotech. 14:303-308.
10. Tyagi S, Bratu DP, Kramer FR 1998. Multicolor molecular beacons for allele discrimination. Nature Biotech. 16:49-53.
11. Virk PS, Pooni HS, Syed NH, Kearsy MJ 1999. Fast and reliable genotype validation using microsatellite markers in *Arabidopsis thaliana*. Theor. Appl. Genet. 98:462-464.
12. Wang DG et al. 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077-1082.
13. Weising K, Winter P, Hüttel B, Kahl G 1998. Microsatellites markers for molecular breeding. J. of Crop Production 1(1):113-143.
14. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res.18:6531-6535.

Table 1. Comparison among several molecular marker systems

	Molecular Marker Systems				
	RFLP	RAPD	AFLP	Microsatellite	SNP
1) Principle	Restriction & Hybridization	PCR with random primers	Selective PCR	PCR of SSR	DNA chip & Hybridization
2) DNA required (μg)	10	0.02	0.5~1.0	0.05~0.10	0.05<
3) PCR-based	No	Yes	Yes	Yes	No
4) Gel-based	Yes	Yes	Yes	Yes	No
5) Genomic abundance	High	Very high	Very high	High	Very high
6) Marker type	Codominant	Dominant	Dominant	Codominant	Codominant
7) Reproducibility	Very high	Fair	Very high	Very high	Very high
8) Sequence information required	No	No	No	Yes	Yes
9) Development	Medium	Low	Low	High	High
10) Easy of use	Labor intensive	Easy	Difficult	Easy	Easy