

벼 microarray를 이용한 유전자발현 profiling 연구동향

윤웅한 · 김연기 · 김창국 · 한장호 · 이태호 · 김동현 · 이강섭 · 박수철 · 남백희

Current status on expression profiling using rice microarray

Ung-Han Yoon · Yeon-Ki Kim · Chang-Kug Kim · Jang-Ho Hahn · Dong-Hern Kim · Tae-Ho Lee · Gang-Seob Lee · Soo-Chul Park · Baek-Hie Nahm

Received: 5 April 2010 / Accepted: 14 April 2010
© Korean Society for Plant Biotechnology

Abstract As the International Rice Genome Sequencing Project (IRGSP) was completed in 2005 and opened to the public, many countries are making a lot of investments in researches on the utilization of sequence information along with system development. Also, the necessity of the functional genomics researches using microarray is increased currently to secure unique genes related with major agricultural traits and analyze metabolic pathways. Microarray enables efficient analysis of large scale gene expression and related transcription regulation.

This review aims to introduce available microarrays made based on rice genome information and current status of gene expression analysis using these microarrays integrated with the databases available to the public. Also, we introduce the researches on the large scale functional analysis of genes related with useful traits and genetic networks. Understanding of the mechanism related with mutual interaction between proteins with co-expression among rice genes can be utilized in the researches for improving major agricultural traits. The direct and indirect interactions of various genes would provide new functionality of rice.

The recent results of the various expression profiling analysis in rice will promote functional genomic researches

in plants including rice and provide the scientists involved in applications researches with wide variety of expression informations.

Keywords rice, microarray, expression profiling

서론

최근 동식물의 유전체완전염기서열분석 연구가 많이 이루어짐과 더불어 과거 수개의 유전자기능 분석을 행하던 수준에서 유전자 발현을 유전체 수준에서 분석하는 transcriptome 연구를 많이 수행하고 있다. 식물 유용형질 유전자 이용 효율을 극대화시키기 위하여 유전자 발현네트워크 분석을 행하여 유전자 발현을 조절하는 상위개념의 조절인자를 활용하는 수준으로 발전되고 있다. 벼의 경우 한국 등 10개국이 참여한 국제 벼 염색체 완전해독 연구 (IRGSP 2005, <http://rgp.dna.affrc.go.jp/IRGSP/>)가 완료되고 이들 염기서열정보가 공공부분에 공개됨에 따라 이를 활용할 수 있는 시스템 구축과 농업상 중요한 농업형질 유전자 탐색에 집중하고 있다 (Jung et al. 2008a; Park et al. 2009)

농업상 중요한 형질 유전자 분리를 위하여 발현유전자의 확보와 발현유전자의 조직내에서의 발현양상 해석 등은 매우 중요한 위치를 차지하고 있다. 이러한 유전자의 발현양상분석과 형질유전자의 기능해석은 최근 개발되어 활용되고 있는 것이 microarray에 의한 유전체 수준에서의 전사체 분석이다. Microarray (일명 DNA chip)는 유전자의 발현양상을 대량으로 분석 가능하게 하여 기존 종래의 단일 유전자 발현에 분석에 비하여 매우 빠른 속도로 분석을 가능하게 하였다. 벼 microarray 제작을 위하여

U.-H. Yoon · C.-K. Kim · J.-H. Hahn · D.-H. Kim · G.-S. Lee · S.-C. Park
국립농업과학원
(Genomics Division, National Academy of Agricultural Science, RDA, Suwon, Korea)

Y.-K. Kim · T.-H. Lee · B.-H. Nahm (✉)
명지대학교
(Bioscience and Bioinformatics Division, MyongJi University, Yongin, Korea)
e-mail: bhnaahm@mju.ac.kr

벼 발현유전자정보가 필요하며 다양한 조직에서 발현하는 발현유전자 정보 분석이 필요하다. 벼 발현유전자 정보 분석의 경우 일본 농업생물자원연구소에서는 벼 염색체 염기서열정보를 기초로 발현유전자영역을 분석하는 프로그램을 이용하여 32,745개의 발현정보를 분석하여 IRGSP Build4를 통하여 발현정보를 제공하고 있다. 또한 미국의 경우 미시간대학 벼 게놈 annotation 프로젝트에서 NCBI의 dbEST 벼 유전체 정보를 종합화하여 pseudomolecule Ver6.1을 만들고 56,797개의 발현유전자 정보를 공개하고 있다. 현재 벼 발현유전자 정보를 이용하여 44K microarray (Agilent), 45K microarray (NSF), 135K microarray (GreenGene Biotech), 300K microarray (GreenGene Biotech), GeneChip array (Affymetrix) 등이 제작되어 활용되고 있다. 초기에는 베이징 게놈연구소의 발현정보를 이용한 60K microarray (GreenGene Biotech)가 제작 상용화되어 대량유전자 발현 분석 등 기능유전체 연구를 위한 도구로 사용되어 왔다.

Microarray를 이용한 벼의 주요농업형질관련 유전자 발현 profiling 분석으로 유전자 발현을 조절하는 2,700여개 전사조절인자의 조직부위별 발현을 검정하고 있으며 호르몬, 내재해성, 종자발달단계 등에서의 유전자 발현을 검정하는 광범위한 실험이 수행되고 있다 (Sato et al. 2007; Jain et al. 2009; Degenkolbe et al. 2009; Yamaguchi et al. 2006; Furutani et al. 2006). 국내에서 벼 microarray를 이용한 유전자 대량발현분석 연구의 경우는 국내에서 대량으로 확보하고 있는 Ac/Ds tagging 삽입변이체 및 T-DNA tagging 삽입변이체 재료를 이용하여 유전자발현 profiling 분석과 유전자간 상호작용을 분석 하고 있다 (Jung et al. 2008a; Lee et al. 2009). 특히 식물의 특정형질 유전자들은 식물의 발달과 생장, 환경스트레스 및 종자의 발달 등에 특이하게 작용할 것으로 유추되며, 이들의 발현분석에는 microarray에 의한 transcriptome profiling의 분석이 필수불가결하다 (Rensink et al. 2005). 또한 유전자 상호간 및 단백질간의 상호작용의 이해는 여러 유전자가 복합적으로 작용하는 주요 수량, 재배, 발달, 미질 관련 형질의 개선에 크게 기여할 것으로 기대되며 형질유전자들 간의 발현 network 구축은 필수적이다 (Briggs et al. 2005).

따라서, 최근 microarray를 이용한 대량유전자발현분석 및 이들 정보를 활용할 수 있는 시스템의 구축과 실제 유전자 발현네트워크를 이용한 유용형질 유전자 분리 연구에 많은 관심이 고조되고 있으며, 벼의 주요 농업형질의 개선연구에 이용하고 있다 (Shinozaki et al. 2007). 따라서 향후 개발되는 21세기의 환경 친화적 신행질 벼의 개발을 위한 관련 유전자 분석 등에 microarray 기술이 크게 이용될 것이며, 이러한 microarray에 의한 유전자의 대량 발현분석은 유전자의 전반적인 발현 조절기작 및 주요농업형질 관련 유전자 기능분석 연구에도 크게 이용될 수

있다. 본 리뷰에서는 유전자발현 profile을 이용한 벼 발현유전자 특성분석에 대한 최근 연구동향을 소개하고자 한다.

벼 발현유전자정보해석

벼의 농업형질유전자의 기능을 해석하기 위하여 완전장 발현유전자 분리 및 염색체구조분석결과를 이용한 유전자 발현영역분석이 필수적이다. 최근까지 다양한 벼 품종에서 대량발현유전자 염기서열분석 및 완전장 유전자분리 연구가 수행되고 있다. 일본 농업생물자원연구소그룹 (Kikuchi et al. 2003; Satoh et al. 2007)은 japonica 품종인 Nipponbare 벼 품종을 이용하여 대량발현유전자 염기서열분석 및 완전장 cDNA 분리에 관한 연구를 수행하였다. 완전장 발현유전자분리는 우선 조직부위별 및 발현시기별로 제작된 완전장 cDNA 581,446개를 분석하여 32,127개의 완전장 cDNA를 분리하였으며 KOME DB (<http://cdna01.dna.affrc.go.jp/cDNA/>)를 통하여 관련연구자들에게 발현유전자 정보를 공개하고 있다. 이러한 완전장 발현유전자분석 결과는 염색체완전염기서열분석결과를 활용한 발현유전자 annotation 분석 프로젝트 (The Rice annotation Project 2008) 등에 적극 활용되고 있으며 RAP-DB (<http://rapdb.lab.nig.ac.jp/>)를 통해 유전자 해석정보를 공개하고 있다 (표 1). 또한 미국의 미시간대학 벼 게놈 명명 프로젝트에서는 벼 염색체완전염기서열분석 정보와 NCBI의 dbEST 정보를 기반으로 2006년 TIGR Pseudomolecules ver. 4를 작성하였으며 이후 2009년 TIGR Pseudomolecules ver. 6.1을 작성하여 56,797개의 벼 발현유전자 정보 (33,800개 유전자 모델)를 공개하고 있다 (Ouyang et al. 2007).

중국의 벼 연구그룹 (Lu et al. 2008)은 indica 품종인 Guangluai 4와 Minghui 63을 이용하여 40,000개의 발현유전자를 분석하였으며 20,000여개의 완전장 cDNA를 분리하여 RICD DB를 통해 정보를 공개하고 있다 (<http://www.ncgr.ac/ricd>). 벼 종자 발현유전자의 경우 국립농업과학원에서 벼 종자 발달단계 발현 유전자 73,000개의 분석을 완료하여 벼 종자 발현 유전자 DB를 구축하여 관련연구자들에게 정보를 제공하고 있다 (Yoon et al. 2009). 최근까지 NCBI dbEST에 등록된 발현유전자수는 1,249,110개이며 unigene으로는 40,978개의 유전자를 포함하고 있다. 이러한 벼 발현유전자 정보와 유전자들을 이용하여 세계 각국은 농업상 중요한 형질 유전자의 대량기능분석 (Nakamura et al. 2007) 및 microarray 제작 등에 적극 활용하고 있다.

벼 microarray 종류 및 특성

벼 유전자 기능분석을 위하여 수개의 유전자 발현분석

Table 1 FL-cDNA clones mapped to five rice genome assemblies (Sato et al. 2007)

		TIGR4	IRGSP4	IRGSP3	Syngenta	93-11 (indica)
FL-cDNA	35,187	32,775	32,745	32,640	31,928	30,354
5'end FL-EST	241,854	212,598	212,539	211,564	208,606	199,001
3'end FL-EST	536,885	483,657	484,358	482,909	482,665	465,775
FL-cDNA locus	Chr01	4,026	4,021	4,039	4,050	3,940
	Chr02	3,196	3,198	3,215	3,186	3,153
	Chr03	3,569	3,567	3,566	3,597	3,607
	Chr04	2,531	2,530	2,534	2,477	2,493
	Chr05	2,313	2,305	2,310	2,338	2,329
	Chr06	2,292	2,293	2,290	2,262	2,266
	Chr07	2,183	2,185	2,193	2,165	2,021
	Chr08	1,933	1,934	1,939	1,912	1,827
	Chr09	1,605	1,605	1,574	1,545	1,515
	Chr10	1,538	1,528	1,536	1,502	1,416
	Chr11	1,685	1,683	1,675	1,486	1,333
	Chr12	1,693	1,692	1,705	1,523	1,435
	Chr0 (*)				434	497
	Total	28,564	28,541	28,576	28,477	27,832

(*) sequence assembled contigs that were not localized to one of the 12 chromosomes

Table 2 Rice microarrays developed up to present and their characteristics

	<i>Oryza sativa</i> 3' tilling 135K Microarray	<i>Oryza sativa</i> 3' tilling 300K Microarray	60K Microarray	NSF 45K Microarray	Agilent 44K Microarray	GeneChip Rice Genome Array
Number of Arrays per slide	12	1	1	1	4	2-30
Source	The Rice Annotation Project Database 2	The Rice Annotation Project Database 1	Beijing Genome Institute	TIGR Ver3	NIAS Sequence and UniGene	NCBI UniGene Build #52, (May 7, 2004)
Number of Genes	31,439	27,448	58,417	40,279	43,803	51,279
Total Number of Probes	12 X 135,000	385,000	60,000	43,482	43,803	multiple
Probe Length	60mer	60mer	70mer	50-70 mer	60 mer	25 mer
Number of Probe sets per Genes	4	10	1	1	1	multiple
Probe Formation	In situ synthesis oligomer	In situ synthesis oligomer	spotted oligomer	spotted oligomer	Agilent's 60-mer SurePrint technology	Photolithographic synthesis

을 행하는 수준에서 최근 비약적으로 발전한 microarray 기술을 이용하면 대량의 유전자 발현 양상을 검정할 수 있다. 벼 microarray 제작은 크게 네 가지로 형태로 만들어져 발현분석에 사용되어지고 있다 (표 2). 처음으로 개발된 spot형 microarray는 초창기 유전자수가 적은 array를 제작할 경우와 연구자 자신이 디자인한 array제작에 많이 쓰인 방식이다. 일본 농업생물자원연구소의 Yazaki 등 (2000)은 벼 발현유전자 9,000개를 이용하여 cDNA array를 제작하여 각종 호르몬실험 등에 사용하였다. 그 이후 대량유전자의 집적은 Affymetrix사의 GeneChip 시스템으로

photolithography와 solid-phase chemistry방식을 이용하여 슬라이드글라스위에 oligonucleotide를 하나씩 합성하는 방법이다. 벼 GeneChip은 NCBI의 unigene 정보 중에서 *japonica* 48,564개 유전자와 *indica* 1,260개 유전자 등 총 51,279개 유전자를 이용하여 25mer oligonucleotide를 합성하여 array를 제작하였으며 이것을 사용하여 각종 유전자발현분석 실험에 이용되어 왔다 (Zhu 2003; Narsai et al. 2009). 또한 NimbleGen사에서는 디지털 마이크로미러 방식을 사용하여 60mer oligonucleotide를 합성하는 방법으로 31,439개 유전자 정보를 바탕으로 제작된 135K microarray와 27,448개

유전자 정보로 제작된 300k microarray등이 있다. 이외에 inkjet 프린트 방식을 이용한 Agilent사의 44K microarray로 일본농업생물자원연구소 완전장유전자정보와 unigene 정보를 이용하여 제작되었으며 43,803 유전자로 구성되어 있다 (Frei et al. 2010). 또한 미국 UC-Davis 대학 Pam Ronald 박사는 TIGR 연구소의 유전자명명 데이터 (TIGR Ver3)를 이용하여 40,279개 유전자를 분석하고 50-70mer oligonucleotide로 구성된 NSF 45K microarray를 제작하였다 (Chou et al. 2004; Jung et al. 2008b). 국내에서는 2000대 초기에는 GreenGene Biotech에서 베이징 계놈연구소의 벼 유전체정보를 이용하여 58,417개 유전자의 70mer oligonucleotide를 슬라이드그라스에 스폿한 60K microarray를 제작하여 시험을 수행하여 왔다 (Lee et al. 2009; Sakar et al. 2009). 그러나 최근에는 27,448개 유전자 정보로 제작된 10개의 중첩된 probe로 구성된 300k microarray와 31,439개 유전자 정보를 바탕으로 제작된 4개의 중첩된 probe로 구성된 135K microarray와 등이 실험에 사용되고 있다. 이밖의 microarray 기술로는 반도체 기술을 응용한 CombiMatrix 시스템이 있으며 소량의 유전자발현분석에 적합하며 array의 재활용이 가능한 것으로 보고되고 있다. 이러한 다양한 microarray의 제작과 활용은 새로운 생물학분야의 발전에 크게 기여 할 것으로 생각된다.

벼 microarray를 이용한 유전자발현 profiling 분석

벼 조직별 발현유전자 profile 분석

벼 60K 올리고머 microarray를 효과적으로 이용하면 벼 유전자의 발현을 유전체 수준에서 분석할 수 있다. 벼의 callus, 재분화 callus, 발아종자, 잎, 뿌리, 출수기, 꽃 등의 조직들로부터 total RNA를 분리하여 그 발현 양상을 살펴 보면 어느 특정 조직이나 발달 단계에서 필수적으로 발현하는 유전자들을 알아 낼 수 있다. 조직발현분석을 위하여 미분화조직인 callus mRNA를 Cy3로 labelling 하고 목표하는 조직으로부터 얻은 mRNA는 Cy5로 labelling 하여, callus를 대조군으로 하는 상대적인 조직 특이 발현 유전자의 발현을 분석한 결과 6개의 실험 세트 중에서 1.6 배 이상 발현되거나 감소된 특이성을 보이는 유전자들만을 선별하여 hierarchical clustering 분석을 실시하였다. 이들 유전자들을 미분화조직인 callus와 비교하여 보았을 때 발아종자에서 발현이 많은 것이 1,597개, 뿌리에서 1,375개, 잎에서 1,943개, 출수기에서 2,748개, 꽃에서 2,720개, 재분화 callus에서 1,322개의 유전자가 callus에 비해서 많이 발현되었다 (그림 1). 분석에 쓰인 유전자중 대부분은 각 조직에서 공통적으로 발현하거나 혹은 몇몇 조직에 걸쳐 발현하는 것이 대부분 이었다. 이러한 유전자는 MADS box, bHLH 전사조절인자와 같이 그 종류가 다양하고 다양한 역할로 작용하는 것들이 대부분이었으며,

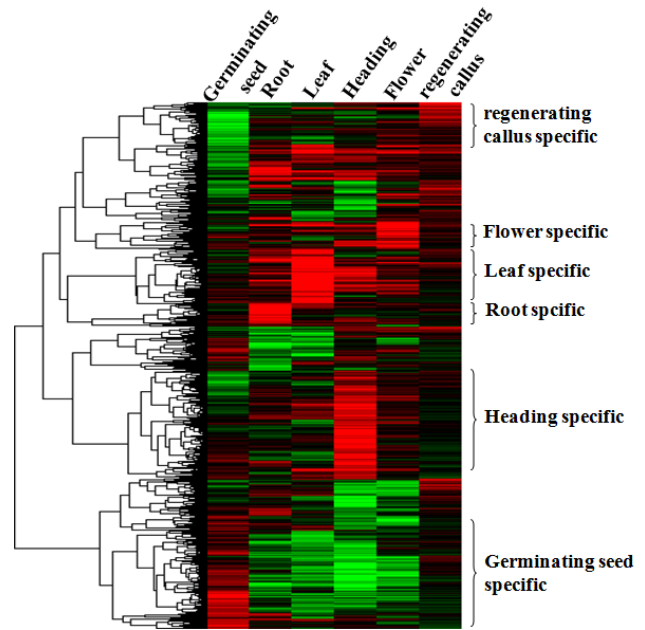


Fig. 1 Hierarchical clustering of organ-specifically expressed genes using rice 60K microarray (data from Baek-Hie Nahm) Gene expression clustered into 6 groups as germinating seed (1,597), root (1,375), leaf (1,943), heading (2,748), flower (2,720), regenerating callus (1,322)

벼의 조직 발달과 그 조절기작에 관한 연구를 위해 각 조직에서 조직특이적으로 발현 하는 유전자의 선별을 행한 결과 발아종자에서 480개, 뿌리에서 90개, 잎에서 198개, 출수기에서 440개, 꽃에서 182개, 재분화 callus에서 348개 유전자를 선별하였다. 조직 특이 발현 유전자의 대부분은 COG 분석에서 signal transduction (T)과 posttranslational modification, protein turnover, chaperone (O)에 속하는 유전자였으며 발아의 특이발현 유전자 480개중 cytochrome c 등의 에너지대사에 관여하는 유전자와 phosphoglycerate mutase, pyruvate kinase 등의 탄수화물대사에 관여하는 유전자의 발현이 높은 것을 살펴 볼 수 있었다. 뿌리에서 발현 하는 특이유전자의 개수는 90개로 6개의 조직들 중 가장 적은 수의 특이 발현 유전자가 있었다. 이러한 결과는 뿌리에서 발현하는 유전자의 상당수가 다른 조직에서도 발현하고 있다는 것을 시사한다.

수술 조직 특이적 유전자발현을 분석하기위하여 중국의 페깅대학의 Xiao-Chun Lu 그룹은 벼 10K cDNA microarray를 이용하여 분석한 결과 DNA polymerase 등 26개 유전자가 초기 수술조직에서 up-regulate되는 것을 확인하였다 (Lu et al. 2006). 또한 중국과학원 (CAS)의 Mei-Na Li 등은 벼 암술조직 특이적 유전자 발현을 검정하기 위하여 벼 10K cDNA microarray와 57K GeneChip을 이용하여 암술을 포함하는 뿌리 등 7개 조직에서의 발현을 검정하여 암술 발현 548개 유전자를 분리하여 qRT-PCR을 통하여 암술 특이 34개 유전자를 분리 하였다. 이와 같은 결과는 조직

특이발현 유전자 profiling 분석을 위하여 cDNA microarray, GeneChip array, qRT-PCR 등이 효율적으로 사용되어 질수 있음을 시사한다.

전사조절인자 발현 profiling 분석

벼에서 유전자 발현을 조절하는 전사조절인자를 TIGR의 유전체해석 Ver 6에서 분석한 결과 2,722개를 분석하였으며 79개 그룹으로 나누었다 (TIGR Ver6, Perez et al. 2010). 특히 벼에 존재하는 전사조절인자는 발달, 환경스트레스에 의한 신호전달, 종자의 발달 등에 특이하게 작용하는 것으로 알려져 있으며 이들의 발현에 의한 하위 유전자의 발현분석에는 microarray에 의한 transcriptome profiling의 분석이 필요하다.

벼 60K microarray에는 2,286개의 전사조절인자에 해당하는 oligomer가 디자인 되어 있으며 이들 전사조절 인자들의 발현 형태를 살펴 보기위해서 각 조직에서 callus에 비해 많이 발현된 유전자들을 선별하여 hierarchical clustering을 실시한 결과 각각의 조직에서 특이한 양상으로 발현하고 있었다. 각 조직에서 발현하는 전사조절인자 유전자의 수는 발아종자에서 109개, 뿌리에서 87개, 잎에서 89개, 출수기에서 184개, 꽃에서 81개, 재분화 callus에서 47개였다. 이러한 다양한 전사조절인자 유전자의 발현은 조직별 발달에 있어 상호 중요한 역할을 하는 것을 시사한다. 꽃 형성과 관련 되어지는 출수기와 꽃에서의 전사조절 관련 유전자들의 발현을 살펴 본 결과 MADS box의 공통발현이 많은 걸 확인할 수 있었다. MADS box는 애기장대에서 꽃의 기관 발달에 관여한다는 사실이 먼저 알려져 있고 벼에서도 적어도 90종류 이상이 존재하는 것으로 알려지고 있다. 또한 출수기에서 애기장대의 꽃 기관 발달에 직접적으로 관여하는 AP2 유전자의 발현이 많았다. 이러한 결과는 꽃의 발달시기 전후에 발현하는 AP2, MADS box 유전자가 꽃의 발달에 긴밀히 연관되어 있음을 시사한다. 또한 잎과 출수기에서 bHLH의 빈도가 높게 나타났다. 발아종자 특이적으로 발현하는 전사조절인자의 경우 SNF2 유전자의 발현이 가장 많았다. 앞으로 조직특이 발현 전사조절인자 발현분석 및 상호작용분석은 벼의 발달 단계를 이해하는데 중요한 역할을 하며 microarray 실험의 반복실험을 통하여 가능할 것으로 생각된다.

최근에는 microarray를 이용한 TIFY, Myb, RING finger, Leucine zipper, MADS box 등의 다양한 벼 전사조절인자의 발현 profiling 분석 및 다른 식물과의 연관분석연구를 수행하고 있다 (Arora et al. 2007; Yanhui et al. 2006; Nijhawan et al. 2008; Ye et al. 2009; Lim et al. 2010). 한편으로는 microarray를 이용하여 전사조절인자의 발현을 분석할 경우 대량의 유전자의 발현을 한번에 측정할 수 있는 장점이 있는 반면 조직에서의 발현양이 적은 전사조절인자의 경우 발현을 정확히 측정하기 힘든 단점을 가지고 있다. 독일 막스플랑크 연

구소에서는 전사조절인자 DB (<http://ricetfdb.bio.uni-potsdam.de/>)의 정보를 바탕으로 벼 전사조절인자 2,500개에 대한 qPCR primer를 제작한 후 발현양이 적은 전사조절인자의 발현을 정확히 측정할 수 있는 방법을 확립하였다 (Caldana et al. 2007). 이러한 방법은 다른 동·식물의 전사조절인자 발현측정에 기초 자료로 많이 사용될 수 있을 것으로 생각된다.

형질 및 대사관련 유전자 발현 profiling 분석

벼 및 식물의 주요 형질인 수량, 재배, 발달, 미질 등의 농업주요형질유전자들의 발현은 서로 연관되어 복합적으로 발현조절을 하는 것으로 알려져 있다. 이러한 농업형질의 유전자발현을 대량으로 분석하기 위하여 microarray 기술이 유용하게 사용될 수 있다. 식물의 성장 발달에 관여하는 형질유전자의 발현조절 기작을 연구하기 위하여 성장과 발달을 조절하는 Homeobox 유전자의 발현분석을 행하였다 (Jain et al. 2008). Homeobox 유전자는 벼에 107종 존재하며 이들 유전자의 발현을 microarray를 통하여 분석한 결과 37종은 엽과 저온 스트레스 상태에서 발현하였으며 4종 (Os01g62310, Os09g21180, Os01g57890, Os02g35770)은 종자형성시기에 발현하는 것을 확인 하였다. 각종 스트레스를 주었을 때 발현하는 유전자의 발현분석을 위하여 Zhou 등은 건조와 염 스트레스를, Degenkolbe 등은 장시간의 건조 스트레스를, Kawasaki 등은 염스트레스때의 유전자발현분석을 검정하였다 (Kawasaki et al. 2001; Zhou et al. 2007; Degenkoble et al. 2009). 이러한 다양한 스트레스 발현유전자의 기능을 해석함으로써 농업중요형질 개선에 중요한 연구기반을 구축 할 수 있을 것으로 생각된다. 또한 작물 생산량 및 미질은 종자형성기에 온도의 영향을 받는 것으로 알려져 있으며 고온에서의 종자 내 발현 유전자들의 발현양상을 microarray를 이용하여 검정하였다 (Yamakawa et al. 2007). 그 결과 고온에서는 waxy 유전자, 전분분지효소 등이 상대적으로 낮게 발현하였으며 α -amylase와 열충격단백질 유전자 등은 높은 발현을 나타내었다. 이러한 유전자의 발현은 종자내 전분의 점도를 낮추어 미질을 나쁘게 함과 아울러 종자의 무게 감소를 나타낸다. 13KDa prolamin 단백질의 함량 감소도 나타났다. 종자내의 전분 합성 및 분해는 종자의 미질 및 발아에 관여한다. 이들 유전자의 발현을 검정하기 위하여 135K oligo microarray를 이용하여 종자내 전분합성 및 분해에 관여하는 유전자 ADP-glucose pyrophosphorylase 유전자 등 72종의 유전자의 발현을 종자, 잎, 뿌리 등 14종의 조직에서의 발현을 검정한 결과 종자 특이적으로 발현하는 ADP-glucose pyrophosphorylase 유전자 및 amylase 유전자 등 12종의 발현을 검정하였다. 또한 opaque 형질을 나타내는 Ac/Ds 삽입변이체를 재료로 microarray를 이용하여 유전자의 발현을 검정하였다. 그 결과 8종의 opaque 계통 Ds 삽입변이체에서 열충격단백질의 발현이 공통적

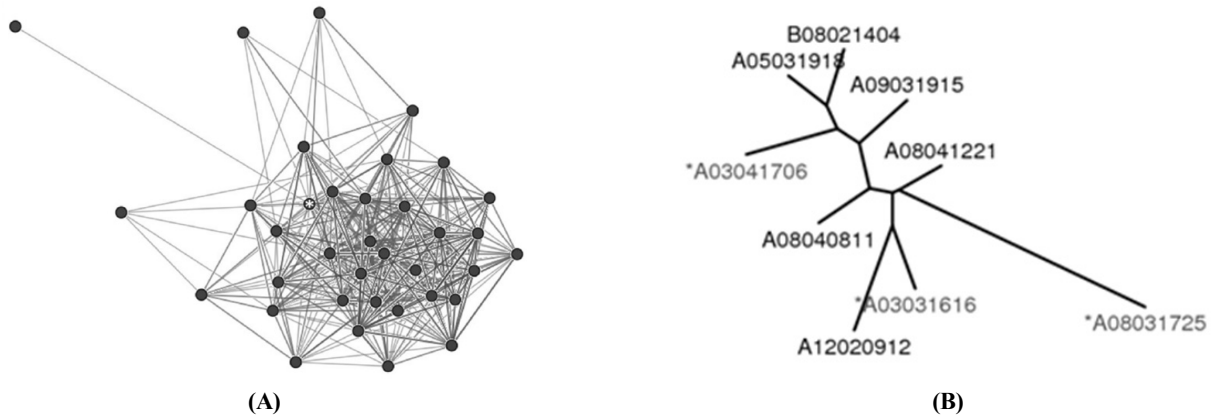


Fig. 2 Correlation analysis of genes in RiceArrayNet (Lee et al. 2009)
 (A) Graphical presentation of coexpressed genes with a ribosomal protein. (B) Correlation tree of trehalose and relating genes. Trehalose-6-phosphate synthase (A03031616), Trehalose synthase (A03041706, A08031725)

으로 감소하는 것을 확인하였다. 이와 같은 결과는 열충격단백질이 opaque 형성에 관여하는 것을 시사한다. 최근 Sakar 등은 몇 종의 벼 열충격단백질이 스트레스와 발달에 관여하는 것을 확인하였다 (Sarkar et al. 2009). 오존농도는 벼의 수량에 영향을 미치는데 Frei 등은 오존반응성 유전자발현분석을 위하여 오존저항성 벼와 감수성 벼를 이용하여 microarray 실험을 수행하였다 (Frei et al. 2010). 그 결과 오존처리에 의하여 ascorbate oxidase 유전자발현이 높게 나타났으며 이 유전자는 오존저항성 QTL 부근에 위치하고 있었다. 오존저항성은 ascorbic acid 대사와 관련 있는 것으로 추정하였다. 이와 같이 최근 microarray 연구는 다양한 형질유전자의 발현양상분석과 QTL 지도를 이용한 형질유전자의 분리를 시도하는 많은 연구들이 이루어지고 있다.

벼 유전자발현 network 분석 및 DB 구축현황

유전자와 기능을 연계 지을 수 있는 하나의 방법인 유전자들 간의 상호발현 network 분석연구는 대장균 및 효모를 모델로 이용하여 그간 많은 연구가 이루어졌다. 그리고 최근 벼와 애기장대에서 각종 microarray 실험 및 생물정보분석 연구가 많이 이루어져 식물유전자들 간의 상호발현분석을 통하여 대사관련 및 형질유전자들 간의 기능을 유추하는 연구가 활발히 이루어지고 있다 (Aoki et al. 2007; Shinozaki et al. 2007; Ludwikow et al. 2008; Lim et al. 2010). 벼에서 각종스트레스와 종자발달에 관련된 유전자들 간의 network분석 연구를 수행한 결과 스트레스와 종자발달에 OsERP유전자 (GenBank AF090698) 등 200여개의 유전자가 상호작용하는 것으로 밝혀졌다 (Cooper et al. 2003). 또한 이들 유전자중 PR1 유전자 (GenBank gi7442184) 등 5개는 애기장대에서 병저항성을 나타내는 유전자로 밝혀졌다. 식물 WRKY 전사조절인자는 식물의 대사, 건조, 발달, 호르몬 및 병저항성에 관여 하는 것으

로 알려져 있다. 벼와 애기장대에서 WRKY 전사조절인자의 상호발현을 검정한 결과 상동성이 있는 20종 가운데 8종이 상호발현하는 것을 발견하였다 (Berri et al. 2009). 이와 같은 결과는 단자엽식물인 벼의 네트워크분석결과가 애기장대와 같은 쌍자엽식물 유전자 기능분석에 사용될 수 있음을 시사하고 있다. 최근 그린진바이오에서는 벼 60K microarray를 이용하여 20개 조직에서 실험한 183개의 발현정보를 종합화 하여 유전자들 간의 상호발현분석을 효율적으로 수행하는 RiceArrayNet (RAN) 프로그램 (<http://www.ggbio.com/arraynet/>)을 개발하여 정보를 공개하고 있다 (Lee et al. 2009). 이 프로그램을 이용하여 유전자들 간의 상호발현분석을 행한 결과 평균 214개의 유전자가 상호발현하는 것으로 밝혀졌다 (그림 2). 그 외 벼 등 10종의 동식물의 microarray 데이터를 기반으로 유전자 상호발현을 분석할 수 있는 STARNET2 프로그램 (<http://vanburenlab.medicine.tamhsc.edu/starnet2.html/>)이 공개되고 있다 (Jupiter et al. 2009). 현재 microarray를 이용하여 유전자간의 상호작용 분석뿐만 아니라 DNA 및 RNA와 상호작용하는 단백질을 분석할 수 있는 ChIP-chip 및 RIP-chip 등의 기술발전이 이루어지고 있다. 유전자간 network 분석은 단백질-단백질 간의 상호작용에서 나타나는 기작을 포함하고 있지 않아서 효모의 two-hybrid 계를 이용한 추가실험 및 단백질체 연구결과와의 공유가 필요하다. 최근 많은 식물에서 microarray를 이용한 유전자발현분석을 행하여 database를 구축 공개하고 있다 (표 3). 또한 식물 microarray 실험 결과를 이용한 유전자기능분류 및 유전자네트워크분석 등의 여러 가지 분석기술이 발달하고 있으며 Coulibaly (2008) 등은 이들 데이터베이스를 종합화하여 리뷰에 소개하였다. 그러나 벼의 경우 실험이 이루어진 microarray 정보가 아직 다양하지 않아 데이터베이스 접근에는 제한적인 상황이다, 세부적인 유전자발현 상호관련 정보를 얻기 위하여는 보다 폭넓은 다양한 연

Table 3 Plant microarray expression database

Database name	URL
Expression database	
Rice Microarray Opening Site	http://cdna01.dna.affrc.go.jp/RMOS/main.html
Gene Expression Omnibus	http://www.ncbi.nlm.nih.gov/geo/
PLEXdb	http://www.plexdb.org/
Genevestigator	https://www.genevestigator.com/gv/index.jsp
AtGenExpress	http://www.weigelworld.org/resources/microarray/AtGenExpress/
ArrayExpress	http://www.ebi.ac.uk/arrayexpress
CDB.DB	http://csbdb.mpimp-golm.mpg.de/index.html
Botany Array Resource	http://bbc.botany.utoronto.ca/ntools/BAR_instructions.html
GeneCat	http://genecat.mpg.de/
PED	http://biocluster.ucr.edu/projects/Unknowns/external/express.html
ATTED-II	http://atted.jp/
Array Project and Database	http://www.ricearray.org
Solanaceae Expression Database	http://solanaceae.plantbiology.msu.edu/
Tomato Expression Database	http://ted.bti.cornell.edu/
Coexpression database	
RiceArrayNet (RAN)	http://www.ggbio.com/arraynet/
Arabidopsis Co-expression Tool	http://www.arabidopsis.leeds.ac.uk/act/
Cress-express	http://www.cressexpress.org/

구가 이루어져야 할 것이다. 또한 이를 보완하기 위하여 벼의 경우 조직발현, 삽입변이체 발현분석, 과발현체 발현분석, 상호발현분석 등의 데이터 축적 및 정보종합화가 요구되어진다. 향후 microarray 데이터분석 공개 프로그램의 보급 및 microarray 실험 결과 공개는 벼 유전자 기능분석 연구 발전에 크게 기여 할 것이다.

결론

최근 벼 구조유전체 연구의 완료와 더불어 이들 정보를 이용한 주요농업형질관련 고유유전자 확보 및 대사경로를 해석하기 위하여 microarray를 이용한 기능유전체 연구의 필요성이 한층 높아지고 있다. 이러한 벼 유전자 대량발현분석 및 상호작용분석을 위하여 microarray는 유용하게 사용되어 지고 있다.

본 리뷰에서 microarray를 이용한 유전자발현 profiling 분석현황 및 유전자 네트워크 작성과 이들 정보를 종합화하여 공개하고 있는 데이터베이스 구축현황을 소개하였다. 벼 유전자들 간의 상호발현에 관한 기작의 이해는 다양한 유전자가 직간접으로 상호작용하여 나타나는 벼의 주요 농업형질의 개선 연구에 크게 활용되어, 미래 환경친화적 벼의 새로운 기능성 부여는 물론 주요농업형질 관련 유전자 기능분석에 크게 이용될 수 있다. 최근에 이루어진 다양한 형태의 벼 유전자발현 profiling 분석 결과는 농업형질 개선을 위한 벼 및 식물전반의 기능유전체 연구 활성화에 기여하고, 실용화 연구를 수행하는 연구자들에게 보다 유익한 정보를 제공할 수 있을 것으로 생각된다.

사사

이 연구는 농촌진흥청 어젠다 2-7-11 “벼 변이집단 및 생물정보를 이용한 유용 농업형질 유전자탐색” 과제, “벼·배추의 생장 및 대사관련 유전자발현 네트워크구축” 과제, 작물유전체사업 CG3310 “형질전환 작물의 유전분석 및 실용화”과제 및 “바이오그린21사업” 과제지원에 의해 수행되었다.

인용문헌

- Aoki K, Ogata Y, Shibata D (2007) Approaches for extracting practical information from gene co-expression networks in plant biology. *Plant Cell Physiol* 48:381-390
- Arora R, Agarwal P, Ray S, Singh AK, Singh VP, Tyagi AK, Kapoor S. (2007) MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genomics* 8:242
- Berri S, Abbruscato P, Faivre-Rampant O, Brasileiro AC, Fumasoni I, Satoh K, Kikuchi S, Mizzi L, Morandini P, Pè ME, Piffanelli P. (2009) Characterization of WRKY co-regulatory networks in rice and Arabidopsis. *BMC Plant Biol* 9:120
- Briggs SP, Singer T. (2005) Genetic networks. *Plant Physiol* 138:542-4
- Caldana C, Scheible WR, Mueller-Roeber B, Ruzicic S. (2007) A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. *Plant Methods* 3:7
- Chou HH, Hsia AP, Mooney DL, Schnable PS. (2004) Picky: oligo microarray design for large genomes. *Bioinformatics* 20:2893-902

- Cooper B, Clarke JD, Budworth P, Kreps J, Hutchison D, Park S, Guimil S, Dunn M, Luginbühl P, Ellero C, Goff SA, Glazebrook J. (2003) A network of rice genes associated with stress response and seed development. *Proc Natl Acad Sci U S A* 100:4945–50
- Coulibaly I, Page GP. (2008) Bioinformatic Tools for Inferring Functional Information from Plant Microarray Data II: Analysis Beyond Single Gene. *Int J Plant Genomics* 2008: 893–941
- Degenkolbe T, Do PT, Zuther E, Reipsilber D, Walther D, Hinch DK, Köhl KI. (2009) Expression profiling of rice cultivars differing in their tolerance to long-term drought stress. *Plant Mol Biol* 69:133–53
- Frei M, Tanaka JP, Chen CP, Wissuwa M. (2010) Mechanisms of ozone tolerance in rice: characterization of two QTLs affecting leaf bronzing by gene expression profiling and biochemical analyses. *J Exp Bot* 2010
- Furutani I, Sukegawa S, Kyoizuka J. (2006) Genome-wide analysis of spatial and temporal gene expression in rice panicle development. *Plant J* 46:503–11
- International Rice Genome Sequencing Project. (2005) The map-based sequence of the rice genome. *Nature* 436:793–800
- Jain M, Tyagi AK, Khurana JP. (2008) Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. *FEBS J* 275:2845–61
- Jain M, Khurana JP. (2009) Transcript profiling reveals diverse roles of auxin-responsive genes during reproductive development and abiotic stress in rice. *FEBS J* 276:3148–62
- Jung KH, An G, Ronald PC. (2008a) Towards a better bowl of rice: assigning function to tens of thousands of rice genes. *Nat Rev Genet* 9:91–101
- Jung KH, Lee J, Dardick C, Seo YS, Cao P, Canlas P, Phetsom J, Xu X, Ouyang S, An K, Cho YJ, Lee GC, Lee Y, An G, Ronald PC. (2008b) Identification and functional analysis of light-responsive unique genes and gene family members in rice. *PLoS Genet* 4:e1000164
- Jupiter D, Chen H, VanBuren V. (2009) STARNET 2: a web-based tool for accelerating discovery of gene regulatory networks using microarray co-expression data. *BMC Bioinformatics* 10:332
- Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ. (2001) Gene expression profiles during the initial phase of salt stress in rice. *Plant Cell* 13:889–905
- Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, Kishimoto N, Yazaki J, Yasunishi A. (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science* 301:376–9
- Lee TH, Kim YK, Pham TT, Song SI, Kim JK, Kang KY, An G, Jung KH, Galbraith DW, Kim M, Yoon UH, Nahm BH. (2009) RiceArrayNet: a database for correlating gene expression from transcriptome profiling, and its application to the analysis of coexpressed genes in rice. *Plant Physiol* 151:16–33
- Lim SD, Yim WC, Moon JC, Kim DS, Lee BM, Jang CS. (2010) A gene family encoding RING finger proteins in rice: their expansion, expression diversity, and co-expressed genes. *Plant Mol Biol* 72:369–80
- Lu T, Huang X, Zhu C, Huang T, Zhao Q, Xie K, Xiong L, Zhang Q, Han B. (2008) RICD: a rice indica cDNA database resource for rice functional genomics. *BMC Plant Biol* 8:118
- Lu XC, Gong HQ, Huang ML, Bai SL, He YB, Mao X, Geng Z, Li SG, Wei L, Yuwen JS, Xu ZH, Bai SN. (2006) Molecular analysis of early rice stamen development using organ-specific gene expression profiling. *Plant Mol Biol* 61:845–61
- Ludwikow A, Sadowski J. (2008) Gene networks in plant ozone stress response and tolerance. *J Integr Plant Biol* 50:1256–67
- Nakamura H, Hakata M, Amano K, Miyao A, Toki N, Kajikawa M, Pang J, Higashi N, Ando S, Toki S, Fujita M, Enju A, Seki M, Nakazawa M, Ichikawa T, Shinozaki K, Matsui M, Nagamura Y, Hirochika H, Ichikawa H. (2007) A genome-wide gain-of-function analysis of rice genes using the FOX-hunting system. *Plant Mol Biol* 65:357–71
- Narsai R, Howell KA, Carroll A, Ivanova A, Millar AH, Whelan J. (2009) Defining core metabolic and transcriptomic responses to oxygen availability in rice embryos and young seedlings. *Plant Physiol* 151:306–22
- Nijhawan A, Jain M, Tyagi AK, Khurana JP. (2008) Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. *Plant Physiol* 146:333–50
- Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, Orvis J, Haas B, Wortman J, Buell CR. (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res* 35(Database issue):D883–7
- Park DS, Park SK, Han AI, Wang HJ, Jun NS, Manigbas NL, Woo YM, Ahn BO, Yun DW, Yoon UH, Kim YW, Lee MC, Kim DH, Nam MH, Han CD, Kang HW, Yi G. (2009) Genetic variation through Dissociation(Ds) insertional mutagenesis system for rice in Korea : progress and current status. *Molecular Breeding* 24:1–15
- Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LG, Rensing SA, Kersten B, Mueller-Roeber B. (2010) PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res* 38(Database issue):D822–7
- Rensink WA, Buell CR. (2005) Microarray expression profiling resources for plant genomics. *Trends Plant Sci* 10:603–9
- Rice Annotation Project. (2008) The Rice Annotation Project Database (RAP-DB). *Nucleic Acids Res* 36(Database issue): D1028–33
- Sarkar NK, Kim YK, Grover A. (2009) Rice sHsp genes: genomic organization and expression profiling under stress and development. *BMC Genomics* 10:393
- Satoh K, Doi K, Nagata T, Kishimoto N, Suzuki K, Otomo Y, Kawai J, Nakamura M, Hirozane-Kishikawa T, Kanagawa S, Arakawa T, Takahashi-Iida J, Murata M, Ninomiya N, Sasaki D, Fukuda S, Tagami M, Yamagata H, Kurita K, Kamiya K, Yamamoto M, Kikuta A, Bito T, Fujitsuka N, Ito K, Kanamori H, Choi IR, Nagamura Y, Matsumoto T, Murakami K, Matsubara K, Carninci P, Hayashizaki Y, Kikuchi S. (2007) Gene organization in rice revealed by full-length cDNA

- mapping and gene expression analysis through microarray. *PLoS One*. 2:e1235
- Shinozaki K, Yamaguchi-Shinozaki K. (2007) Gene networks involved in drought stress response and tolerance. *J Exp Bot* 58:221-7
- Yamakawa H, Hirose T, Kuroda M, Yamaguchi T. (2007) Comprehensive expression profiling of rice grain filling-related genes under high temperature using DNA microarray. *Plant Physiol* 144:258-77
- Yamaguchi-Shinozaki K, Shinozaki K. (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev Plant Biol* 57:781-803
- Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q, Yunping S, Li Z, Xiaohui D, Jingchu L, Xing-Wang D, Zhangliang C, Hongya G, Li-Jia Q. (2006) The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Biol* 60:107-24
- Yazaki J, Kishimoto N, Nakamura K, Fujii F, Shimbo K, Otsuka Y, Wu J, Yamamoto K, Sakata K, Sasaki T, Kikuchi S. (2000) Embarking on rice functional genomics via cDNA microarray: use of 3' UTR probes for specific gene expression analysis. *DNA Res* 7:367-70
- Ye H, Du H, Tang N, Li X, Xiong L. (2009) Identification and expression profiling analysis of TIFY family genes involved in stress and phytohormone responses in rice. *Plant Mol Biol*. 71:291-305
- Yoon UH, Lee GS, Lee JS, Hahn JH, Kim CK, Kikuchi S, Satoh K, Kim JA, Lee JH, Lee TH, Kim YH. (2009) Structural Analysis of Seed Developmental Stage ESTs in *Oryza sativa* L. *Korean J of Plant Biotechnology* 36:130-136
- Zhou J, Wang X, Jiao Y, Qin Y, Liu X, He K, Chen C, Ma L, Wang J, Xiong L, Zhang Q, Fan L, Deng XW. (2007) Global genome expression analysis of rice in response to drought and high-salinity stresses in shoot, flag leaf, and panicle. *Plant Mol Biol*. 63:591-608
- Zhu T. (2003) Global analysis of gene expression using GeneChip microarrays. *Curr Opin Plant Biol*. 6:418-25