벼 유전체 해독 정보를 이용한 야생벼속 식물의 비교 유전체 연구동향

김 혜 란(Arizona Genome Institute)

I. 서 론

벼는 생산량의 100%가 인간에 의해서 소비되는 가장 중요한 식량자원으로서, 전 세계 인구 50% 이상의 주식으로서 이용되고 있다. 또한 곡류 작물 (cereal crop)로서 유전체 크기가 가장 작아 (389 Mb) 'Crop Circle (grass family 의 consensus synteny map Moore *et al*, 1995)'에서 가장 중심에 존재하는 모델 단자엽 식물로서도 그 중요성을 가진다. 주요 식량자원이면서 학문적으로도 중요한 벼의 가치는 쌍자엽 식물의 모델이 되는 *Arabidopsis*에 이어, 단자엽 식물의 모델로서 전체 유전체 해독 연구를 출범시켰고, 현재 whole-genome shotgun 방법에 의한 103,044 개의 scaffold로 구성된 466 Mb의 indica (*Oryza sativa* L. ssp. *indica* cv. 93-11) draft sequence (Yu *et al.* 2002) 와 42,109 개의contig로 구성된 390 Mb의 *japonica* (*Oryza sativa* L. ssp. *japonicacv.* Nipponbare) draft sequence (Goff *et al.* 2002), 그리고 370 Mb 의 보다 정확하고 contiguous 한 International Rice Genome Sequencing Project (IRGSP) pseudo-molecule이 발표되어 있다 (Table 1). 이중 IRGSP pseudo-molecule은 벼지놈의 물리지도에 위치한 PAC (P1-derivedArtificial Chromosome)과 BAC (Bacterial Artificial Chromosome)을 이용한 finished sequence로서 95% 이상의 벼지놈을 정확한 순서와 위치로 나타내 주고 있다. IRGSP sequence와 draft sequence를 비교해보면 *indica*와 *japonica* 의 draft sequence는 각각 전체 지놈의 ~70% 정도를 보여주고 있는 것을 알 수 있다. 또한 이draft sequence들은 repeat, centromere, exogenous염기서열과 같은 생물학적으로 중요한 의미를 가진 부분들의 정보가 결여되어 있고, 너무 많은 물리적인 gap과 misassembly 가능성과 같은 문제점을 가지고 있다. 이에 비해 물리지도를 이용해 'clone-by-clone' 방법으로 제작된 IRGSP sequence는 전체 지놈의 대부분을 염기서열의 내용과 관계없이 실제 지놈상의 상태대로 보여주며, 좀 더 신빙성 있는 유전자의 annotation 데이타를 내포하고 있어 그 우수성을 한눈에 알 수 있다. 보다 정확한 IRGSP 지놈 sequence는 벼지놈의 청사진으로서 1,184,706 개

Table 1. 벼의Whole-genome Sequence 별 결과 비교

	indica draft sequence	*japonica* draft sequence	IRGSP sequence
Method	Whole-genome shotgun	Whole-genome shotgun	Map based clone-by-clone
Covered size	466 Mb	420 Mb	370 Mb
No. of pieces	103,044	42,109	12 pieces with 62 gaps
Accuracy	-	98%	99.99%
Quality of sequence	Phred Q 20	-	Phred Q 40
# of genes predicted	46,022-55,615	32,000-50,000	37,544 (non-transposable-element related protein-coding sequences)
Reference	Yu *et al.* 2002	Goff *et al.* 2002	IRGSP 2005

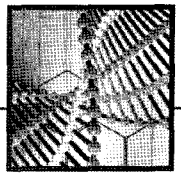


Table 2. Some wild species of Oryza and their useful traits (unpublished; Dr. Dashan Brar, IRRI)

SPECIES	GENOME	USEFUL TRAITS
O. sativa	**AA**	**Cultivated worldwide**
O. glaberrima	**AA**	**Cultivated in West Africa; increased tolerance to biotic and abiotic stresses; weed competitiveness**
O. longistaminata	**AA**	**Resistance to BB, stemborer; drought avoidance**
O. rufipogon	**AA**	**Resistance to BB, sheathblight, abiotic stresses, CMS, high iron and zinc (55 mg vs 22)**
O. minuta	**BBCC**	**Resistance to BPH, BB, blast, sheath blight**
O. officinalis	**CC**	**Resistance to BPH, BB, GLH**
O. latifolia	**CCDD**	**Resistance to BPH, high biomass, high iron and calcium(240 mg vs 78)**
O. australiensis	**EE**	**Resistance to BPH, drought avoidance**
O. brachyantha	**FF**	**Resistance to BB, stemborer**
O. granulata	**GG**	**Shade tolerance, adaptation to aerobic soil conditions**
O. ridleyi	**HHJJ**	**Resistance to BB, blast, stemborer**
O. coarctata	**HHKK**	**Salt tolerance**

*BPH=brown planthopper; WBPH=white backed planthopper; Bl=blast; BB=bacterial blight
* GSV=grassy stunt virus; RTV=rice tungro virus; RSV= rice stripe virus

의 EST (dbEST release 05/19/2006)와 32,000개의 full length cDNA (Kikuchi *et al.* 2003) 와 함께 벼의 기능 유전체 연구에 필수 불가결한 자료임이 분명하다.

미생물에서 고등 동식물에 이르기까지 여러 생물체의 전체 지놈 sequence가 밝혀진 이 시대를 우리는 포스트 제놈 시대 (post-sequencing era) 라고 부르며 이전과는 달리 유전체 전체의 염기서열을 이용한 massive한 형태의 새로운 연구들이 급속도로 발전하고 있다. 대표적인 post-sequencing era 연구테마로는 염기서열이 밝혀진 유전자들의 기능과 상호 관계에 대한 연구 (Functional Genomics), 다른 종들과의 비교 연구 (Comparative Genomics)와 다양한 유전자들의 활용을 꼽을 수 있을 것이다. 현재 기능 유전체 연구서는, T-DNA 나 transposon 등에 의해 유기된 mutant 들을 이용해 유전자의 기능을 증명하거나 밝히는 연구(loss-of-function analysis 또는 gain-of-function analysis)와 microarray 기술을 이용한 profiling 연구 (expression profiling, transcript profiling, metabolite profiling, phenotypic profiling) (Kjemtrip *et al.* 2003)가 대표적이다. 최근에 chemical mutant의 표현형 변이를 single sequence motif 수준에서 검출하는 TILLING (Targeting Induced Local Lesions in Genome) 방법과 RNAi와 같은 유전자의 knock-out 방법도 새로운 유전자들의 기능연구에 각광을 받는 분

야로 주목을 받고 있다.

기능 유전체 연구와 함께 밝혀진 유전체의 구조를 이용한 또 하나의 연구 분야로는 비교유전체 연구를 꼽을 수 있다. 벼는 주요 곡류 작물들과 진화적인 역사를 함께synteny의 출발점에 위치해 있다. 이와 같은 특성으로 벼는 곡류작물의 생물학적 모델로서 기존에 밝혀진 벼의 유전체의 구성과 유전자 정보들은 다른 곡류작물간의 비교에서 출발점 제공하고 이를 용이하게 해준다. 이런 작물간의 비교 연구의 결과는 새로운 유전자의 검출과 규명은 물론 진화, 종분화, 순화와 같은 연구에 박차를 가할 것이다. 또 현존하는 방대한 양의 벼의 정보를 다른 주요작물에 손쉽고 빠르게 효율적으로 적용할 수 있게 해주며 이는 차기 연구를 놀라운 속도로 발전 시켜줄 것이다.

벼의 전체 유전체 분석이 끝나갈 무렵인 2004년9월에 Arizona Genomics Institute, Purdue University 그리고 Cold Spring Harbor Laboratory의 세개의 기관은 벼의 전체 지놈sequence와 야생벼 (*Oryza* genus)를 이용한 비교유전체 연구인Oryza Map Alignment Project (OMAP)를 착수했다 (Wing *et al.* 2005). 이 프로젝트는 지놈프로젝트의 결과를 이용한 큰 규모 (NSF#DBI-0321678; $9,743,546)의 차기 연구라는데 의의 있으며, IRGSP의 우수한sequence와 그 구조적, 기능적 정보가 이 연구의 탄탄한 시작점이 될 것이다. 이글을 통해 필자는 OMAP 프로젝트의 직접적인 수행 맴버로서 이 프로젝트의 개요와 현재까지의 진행상황과 성과, 그 결과를 이용한 최근의 야생벼에 대한 연구를 소개하여 post-sequencing 시대의 연구 동향과 전망을 그려보고자 한다.

II. *Oryza* 속과 야생벼

Oryza 속은 작은 규모의 속 (genus) 으로서 2개의 재배종과 22개의 야생종으로 분류된다 (Ge *et al* 1999). 전체 24개의 종은 표현형, 핵형, 분자적 차이로

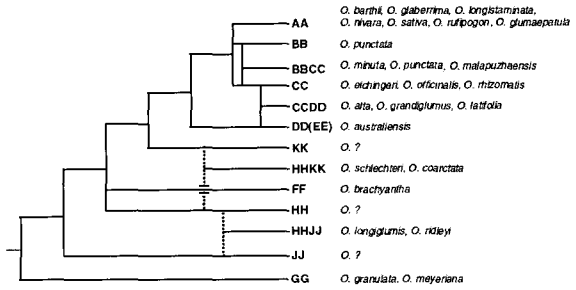
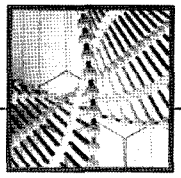


Figure1. Phylogenetic tree of the genus Oryza and the species (Wing et al. 2005 based on Ge *et al.* 1999)

10개의 유전자형으로 다시 나눌 수 있고 (Aggarwal *et al.* 1997 Khush 1997 Ge *et al.* 1999) 그 유전자형은 AA, BB, CC, EE, FF, GG의 이배체형 (2n=24) 과 BBCC, CCDD, HHJJ, HHKK의 이질4배체 (allotetraploid) (2n=48)로 명명이 되어 있다. *O. sativa* 와 *O. glaberrima* 는 유일한 재배종으로서 AA 유전자형의 이배체로 분류되며 같은 유전자형을 가진 야생종으로는 6개의 종이 보고되어 있다 (그림 1). 그리고 나머지 16개의 종은 야생벼로서 9개의 다른 유전형으로 분류가 된다. 그림 1은 두개의 핵 유전자와 한 개의 엽록체 유전자를 이용한 *Oryza* 속의 진화적 계통 관계와 유전자형별 분류를 보여준다.

Oryza 속의 종들은 유전체의 크기가 362 Mb 에서 1278 Mb에 이르기까지 다양하며 (Ammiraju *et al* 2006), 배수체가 존재하며, 염색체 구조 변화에 대한 다양한 기존 연구결과가 존재해서 유전체의 크기 진화와 염색체의 구조 변화 및 진화 연구에도 좋은 재료이다.

야생벼는 수량, 각종 스트레스 저항성, 내병성, 내충성과 같은 농업적으로 유용한 유전자들의 보고로서 그 중요성을 더 한다 (Table 2). 야생벼의 수량증진 유전자(hybrid rice revolution, Xiao *et al.* 1996, 1998)나 생물학적 (biotic) 비생물학적 (abiotic) 저항성 유전자의 도입으로 인한 재배벼의 개량 (Brar and Khush 1997)은 이미 널리 보고되어 있다. 이러한 유용한 형질들은 현재 재배벼의 농업적 문제점을 보다



효율적으로 개선하고 그 형질을 향상시키는데 중요한 요소로 역할을 할 것이다.

Ⅲ. *Oryza* Map Alignment Project (OMAP)

10개국의 국제 공동 연구로서 7년에 걸친 벼 유전체의 염기서열 분석 노력과 함께 야생벼를 이용한 재배벼의 개량과 같은 연구가 지속적으로 진행이 되어 왔음에도 불구하고, 야생벼의 유전자풀 (gene pool)의 유전적 변이나 진화적 상관관계에 대한 연구는 미개척 분야였었다. 이에 OMAP 그룹은 IRGSP 지놈 sequence를 기초로 하여 야생벼 간의 비교 유전체 연구 플랫폼을 갖추려는 *Oryza* Map Alignment Project (OMAP) 를 착수하였다. 이는 단일 genus 에 대해 진화, 발달, 유전체의 구조, 배수체화, 순화, 유전자의 조절적 네트웍 및 작물의 개량과 같은 연구을 위한 유전자 수준의 closed experimental system을 구축하는것을 목표로 하고 있다. 보다 세분화된 연구 목표를 살펴보면 1) 10개의 다른 유전자형을 대표하는 11개의 야생벼와 1개의 아프리카 재배벼의 10X의 genome coverage를 가지는 BAC 라이브러리 제작, 2) 제작된 12개의 BAC 라이브러리에 대한 fingerprinting과 BAC end sequence (BES) 데이타베이스 확립, 3) 제작된 fingerprinting 데이타를 이용한 12개의*Oryza* 종들의 물리지도 작성 및 작성된 지도의 IRGSP genome sequence로의 align, 4) 벼유전체 염기서열을 이용한 12개의 *Oryza* 종들의 염색체 1, 3, 10번 재건 (reconstruction)으로 요약할 수 있다. 이 프로젝트는 2004년 9월에 시작되었으며, 지금부터는 OMAP의 진행상황과 결과를 이용한 연구들에 대해 언급해 보고자 한다.

Ⅳ. The first outcome of OMAP and the Oryza research

OMAP 그룹은 비교유전체학적 관점에서 벼 지놈의 기능적 형질을 밝히기 위한 첫 번째 단계로 10개의 다른 유전자형을 대표하는 12개의 *Oryza* 종들의 BAC 라이브러리를 제작하여 약 100만개의 클론을 확보하였으며 그 특성과 우수성을 보고하였다 (Ammiraju *et al* 2006). 12 종의 라이브러리의 삽입 DNA (insert) 크기는 123 Kb 에서 161 Kb로 분포했고, 세포질 DNA 오염율은 평균 0.4% - 4.1% 로 검출되었다. 벼의 RFLP 마커 12개 (1 마커/염색체)를 이용한 hybridization 실시로 라이브러리의 randomness와genome coverage도 검증되었다 (Table 3). 또한 각 라이브러리의 0.1X의genome coverage에 해당하는

Table 3. The summary of OMAP resources

Project	Material		BAC library[a]			BAC end sequencing				Fingerprinting	
	Genome type	Accession No.	No. of clones	Average insert size (Kb)	New genome size	No. of GenBank submissions	Avg length after trim (in Genbank)	Total sequenced length (in Genbank)	Genome coverage	No. of contigs	No. of singletons
O. nivara	AA	W0106	55,296	161	448	106,124	665 bp	~ 71 Mb	16%	340*	2,356
O. rufipogon	AA	105491	64,512	134	439	70,982	704 bp	~ 50 Mb	11%	327*	1,305
O. glaberrima	AA	96717	55,296	130	357	66,821	590 bp	~ 39 Mb	11%	167*	2,098
O. punctata	BB	105690	36,864	142	425	68,384	710 bp	~ 49 Mb	11%	210*	1,482
O. officinalis	CC	100896	92,160	141	651	103,251	717 bp	~74 Mb	11%	310*	2,052
O. minuta	BBCC	101141	129,024	125	1,124	169,651	559 bp	~ 95 Mb	8%	3,962	9,576
O. alta	CCDD	105143	92,160	133	1,008	128,732	586 bp	~ 75 Mb	7%	2,492	3,111
O. australiensis	EE	100882	92,160	153	965	128,599	676 bp	~ 87 Mb	9%	1,409	2,163
O. brachyantha	FF	101232	36,864	131	362	67,364	672 bp	~ 45 Mb	13%	225*	1,805
O. granulata	GG	102118	73,728	134	882	138,171	674 bp	~93 Mb	11%	2,358	3,032
O. ridleyi	HHJJ	100821	129,024	127	1283	204,729	632 bp	~129 Mb	10%	1,250	1,810
O. coarctata	HHKK	104502	147,456	123	ND	195,285	661 bp	~129 Mb	ND (>10%)	2,190	5,169
Total/Avg			1,004,544	136		1,448,093	654 bp	~ 937 Mb	11%		

[a] BAC libraries, high density filters can be ordered form the AGI BAC/EST Resource center (www.genome.arizona.edu)
* highly manually edited (HME)

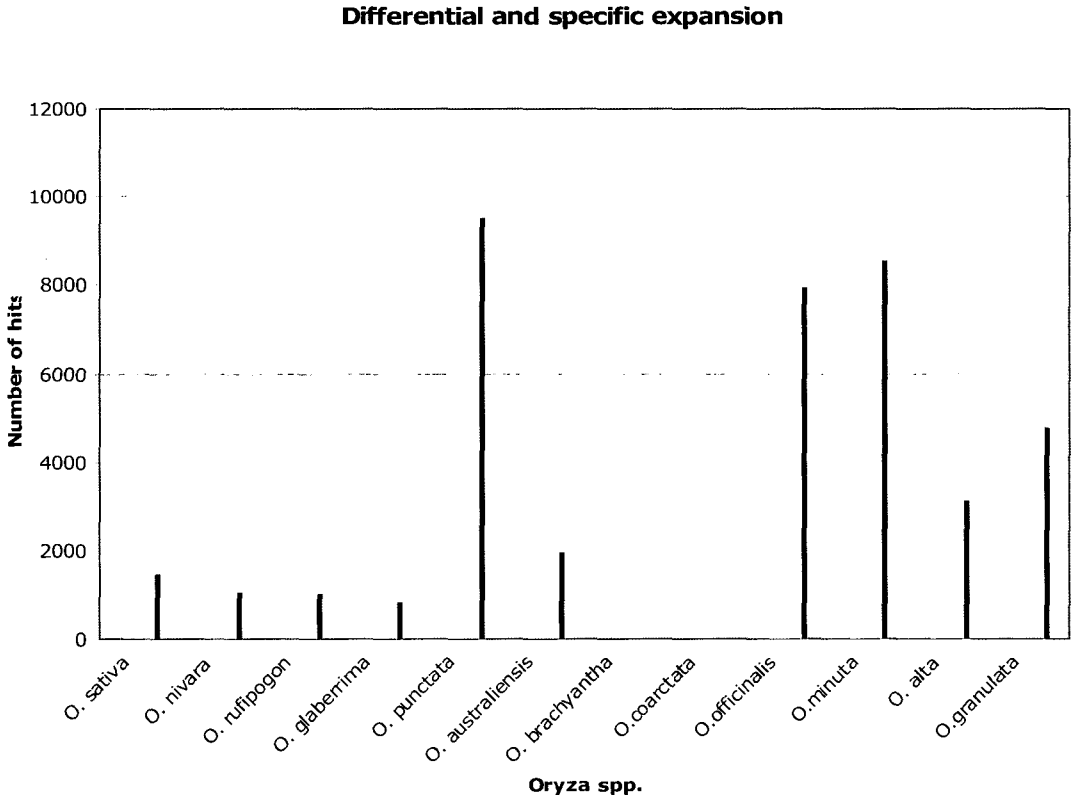


Figure 2. Distribution of Atlantys elements in the Oryza species.

survey sequence를 제작하여 12개 종간의 repeat 양과 종류의 비교 결과도 발표되었다. 더 정확한 실험 설계를 위해 flow cytometry를 이용하여 12종의 지놈 크기를 재측정 하였으며 각 라이브러리의 10X genome coverage 에 해당하는 클론들의 end sequencing 과 fingerprinting을 수행하였다. 그 결과 현재 약 140만 개의 OMAP BES를 GenBank에 등록하였고 (전체 약 937 Mb에 해당) 12개 종의 비교 물리지도를 작성, 발표하여 *Oryza* genus의 비교 유전체학에 필요한 resource를 확보하였다 (Kim H *et al*, *in preparation*). 모든 OMAP 데이타는 www.omap.org에 수록되어 있고, 또 Table 3에 요약되어 있다. 이와 같은 OMAP 데이타는 첨단 생산 기술에 의해 만들어진 우수한 질의 데이타로도도 중요성을 갖는다. BES 제작은 384-well 형식의 1/16 반응 (BigDye 0.5μl reaction)으로 되었고 물리지도 작성을 위한 fingerprinting 은 보다 효율적인 HICF (High Information Content Fingerprinting Luo *et al*, 2003) 방법을 이용하였다. 작성된 물리지도는 지놈 크기가 클수록 보다 많은 contig 수를 갖는 것을 볼 수 있었고, singleton의 양은 *O. minuta*를 제외하고는 전체 fingerprints의 5% 내외로 분포했다. 작성된 물리지도 (Phase 1 map)와, BES 데이타는 SyMap (Synteny Mapping and Analysis Program) 소프트웨어 (Cari Soderlund, *in preparation*)를 이용해 reference 지놈 (IRGSP 지놈)에 align이 되고, align 데이타와 FPC 소프트웨어의 기능을 이용해 phase 1 map은 HME (Heavily Manually Edited) map으로 editing이 되며, 현재 5종 (Table 3)의 HME map이 완성되었다.

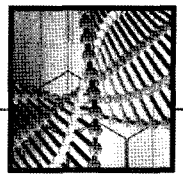


Table 4. OMAP cross combinations in Arizona Genomics Institute (AGI)

AA x AA	Reciprocal crossess of 9311(Indica) x Nipponbare (Japonica)
	Nipponbare x (*O. glaberrima, O. glumaepatula, O. nivara, O. barthii,* O. *meridionalis, O. longistaminata, O. ru*
	9311 x (*O. glaberrima, O. glumaepatula, O. nivara, O. barthii, O. meridionalis, O. longistaminata, O. rufipogo*
BB x BB	*O. punctata* (IRGC 105690) x *O. punctata* (Acc from Africa)
CC x CC	O. officinalis (IRGC 100896) x O. officinalis (Acc from Asia)
	O. officinalis (IRGC 100896) x *O. eichingeri* (Acc from Africa)
	O. eichingeri (Acc from Africa) x *O. eichingeri* (Acc from Asia)
BBCC x BBCC	O. minuta (IRGC 101141) x O. malaphuzaensis (Acc from India)
	O. minuta (IRGC 101141) x *O. punctata* (Acc from Africa)
	O. minuta (IRGC 101141) x *O. minuta* (Acc. from Asia)
CCDD x CCDD	O. alta (IRGC105143) x O. latifolia (Acc from South America)
	O. alta (IRGC105143) x *O. grandiglumis* (Acc from South America)
GG x GG	*O. granulata* (IRGC 102118) x *O. meyeriana* (Acc from Asia)
	O. granulata (IRGC 102118) x *O. neocaledonica* (Acc from New Caledonia)
HHJJ x HHJJ	*O. ridleyi* (IRGC 100821) x *O. longiglumis* (Acc from Africa)
	O. ridleyi (IRGC 100821) x *O. ridleyi* (Acc from Asia)
HHKK x HHKK	*P. coarctata* (IRGC 104502) x *O. schechteri* (Acc from Indonesia)
	P. coarctata (IRGC 104502) x *O. schechteri* (Acc from Papua New Guinea)
EE x EE	*O. australiensis* (IRGC 100882) x Annual and Perennial *O. australiensis*

이러한 OMAP resource를 이용해 OMAP 그룹에서 진행하는 연구는 다음과 같다.

1) *Oryza* 종간 repeats 비교: 평균 11%에 해당하는 각 지놈을 대표하는 BES와 각 지놈의 5%정도를 나타내는 random genomic shotgun 라이브러리를 이용해 *Oryza* 종간의 주요 repeat의 종류와 양을 비교 분석하고 있다. LTR(long terminal repeat)의 유사성을 이용해 12개의 종에 공통적으로 풍부하게 분포하는 element나 종별로 특별하게 분포하는 element, 또는 종별로 다른 양으로 존재하는 element를 밝힘으로 repeat과 speciation의 관계를 연구하는데 그 목적을 둔다. 그 결과의 한 예로서 Atlantys element의 경우는 12개종에 모두 분포하며, BB, CC, BBCC, CCDD, GG 유전자형의 종에 현저하게 많은 양으로 분포하는 것을 볼 수 있었다 (Fig. 2). 이element는 종별로 다른 정도의 expansion의 양상을 보이며, 대부분의 경우 expansion이 speciation 보다 앞서는 것으로 나타났다. 이 연구는 더 많은 종류의 *Oryza* 의 주요 repeat element를 종간, 종내의 수준에서 분석하는 방향으로 향후 확대될 예정이다.
2) 종간의 variation 연구: 12개 종의 BES 와 reference 지놈과의 alignment 데이타를 이용해 SSR, SNP 마커 개발, INDEL, rearrangement Index 제작과 같은 *Oryza*의 종간의 차이를 유전체의 구조적 수준에서 밝히려는 연구가 진행 중이며, 모든 데이타는 Gramene (www.gramene.org) 에 deposit 될 계획이다.
3) 지놈 레벨의 종간 비교: reference 지놈에 align 된 12종의 물리지도를 통해 각종의 유전체 수준의 마크로 한 구조적 변이를 검출한 후, 각 종간에 변이가 심한 부분과 보존적인 부분들을 중점으로 DNA 수준의 구조와 기능, 그리고 *Oryza*의 진화에 대한 연구가 진행중이다.
4) 진화적으로 중요한 유전자 locus의 종간 비교: 곡류 작물 간에 연구가 가장 많이 된 Adh1-Adh2 locus 와 벼의 대표적인 domesticated 유전자인 Hd1 (Heading date 1 Yano *et al*, 2000) locus, 종간 구조적 변이를 보이는 locus, centromere 8 의 염기서열상의 변이에 기초한 *Oryza* 속의 진화 연구도 진행중이다.
5) Mapping Population 개발 및 map based cloning: 식물 육종에 있어 OMAP 가치를 더하기위한 방안으로 *Oryza* 종간, 종 내의 mapping population

이 개발되고 있다. Table 1에서 보여지듯 현재 AGI에서는 9개의 다른 유전자형을 갖는 종들 간의 교배가 시작되었고, 12종의 유전자 지도 작성은 물론 표현형적 변이를 보이는 계통을 이용해 유용 유전자를 map based cloning 할 계획이다. 이때 OMAP으로 인해 이미 작성된 12종의 물리지도는map based cloning의 초석으로 기능을 할 것이고, 이로 인한 결과는 *Oryza*의 연구에 있어 또 다른 진화론적 접근을 가능하게 할 것이다.

V. *Oryza* 연구의 전망

OMAP은 막대한 양의 *Oryza*의 새로운 생물학적인 정보를 제공하는 것은 물론, 보다 경제적이고 효율적인 염기서열 분석 기법과 물리지도 작성법과 같은 첨단기술의 개발 및 구축으로도 매우 중요한 의의를 가진다. 이렇게 확립된 기술과 연구 시스템 구축 방법은 기존의 염기서열이 밝혀진 Arabidopsis나 현재 지놈 수준의 염기서열 분석중인 medicago나 토마토와 같은 reference 지놈이 있는 genus들의 비교 유전체 연구 및 진화연구에 직접적으로 적용이 가능하다고 사료된다.

*Oryza*는 전체 지놈의 염기 서열이 밝혀진 재배벼가 속한 genus로서 인간에게 있어 가장 중요한 곡류 작물 및 단자엽 식물의 생물학적, 진화학적 지표로서 그 중요성이 매우 높이 평가된다. *Oryza*속내의 비교 유전체 연구에 필수 요소이자 박차를 가할 시스템을 구축하는 OMAP 프로젝트는 *Oryza* 연구 뿐만 아니라 모든 생물의 체계적인 구조적, 기능적 진화 연구에 새로운 장을 열어줄 것이다.

OMAP은 12종의 우수한 BAC 라이브러리와, BES 데이타베이스, 물리지도의 제작에 이어 종간의 비교 물리지도, *Oryza* 종간의 rearrangement index, 각종 유전적 마커의 virtual map, *Oryza* 특별 repeat 데이터베이스 등을 제공할 것이며, 이는 speciation, domestication, 진화와 분화, 기능 유전체연구 등과 같은 전반적인 생물학에 새로운 조명을 비춰줄 것이다.

OMAP그룹은 이 프로젝트를 통한 연구 결과를 궁극적으로 재배벼의 개량과 육종에 도입할 계획이다. 이를 위해서는 자연히 육종가의 전문적인 지식이 요구되며 육종가와의 공동연구가 OMAP 결과의 해석과 실용화에 최상의 상승효과를 줄 것으로 믿어 의심치 않는다.

참고문헌

Aggarwal *et al.*1997. Two new genomes in the Oryza complex identified on the basis of molecular divergence analysis using total genomic DNA hybridization. Mol. Gen. Genet. 254: 1-12

Ammiraju *et al.* 2006 The Oryza bacterial artificial chromosome library resource: construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus Oryza. Genome Res.Jan;16(1):140-7.

Brar, D.S. and Khush, G.S. 1997. Alien introgression in rice. Plant Mol. Biol. 35: 35-47

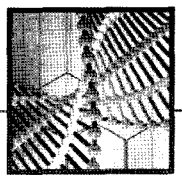
Ge *et al.* 1999 Phylogeny of rice genomes with emphasis on origins of allotetraploid species Proc. Natl. Acad. Sci. USA 96: 14400-14005

Goff *et al.*2002A draft sequence of the rice genome (Oryza sativa L. ssp. japonica).Science. 296(5565):92-100.

International Rice Genome Sequencing Project The map-based sequence of the rice genome Nature 436, 793-800

Khush, G.S. 1997. Origin, dispersal, cultivation and variation of rice. Plant Mol. Biol. 35: 25-34

Kikuchi *et al.* 2003 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. Science 301(5631):376-9. Erratum in: Science



301(5641):1849.

Kjemtrip *et al* 2003 Growth stage-based phenotypic profiling of plants. In: E. Grotewold (Ed.), Methods in Molecular Biology, Vol. 236. Plant Functional Genomics: Methods and Protocols, pp. 427 - 441

Luo *et al* 2003 High-throughput fingerprinting of bacterial artificial chromosomes using the snapshot labeling kit and sizing of restriction fragments by capillary electrophoresis. Genomics 82: 378 - 389

Moore *et al* 1995 Cereal genome evolution. Grasses, line up and form a circle. Curr Biol. 5(7):737-9

Wing *et al* 2005 The oryza map alignment project: the golden path to unlocking the genetic potential of wild rice species. Plant Mol Biol. 59(1):53-62.

Xiao *et al.* 1996. Genes from wild rice improve yield. Nature 384: 223-224

Xiao *et al.* 1998. Identification of trait-improving quantitative trait loci alleles from a wild rice relative, Oryza rufipogon. Genetics 150: 899-909

Yano *et al* 2000 Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. Plant Cell. 12(12):2473-2484.

Yu *et al* 2002 A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science. 296(5565):79-92.