

벼 glutelin 유전자 구조 및 발현특성분석

윤웅한 · 김창국 · 이강섭 · 한장호 · 이정화 · 김연기 · 지현소 · 문정환 · 이태호 · 김태호

Structural and expression analysis of glutelin genes in *Oryza sativa* L.

Ung-Han Yoon · Chang-Kug Kim · Gang-Seob Lee · Jang-Ho Hahn · Jeonghwa Lee · Yeon-ki Kim · Hyeon-So Ji · Jeong-Hwan Mun · Tae-Ho Lee · Tae-Ho Kim

Received: 25 May 2011 / Accepted: 3 June 2011
© Korean Society for Plant Biotechnology

Abstract Rice is one of the most important crop in the world, in particular for food resources. With its small genome size of 383 Mb, the *Oryza sativa* is a model plant for genome research. Indeed, it's grain provides human with a source of carbohydrates and proteins. Rice grain has relatively low protein contents (around 8%) compared to other legume seeds (around 40%). Osborne classified seed proteins into water soluble albumin, salt soluble globulin, alcohol soluble prolamin and acidic/alkaline solution soluble glutelin. Glutelin and prolamin are the major storage proteins in rice.

For the gene expression study of seed storage proteins, we analyzed 33,192 EST clones at immature stages in a rice cultivar (*Oryza sativa* L. cv. 'Ilpum'). Based on the expression analysis, we cloned 11 glutelin genes and figured out the 8 genes are located on Chromosome 2. The expression of glutelin genes appears to be about 28.2% of total level in immature seeds. Interestingly, glu-04 is duplicated as inverted sequences on the same chromosomes as far 4.5 kb. Our results indicate that glutelin genes, evolutionarily, were replicated on the chromosome and thus expressed as specific manners. In a whole protein composition analysis, glu05 (type B7) contains the highest lysin contents (4.51%) among

the 11 rice glutelin genes. It will be an interesting future work to increase lysin contents by the gene overexpressor strategy with the aim of improved diet nutritionally fortified.

서론

벼는 세계인의 주요 식량자원으로 중요시되며 게놈 크기가 작고 중요농업형질유전자 분석 등의 유전적 업적이 많아 작물유전체 연구재료로서 매우 중요시 되고 있다. 벼 염색체완전해독연구(<http://rgp.dna.affrc.go.jp/IRGSP/>)가 2004년 완료된 이후 세계 각국은 벼 유전자 3만여 개의 기능을 밝히고 유용 형질 유전자 분리 및 타 작물 유전자와의 비교 분석 연구를 통한 유용 형질 유전자의 실용화 연구에 박차를 가하고 있다. 특히 벼에서 확인된 중요형질 유전자 기능분석 정보들은 다른 곡물에도 활용 가능하다는 점에서 이들 유용유전자 분리 및 이용의 필요성이 크게 대두되고 있다. 화분과 작물에 있어서 종자는 식물 종의 유지 및 인간에게 식량원으로서 중요한 역할을 하며, 최근 이루어진 화분과 작물의 염기서열 정보를 기반으로 종자 내에서 발현하는 유전자 비교분석에 관한 기초 연구가 많이 이루어지고 있다 (Paterson et al. 2008). 이러한 연구 결과들은 분자진화, 생물종의 다양성 그리고 농업생산성 등의 연구에 있어 새로운 방향성을 제시하고 있다. 화분과 식물 종자의 주요 성분인 전분, 단백질, 지질 등은 등숙기 중 배유 조직에 축적되며 이러한 저장 물질은 종자가 발아 할 때에 저장물질이 분해되어 조직발달을 위한 에너지원으로 사용된다. 또한 벼 종자의 중요 구성성분인 탄수화물과 단백질은 인간에게 주요 영양분 공급원으로서 중요한 역할을 한다. 일반적으로 콩과류 식물의 종자내에서의 저장단백질 함량은 약 40%를 차지하며 곡류 종자의 단백질 함량은 10~12%를 차지

U.-H. Yoon (✉) · C.-K. Kim · G.-S. Lee · J.-H. Hahn · J. H. Lee · H.-S. Ji · J.-H. Mun · T.-H. Kim
국립농업과학원
(Genomics Division, National Academy of Agricultural Science, RDA, Suwon, Korea)
e-mail: uhyoon@korea.kr

Y.-K. Kim · T.-H. Lee
명지대학교
(Bioscience and Bioinformatics Division, MyongJi University, Yongin, Korea)

하는 것으로 보고되어지고 있다 (Shewry and Halford 2002). 또한 벼 종자 배유의 주요 성분은 90% 이상의 전분과 다른 주요 곡물보다 낮은 6~8%의 단백질로 구성되어 있으며 이러한 전분과 단백질의 조성은 쌀의 품질에 중요한 영향을 미친다 (Duan and Sun 2005; Matsuo et al. 1995). Choi 등 (2002)의 연구에 의하면 쌀의 저장단백질은 주로 쌀알의 외층과 전분입자 사이에 분포하여 쌀의 수분흡수 및 취반시에 투수성이나 전분입자의 호화팽창에 영향을 미치는 것으로 알려져 있다. 종자 저장 단백질의 특성 및 유전자 발현 기작을 이해하는 것은 영양성 향상 등 종자형질의 분자유종을 위한 연구에 매우 중요하다.

오스본의 종자 저장단백질 분류에 의하면 각각의 용해도에 따라 수용성의 albumin, 염용해성의 globulin, 약산성 및 알카리 용해성의 glutelin 그리고 알콜 용해성의 prolamins 등 네 종류로 분류를 하였다 (Osborn 1924). 벼 종자 저장단백질의 60~80%는 glutelin으로 구성되어 있으며 20%는 prolamins, globulin으로 구성되어 있다. 이와 달리 옥수수, 보리 등의 종자 저장단백질 경우는 prolamins이 glutelin보다 많이 함유되어 있으며 귀리의 경우 globulin 함량이 가장 높았다 (Bewley and Black 1985). 따라서 벼를 제외한 대부분의 곡류 저장단백질의 주성분은 prolamins으로 분류되어진다 (Shewry et al. 1995). 종자 배유에서 만들어진 glutelin과 prolamins은 단백질과립 (protein body) 중에 저장되며 특히 prolamins은 소포체유래 단백질과립-1에 glutelin과 globulin은 액포형 단백질과립-2에 각각 저장되어진다 (Tanaka et al. 1980). 이후 종자 저장단백질은 종자 발아 도중에 분비되는 단백질분해 효소에 의해 분해되어 식물생장을 위한 영양분으로 이용되어진다.

벼 glutelin 단백질의 구조를 분석한 결과 35~37 kDa의 산성 서브유닛과 20~25 kDa의 염기성 서브유닛으로 구성되어 있으며 각각의 서브유닛이 S-S 결합을 통하여 단백질을 형성하였다. 이들 glutelin 단백질을 이용하여 2차원 폴리아크릴아마이드 겔 전기영동을 행한 후 12개의 산성 서브유닛과 9개의 염기성 서브유닛을 분리하였다 (Wen and Luthe 1985, Abe T et al. 1996). 벼 종자 저장단백질은 구조가 비슷한 폴리펩타이드로 구성되어 전기영동으로는 각각의 단백질을 분리하기 어려운 점이 있으며 유전자수준에서의 발현조절을 통한 단백질 합성조절을 행하기 위하여 유전자구조분석 연구가 많이 수행되었다 (Takaiwa et al. 1986; Takaiwa et al. 1989; Okita et al. 1989; Takaiwa et al. 1991; Takaiwa et al. 1996; Katsube-Tanaka et al. 2004). 또한 벼의 경우 저장단백질 합성과 관련된 유전자들의 발현은 등숙과정에서 종자 특이적으로 대량으로 발현하는 특성을 나타내므로 이들 유전자의 프로모터를 분리하여 종자특이발현벡터 제작 및 종자저장 단백질 발현조절 기작 연구를 수행하고 있다 (Wu et al. 1998; Qu and Takaiwa 2004; Qu et al. 2008; Wakasa et al. 2009; Liu

et al. 2010). Kawakatsu 등 (2010)은 이들 종자저장단백질 유전자들의 발현을 조절함으로써 종자의 품질 및 영양성을 조절하는 연구를 수행하였다. 그러나 식물염색체에는 약 14% 정도의 반복적 유전자 염기 서열이 존재하기 때문에 (IRGSP 2005), 특정 단백질합성 유전자의 효과적인 발현조절 연구 수행을 위해서는 그 유전자의 구조특성분석이 요구된다. 본 연구는 미숙종자의 발현유전자 대량 염기서열 분석 및 종자저장단백질인 glutelin 유전자의 발현 특성분석을 목적으로 수행 하였다. 이러한 glutelin 유전자의 발현특성분석 결과는 식물전반의 종자저장단백질 유전자 구조 및 기능을 분석하는 연구자들에게 보다 유익한 정보를 제공할 수 있을 것이며, 종자 형질개선 등의 실용화 연구에 기여할 수 있을 것으로 기대되어 진다. 특히 지구온난화에 따른 고온 등숙기의 벼 품질에 영향을 미치는 저장단백질의 생합성, 저장기구를 분자수준에서 이해하는데 많은 기여를 할 것으로 생각된다.

재료 및 방법

벼 미숙종자 발현 유전자 염기서열분석

본 연구의 공시 품종은 일품벼이며 출수 후 20일이 경과한 미숙종자를 재료로 사용하였다. Yoon 등 (2009)이 발표한 바와 같이 일품벼 미숙종자 RNA 분리는 우선 Trizol 시약을 이용하여 total RNA를 분리한 후 PolyA Tract mRNA kit (Promega)를 사용하여 Poly(A) mRNA를 분리하였다. 미숙종자 cDNA 은행제작은 ZAP-cDNA Gigapack III 골드 클로닝 kit 및 Gigapack III Gold Packaging Extract를 사용하여 cDNA 은행제작을 하였다. 개별 유전자 클론의 플라스미드 DNA 분리는 96 well 밀리포어 멀티스크린 (Millipore MANANLY50, MAFBNOB50)을 이용하여 대량으로 플라스미드 DNA를 분리하였다. 미숙종자 발현유전자 각각의 염기서열 결정은 BigDye terminator ver. 3.1 시약을 이용하여 5'말단은 SK 프라이머 (5'-cgctctagaactagtggatc-3'), 3'말단은 M13F 프라이머 (5'-gtttcccagtcacgac-3')로 염기서열 분석반응을 수행한 후 ABI3100 및 ABI3730 유전자 염기서열분석기 (Applied Biosystems)를 사용하여 벼 미숙종자 발현유전자에 대한 염기서열을 결정하였다.

벼 미숙종자 unigene 분석 및 glutelin 유전자 발현특성분석

미숙종자 EST 염기서열 정보분석을 위하여 각 클론의 발현유전자 염기서열을 결정한 후 SeqMan 프로그램 (DNASTAR Inc)을 사용하여 벡터부위를 포함하는 염기서열부분과 불명확한 염기서열을 제거하였다. 미숙종자에서 발현하는 unigene 선별은 SeqMan 프로그램을 이용하여 100 bp씩 90%

의 상동성으로 유전자 염기서열들을 연결하여 얻어진 singleton 유전자와 두개 이상의 유전자로 cluster된 contig 수를 합하여 unigene 수로 결정하였다 (Kikuchi et al. 2003, Yoon et al. 2009a, Yoon et al. 2009b). 미숙종자내에 발현하는 glutelin 유전자의 발현분석을 위하여 국립농업과학원 농업생명공학정보센터 (<http://www.niab.go.kr>)에 설치된 web blast 및 NCBI의 BLASTN, BLASTX 프로그램 (<http://www.ncbi.nlm.nih.gov>)을 이용하여 상동성 검사를 수행하였으며, 상동성 검사 결과 얻어진 유전자정보를 바탕으로 각각의 glutelin 유전자들의 발현수를 분석하였다.

Glutelin 유전자의 완전장 결정 및 염색체 mapping

일품벼 glutelin 완전장유전자들의 염기서열결정은 glutelin 유전자를 포함하는 5'부위의 contig 염기서열정보와 poly A 부위의 3'부위 contig 염기서열을 연결하여 완전장 유전자 염기서열을 결정하였다. 아울러 유전자부위에 gap이 있는 클론들은 각 말단의 primer를 제작하여 염기서열을 분석한 후 완전장을 결정하였다. 완전장이 결정된 각각의 일품벼 glutelin 유전자들을 GenBank에 등록하였다.

벼 glutelin 유전자들의 벼 염색체내 위치를 결정하기 위하여 일본의 벼 명명 프로젝트에서 구축한 데이터베이스 (The Rice Annotation Project DB, <http://rapdb.dna.affrc.go.jp/>)와 미국 미시간대학교 벼 게놈 명명프로젝트 데이터베이스 (Rice Genome Annotation Project DB, <http://rice.plantbiology.msu.edu/>)의 Rice Pseudomolecule ver. 6.1 정보를 이용하였다 (Ouyang et al. 2007). 벼 glutelin 완전장 유전자 11개를 이용하여 RAP DB와 RGAP DB의 BLAST 분석을 통해 99% 이상의 염기서열 유사성을 나타내는 염색체 내 지역을 판별하여 각 유전자의 염색체 내 위치를 표시하였다.

Glutelin 유전자의 구조분석 및 계통발생학적 분석

벼 glutelin 유전자 genomic DNA의 구조분석은 Rice Pseudomolecule ver. 6.1의 glutelin 게놈 염기서열정보와 일품벼 glutelin 완전장발현유전자 염기서열정보를 비교하여 각각의 유전자들의 exon과 intron 영역을 분석하여 glutelin 유전자들의 구조를 나타내었다. 일품벼 glutelin 유전자의 아미노산 분석은 DNASTAR의 Protean 프로그램의 전체 단백질 구성분석 도구를 사용하여 20개 아미노산 조성을 분석하였다. 일품벼 glutelin 유전자 11개의 상동성 분석은 NCBI의 BLASTX 프로그램을 이용하여 각각의 glutelin 과의 상동성을 분석하였다. 단백질 계통발생학적 분석은 Molecular Evolutionary Genetics Analysis (MEGA4, <http://www.megasoftware.net/>) 프로그램을 이용하여 UPGMA 방법으로 계통발생학적 분석을 수행하였다 (Kumar et al. 2008).

결과 및 고찰

벼 미숙종자 발현 유전자 염기서열분석

등숙기 중 배유 조직에 축적되는 벼 종자의 단백질과 전분은 인간에게 주요 영양 공급원으로 중요한 역할을 한다. 미숙종자 발달시기에 발현하는 유전자들의 발현특성을 이해하기위해 Yoon 등 (2009a)은 앞서 25,668개 일품벼 미숙종자 EST의 구조분석 및 정보해석을 행하고 유전자 기능분류를 행하였다. 또한 일본 농업생물자원연구소의 Kikuchi 등 (2003)은 Nipponbare 벼 품종에서 32,127개 완전장 발현유전자의 구조분석을 행한 후 KOME 데이터베이스 (<http://cdna01.dna.affrc.go.jp/cDNA/>)를 통하여 정보제공 및 기능분석을 위한 유전자 클론을 분양하고 있다.

본 실험에서는 추가적인 7,533개의 미숙종자 EST 분석을 수행하여 33,201개의 발현유전자 염기서열을 얻었다. 미숙종자 발현유전자 대량 염기서열 정보 분석 및 종자 발현유전자분리를 위하여 미숙종자 발현 unigene을 분리하였다. 미숙종자에서 발현되는 33,201개 클론 중 중복클론을 제외한 unigene 분리 및 발현유전자 contig 작성을 위하여 DNASTAR의 SeqMan 프로그램을 이용하여 basecalling과 clustering을 실시하였다. 그 결과 2개 이상의 클론이 중복된 contig 3,713개와 singleton 5,110개 등 미숙종자에서 발현하는 unigene 8,823개를 얻었다 (Table 1).

미숙종자 33,201개 클론으로부터 분석된 전체 염기서열 길이는 21,364,765 bp로 1개 클론의 평균 길이는 643 bp 였으며 singleton을 제외한 유전자 중복클론은 28,091개로 나타났다. 100개 이상의 클론으로 구성된 cluster는 24개였으며 이들 완전장 유전자 염기서열을 결정 후 GenBank에 등록 하였다 (Table 2). 그중에서 glutelin 유전자를 제외한 유전자들은 starch branching enzyme (Os06g0726400), sucrose synthase (Os07g0616800, Os06g0194900), calcium de-

Table 1 Summary of assembly and clustering of cDNA clones from immature seed of rice cv. *Ilpumbyeo*.

To estimate the level of redundancy, the *Ilpumbyeo* immature seed ESTs were clustered together by using the SeqManII program (DNASTAR), which considers two sequences originated from the same transcript when they have 90% nucleotide identity over a minimum of 100 bp. Poor quality sequences (<100 bp) and vector sequences were eliminated. The total number of unique clones is 8,823

Groups	Records
Number of sequenced clones	33,201
Total sequence length(bp)	21,364,765
Average length(bp)	643
Singletons	5,110
Clustering	3,713
A total Unigene	8,823

Table 2 The list of abundantly expressed genes from immature seed of rice cv. *Ilpumbyeo*.

Multiple sequence assembly were performed with SeqMan program and EST over than 100 bp are identified. Most genes are major seed storage proteins (glutelin and prolamin) and starch synthesis related one; they are highly expressed

Clone No.	Gene name	Length(aa)	GenBank Accession No.	RAP DB Gene locus
KCS045E07	Glutelin type-B1	499	EF122460	LOC_Os02g0249800
KCS224H01	Glutelin type-A1	499	EF122457	LOC_Os01g0762500
KCS129F04	Glutelin type-B5	500	FJ940203	LOC_Os02g0268100
KCS334A11	Glutelin type-C1	510	EF122463	LOC_Os02g0453600
KCS065G07	Glutelin type-A2	499	EF122456	LOC_Os10g0400200
KCS262D12	Glutelin type-B2	495	EF122466	LOC_Os02g0249600
KCS307D06	Glutelin type-B4	500	FJ940204	LOC_Os02g0268300
KCS349H11	Glutelin type-A3	496	EF122459	LOC_Os03g0427300
KCS014A11	Glutelin type-B7	499	EF122462	LOC_Os02g0242600
KCS318A05	Starch branching enzyme	820	EF122470	LOC_Os06g0726400
KCS337A09	Glutelin type-B8	484	EF122464	LOC_Os02g0249000
KCS336B04	Sucrose synthase 3	816	EF122474	LOC_Os07g0616800
KCS336G08	Calcium-dependent protein kinase	534	EF122475	LOC_Os10g0539600
KCS315C06	Glyceraldehyde 3-phosphate dehydrogenase	337	EF125181	LOC_Os02g0601300
KCS213A09	Prolamin	150	EF122440	LOC_Os05g0329300
KCS204B08	Prolamin	156	EF122450	LOC_Os07g0206500
KCS140F05	Prolamin	156	EF122447	LOC_Os07g0206400
KCS319H09	Glyceraldehyde 3-phosphate dehydrogenase	337	EF122472	LOC_Os08g0126300
KCS326G03	AGPase large subunit	518	EU267956	LOC_Os01g0633100
KCS348H05	OsNAC6 Protein	252	EF122479	LOC_Os02g0214500
KCS315B08	Fructose-biphosphate aldolase	358	EF122473	LOC_Os01g0905800
KCS335B07	alpha-globulin	737	EF122478	LOC_Os05g0499100
KCS091H04	Sucrose synthase 2	808	EF122480	LOC_Os06g0194900
KCS024B02	Prolamin	134	EF122438	LOC_Os03g0766100

pendent protein kinase (Os10g0539600), glyceraldehyde 3-phosphate dehydrogenase (Os02g0601300, Os08g0126300), prolamin (Os05g0329300, Os07g0206500, Os07g0206400, Os03g0766110), AGPase (Os01g0633100), fructose-bisphosphate aldolase (Os01g0905800), alpha globulin (Os05g0499100) 등으로 구성되어 있었다 (Table 2). 이와 같은 유전자들의 발현이 현저히 높은 것은 종자발달시기에 단백질과 전분합성이 왕성하다는 사실을 의미한다. Yoon 등 (2010)은 벼 60K microarray 와 135K microarray를 이용하여 벼의 callus, 수잉기, 출수기, 미숙종자, 발아종자, 잎, 뿌리, 줄기에서 조직 특이 발현유전자의 발현을 분석하였다. 그 결과 callus에서 299개, 수잉기에서 285개, 출수기에서 277개, 미숙종자에서 241개, 뿌리에서 300개, 잎에서 447개의 유전자가 조직 특이적으로 발현하는 것을 확인 하였다. 이와 같은 결과는 미숙종자에서의 unigene 8,823개 유전자 중 단백질 합성 및 전분대사 등에 관여하는 241개 유전자를 제외한 대다수의 유전자들이 조직 공통적으로 발현하고 있음을

나타내고 있다.

Glutelin 유전자의 발현특성 분석

일품벼 미숙종자에서 발현하는 종자 저장단백질 glutelin 유전자들의 발현특성 검정 및 염색체 mapping을 위하여 glutelin 유전자들의 완전장 염기서열을 결정하였다. 완전장 염기서열 결정은 33,201개의 EST 클론의 염기서열을 DNASTAR SeqMan 프로그램으로 assembly 한 후 glutelin 유전자를 포함하고 있는 contig 염기서열을 얻었다. 이 후 glutelin 유전자들을 포함하는 contig 염기서열에서 5'부위의 ATG coding 서열과 3'부위의 종결 서열을 포함하는 클론을 선발하여 완전장을 결정하였다 (Table 3). 완전장 염기서열이 결정된 11개 일품벼 종자발현 glutelin 유전자들의 염기서열을 GenBank에 등록 하였다 (EF122456, EF122457, EF122459, EF122460, EF122462, EF122464, EF122466, FJ940205, FJ940204, FJ940203).

Table 3 The list of glutelin genes in immature seed. The Rice Annotation Project DB and Rice Pseudomolecule ver. 6.1 of Rice Genome Annotation Project DB are used in order to determine the location of rice glutelin genes on chromosomes

Clone Name	Gene name	Number of EST clones	Length (aa)	GenBank Accession No.	RAP DB (Ver.5) Gene Locus	RGAP DB (Ver.6.1) Gene Locus
KCS065GO7	Glu01 (Glutelin type A2)	778	499aa	EF122456	LOC_Os10g0400200	LOC_Os10g26060 (Exon : C13428139-13426365)
KCS224H01	Glu02 (Glutelin type A1)	1582	499aa	EF122457	LOC_Os01g0762500	LOC_Os01g55690 (Exon : W32075725-32077498)
KCS349H11	Glu03 (Glutelin type A3)	532	496aa	EF122459	LOC_Os03g0427300	LOC_Os03g31360 (Exon : W17863093-17864851)
KCS045E07	Glu04 (Glutelin type B1)	2401	499aa	EF122460	LOC_Os02g0249800	LOC_Os02g15169 (Exon : C8460590-8458771)
					LOC_Os02g0249900	LOC_Os02g15178 (Exon : W8465107-8466920)
KCS014A11	Glu05 (Glutelin type B7)	325	499aa	EF122462	LOC_Os02g0242600	LOC_Os02g14600 (Exon : W8057685-8059532)
KCS334A11	Glu06 (Glutelin type C1)	869	510aa	EF122463	LOC_Os02g0453600	LOC_Os02g25640 (Exon : C14987791-14985927)
KCS337A09	Glu07 (Glutelin type B8)	215	484aa	EF122464	LOC_Os02g0249000	LOC_Os02g15090 (Exon : C8408651-8406836)
KCS262D12	Glu08 (Glutelin type B2)	753	495aa	EF122466	LOC_Os02g0249600	LOC_Os02g15150 (Exon : W8446657-8448470)
KCS304H02	Glu09 (Glutelin type B6)	31	495aa	FJ940205	LOC_Os02g0248800	LOC_Os02g15070 (Exon : C8403191-84013888)
KCS307D06	Glu10 (Glutelin type B4)	747	500aa	FJ940204	LOC_Os02g0268300	LOC_Os02g16830 (Exon : C9587970-9586135)
KCS129F04	Glu11 (Glutelin type B5)	1143	500aa	FJ940203	LOC_Os02g0268100	LOC_Os02g16820 (Exon : W9580524-9582359)

종자 glutelin 유전자들의 발현특성을 분석하기 위하여 미숙종자에서 발현하는 glutelin 유전자들의 EST 분석을 수행하였다. 국립농업과학원 농업생명공학정보센터 (<http://www.niab.go.kr>)에 설치된 web blast 및 NCBI의 BLASTN, BLASTX 프로그램 (<http://www.ncbi.nlm.nih.gov>)을 이용하여 각각의 발현유전자에 대하여 상동성 검사를 수행하였다. 상동성 검사 결과 얻어진 유전자정보를 바탕으로 미숙종자내에 발현하는 11개 glutelin 유전자들의 EST 수를 분석하여 발현특성을 분석하였다 (Table 3). 그 결과 33,201개 미숙종자 EST 중에서 Glu04 (glutelin type B1, LOC_Os02g0249800, LOC_Os02g0249900) 유전자가 2401개, Glu02 (type A1, LOC_Os01g0762500) 유전자가 1582개, Glu11 (type B5, LOC_Os02g0268100) 유전자가 1143개, Glu06 (type C1, LOC_Os02g0453600) 유전자가 869개, Glu01 (type A2, LOC_Os10g0400200) 유전자가 778개, Glu08 (type B2, LOC_Os02g0249600) 유전자가 753개, Glu10 (type B4, LOC_Os02g0268300) 유전자가 747개, Glu03 (type A3, LOC_Os03g0427300) 유전자가 532개, Glu05 (type B7, LOC_Os02g0242600) 유전자가 325개, Glu07 (type B8, LOC_Os02g0249000) 유전자가 215개, Glu09 (type B6,

LOC_Os02g0248800) 유전자가 31개 등의 발현양을 나타내었으며 총 9376개의 glutelin 유전자 EST가 발현하여 미숙종자 전체 발현양의 28.2%를 차지하였다 (Table 3). 이와 같은 결과는 미숙종자발달시기에 단백질합성이 가장 왕성하게 이루어지고 있음을 시사한다. 또한 종자의 단백질 함량은 미질과 영양분의 측면에서 매우 중요할 뿐만 아니라, 최근 이들 단백질합성 유전자의 프로모터는 식물형질전환 응용연구에 사용되고 있다. 종자 glutelin 유전자들의 발현특성 분석 결과를 토대로 종자 중의 단백질 함량을 높이는 연구를 위해 종자에서 발현양이 많은 Glu04, Glu02, Glu11, Glu06 유전자를 이용한 과발현체 작성이 요구되어진다. Qu 등(2008)은 벼 genomic DNA에서 PCR방법으로 6개의 glutelin 유전자 프로모터를 분리한 후 GUS 발현분석으로 종자발현을 분석한 결과 glu06-type C1 유전자의 프로모터가 종자 전체적으로 발현하는 것을 확인 할 수 있었다. 또한 유전자발현 억제를 위한 RNAi 형질전환체 작성에는 glutelin 유전자 염기서열을 공통적으로 포함하는 RNAi 벡터 작성을 통하여 보다 효율적인 유전자발현 억제가 가능할 것으로 생각된다.

Glutelin 유전자의 구조분석 및 염색체 mapping

일품벼 미숙종자발현 11개의 glutelin 유전자 구조분석은 glutelin 발현유전자 cDNA의 완전장 염기서열정보와 미국 미시간대학교 벼 게놈 명명프로젝트 데이터베이스 (RGAP DB, <http://rice.plantbiology.msu.edu/>)의 Rice Pseudomolecule ver. 6.1에서 glutelin 유전자의 게놈 염기서열정보를 비교분석하여 exon과 intron으로 나누어 각 유전자의 구조를 분석하였다 (Fig. 1). 진핵생물 유전자 예측 프로그램인 FGENESH (<http://linux1.softberry.com/berry.phtml>) 프로그램을 이용하여 벼 glutelin 유전자의 구조를 분석한 결과, glutelin 유전자들은 4개의 exon과 3개의 intron으로 구성되어 있었으며 intron의 길이는 77 bp에서 139 bp사이의 짧은 염기서열로 구성되어 있었다 (Fig. 1). 이와 같은 유전자구조는 glutelin의 경우 진화학적으로 염색체내에서 복제되어 공통적으로 짧은 intron을 가지게 되었고, 벼 종자저장단백질 유전자의 특성으로 생각되어진다. 같은 벼 종자 저장단백질인 prolamins 유전자의 경우는 intron이 없는 1개의 exon으로 구성되어 있다. 벼 glutelin 게놈 유전자들의 염기서열에서 exon과 intron의 경계 염기서열은 AG/GT의 공통적인 염기서열을 나타내고 있었다. Glu09-type B6 게놈 유전자의 intron서열은 75.5%의 높은 A, T를 포함하고 있었으며 exon에서는 57.9%의 A, T를 나타냈다. 또한 Glu06-type C1 게놈 유전자의 intron에서는 71%의 A, T를 포함하고 있었으며 exon에서는 50.2%의 A, T를 나타냈다. 대다수 glutelin 게놈 유전자의 intron 서열은 69-75%

범위의 비교적 높은 A, T를 포함하는 구조를 가지고 있었다.

벼 종자 저장단백질의 60-80%를 차지하는 glutelin 단백질의 아미노산 함량은 종자의 영양성분으로 매우 중요한 역할을 한다. 특히 벼 종자에는 옥수수에 비하여 lysin 함량이 낮아 영양성이 떨어지는 것으로 알려져 있다. 벼 미숙종자발현 11개 glutelin 유전자의 아미노산서열을 이용하여 각 단백질의 lysin 함량을 검토한 결과 glu02-type A1은 2.73%, glu01-type A2는 2.73%, glu03-type A3은 3.43%, glu04-type B1는 3.63%, glu08-type B2는 3.66%, glu10-type B4는 3.38%, glu11-type B5는 3.38%, glu09-type B6은 3.17%, glu05-type B7은 4.51%, glu07-type B8은 3.97%, glu06-type C1은 3.37%의 함량 비율을 나타내었다 (Table 4). 벼 glutelin type A는 2.73-3.43% 범위의 lysin 함량을 나타내었으며 glutelin type B는 3.17-4.51% 범위의 lysin 함량을 나타내어 type A에 비하여 상대적으로 많은 lysin 함량을 나타내었다. 이에 비하여 벼 종자 prolamins 유전자의 아미노산서열 중 lysin 함량은 0.7-1.75%를 나타내었다 (Yoon et al. in preparation). 이와 같은 결과는 벼 종자의 lysin 함량을 높이기 위해서는 lysin 함량이 3.97-4.51%로 상대적으로 높은 glu07-type B8과 glu05-type B7 유전자의 과발현체를 이용한 lysin 함량을 높이는 영양성 강화 연구가 요구되어진다.

벼 glutelin 완전장 유전자 11개의 염기서열 유사성 분석 (재료 및 방법 참조)을 바탕으로, 벼 glutelin 유전자들의 벼 염색체내 위치를 결정하였다. 벼 glutelin 유전자는

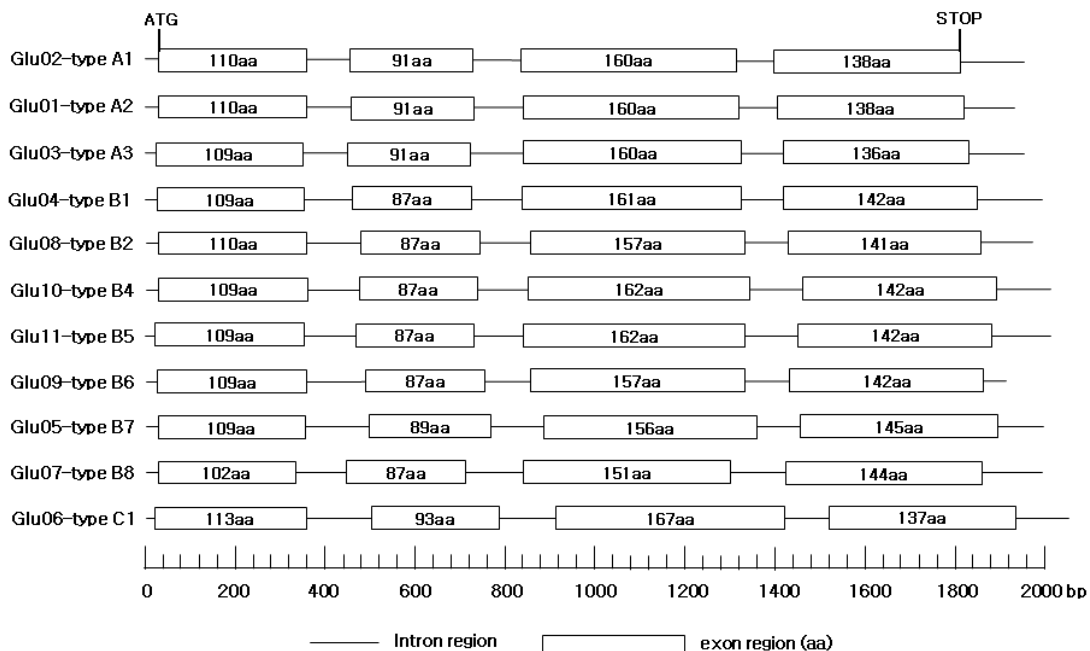


Fig. 1 Gene structure of rice glutelin proteins. Coding exons were represented by boxes, and the numbers indicate the deduced amino acid size. Lines between two boxes represent introns

Table 4 Predicted whole protein composition from deduced amino acids of rice glutelin.

Amino acid analysis of rice glutelin were performed by using whole protein composition analysis protocol in Protean program (DNASTAR). Each amino acid content was evaluated as a percentage by weight

	Glu02 (type A1)	Glu01 (type A2)	Glu03 (type A3)	Glu04 (type B1)	Glu08 (type B2)	Glu10 (type B4)	Glu11 (type B5)	Glu09 (type B6)	Glu05 (type B7)	Glu07 (type B8)	Glu06 (type C1)
A Ala	4.30	4.04	3.93	4.40	4.44	4.63	4.63	3.39	3.63	4.79	5.35
C Cys	1.83	1.83	1.84	1.28	1.29	1.45	1.45	1.46	1.27	1.32	1.08
D Asp	3.07	3.27	4.52	2.44	2.26	2.03	2.03	3.05	2.43	3.57	3.63
E Glu	6.20	5.96	5.07	6.16	6.45	7.27	7.27	6.85	6.36	5.88	5.88
F Phe	6.02	6.27	6.04	7.29	7.62	7.25	7.25	7.02	6.73	6.71	6.70
G Gly	3.55	3.75	3.56	3.33	3.26	3.51	3.51	3.03	3.51	3.54	3.40
H His	2.19	2.44	3.18	2.67	2.45	2.90	2.90	2.67	2.90	3.25	3.84
I Ile	4.63	4.82	5.05	5.20	5.25	4.78	4.78	6.20	7.17	6.19	4.56
K Lys	2.73	2.73	3.43	3.63	3.66	3.38	3.38	3.17	4.51	3.97	3.37
L Leu	7.85	7.64	8.28	7.00	6.86	8.56	8.76	7.80	7.97	8.46	8.52
M Met	0.70	1.17	1.17	1.39	1.40	1.62	1.62	1.39	1.39	0.72	2.30
N Asn	6.90	6.89	5.91	7.06	7.13	8.03	8.03	7.06	9.04	6.45	5.39
P Pro	3.63	3.79	3.47	3.78	3.64	4.10	3.76	4.12	3.93	3.72	4.25
Q Gln	12.30	11.83	11.67	12.46	12.35	11.45	11.95	11.32	10.37	10.74	10.77
R Arg	11.11	10.54	10.04	9.39	8.92	9.62	9.62	9.11	10.17	10.53	9.57
S Ser	6.35	5.88	5.75	6.16	6.53	4.29	4.44	4.93	3.98	4.44	5.79
T Thr	3.60	3.77	3.61	3.04	3.61	2.85	2.85	4.65	3.91	3.32	3.54
V Val	6.52	6.51	6.90	6.49	6.01	5.58	5.58	5.96	5.41	5.96	6.94
W Trp	0.99	1.32	1.00	1.32	1.33	0.98	0.98	1.32	0.98	1.36	1.96
Y Tyr	5.51	5.51	5.53	5.48	5.53	5.17	5.17	5.48	4.31	5.06	3.14

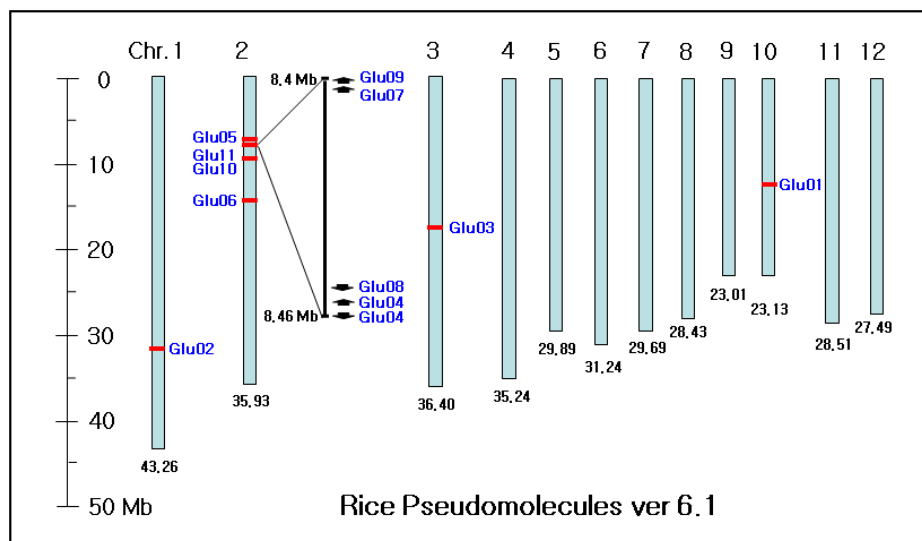


Fig. 2 Chromosomal location of glutelin genes in rice. In the basis of gene expression analysis for seed storage proteins, 11 glutelin genes were cloned and the 8 genes were located on chromosome 2. They showed 28% of a total expression. These results showed storage protein were duplicated evolutionally on a specific chromosome

1번, 2번, 3번, 10번 염색체에 mapping 되었으며 특히 2번 염색체 8.4 Mb 위치에 5개의 glutelin 유전자가 위치하고

있었다 (Table 2, Fig. 2). 이와 같은 결과는 glutelin 유전자의 경우 2번 염색체 특이적으로 위치하고 있으며 진화학

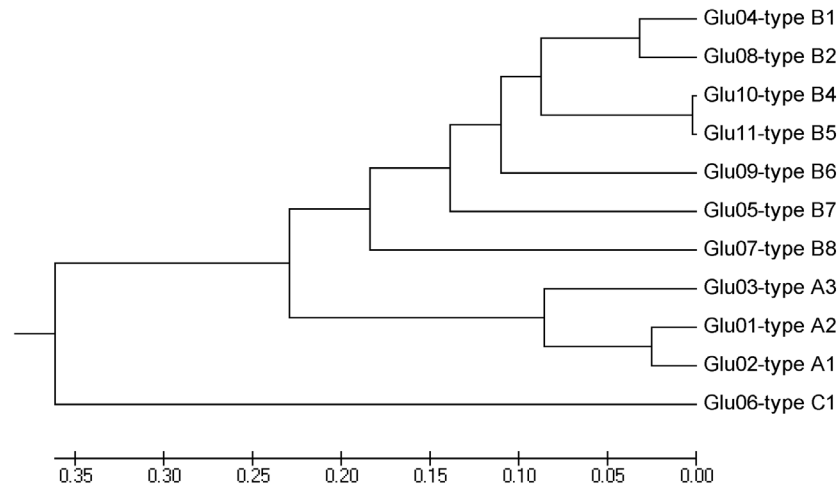


Fig. 3 Phylogenetic tree of the glutelin genes in rice. The phylogenetic trees were constructed by using MEGA 4.0. The graphic representation shows the results by the unweighted pair group with arithmetic mean method (UPGMA). The same analysis also was performed with the neighbor-joining method (data not shown)

B5, Glu09-type B6, Glu05-type B7, Glu07-type B8), glu-type C (Glu06-type C1)의 3개 그룹으로 나누어지는 것을 보여 주었다 (Fig. 3). 이와 같은 결과는 glutelin 유전자의 계통 구조를 분석한 그림 1의 exon 부위 구조와 일치하는 경향을 나타내고 있다.

적 요

벼는 세계에서 가장 중요한 작물이며 크기가 383Mb로 계통연구 모델 작물로 이용되고 있다. 또한 그 종자는 인간에게 탄수화물과 단백질 영양원을 제공한다. 벼 종자의 단백질은 약 8%를 차지하며 40%를 차지하는 콩 종자의 단백질 양에 비하여 상대적으로 적은 양을 나타낸다. 오스본의 분류에 의하면 종자 단백질은 수용성의 albumin, 염용해성 globulin, 알코올 용해성 prolamin 그리고 약산 또는 알카리 용해성 glutelin으로 나누어진다. Glutelin과 prolamin은 벼의 주요 저장단백질이다.

벼 glutelin 저장단백질 유전자의 발현분석을 위하여 일 품벼 미숙종자의 발현유전자 (EST) 분석을 행하였다. 그 결과 11종의 미숙종자 발현 glutelin 유전자를 분리 하였으며 8개의 유전자는 염색체 2번에 위치하였다. Glutelin 유전자 발현양은 전체 미숙종자 발현유전자의 약 28.2%를 차지하였다. 또한 glu-04의 경우 같은 염색체 상에서 4.5 kb 떨어진 곳에 역방향의 같은 염기서열로 복제되어 있었다. 이와 같은 결과는 glutelin 유전자는 진화학적으로 복제되어 염색체 특이적으로 발현하는 것을 나타낸다. 종자 11개 glutelin 유전자들의 아미노산서열분석을 통하여 lysin 함량을 조사한 결과 glu05-type B7에서 4.51%의 높은 lysin 함량을 나타내었다. 향후 유전자의 과발현

체를 이용한 lysin 함량을 높이는 영양성 강화 연구가 요구되어진다.

사 사

본 연구는 농촌진흥청 국립농업과학원 농업과학기술 연구개발사업 (과제번호: PJ006108, PJ006758) 및 2011년도 농촌진흥청 국립농업과학원 박사후연수과정지원사업의 지원에 의해 이루어진 것임

인용문헌

- Abe T, Gusti RS, Ono M, Sasahara T. (1996) Variations in glutelin and high molecular weight endosperm proteins among subspecies of rice (*Oryza sativa* L.) detected by two-dimensional gel electrophoresis. *Genes Genet Syst* 71:63-8
- Bewley DJ, Black M. (1985) *Seeds, Physiology of development and germination*, Plenum press, New York, pp 1-27
- Choi H. (2002) Current status and perspectives in varietal improvement of rice cultivars for high-quality and value-added products. *Korean J Crop Sci* 47:15-32
- Duan M, Sun SS. (2005) Profiling the expression of genes controlling rice grain quality. *Plant Mol Biol* 59:165-78
- 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
- Katsube-Tanaka T, Duldulao JB, Kimura Y, Iida S, Yamaguchi T, Nakano J, Utsumi S. (2004) The two subfamilies of rice glutelin differ in both primary and higher-order structures.

Biochim Biophys Acta 1699:95-102

- Kawakatsu T, Hirose S, Yasuda H, Takaiwa F. (2010) Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. *Plant Physiol* 154:1842-54
- Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, Kishimoto N, Yazaki J, Ishikawa M, Yamada H, Ooka H, Hotta I, Kojima K, Namiki T, Ohneda E, Yahagi W, Suzuki K, Li CJ, Ohtsuki K, Shishiki T; Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otomo Y, Murakami K, Iida Y, Sugano S, Fujimura T, Suzuki Y, Tsunoda Y, Kurosaki T, Kodama T, Masuda H, Kobayashi M, Xie Q, Lu M, Narikawa R, Sugiyama A, Mizuno K, Yokomizo S, Niikura J, Ikeda R, Ishibiki J, Kawamata M, Yoshimura A, Miura J, Kusumegi T, Oka M, Ryu R, Ueda M, Matsubara K; RIKEN, Kawai J, Carninci P, Adachi J, Aizawa K, Arakawa T, Fukuda S, Hara A, Hashizume W, Hayatsu N, Imotani K, Ishii Y, Itoh M, Kagawa I, Kondo S, Konno H, Miyazaki A, Osato N, Ota Y, Saito R, Sasaki D, Sato K, Shibata K, Shinagawa A, Shiraki T, Yoshino M, Hayashizaki Y, Yasunishi A. (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science* 301:376-9
- Kumar S, Nei M, Dudley J, Tamura K. (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.* 9:299-306
- Liu WX, Liu HL, Chai ZJ, Xu XP, Song YR, Qu le Q. (2010) Evaluation of seed storage-protein gene 5' untranslated regions in enhancing gene expression in transgenic rice seed. *Theor Appl Genet* 121:1267-74
- Matsuo T, Kumazawa K, Ishii R, Ishihara K, Hirata H. (1995) *Science of the rice plant volume two physiology*, Nobunkyo press, Tokyo, pp 1063-97
- Okita TW, Hwang YS, Hnilo J, Kim WT, Aryan AP, Larson R, Krishnan HB. (1989) Structure and expression of the rice glutelin multigene family. *J Biol Chem* 264:12573-81
- Osborn T. (1924) *The vegetable proteins*. London, FL: Longmans, Green and Co
- 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
- Paterson AH, Freeling M, Sasaki T. (2008) Grains of knowledge: genomics of model cereals. *Genome Res* 15:1643-50
- Qu le Q, Takaiwa F. (2004) Evaluation of tissue specificity and expression strength of rice seed component gene promoters in transgenic rice. *Plant Biotechnol J* 2:113-25
- Qu le Q, Xing YP, Liu WX, Xu XP, Song YR. (2008) Expression pattern and activity of six glutelin gene promoters in transgenic rice. *J Exp Bot* 59:2417-24
- Shewry PR, Napier JA, Tatham AS. (1995) Seed storage proteins: structures and biosynthesis. *Plant Cell* 7:945-56
- Shewry PR, Halford NG. (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot* 53:947-58
- Takaiwa F, Kikuchi S, Oono K. (1986) The structure of rice storage protein glutelin precursor deduced from cDNA. *FEBS Lett* 206:33-35
- Takaiwa F, Kikuchi S, Oono K. (1989) The complete nucleotide sequence of new type cDNA coding for rice storage protein glutelin. *Nucleic Acids Res* 17:3289
- Takaiwa F, Oono K, Wing D, Kato A. (1991) Sequence of three members and expression of a new major subfamily of glutelin genes from rice. *Plant Mol Biol* 17:875-85
- Takaiwa F, Yamanouchi U, Yoshihara T, Washida H, Tanabe F, Kato A, Yamada K. (1996) Characterization of common cis-regulatory elements responsible for the endosperm-specific expression of members of the rice glutelin multigene family. *Plant Mol Biol* 30:1207-21
- Tanaka K, Sugimoto T, Ogawa M, Kasi Z. (1980) Isolation and characterization of two types of protein bodies in the rice endosperm. *Agric. Biol. Chem* 44:1633-39
- Wakasa Y, Yang L, Hirose S, Takaiwa F. (2009) Expression of unprocessed glutelin precursor alters polymerization without affecting trafficking and accumulation. *J Exp Bot* 60: 3503-11
- Wen TN, Luthe DS. (1985) Biochemical characterization of rice glutelin. *Plant Physiol* 78:172-77
- Wu CY, Suzuki A, Washida H, Takaiwa F. (1998) The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants. *Plant J* 14:673-83
- Xu JH, Messing J. (2008) Organization of the prolamin gene family provides insight into the evolution of the maize genome and gene duplications in grass species. *Proc Natl Acad Sci U S A* 105:14330-5
- Xu JH, Messing J. (2009) Amplification of prolamin storage protein genes in different subfamilies of the Poaceae. *Theor Appl Genet* 119:1397-412
- Yoon UH, Lee GS, Lee JS, Hahn JH, Kim CK, Kikuchi S, Satoh K, Kim JA, Lee JH, Lee TH, Kim YH. (2009a) Structural analysis of seed developmental stage ESTs in *Oryza sativa* L. *Korean J of Plant Biotechnology* 36:130-136
- Yoon UH, Lee GS, Kim CK, Lee JS, Hahn JH, Yun DW, Ji HS, Lee TH, Lee JH, Park SH, Kim GW, Seo MS, Kim YH (2009b) Analysis of germination seed stage expressed sequence tags in *Oryza sativa* L. *Korean J of Plant Biotechnology* 36:281-288
- Yoon UH, Kim YK, Kim CK, Hahn JH, Kim DH, Lee TH, Lee GS, Park SC, Nahm BH (2010) Current status on expression profiling using rice microarray. *Korean J of Plant Biotechnology* 37: 144-152