벼 핵심집단의 마그네슘 함량 관련 통합 전장유전체 분석

허은범¹ · 유지민¹ · 이원도² · 추상호³ · 김규원³ · 조유현² · 박용진^{1,3*}

¹공주대학교 산업과학대학 식물자원학과, ²주식회사 시드피아, ³공주대학교 작물 오믹스지능육종 연구센터

Integrated Genome-Wide Association Studies to Dissect Natural Variation for Magnesium Ion Contents in Rice Germplasm

Eun-Beom Heo¹, Ji-Min Yoo¹, Won-Do Lee², Sang-Ho Chu³, Kyu-Won Kim³,
Yoo-Hyun Cho², and Yong-Jin Park^{1,3*}

¹Department of Plant resources, Kongju National University, Yesan 32439, Korea

²Seedpia Inc., 85 Maesil-ro, Gwonseon-gu, Suwon-si, Gyeonggi-do, Korea

³Center of Crop Breeding on Omics and Artificial Intelligence, Kongju National University, Yesan 32439, Korea

Abstract : Magnesium is important not only for the growth of rice itself, but also as an essential micronutrient component of half of the world population who are supported by rice. Here, we performed genome-wide association study (GWAS) with high-resolution density SNPs to identify natural allelic variation in Mg^{2+} increase from rice set, which is derived from a total 24,368 rice germplasms. The range of the concentration and distribution of Mg^{2+} in 295 core accessions of brown rice grain were wide, from 18.17mg/L to 57.11mg/L, with mean 39.71mg/L. In particular, GWAS result shows that the high peak found on chromosomes 3 and 11. The new natural variants identified through haplotyping analysis would be useful to develop new rice varieties with improved storage ability of the valuable mineral through the future molecular breeding.

Keywords : GWAS, Core-collection, Rice, Mg^{2+} , Haplotype

서 언

벼는 국민 일일 에너지 공급의 약 40%를 차지하고 있으며, 국내 농업소득의 70%에 달하는 것으로 알려져 있으나 식생활의 서구화 및 다양화로 인하여 소비량이 감소 추세에 있는 실정이다(Kim, 2011). 반면, 국내 쌀시장에서 건강기능성 특수미 및 고품질 벼 품종개발의 요구도가 커지고 있어(Choi, 2002), 국내에서는 고품질, 유용한 물질이 고 함유되어 있는 기능성 쌀을 육종하고 있다(Park et al. 2015, Kwon et al. 2011). 벼에는 3대 영양소가 포함되어 있으며 다양한 무기성분이 포함되어 소비자들에게 필요한 에너지를 공급해 줄 뿐만 아니라(Choi, 2002), 벼는 다른 곡류에 비해 영양학적으로 우수한 것으로 알려져 있어 다양한 성인병을 예방하는 효과가 있는 것으로

알려져 있다(Son, 2001). 또한 쌀은 세계적인 주곡 작물로 아시아에서는 중요 영양공급원으로서 보고되고 있으나, 벼의 도정에 따라 영양소가 감소하게 되는 경향을 보이는 것으로 보고되고 있다. 현재까지 보고된 바로는 쌀의 품종 별 백미와 현미의 영양성분의 조성을 비교해 주요 재배장려품종인 6종의 백미와 현미종의 영양성분을 분석해 도정에 따른 무기질의 변화를 보고하였는데 도정률이 증가할수록 인, 철, 아연, 마그네슘 함량이 감소하는 경향을 보일 뿐만 아니라, 인체에 필수 영양소인 무기성분이 감소하는 것으로 보고하였다(Choe et al. 2002, Juliano & Bechtel, 1985).

그 중 마그네슘(Mg)은 인체에 중요한 역할을 하며 가벼운 Mg 결핍은 골다공증의 위험인자로 작용할 수 있고, 저칼슘혈증, 심장부정맥, 고혈압, 동맥경화성 혈관 질환면에서도 영향을 주며, 칼슘대사에 관련하여도 작용을 하는 것으로 보고되고 있다(Fatemi et al. 1991, Rude, 1998). 그러나 이와 관련하여, 차세대 염기서열 결정 방법(NGS, Next Generation Sequencing)

*Corresponding Author (E-mail: yjpark@kongju.ac.kr, Tel: +82-41-330-1201)

(Received on May 11, 2017. Accepted on August 2, 2017.)

으로 결정된 유전체 정보를 이용하여 Mg^{2+} 함량과 연관 유전자 발굴 연구는 미흡한 실정이다.

본 연구에서는 한국 벼 핵심집단 KRICE_CORE 137점(Kim et al. 2016)에 대해 ICP-OES(Inductively Coupled Plasma Optical Emission Spectroscopy)를 이용하여 Mg^{2+} 함량을 분석한 뒤 전장유전체연관분석(GWAS, Genome-Wide Association Study)를 수행하여 Mg^{2+} 함량 연관 유전자를 탐색하고, 향후 Mg^{2+} 고함량 품종 육성의 기초자료로 활용하고자 수행하였다.

재료 및 방법

공시재료

본 연구에 공시한 벼는 농촌진흥청 국립농업과학원 유전자원 센터에서 보유하고 있는 약 25,604점의 유전자원 중에 특성정보가 있는 자원을 선별한 KRICE_CORE (Kim et al. 2016) 137점과 국내 육성종 158점으로 이루어진 국내 벼 핵심집단 295점을 사용하였다. KRICE_CORE는 해외 도입종(introduced) 893점, 국내에서 육성되고있는 유전자원 중 수집된 육성종(breeding line) 1,116점, 예전부터 전하여 내려오는 재래종(landrace) 394점, 고유의 품종이 아닌 재배논에서 산발적으로 발생하는 잡초성벼(weedy) 1,955점 및 International Rice Research Institute(IRRI) 자원 중 선별한 핵심집단 48점, 총 4,046점에 대해 PowerCore(Kim et al. 2007)를 이용하여 최종적으로 작성된 137점의 벼 핵심집단이다. 또한, 본 연구에서는 유전체 연관분석을 위하여 핵심집단 295점의 유전체 재분석 정보를 이용하였다. 유전체 재분석에 의하면 핵심집단 295점은 Indica 64점, Japonica 217점, Aus 8점, Aromatic 2점, Admixture 4점으로 구성되어 있다.

국내 벼 핵심집단의 DNA 추출 및 유전체 재분석

295 품종으로부터 1주에서 재배된 잎을 사용하여 총 genomic DNA는 DNeasy® Plant Mini Kit(QIAGEN)를 이용해 채취하여 HiSeq 2500(Illumina) 장비를 이용하여 short read sequence 생산하였다. 생산된 염기단편을 IRSGP 1.0 참조유전체에 정렬하여 서열변이를 확인하였다(Shearer et al. 2011). Sequencing을 통하여 얻어진 염기단편을 정렬하기 위해서 BWA v0.7.5a(Li & Durbin, 2009)을 이용하였고 GATK v2.3.9 Lite(McKenna et al. 2010)을 이용하여 SNP를 발굴하였다.

국내 벼 핵심집단에 대한 ICP-OES분석

국내 벼 핵심집단 295점에 대하여 국립공주대학교 시험포장

에서 2014년도에 종자증식을 수행하였다. Mg^{2+} 분석을 위해 수확 후 상온에서 풍건하여 5°C 저온저장고에 보관중인 종자를 현미상태로 도정하여 이용하였다. ICP-OES 분석을 위하여 vessel안에 현미 0.5g HNO_3 3mL, H_2O_2 2mL를 함께 넣어준 뒤 microwave 3000(Anton paar, USA)을 이용하여 전처리를 실시하였으며, 1%의 HNO_3 을 이용하여 최종 volume을 14mL로 맞추었고, ICP-OES 분석을 실시해 현미에 함유되어 있는 Mg^{2+} 함량을 측정하였다.

Mg 함량에 대한 전장유전체연관분석(GWAS, Genome-Wide Association Study)

식물의 양적 형질에 적용할 수 있도록 개발된 GAPIT(Lipka et al. 2012)를 GWAS분석에 사용하였다. 혼합 선형 모형(mixed linear model)을 사용하였다. 혼합 선형 모형은 기존 선형 모형에 SNP들의 임의 효과(random effect)를 추가한 모델이다. 표현형데이터와 유전자형데이터를 입력 후 GWAS분석을 진행하였다.

결 과

GWAS(Genome-Wide Association Study) 분석

유전체 분석자원 295점에 대하여 GAPIT를 사용하여 GWAS 분석을 한 결과 염색체 3번과 염색체 11에서 $p\text{-value} \leq 4.21 \times 10^{-6}$, $FDR < 0.011$ 로 유의수준 $\alpha = 0.05$ 에서 유의한 연관을 나타내었다. 3번 염색체에서는 1,704,544bp에서 $p\text{-value}$ 가 최소값(1.6×10^{-7})을 보였으며, 11번에서는 22,990,844bp에서 $p\text{-value}$ 가 최소값(1.86×10^{-7})을 보였다. 유전체 분석자원 295점의 GWAS분석 결과 중 Q-Q 플롯에서 예상되는 값보다 실제 관찰되는 값이 커지는 것을 관찰되었다. 따라서 관찰된 값에 의해서 형질과 연관성이 있다고 예상할 수 있다(Table 1, Table 2, Fig. 1). 6번 염색체의 연관 후보 유전자의 영역은 높은 연관 영역이 8Mb-22Mb에 광범위하게 흩어져 있어 유전자를 특정하기 어려우므로 분석에서 제외하였다.

GWAS 분석 결과 기반 후보유전자 동정

전장유전체 재분석된 핵심집단 295점의 표현형(phenotype data) 데이터와 유전자형(genotype data) 데이터를 GWAS 분석을 통해 Mg^{2+} 과 연관된 후보유전자를 선별하였다. 후보 유전자는 Rap-db(<http://rapdb.dna.affrc.go.jp>)에서 검색하였고, 염색

Table 1. SNP genes position associated with the phenotypes.

SNP	Chromosome	Position	MAF ^z	P-value	FDR
chr3_1130321	3	1,130,321	0.3698	4.21E-06	0.01157
chr3_1212678	3	1,212,678	0.221	1.03E-06	0.01096
chr3_1478610	3	1,478,610	0.2417	4.14E-06	0.01143
chr3_1678262	3	1,678,262	0.4318	1.81E-06	0.01096
chr3_1704544	3	1,704,544	0.3409	1.61E-07	0.01096
chr3_1704841	3	1,704,841	0.4525	3.49E-06	0.01096
chr3_1759891	3	1,759,891	0.4504	3.37E-06	0.01096
chr3_1763199	3	1,763,199	0.2252	4.99E-07	0.01096
chr3_1820186	3	1,820,186	0.2252	9.60E-07	0.01096
chr3_1820245	3	1,820,245	0.4421	2.60E-06	0.01096

^zMinor allele frequency**Table 2.** SNP genes position associated with the phenotypes.

SNP	Chromosome	Position	MAF ^z	P-value	FDR
chr11_22938644	11	22,938,644	0.2128	3.79E-07	0.01096
chr11_22938660	11	22,938,660	0.2045	4.85E-07	0.01096
chr11_22954976	11	22,954,976	0.1942	4.26E-07	0.01096
chr11_22977054	11	22,977,054	0.186	4.11E-07	0.01096
chr11_22977058	11	22,977,058	0.186	4.11E-07	0.01096
chr11_22990844	11	22,990,844	0.1942	1.86E-07	0.01096
chr11_22992736	11	22,992,736	0.1983	5.28E-07	0.01096
chr11_22995258	11	22,995,258	0.2004	4.28E-07	0.01096
chr11_23002526	11	23,002,526	0.1963	4.29E-07	0.01096
chr11_23008234	11	23,008,234	0.2025	2.76E-07	0.01096

^zMinor allele frequency **Fig. 1.** Manhattan plot and QQ plot of genome-wide association scan for Mg contents in 295 accessions.

체 3번에서 p-value가 가장 높은 1,704,544 위치로부터 후보 유전자를 $\pm 40\text{kb}$ 범위에서 찾았고, 염색체 11번에서 p-value가 가장 높은 22,990,844 위치로부터 후보 유전자를 $\pm 40\text{kb}$ 범위에서 찾았다. 염색체 3번에서 찾은 첫 번째 후보 유전자는 Os03g0130100이며 1,706,053 ~ 1,708,135(+strand)에 위치하고 있으며 Acyl-activating enzyme 11과 유사한 유전자로

annotation 되어 있었다. 두 번째 후보 유전자는 Os03g0130200이며 1,706,399 ~ 1,711,755(-strand)에 위치하고 있으며 다른 종에서도 상동유전자가 존재하며 단백질을 암호화하는 유전자로 annotation 되어 있었다. 염색체 11번에서 찾은 세 번째 후보 유전자 Os11g0599466이며 22,948,353 ~ 22,954,688 (-strand)에 위치하고 있으며, 유전자의 기능을 할 것으로 annotation 되어

Table 3. Significantly associated with Mg^{2+} contents in 295 accessions.

Name	Description	Position	Length(bp)
Os03g0130100	Similar to Acyl-activating enzyme 11 (Os03t0130100-01)	chr03:1706053..1708135 (+ strand)	2,083
Os03g0130200	Conserved hypothetical protein (Os03t0130200-00)	chr03:1706399..1711755 (- strand)	5,357
Os11g0599466	Hypothetical gene (Os11t0599466-01)	chr11:22948353..22954688 (- strand)	6,336
Os11g0599500	Similar to RNA helicase (Fragment) (Os11t0599500-00)	chr11:22949431..22954549 (+ strand)	5,119

Chromosome	3	3	3	Mean	Std. Deviation	Std. Error	No. of Accession
Position	1,707,046	1,707,451	1,707,733				
Exon No.	2	2	2				
Reference	A	A	C				
Alternative	G	C	T				
Codon	CAA→CAG	GAG→GAC	ACC→ACU				
Amino Acid	Gln→Gln	Glu→Asp	Thr→Thr				
Haplotype1	A	A	C	39.63	6.52	0.45	214
Haplotype2	G	C	C	38.60	5.50	1.74	10
Haplotype3	G	C	T	41.34	5.18	1.44	13

Fig. 2. Haplotype variation in Os03g0130100 genes related Mg^{2+} content among the 295 accessions of rice core set.

있었다. 네 번째 후보 유전자는 Os11g0599500이며 22,949,431 ~ 22,954,549 (+ strand) 위치하고 있으며 RNA helicase와 유사한 기능을 할 것으로 annotation 되어 있었다(Table 3). 또한 각 염색체에서 찾은 후보 유전자에 대하여 haplotyping을 실시하였으며, haplotyping 결과와 표현형 데이터를 비교하여 후보 유전자가 벼의 Mg 함량에 관여하는지 검정하였다.

전장유전체 재분석 핵심집단 295점을 GWAS 분석 방법을 통하여 분석한 결과 후보 유전자 Os03g0130100, Os03g0130200, Os11g0599466, Os11g0599500에 대하여 haplotyping을 하였으며, haplotyping 결과와 마그네슘 분석 데이터를 비교 분석하여 후보 유전자와 벼의 Mg^{2+} 함량 사이의 관련 여부에 대하여 검정을 실시하였다.

염색체 3번 Os03g0130100 유전자내에서는 exon 2번에 3개의 SNP가 있으며, 2개는 synonymous SNP, 1개는 non-synonymous SNP였다. Haplotyping결과 haplotype 그룹이 3개로 나타났다. 각 haplotype 그룹의 Mg^{2+} 함량의 평균은 39.63mg/L, 38.60mg/L, 41.34mg/L 으로 나타났다. Os03g0130100 유전자 1,707,046 위치에서 염기서열이 A에서 G로 바뀐 그룹이 있으며 코돈은 CAA에서 CAG로 바뀌었지만 아미노산은 바뀌지 않았다. Os03g0130100 유전자 1,707,451 위치에서 염기서열이 A에서 C로 바뀐 그룹이 있으며 코돈은 GAG에서 GAC로 바뀌어 아미

노산이 Glutamic acid에서 Aspartic acid로 바뀌었다. Os03g0130100 유전자 1,707,733 위치에서 염기서열이 C에서 T로 바뀐 그룹이 있고 코돈은 ACC에서 ACU로 바뀌었지만 아미노산은 바뀌지 않은 결과를 보였다(Fig. 2).

Chromosome 3번 Os03g0130200 유전자는 2개 exon에서 총 4개의 SNP를 가지고 있으며 그 중 1개는 synonymous SNP, 3개는 non-synonymous SNP였다. Haplotyping결과 haplotype 그룹이 3개로 나타났다. 각 haplotype 그룹의 Mg^{2+} 함량의 평균은 39.04mg/L, 39.55mg/L, 41.86mg/L 으로 나타났다. Os03g0130200 유전자 1,707,046 위치에서 염기서열이 A에서 G로 바뀐 그룹이 있으며 코돈은 CAG에서 CGG로 바뀌어 아미노산이 Glutamine에서 Arginine로 바뀌었다. Os03g0130200 유전자 1,707,451 위치에서 염기서열이 A에서 C로 바뀐 그룹이 있으며 코돈은 GAG에서 GCG로 바뀌어 아미노산이 Glutamic acid에서 Alanine로 바뀌었다. Os03g0130200 유전자 1,711,493 위치에서 염기서열이 C에서 T로 바뀐 그룹이 있으며 코돈은 GAC에서 GAU로 바뀌었지만 아미노산은 바뀌지 않았다. Os03g0130200 유전자 1,711,733 위치에서 염기서열이 C에서 A로 바뀐 그룹이 있으며 코돈은 UAC에서 UAA로 바뀌어 아미노산이 Tyrosine에서 Stop codon으로 바뀌는 결과를 보였다(Fig. 3).

Chromosome	3	3	3	3	Mean	Std. Deviation	Std. Error	No. of Accession
Position	1,707,046	1,707,451	1,711,493	1,711,733				
Exon No.	1	1	2	2				
Reference	A	A	C	C				
Alternative	G	C	T	A				
Codon	CAG→CGG	GAG→GCG	GAC→GAU	UAC→UAA				
Amino Acid	Gln→Arg	Glu→Ala	Asp→Asp	Tyr→*				
Haplotype1	A	A	C	A	39.04	2.06	0.92	5
Haplotype2	A	A	C	C	39.55	6.69	0.48	197
Haplotype3	G	C	T	A	41.86	4.81	1.39	12

*Stop codon

Fig. 3. Number of haplotype variation in Os03g0130200 genes related Mg^{2+} content among the 295 accessions of rice core set, Nucleotide variation in the coding region.

Chromosome	11	11	11	11	Mean	Std. Deviation	Std. Error	No. of Accession
Position	22,949,610	22,950,090	22,950,159	22,950,184				
Exon No.	1	1	1	1				
Reference	C	C	C	A				
Alternative	G	T	G	G				
Codon	UCC→UCG	GGC→GGU	CCC→CCG	AUG→GUG				
Amino Acid	Ser→Ser	Gly→Gly	Pro→Pro	Met→Val				
Haplotype1	C	C	C	A	40.05	7.02	0.56	155
Haplotype2	C	C	G	G	41.29	4.07	0.96	18
Haplotype3	C	T	C	A	37.66	4.90	0.72	47
Haplotype4	G	C	G	G	40.75	5.64	1.56	13

Fig. 4. Number of haplotype variation in Os11g0599466 genes related Mg^{2+} content among the 295 accessions of rice core set, Nucleotide variation in the coding region.

염색체 11번 Os11g0599466 유전자내 1번 exon에서 4개의 SNP가 탐색되었으며 그 중 3개가 synonymous SNP, 1개가 non-synonymous SNP였다. Haplotyping결과 haplotype 그룹이 4개로 나타났다. 각 haplotype 그룹의 Mg^{2+} 함량의 평균은 40.05mg/L, 41.29mg/L, 37.66mg/L, 40.75mg/L으로 나타났다. Os11g0599466 유전자 22,949,610 위치에서 염기서열이 C에서 G로 바뀐 그룹이 있으며 코돈은 UCC에서 UCG로 바뀌었지만 아미노산은 바뀌지 않았다. Os11g0599466 유전자 22,950,090 위치에서 염기서열이 C에서 T로 바뀐 그룹이 있으며 코돈은 GGC에서 GGU로 바뀌었지만 아미노산은 바뀌지 않았다. Os11g0599466 유전자 22,950,159 위치에서 염기서열이 C에서 G로 바뀐 그룹이 있으며 코돈은 CCC에서 CCG로 바뀌었지만 아미노산은 바뀌지 않았다. Os11g0599466 유전자 22,950,184 위치에서 염기서열이 A에서 G로 바뀐 그룹이 있으며 코돈은 AUG에서 GUG로 바뀌어 아미노산이 Methionine(start

codon)에서 Valine로 바뀌는 결과를 보였다(Fig. 4).

Chromosome 11번 Os11g0599500 유전자의 haplotyping결과 haplotype 그룹이 6개로 나타났다. 각 haplotype 그룹의 Mg^{2+} 함량의 평균은 41.25mg/L, 40.68mg/L, 39.97mg/L, 41.23mg/L, 37.57mg/L, 40.75mg/L 으로 나타났으며, 유전자내 8개의 위치에서 염기서열 변화가 있었고 4개의 exon에서 변화를 보였다. Os11g0599500 유전자 22,949,610 위치에서 염기서열이 C에서 G로 바뀐 그룹이 있으며 코돈은 CCU에서 CGU로 바뀌어 아미노산이 Proline에서 Arginine로 바뀌었다. Os11g0599500 유전자 22,950,090 위치에서 염기서열이 C에서 T로 바뀐 그룹이 있으며 코돈은 GCC에서 GUC로 바뀌어 아미노산이 Alanine에서 Valine로 바뀌었다. Os11g0599500 유전자 22,950,159 위치에서 염기서열이 C에서 G로 바뀐 그룹이 있으며 코돈은 CCA에서 CGA로 바뀌어 아미노산이 Proline에서 Arginine로 바뀌었다. Os11g0599500 유전자 22,951,285 위치에서 염기서

Chromosome	11	11	11	11	11	11	11	11	Mean	Std. Deviation	Std. Error	No. of Accession
Position	22,949,610	22,950,090	22,950,159	22,951,285	22,952,950	22,953,028	22,953,517	22,953,724				
Exon No.	1	1	1	3	5	5	6	6				
Reference	C	C	C	G	T	T	T	G				
Alternative	G	T	G	C	C	C	C	C				
Codon	CCU→CGU	GCC→GUC	CCA→CGA	CCG→CCC	UUU→UUC	GGU→GGC	GUU→GUC	GGG→GGC				
Amino Acid	Pro→Arg	Arg→Pro	Ala→Val	Pro→Arg	Phe→Phe	Gly→Gly	Val→Val	Gly→Gly				
Haplotype1	C	C	C	G	C	C	C	C	41.25	3.60	1.27	8
Haplotype2	C	C	C	G	C	T	T	C	40.68	3.26	1.15	8
Haplotype3	C	C	C	G	T	T	T	G	39.97	7.52	0.66	128
Haplotype4	C	C	G	C	C	T	T	C	41.23	4.19	1.02	17
Haplotype5	C	T	C	G	C	T	T	C	37.57	4.87	0.71	47
Haplotype6	G	C	G	C	C	T	T	C	40.75	5.64	1.56	13

Fig. 5. Number of haplotype variation in Os11g0599500 genes related Mg^{2+} content among the 295 accessions of rice core set, Nucleotide variation in the coding region.

Table 4. F-tests by haplotypes of candidate genes related with Mg^{2+} contents.

Gene	Term	Df	Sum of Squares	Mean Square	F value	P-value
Os03g0130100	Group	2	48	23.94	0.583	0.559
	Residuals	235	9646	41.05		
Os03g0130200	Group	2	62	31.16	0.726	0.485
	Residuals	211	9053	42.9		
Os11g0599466*	Group	3	275	91.78	2.247	0.0836
	Residuals	229	9355	40.85		
Os11g0599500	Group	5	304	60.77	1.436	0.212
	Residuals	215	9097	42.31		

열이 G에서 C로 바뀐 그룹이 있으며 코돈은 CCG에서 CCC로 바뀌었지만 아미노산은 바뀌지 않았다. Os11g0599500 유전자 22,950,159 위치에서 염기서열이 C에서 G로 바뀐 그룹이 있으며 코돈은 CCA에서 CGA로 바뀌어 아미노산이 Proline에서 Arginine로 바뀌었다. Os11g0599500 유전자 22,952,950 위치에서 염기서열이 T에서 C로 바뀐 그룹이 있으며 코돈은 UUU에서 UUC로 바뀌었지만 아미노산은 바뀌지 않았다. Os11g0599500 유전자 22,953,028 위치에서 염기서열이 T에서 C로 바뀐 그룹이 있으며 코돈은 GGU에서 GGC로 바뀌었지만 아미노산은 바뀌지 않았다. Os11g0599500 유전자 22,953,517 위치에서 염기서열이 T에서 C로 바뀐 그룹이 있으며 코돈은 GUU에서 GUC로 바뀌었지만 아미노산은 바뀌지 않았다. Os11g0599500 유전자 22,953,724 위치에서 염기서열이 G에서 C로 바뀐 그룹이 있으며 코돈은 GGG에서 GGC로 바뀌었지만 아미노산은 바뀌지 않은 결과를 보였다 (Fig. 5).

위의 4개의 유전자의 각 haplotype간 Mg^{2+} 함량의 차이를 확인하기 위해 F-test를 실시하였다. 모든 haplotype 간 유의한 Mg^{2+} 함량의 차이는 없었으나 Os11g0599466 유전자의 haplotype2와 haplotype3간 t-test를 수행한 결과 t value = 3.0367, p-value = 0.004371로 Mg^{2+} 함량의 유의한 차이를 보였다(Table 4).

*Haplotype2와 Haplotype3 사이의 t-test 결과 t value = 3.0367, p-value = 0.004371로 haplotype간 유의한 차이를 보였다.

벼 핵심집단 Mg^{2+} 검정

전장유전체 재분석 핵심집단 295점에 대하여 ICP-OES 분석을 실시한 결과, 전체의 평균은 39.70mg/L이며, KNURWG-254 품종이 57.11mg/L으로 최댓값으로 측정되었고, KNURWG-226 품종이 18.16mg/L으로 최솟값으로 측정되었다. 핵심집단(HS) 137품종의 평균은 38.51mg/L이며, KNURWG-060 품종이 51.47mg/L으로 최댓값으로 측정되었고, KNURWG-027 품종이 21.12mg/L으로 최솟값으로 측정되었으며, 국내육성종 158품종

Table 5. Mg^{2+} contents by population. HS: KRICE_CORE, KB: breeding line

Population	No. of accessions	Mean (mg/L)	Max (mg/L)	Min (mg/L)	Std. Deviation	Std. Error
Total	295	39.71	57.11	18.17	6.38	0.41
HS	137	38.51	51.47	21.12	5.57	0.58
KB	158	40.44	57.11	18.17	6.74	0.55

Fig. 6. Mg^{2+} content by sub-species. Bars represent standard deviation. n: population size.

의 평균은 40.44mg/L이며, KNURWG-254 품종이 57.11mg/L으로 최댓값으로 측정되었고, KNURWG-226 품종이 18.17mg/L으로 최솟값으로 측정되었다(Table 5).

유전체 분석자원 295점 측정값의 최댓값과 최솟값의 차이는 3.1배의 차이를 보이며 핵심집단, 육성계통, 인디카 그룹, 자포니카 그룹에서는 1.7배~3.1배의 차이를 보여주는 결과를 보였다(Table. 5, Fig. 6). Indica 그룹의 평균은 40.98mg/L (최대 51.47mg/L, 최소 30.25mg/L)이었으며 Japonica 그룹의 평균은 39.50mg/L(최대 57.11mg/L, 최소18.17mg/L)로 나타났다(Fig. 6).

고 찰

작물분야에서 유전체 분석은 NGS 기법을 이용하여, 기존의 기술보다 효율적이고 경제적으로 유전체에 대하여 깊은 연구를 수행하게 되었다(Varshney et al. 2009). 특히, re-sequencing을 이용한 방법은 벼에 국한 되지 않고 다양한 작물에서 이루어지는 실정이며, Lam et al. (2010)은 31점의 재래종 및 육성종 대두에 대해 실시하고, Zheng et al. (2011)은 수수 3개의 품종에 대하여

resequencing을 진행하였다.

현재 벼는 IRGSP에서 진행한 벼 전장 유전체 분석을 ‘Nipponbare’를 기준으로 전장 유전체 연구를 진행하였고(Project, 2005), 세계 각국에서 벼 연구를 진행 중인 연구가, 육종가 등의 벼와 관련된 전문가들에게 Rap-DB 홈페이지에 *Oryza sativa* ssp. japonica의 정보를 제공하는 중이다(Ohyanagi et al. 2006). 또한 벼에서 resequencing은 많은 연구가 이루어지며, Xu et al. (2012)은 벼 품종 50점을 평균 depth 11.8 X로 re-sequencing하여 6.5 백만 개의 SNP를 밝혀내었고, Arai-Kichise et al. (2011)은 분석한 mapping rate 89.7%, SNP 개수 132,462개, Insertion 16,448개 및 deletion19,318개를 찾았다.

마그네슘 함량은 단백질함량, 현미 립장, 정조 립장, 간장, 이삭수 및 이삭 당 종자 수와 미숙립 비율에 연관성이 있으며, 그 중 단백질 함량에 영향을 주는 것으로 보고된 바가 있으며(Zeng et al. 2005), 마그네슘 함량과 관련하여 현재까지 밝혀진 유전자로는 Gardner가 보고한 바와 같이 MRS2 family가 있으며, 이 유전자는 마그네슘의 이동에 연관이 있는 것으로 밝혀졌다(Gardner, 2003). 또한 Kang et al. (1998)은 벼의 화학적 조성에 관하여 RIL(Recombination Inner Line) 집단으로 QTL 분석을

실시해 염색체 7번에 있는 마그네슘과 칼륨의 비율에 관여하는 MgK1을 찾아내었다. Garcia-Oliveira et al. (2009)은 85점의 ILS(Introgression Lines)로 철, 아연, 망간, 구리, 칼슘, 마그네슘, 인, 철 그리고 칼륨의 함량에 관여하는 QTL 분석을 실시하였으며, 염색체 1, 3, 5, 9 및 12번에서 각각 qMG1-1과 qMG3-1, qMG5-1, qMG9-1, qMG12-1이 마그네슘 함량에 관여하는 것을 밝혔다.

본 연구에서 GWAS 분석에 의해 기존 연구에서 연관관계가 밝혀지지 않은 후보 유전자 Os03g0130100, Os03g0130200, Os11g0599466, Os11g0599500를 동정하였다. Haplotype 그룹 간 분산분석을 수행한 결과 유의한 Mg^{2+} 함량차이를 보이지 않았으나 Os11g0599466의 haplotype2와 haplotype3 사이에서 t-test를 수행한 결과 유의한 Mg^{2+} 함량차이를 발견할 수 있었다. 본 실험에서는 후보유전자들의 SNP를 이용한 haplotyping을 수행하였는데, 유전자 영역의 INDEL 또한 유전자 기능에 영향을 미칠 수 있어 앞으로의 연구에서 INDEL을 이용한 haplotyping을 수행할 필요가 있다. 본 연구를 통하여 haplotype 결과에 따른 벼 Mg^{2+} 함량 관련 후보 유전자인 Os03g0130100을 포함한 4개의 유전자들 중 특이 염기서열은 차후 분자마커 제작이나 유전자편집을 이용한 마그네슘 고품종 품종육종 분야 및 기능성 품종 육성에 기초자료로 유용하게 이용될 것으로 사료된다.

적 요

본 연구는 유전체 재분석 국내 벼 핵심집단 295점에 대하여 벼의 Mg^{2+} 함량에 관여하는 유전자 발굴 및 육종을 위한 소재 개발 또는 DNA마커 제작에 활용할 수 있는 유전체 육종의 기초 정보를 제공하고자, 수확 후 보관중인 종자를 ICP-OES장비를 이용하여 Mg 함량 데이터를 얻은 후, GWAS 분석을 통해 Mg^{2+} 함량 연관 유전자를 동정하였다. 또한 연관 유전자에 대하여 haplotyping을 수행하여 Mg^{2+} 함량과 연관된 특정 변이를 추적하였다. 그 결과를 요약하면 다음과 같다.

Mg 함량을 측정하기 위하여 ICP-OES 분석을 실시한 결과, 전체의 평균은 39.70mg/L이며, KNURWG-254 품종이 57.11mg/L으로 최댓값으로 측정되었고, KNURWG-226 품종이 18.16mg/L으로 최솟값으로 측정되었다.

전체 유전체 재분석 핵심집단 295점의 표현형 데이터와 유전자형 데이터를 GWAS 분석을 통해 Mg^{2+} 함량과 연관된 후보유전자를 선별하였다. 후보 유전자는 3번, 11번 염색체에 존재하였다.

사 사

본 논문은 농림축산식품부의 재원으로 농림수산식품기술기획평가원의 농생명산업기술개발사업의 지원을 받아 연구되었음(315025-03-2-HD030)

REFERENCES

1. Arai-Kichise Y, Shiwa Y, Nagasaki H, Ebana K, Yoshikawa H, et al. 2011. Discovery of genome-wide DNA polymorphisms in a landrace cultivar of Japonica rice by whole-genome sequencing. *Plant and cell physiology* 52: 274-282.
2. Choe J-S, Ahn H-H, Nam H-J. 2002. Comparison of nutritional composition in Korean rices. *Journal-Korean Society of Food Science and Nutrition* 31: 885-892.
3. 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.
4. Fatemi S, Ryzen E, Flores J, Endres DB, Rude RK. 1991. Effect of experimental human magnesium depletion on parathyroid hormone secretion and 1, 25-dihydroxyvitamin D metabolism. *The Journal of Clinical Endocrinology & Metabolism* 73: 1067-1072.
5. Garcia-Oliveira AL, Tan L, Fu Y, Sun C. 2009. Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain. *Journal of Integrative Plant Biology* 51: 84-92.
6. Gardner RC. 2003. Genes for magnesium transport. *Current opinion in plant biology* 6: 263-267.
7. Juliano B, Bechtel D. 1985. *Rice Chemistry and Technology*. 2nd ed. The American Association of Cereal Chemists, St. Paul, MN. pp. 17-57.
8. Kang H-J, Cho Y-G, Lee Y-T, Kim Y-D, Eun M-Y, Shim J-U. 1998. QTL mapping of genes related with grain chemical properties based on molecular map of rice. *Korean Journal of Crop Science* 43: 199-204.
9. Kim K-W, Chung H-K, Cho G-T, Ma K-H, Chandrabalan D, et al. 2007. PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. *Bioinformatics* 23: 2155-2162.
10. Kim M-R. 2011. The status of Korea's rice industry and the rice processing industry. *Food Industry and Nutrition* 16: 22-26.

11. Kim T-S, He Q, Kim K-W, Yoon M-Y, Ra W-H, et al. 2016. Genome-wide resequencing of KRICE_CORE reveals their potential for future breeding, as well as functional and evolutionary studies in the post-genomic era. *BMC genomics* 17: 408.
12. Kwon SW, Chu SH, Han SJ, Ryu SN. 2011. A New Rice Variety 'Superjami' with High Content of Cyanidin 3-glucoside. *Korean Journal of Breeding Science* 43: 196-200.
13. Lam H-M, Xu X, Liu X, Chen W, Yang G, et al. 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nature genetics* 42: 1053-1059.
14. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754-1760.
15. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, et al. 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28: 2397-2399.
16. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 20: 1297-1303.
17. Ohyanagi H, Tanaka T, Sakai H, Shigemoto Y, Yamaguchi K, et al. 2006. The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucleic acids research* 34: D741-D744.
18. Park D-S, Hwang U-H, Park S-K, Lee J-H, Han S-I, et al. 2015. A waxy black giant embryo early maturing rice variety 'Nunkeunheugchal'. *Korean J Breed Sci* 47: 68-74.
19. Project IRGS. 2005. The map-based sequence of the rice genome. *Nature* 436: 793-800.
20. Rude RK. 1998. Magnesium deficiency: a cause of heterogenous disease in humans. *Journal of Bone and Mineral Research* 13: 749-758.
21. Son SM. 2001. Rice based meal for prevention of obesity and chronic disease. *Korean J Community Nutr* 6: 862-867.
22. Varshney RK, Nayak SN, May GD, Jackson SA. 2009. Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in biotechnology* 27: 522-530.
23. Xu X, Liu X, Ge S, Jensen JD, Hu F, et al. 2012. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature biotechnology* 30: 105-111.
24. Zeng Y-W, Shen S-Q, Wang L-X, Liu J-F, Pu X-Y, et al. 2005. Correlation of plant morphological and grain quality traits with mineral element contents in Yunnan rice. *Rice Sci* 12: 101-106.
25. Zheng L-Y, Guo X-S, He B, Sun L-J, Peng Y, et al. 2011. Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*). *Genome biology* 12: R114.