

벼종자 미량 단백질의 프로테오믹스 연구를 위한 글루테린 저장 단백질의 제거방법

우선희* · 김세영** · 김태선* · 조성우* · 조 건** · 정근욱*** · 김선림**** · 조용구*
김홍식* · 송범현* · 이철원* · 정승근* · 박영목** · 최중순**†

*충북대학교 식물자원학과, **한국기초과학지원연구원, ***충북대학교 농화학과, ****작물과학원

A New Removal Method of Glutelin Storage Proteins for the Proteome Study of Non-Glutelin Proteins in Rice Seeds

Sun-Hee Woo*, Se-Young Kim**, Tae-Seon Kim*, Seong-Woo Cho*, Kun Cho**, Keun-Yook Chung***
Sun-Lim Kim****, Yong-Gu Cho*, Hong-Sig Kim*, Boem-Heon Song*, Chul-Won Lee*
Seung-Keun Jong*, YoungMok Park**, and Jong-Soon Choi**†

*Department of Crop Science, Chungbuk National University, Cheongju 361-763, Korea

**Proteomics Team, Korea Basic Science Institute, Daejeon 305-333, Korea

***Department of Agricultural Chemistry, Chungbuk National University, Cheongju 361-763, Korea

****National Institute of Crop Science, Suwon 441-857, Korea

ABSTRACT Abundant proteins often cause problems in proteome study. Glutelin family proteins (hereafter referred to glutelin) are present in rice proteome sample as overwhelming constituents with very high abundance. In order to increase the number of identified proteins in rice proteome study, we developed a newly improved method for sample preparation through the removal of glutelin. When the protein samples from rice seed were extracted by the conventional trichloroacetic acid (TCA) extraction method, glutelin accounts for about 60% of total rice seed proteins in SDS gels. Using our new water extraction method, glutelin consists of only about 10% of total proteins. After analyzing on a two-dimensional gel electrophoresis (2-DE), 937 protein spots were detected using the conventional TCA extraction method. On the other hand, 1240 proteins could be seen using the new water extraction method. The selectivity for non-glutelin and less abundant protein by the water extraction method was also confirmed by ESI-Q/TOF mass spectrometry analysis. Thus, the new water extraction method developed here can be efficiently used to study the proteome analysis of rice storage seed.

Keywords : rice, proteome, two-dimensional gel electrophoresis, removal of abundant protein, glutelin

포스트게놈 시대에 생명체들의 유전자에서 발견되는 단백질을 조사하고 이들의 기능을 밝혀내는 프로테오믹스 분야의 연구는 신약개발과 질병예방을 위한 기초연구로서의 중요성을 가진다. 최근 유전체 염기서열의 고속 분석기법이 개발되면서 많은 생물 개체들의 유전체 염기서열이 밝혀지고 있다. 그러나 어떤 생물체의 유전체 염기서열이 밝혀졌다고 해도 최종적으로 완벽한 모양이 갖추어진 단백질을 분석하지 않고는 그 유전자에 대한 연구가 필요하게 된다. 프로테오믹스(proteomics)는 유전자 정보로부터 만들어진 프로테오믹스(단백질체)를 대상으로 유전자의 기능, 단백질의 기능이상 및 구조변형 유무 등을 규명하고 질병 과정 등을 추적하는 분석기술로서, 단백질체학이라고도 한다. 따라서 유전체 연구와는 별도로 단백질체(단백질)와 옴(ome; 전체)의 합성어로서, 게놈이 사람이 지닌 모든 유전정보의 집합체라면 프로테오믹스는 특정 세포나 특수 상황에서 만들어지고 작용하는 단백질의 총집합이다. 따라서 프로테오믹스는 생명체의 전체 유전자인 게놈에 의해 발견되는 모든 단백질의 총합인 프로테오믹스를 다루는 학문으로, 단백질들을 대량으로 분석하고 상호기능관계 지도를 작성하며 구조분석을 통해 궁극적으로 특정 단백질과 이를 만드는 유전자의 기능을 동시에 밝혀내는 것을 목적으로 하는 것이다. 고등 진핵생물의 프로테오믹스 분석과정에서, 고품량 단백질들의 존재는 미량 단백질들을 분석하는데 장애가 되고 있다. 프로테오믹스 연구에 어려

†Corresponding author: (Phone) +82-42-865-3428
(E-mail) jschoi@kbsi.re.kr

음을 주는 고함량 단백질의 예로 인간 플라즈마 시료의 혈청 알부민이 있다(Anderson & Anderson, 2002). 다량의 고함량 단백질들의 존재는 2-DE(2차원 전기영동)를 이용하는 프로테오믹스 연구에서 다음과 같은 많은 문제들을 초래한다. ① IPG 스트립에 로딩되고 분리될 수 있는 총 단백질 양이 5 mg으로 제한되기 때문에 대부분의 미량 단백질들은 2-DE 겔에 의해 분석되기 어렵다. ② 2-DE 겔을 Coomassie 또는 은 염색하면 고함량 단백질들의 밀집 스팟이 넓게 형성되어 다른 단백질 스팟이 파묻힐 수 있고, 따라서 미량 단백질들은 감지되지 않을 것이다. 그러므로 고함량 단백질들을 효과적으로 제거할 수 있는 수단이 없다면 종래의 프로테오믹스 분석방법들은 미량 단백질들을 광범위하게 감지해 내기에 적합하지 않다. 이러한 문제를 해결하기 위해서, 최근 몇 년 동안 프로테오믹스 연구 분야에서 고함량 단백질을 효과적으로 분리하거나 제거하는 방법의 개발에 대한 연구가 행하여졌다. 최근 개발된 고함량 단백질 제거방법들로는 친화성을 이용한 방법(affinity method)으로 예를 들면, 인간 항혈청 알부민 항체 컬럼(Sato *et al.*, 2002), 단백질 A/G(Adkins *et al.*, 2002), 면역 친화 수지(Steel *et al.*, 2003), 음이온 교환이나 분자량-배제(size exclusion) 크로마토그래피(Pieper *et al.*, 2003)등과 연속형 크로마토그래피 등이 있다.

벼는 한국을 포함한 아시아의 주식으로, 이에 대한 유전학적, 생리학적, 단백질체학적, 병리학적 특성의 규명과 품질, 생산수량 등의 개선은 인구의 약 50%를 차지하는 아시아인에 대한 식량의 안정적 제공이라는 점에서 매우 중요한 사안이 되고 있다. 본 연구는 벼종자의 단백질체학(proteomics; 이하 “프로테오믹스”라 함) 연구 방법론에 관한 것이다. 글루테린류 단백질들은 벼종자에 다량 존재하는데, 이러한 고함량 단백질은 다른 다수의 단백질들을 분석하는데 장애가 되고 있다. 발아시 질소원으로 사용되는 글루테린은 벼종자의 주요 저장 단백질로서 총 벼종자 단백질의 약 70%를 차지한다(Takaiwa *et al.*, 1991). 글루테린은 벼종자의 주요 구성요소이며, 비글루테린 단백질, 특히 미량 단백질들의 감지와 동정(identification)을 어렵게 한다. 종래 전형적으로 이용되고 있는 추출법(예를 들면, 세포용해 완충액 추출법 또는 TCA 추출법)에 의한 추출물을 SDS-PAGE 또는 질량 스펙트로메트리로 분석하면 단백질 중 대부분이 글루테린이다. 벼 프로테오믹스 연구를 위한 단백질 시료들을 준비하기 위해서, 두 가지 방법(세포용해 완충액법(O’Farrell, 1975; Hirano *et al.*, 1991)과 TCA 완충액법(Ramani *et al.*, 1997)이 일반적으로 사용되고 있다. 이들 방법은 잎, 뿌리, 줄기와 같은 벼 조직의 프로테오믹스 연구에 사용된다. 그러나

이들 방법은 벼종자에 다량으로 존재하는 고함량 글루테린 저장 단백질을 추출하는 경향이 있기 때문에 벼종자 시료들에 대한 효율적인 방법이 아니다. 따라서, 글루테린을 효율적으로 제거하고 비글루테린 단백질들 특히, 미량 단백질들을 선택적으로 추출하는 새로운 추출법이 요구되고 있다.

본 연구는, 벼 글루테린 저장 단백질을 효과적으로 제거하기 위한, 신속 저렴하고 효율적인 새로운 추출법을 제공하는 것을 목적으로 한다. 또한 상기 추출법을 2-DE PAGE에 의해 분리되고 ESI Q-TOF 질량 스펙트로메트리에 의해 동정된 배아(embryo)와 내배유(endosperm) 단백질들의 동정에 적용하는 방법을 제공하는 것을 목적으로 하였다.

재료 및 방법

벼종자 글루테린 저장 단백질의 제거

한국작물과학원에서 육성한 야생형 벼(*Oryza sativa* L.), 일품벼(Japonica type)를 공시하였다. 벼의 성장과 곡물의 채취를 위해 제어 가능한 환경시설들을 이용하였다. 개화기를 7주(49일) 경과할 때 종자 시료를 채취하고 -80℃에서 저장하면서 사용하였다. 채취된 벼종자를 액체 질소와 함께 30분 이상 분쇄하여 미세한 가루를 만들었다. 가루를 2차 증류수에 현탁하였다. 현탁액을 0℃에서 30초간 3회 초음파 처리하여 단백질들이 용해되도록 하였다. 용해액을 10,000 g로 15분간 2회 원심분리하여 용해되지 않은 물질을 제거하였다. Bradford 법(1976)으로 상등액 중의 단백질 농도를 측정하였다. 한편, 대조를 위해 종래의 TCA 법(Woo *et al.*, 2002)에 따라 단백질을 추출하였다. 즉, 벼종자를 액체 질소와 함께 30분 이상 분쇄하여 미세한 가루를 만들고, 상기 가루를 10% w/v TCA 및 0.07% w/v β-mercaptoethanol이 함유된 -20℃의 아세톤 용액에 1시간 동안 현탁한 다음, 현탁액을 10,000 g로 15분 동안 원심분리 하였다. 분리된 펠릿을 0.07% β-mercaptoethanol이 함유된 냉 아세톤 용액으로 세척하고 -20℃에서 1시간 방치한 후 4℃에서 원심분리 하였다. 이 과정을 3번 반복한 후 펠릿을 동결건조 하였다. 동결건조된 시료를 재수화(rehydration) 완충액(7M urea, 2M thiourea, 4% CHAPS, 0.28% w/v DTT, 0.5% v/v IPG 완충액)에 5초간 4번을 용해시킨 다음 BSA를 표준물질로 하여 Bradford 법(1976)으로 단백질 함량을 측정하였다.

이차원전기영동을 이용한 단백질분리

물추출법 및 TCA 추출법에 의해 추출된 벼종자 단백질들을 2-DE(Amersham Pharmacia Biotech, Uppsala, Sweden)

분석하였다. 이를 위해 먼저, 종래 방법(Woo *et al.*, 2003)을 조금 변경시켜 2-DE를 수행하였다. 실온의 reswelling tray (Amersham Pharmacia Biotech, Uppsala, Sweden)에서 IPG (immobilized pH gradient) 스트립(18cm, pH 3-10)을 재수화 완충액(위와 동일한 조성) 350 μ l로 하룻밤 재수화하였다. 각 IPG 스트립에 각각 단백질 350 μ g을 로딩하였다. Pharmacia Mutiphore II와 Drystrip 키트를 사용하여 20°C에서 Isoelectric focussing을 유도하였다(수행조건; 200V 1시간, 500V 1시간, 1000V 1시간, 최종적으로 8000V 8시간). 이차원 분리에서 수직의 판은 Ethan Dalt 전기영동 키트를 사용한 SDS-PAGE(13% acrylamide monomer, 2.6%의 crosslinker) 겔로 수행하였다. 분석겔상의 단백질 스팟이 가시화되도록 Heudeshoven과 Dernick이 제시한 방법(Heukeshoven and Dernick, 1985)에 따라 은질산염 염색하였다.

추출된 단백질의 분리 및 서열분석

앞에 실시한 물추출법을 이용하여 성숙한 벼 종자에서 추출된 단백질들을 분리하고 서열분석 하였다.

1) 단백질의 분리

앞의 2-DE겔에서, 각 레인의 동일한 위치에 대응되고, 두 추출법 사이에서 뚜렷한 차이를 보이는 3개의 단백질 밴드(Fig. 1A 및 1B에서 화살표 α , β , γ)를 선택하여 절취한 다음 트립신(Promega, Madison, WI, USA)으로 처리하였다. 겔 조각을 10 mM 암모늄 중탄산염과 50% 아세토니트릴로 세척한 다음 50 mM 암모늄 중탄산염, 5 mM CaCl₂, 12.5 ng/ μ l 트립신이 함유된 소화 완충액에 담가 팽창시킨 후, 37°C에서 12~16 시간 방치하였다. 50 mM 암모늄 중탄산염 및 100% 아세토니트릴의 2단계 추출로 서열분석을 위한 시료 펩타이드를 회수하였다. 추출된 펩타이드를 모아 진공 원심분리기에서 동결건조하여 -20°C에서 보관하였다.

2) 단백질 서열분석

앞에서 준비된 각 펩타이드 시료를 ESI Q-TOF MS를 이용한 MS/MS 분석으로 서열 분석하였고, MASCOT 검색 소프트웨어를 사용하여 단백질을 동정하였다. 각 밴드 세트(밴드 α , β , γ)로부터 동정된 단백질을 Table 1에 나타내었다. 구체적으로는, 극 HPLC 시스템 및 nano-ESI source가 장착된 Q-TOF2 질량 스펙트로메터(Micromass, U.K)로 구성된 nano-LC/MS 시스템을 사용하여 펩타이드 서열분석을 위한 MS/MS 실험을 수행하였다. 탈염 및 농축을 위하여, 오토샘플러(FAMOS, LC Packings)로 앞에서 준비된 각각 시

료 10 μ l 씩을 C18 trap 컬럼(I.D 300 μ m, 길이 5 mm, 파티클 사이즈 5 μ m; LC Packings)에 로딩하였다. 이때 유속은 20 μ l/min로 하였다. 컬럼에 흡착되어 있던 펩타이드를 역세한 다음 C18 나노-컬럼(I.D 75 μ m, 길이 150 mm, 파티클 사이즈 5 μ m; LC Packings)상에서 분리하였다. 유속이 150 μ l/min이 되도록 하면서, 0% 아세토니트릴을 10분간, 이어서 0% - 50% 아세토니트릴을 80분 이상 흘려준 다음 50% 아세토니트릴을 10분간 흘려주는 농도구배를 따랐다. Nano-electrospray 이온화 source에서, nano-LC 컬럼의 모세관 말단을 pico-tip silica tubing(I.D 8 μ m; New Objectives, USA)에 연결시켰다. Electrospray를 만들기 위해 액체접합부(liquide junction)에 가해진 적용전압은 2-2.5 kV이고, cone전압은 30V였다. 충돌가스로 아르곤을 10psi가 되도록 가하였다. 충돌에너지가 점진적으로 25, 30 및 35 eV로 증가되도록 하면서, 데이터 의존형 MS/MS 모드 형태의 MS/MS 스펙트럼을 얻었다.

결과 및 고찰

Fig. 1A는 종래의 TCA 추출법(I)과 본 연구에 의한 물추출법(II)에 의한 벼종자 총 단백질의 SDS-PAGE 결과를 나타낸 것이다. 레인 III은 레인 I 및 레인 II의 시료 절반씩을 혼합한 시료이다. 분자량의 기준은 왼쪽에 표시되어 있다. Fig. 1B는 이미지 분석 소프트웨어인 Image Quant TL (Amersham Pharmacia Biotech, Uppsala, Sweden)를 이용하여 얻은 SDS-PAGE의 덴시토그램이다. I, II, III은 위 SDS-PAGE 사진(A)의 레인 I, II, III의 덴시토그램을 나타낸다. 두꺼운 화살표는 글루테린 단백질들을 의미한다. α , β , γ 의 화살표는 레인 I과 II에서 가장 큰 차이를 보여주는 단백질 밴드를 나타낸다.

물추출법에 따라 벼종자의 단백질을 분리하였을 때, SDS-PAGE 겔상에서 글루테린 밴드의 강도가 감소되었다. TCA 추출법에 의하면 벼종자 총 단백질의 약 70%가 글루테린인 것으로 측정된다(Fig. 1A의 레인 I 및 Fig. 1B의 덴시토그램 I). 반면 본 시험에 의한 물추출법에 의하면 글루테린 단백질은 총 단백질의 10% 미만이며(Fig. 1A의 레인 II 및 Fig. 1B의 덴시토그램 II), 더 많은 미량 비글루테린 단백질의 밴드들이 SDS겔 상에서 더 높은 강도로 가시화됨을 확인하였다.

Fig. 2는 종래의 TCA 방법(A)과 본 시험에 의한 물추출법(B)으로 추출된 벼종자 총 단백질의 2-DE겔 결과사진을 나타낸 것이다. 물추출법에 의한 시료에서, 굵은 선으로 표시된 부위 2와 4의 단백질 스팟이 현저하게 감소되었다. 반면에

Table 1. List of proteins identified by ESI Q-TOF MS from 1D-gel slices (thin arrows α , β , and γ in Figure 1A) showing the greatest differences between TCA precipitation and water extraction methods.

Water extraction method

Band α

Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 3023816	Glyceraldehyde 3-phosphate dehydrogenase, Cytosolic	36470	10	Q42977
gi 2130073	Fructose-bisphosphate aldolase (EC 4.1.2.13) isoenzyme C-1, Cytosolic	38799	10	S65072
gi 113622	Fructose-bisphosphate aldolase, Cytoplasmic isozyme	38786	8	P17784
gi 10140741	Cytoplasmic malate dehydrogenase	35546	5	AAG13573
gi 14164433	Putative malate dehydrogenase	35439	7	BAB55686
gi 584706	Aspartate aminotransferase, Cytoplasmic (transaminase A)	44479	6	P37833
gi 3646373	RGP1 protein	39452	5	CAA09469
gi 4097102	Globulin-like protein	51700	4	AAD10375
gi 21263611	Formate dehydrogenase, Mitochondrial precursor (NAD-dependent formate dehydrogenase)	41217	3	Q9SXP2
gi 15290145	Putative prepro-glutelins	40238	2	BAB63836
gi 6015432	OSEYA1	31823	2	BAA85161
gi 113240	ACTIN 2	41976	2	P17298
gi 13603448	Unnamed protein product (<i>Saccharomyces cerevisiae</i>)	191217	1	CAA97554
gi 7489449	Pullulanase	109524	2	T03760
gi 16580747	Glyoxalase I	32519	1	BAB71741
gi 7434944	Pyruvate, Phosphate dikinase (EC 2.7.9.1) precursor	102736	2	T02979
gi 14090214	Putative plastidic aldolase	41968	1	BAB55475
gi 16566704	Putative fructokinase II	35494	1	AAL26573
gi 12082087	UV-damaged DNA binding protein	121868	2	BAB20761
gi 2662341	EF-1 alpha	49262	1	BAA23657
gi 121349	Glutamine synthetase shoot isozyme (Glutamate-ammonia ligase)	39176	2	P14656
gi 14028985	Gst6 (<i>Arabidopsis</i>)	123271	1	CAA64613
gi 121477	Glutelins precursor	56783	1	P14614
gi 7248401	Putative chaperonin 60 beta precursor	64046	2	BAA92724

Band β

Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 14164433	Putative malate dehydrogenase	35439	9	BAB55686
gi 4126809	Glyoxalase I	32533	10	BAA36759
gi 10140741	Cytoplasmic malate dehydrogenase	35546	9	AAG13573
gi 3023816	Glyceraldehyde 3-phosphate dehydrogenase, Cytosolic	36470	7	Q42977
gi 15624063	Putative NAD-malate dehydrogenase	42693	4	BAB68116
gi 730580	60S acidic ribosomal protein P0	34356	5	P41095
gi 4097102	Globulin-like protein	51700	6	AAD10375
gi 2130073	Fructose-bisphosphate aldolase (EC 4.1.2.13) isoenzyme C-1, cytosolic	38799	4	S65072

Table 1. Continued.

Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 1346109	Guanine nucleotide-binding protein beta subunit-like protein (GPB-LR) (RWD)	36209	4	P49027
gi 7209794	Protein disulfide isomerase	33382	3	BAA92322
gi 584706	Aspartate aminotransferase, Cytoplasmic (Transaminase A)	44479	4	P37833
gi 2662343	EF-1 alpha	49248	2	BAA23658
gi 16566704	Putative fructokinase II	35494	2	AAL26573
gi 16566707	Putative fructokinase I	34698	3	AAL26574
gi 5777592	Low molecular weight globulin	17966	2	CAA44001
gi 2499819	Aspartic proteinase oryzasin 1 precursor	54111	1	Q42456
gi 4096173	Early embryogenesis protein	58735	1	AAD10369
gi 2118428	AllergenRA14C precursor	17373	2	S59923
gi 13603448	Unnamed protein product	191217	2	CAA97554
gi 10140799	Putative steroid membrane binding protein	24619	2	AAG13629
gi 169805	Unknown protein	19800	1	AAA72362
gi 121473	Glutelin type I precursor	55808	1	P07729
gi 556403	Glutelin	56205	1	AAA50317
gi 1707986	Glutelin type-A III precursor	56040	1	Q09151
gi 15290145	Putative prepro-glutelin	40238	1	BAB63836
gi 3172048	Starch debranching enzyme	109790	2	BAA28632
gi 5852089	Zwh12.1	112089	2	CAB55396

Band γ

Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 7435021	GF14-b protein-rice	29845	10	T04152
gi 7271253	14-3-3-like protein	29674	8	CAB77673
gi 13702816	Glucose-6-phosphate isomerase	29160	7	CAA61577
gi 7435022	GF14-c protein	28808	8	T04153
gi 1351270	Triosephosphate isomerase, Cytosolic (TIM)	27046	7	P48494
gi 15289953	Putative ribosomal protein S4	28601	9	BAB63648
gi 232161	19 kDa globulin precursor (alpha-globulin)	21036	7	P29835
gi 7435023	GF14-d protein-rice	29244	4	T04154
gi 585337	Adenylate kinase A (ATP-AMP transphosphorylase)	26389	6	Q08479
gi 4126809	Glyoxalase I	32533	4	BAA36759
gi 3123270	40S ribosomal protein S4 (SCAR PROTEIN SS620)	29976	6	P49398
gi 121476	Glutelin precursor	56529	2	P14323
gi 1350986	40S ribosomal protein S3A (CYC07 PROTEIN)	29709	4	P49397
gi 4097102	Globulin-like protein	51700	4	AAD10375
gi 3023816	Glyceraldehyde 3-phosphate dehydrogenase, Cytosolic	36470	2	Q42977
gi 121477	Glutelin precursor	56783	2	P14614
gi 6166140	Elongation factor 1-beta (EF-1-beta)	24847	4	Q40680
gi 5257277	S-adenosyl-L-methionine	28370	2	BAA81776

Table 1. Continued.

Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 585338	Adenylate kinase B (ATP-AMP transphosphorylase)	26658	2	Q08480
gi 6856560	Isopentenyl pyrophosphate : dimethylallyl pyrophosphate isomerase	29009	2	AAF29978
gi 12644044	Proteasome subunit alpha type 7 (20S proteasome alpha subunit D)	27084	3	O04861
gi 2499819	Aspartic proteinase oryzasin 1 precursor	54111	3	Q42456
gi 7435885	Probable C. 3. 4. 25. 1 proteasome endopeptidase complex (EC 3.4.99.46) alpha chain	27279	2	T04300
gi 3894214	Elongation factor 1 beta 2	24639	2	BAA34598
gi 2662343	EF-1 alpha	49248	2	BAA23658
gi 7489449	Pullulanase	109524	2	T03760
gi 12328500	Unnamed protein product	27553	2	BAB21159
gi 548774	60S ribosomal protein L7A	29285	1	P35685
gi 5360230	Ran	25023	1	CAA84843
gi 730580	60S acidic ribosomal protein po	34356	2	P41095
gi 7441868	DnaK-type molecular chaperone BiP	73495	2	T03581
gi 3334279	Oleosin 18 kDa (OSE721)	17269	1	Q40646

TCA extraction method

Band α

Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 544399	Glutelin type-B1 precursor	56640	9	Q02898
gi 544400	Glutelin type-B2 precursor	56190	8	Q02897
gi 100680	Glutelin 22 precursor	55820	8	D34332
gi 121477	Glutelin precursor	56783	7	P14614
gi 2130073	Fructose-bisphosphate aldolase isoenzyme C-1	38799	7	S65072
gi 3023816	Glyceraldehyde 3-phosphate dehydrogenase, Cytosolic	36470	7	Q42977
gi 14587314	Glutelin type I precursor	56212	6	BAB61225
gi 121475	Glutelin type II precursor	56271	6	P07730
gi 4126687	Glutelin	57084	5	BAA36695
gi 14164433	Putative malate dehydrogenase	35439	5	BAB55686
gi 809064	Glutelin	25507	2	CAA29507
gi 10140741	Cytoplasmic malate dehydrogenase	35546	3	AAG135753
gi 7489571	R40c1 protein	38798	3	T03911
gi 9927273	Unnamed protein product	328372	3	BAA96774
gi 584706	Aspartate aminotransferase, Cytoplasmic (Transaminase A)	44479	1	P37833
gi 12002701	Putative cytosine DNA methyltransferase	169482	2	AAG43376

Table 1. Continued.

Band β				
Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 225710	Glutelin	56157	7	1311273A
gi 14587314	Glutelin type I precursor	56212	9	BAB61225
gi 225174	Glutelin precursor	56196	8	1210248A
gi 121476	Glutelin precursor	56529	6	P14323
gi 100680	Glutelin 22 precursor	55820	5	D34332
gi 556403	Glutelin	56205	7	AAA50317
gi 121477	Glutelin precursor	56783	5	P14614
gi 1707986	Glutelin type-A III precursor	56040	5	Q09151
gi 556401	Glutelin	56310	5	AAA50316
gi 2130073	Fructose-bisphosphate aldolase (EC 4.1.2.13) isoenzyme C-1, Cytosolic	38799	5	S65072
gi 809064	Glutelin	25507	4	CAA29507
gi 14164433	Putative malate dehydrogenase	35439	2	BAB55686
gi 10140741	Cytoplasmic malate dehydrogenase	35546	4	AAG135753
gi 113622	Fructose-bisphosphate aldolase, Cytoplasmic isozyme	38786	4	P17784
gi 3023816	Glyceraldehyde 3-phosphate dehydrogenase, Cytosolic	36470	5	Q42977
gi 4097102	Globulin-like protein	51700	4	AAD10375
gi 4126687	Glutelin	57084	3	BAA36695
gi 730580	60S acidic ribosomal protein P0	34356	2	P41095
gi 12328482	Unknown protein	78363	2	BAB21142
gi 1363512	Glutelin	9267	1	PQ0722
gi 15217255	Putative retroelement	163930	2	AAK92599
gi 13603448	Unnamed protein product	191217	2	CAA97554
gi 2130065	Alpha-globulin precursor	21041	1	JC4784
gi 12039395	Putative ankyrin protein	27913	1	AAA933397

Band γ				
Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 121476	Glutelin precursor	56529	12	P14323
gi 544400	Glutelin type-B2 precursor	56190	11	Q02897
gi 14587314	Glutelin type I precursor	56212	9	BAB61225
gi 4126687	Glutelin	57084	7	BAA36695
gi 121477	Glutelin precursor	56783	9	P14614
gi 1707986	Glutelin type-A III precursor	56040	8	Q09151
gi 100680	Glutelin 22 precursor	55820	7	D34332
gi 232161	19 kDa globulin precursor (Alpha-globulin)	21036	6	P29835
gi 72300	Glutelin II precursor	56285	6	FWR22
gi 556403	Glutelin	56205	6	AAA50317
gi 809064	Glutelin	25507	2	CAA29507
gi 1363512	Glutelin 8.0/31K	9267	1	PQ0722

Table 1. Continued.

Genebank Identifier	Protein	MW	Peptide Matches	Protein Accession Number
gi 14028985	Gst6 (Arabidopsis)	123271	1	CAA64613
gi 4097102	Globulin-like protein	51700	1	AAD10375
gi 13603448	Unnamed protein product	191217	3	CAA97554
gi 4680491	Hypothetical protein	98131	2	AAD27671
gi 12328482	Unnamed protein product	78363	1	BAB21142
gi 9558523	Unnamed protein product	144503	2	BAB03441
gi 9558523	Unnamed protein product	144503	2	BAB03441

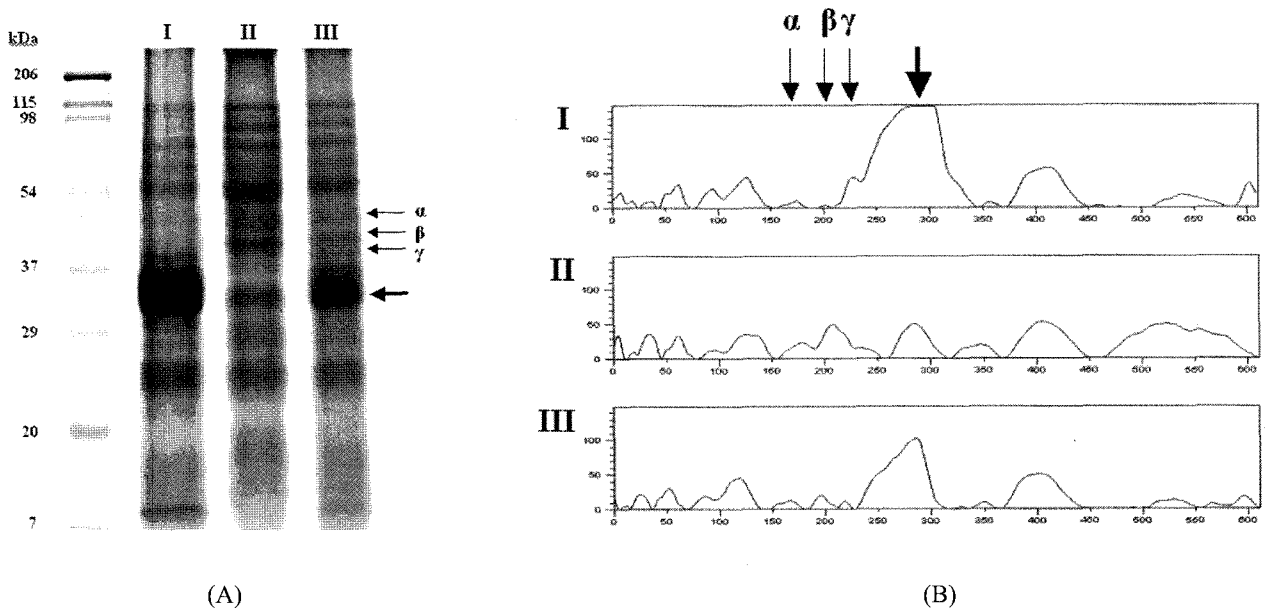


Fig 1. SDS-PAGE images of total rice seed proteins and its scanned densitograms. (A) SDS-PAGE images of total rice seed proteins extracted with TCA (lane I) and with water (lane II). One half of the protein samples shown in lane I and II were mixed together, and are shown in lane III. Molecular weight standards are shown on the left. (B) Scanned densitograms of SDS-PAGE images obtained with the image analysis software, ImageQuant TL (Amersham Pharmacia Biotech, Uppsala, Sweden). Roman numbers I, II, and III represent the densitograms of each lane of the 1-D gel shown in panel (A). The thick arrow represents the glutelin family proteins. Thin arrows with α , β , and γ indicate the protein bands, showing the most prominent differences between lanes I and II. (Novel Mark 8 standards are used to show approximate molecular weights. Lane 1, extracted glutelin storage proteins preparation by TCA method, lane 2, extracted glutelin storage proteins by water method, and lane 3, extracted glutelin storage proteins preparation by both TCA and water methods. Proteins are separated using 13% SDS-PAGE and visualized with silver staining. Among the lane 3, were cut into 3 pieces (5, 6 and 7) of equal size per lane and subjected to in-gel digestion with trypsin prior to ESI-Q-TOF/MS/MS analysis.

가는 점선으로 표시된 부위 1과 3의 단백질 스팟은 물추출법에 의한 시료에서 더 강도가 높아지거나, 물추출법에 의한 시료에서만 관찰되었다(Fig. 1A 및 1B). 글루테린류 단백질들은 부위 2에 위치하였다. 화살표 a와 b는 이들 두 단백질들이 물추출법에 의한 경우에만 관찰되며, ESI Q-TOF

MS에 의해 확인됨을 나타낸다. Fig. 2B에서 볼 수 있듯이, 두 추출법 간에 명백한 차이점이 나타났다. 종래의 TCA법으로 준비된 단백질 시료에서는 937개의 단백질 스팟이 나타났다지만, 본 시험에 의한 물추출법에 의한 단백질 시료에서는 1,240개의 단백질 스팟만이 관찰되었다. 물추출법에 의해 준

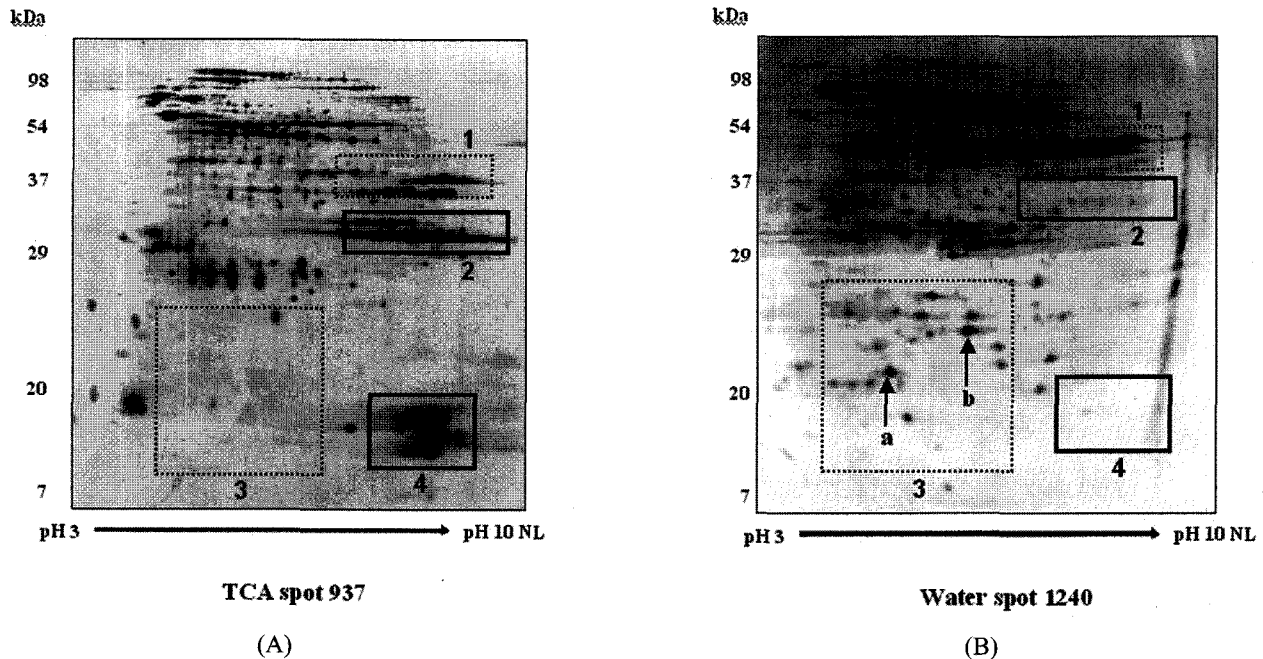


Fig 2. 2-D gel pictures of the total rice seed proteins extracted via the TCA method (A) and via water extraction (B). The proteins were resolved via 2-DE with isoelectro focusing separation in pH 3-10 NL, 18-cm IPG strips in the first dimension, and 12% SDS-PAGE in the second dimension. Proteins were stained with silver-diamine. Glutelin family proteins are located in zone 2. Arrows a and b indicate that these two proteins were observed exclusively in the water extraction samples.

비된 벼종자 단백질 시료의 2-DE겔 두 지점(2, 4)에서 단백질 스팟이 현저하게 감소되었다.

이상 2D-PAGE 방법으로 검토한 결과, 본 시험에 의한 물추출법이 벼종자 단백질 시료 준비에 더 우수함이 확인되었다. 따라서 본 시험에 의한 추출방법은 프로테오믹스 연구에서 더 많은 단백질을 발견·분석할 수 있게 하고, 분석 속도 및 효율성을 증대시킬 수 있을 것으로 판단된다.

단백질 동정을 위하여, NCBI(Perkins *et al.*, 1999)에서 다운로드된 nr DB의 녹색 식물 단백질 DB를 대상으로 MASCOT software program(Matrix science, www.matrixscience.com, U.K)을 활용하여 MS/MS스펙트럼을 검색하였다. 관찰된 대응 펩타이드가 5% 이하 확률의 랜덤 서열일 수 있는 정도의 수준인 대응 펩타이드 목록을 DB로부터 선별하였다.

ESI Q-TOF 질량 스펙트로메터를 이용한 MS/MS 서열분석결과(결과 미제시), TCA 추출법에 의한 샘플에서, 지점 2에 위치한 대부분의 주요 단백질 스팟은 벼 글루테린 저장 단백질($pI=7.0-10.0$, M_r 은 약 40,000)인 반면, 지점 4에 위치한 스팟들은 종자 알레르기 단백질 또는 다른 저장 단백질의 일종인 프롤라민인 것으로 확인되었다. 그러나 물추출법으로 추출된 샘플에서, 글루테린, 종자 알레르기 단백질들 그

리고 프롤라민 스팟이 지점 2와 4에 나타나지 않았다. 나아가 물추출법을 적용한 경우, 단백질 중 글루테린이 대부분 제거되었기 때문에 종래 TCA 추출법 적용시 발견되지 않았던 많은 새로운 스팟이 관찰되었다. ESI Q-TOF 질량 스펙트로메터를 이용한 MS/MS 서열분석결과(결과 미제시), 증가되거나 2-DE겔 상의 다른 두 부위(1과 3)에서 새롭게 나타난 많은 스팟들은 미량 비글루테린 단백질로 확인되었다(Fig. 1).

물추출법에 의해 새롭게 확인된 단백질들의 예를 Fig. 2에 화살표로 나타내었다. 확인된 두 개의 스팟 중 하나는 *Oryza sativa*의 class I 저분자 heat shock 단백질이고(화살표 a), 다른 하나는 *Oryza sativa*의 superoxide dismutase[Cu-Zn]이다(화살표 b).

종래의 TCA 추출법에 의한 시료의 밴드 a는 주요하게 글루테린류 단백질들을 함유하고 있다. MASCOT 검색 프로토콜을 사용하였을 때, 글루테린은 높은 대응도에 의해 주로 동정되었으며, 많은 수의 펩타이드들이 대응되었다. 그러나 본 연구에 의한 물추출법에 의한 시료의 밴드 a는 미량 비글루테린 단백질들을 함유하고 있고, 글루테린 단백질은 단정한 종류 확인되었다. 동정된 많은 단백질들과 함께 동정된

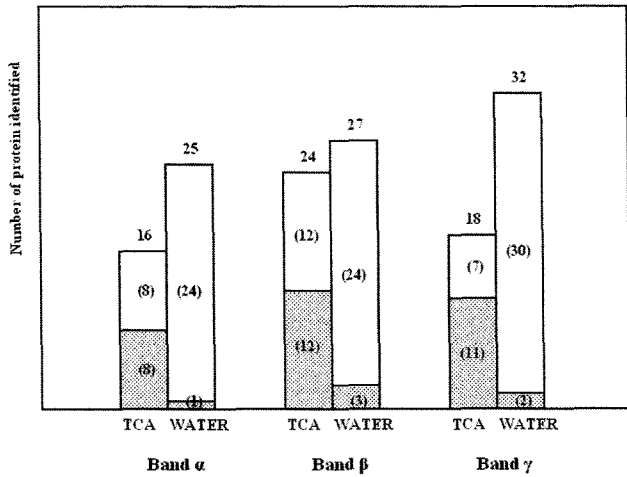


Fig 3. Comparison of the number of proteins identified by ESI Q-TOF MS in the peptide bands (thin arrows α, β, and γ in Figure 1, A), which evidenced the most prominent differences between the TCA precipitation and water extraction methods. The grey and white bars designate highly abundant glutelins and less abundant non-glutelins, respectively.

미량 비글루테린 단백질들도 다수 확인되었다. 이러한 결과는 물추출법이 글루테린이 많이 함유된 벼종자 시료에서 단백질들, 특히 미량 단백질들을 동정하기에 적합한 방법임을 나타낸다고 할 수 있다. 두 추출법 사이의 차이점은 밴드 β와 γ의 분석에서도 확인되었다.

Fig. 3은 각 밴드에서 확인된 단백질들의 숫자라는 측면에서 두 추출방법 사이의 차이점을 나타낸다. 회색봉은 고품량 글루테린류 단백질을, 백색봉은 미량 비글루테린류 단백질을 나타내며, TCA 추출법과 물추출법 사이의 차이점을 대부분 보여주고 있다. 즉, 물추출법에 의해 추출된 시료에서 선택·분석된 모든 밴드에서 대부분의 글루테린류 단백질들이 감소되었다. 한편 다수의 비글루테린 단백질들은 크게 증가하였다. 1차원 SDS겔 상에서 선택된 3개의 밴드들을 분석하여 얻은 결론으로부터, 동일한 겔에서 모든 밴드들이 비슷한 경향을 보인다고 추정할 수 있다. 따라서 물추출법이 글루테린이 많이 함유된 벼종자 시료의 프로테오믹스 분석을 위해 더 효과적인 방법임을 알 수 있다. 전통적인 TCA 추출법으로 벼종자로부터 단백질을 추출하는 경우, SDS겔 상에서 총 벼종자 단백질의 약 70%가 글루테린이다. 그러나 본 연구에 의한 물추출법을 적용하면, 글루테린은 같은 겔 상에서 총 단백질의 10%로 감소한다. 또한 전통적인 TCA 추출법으로 추출된 샘플을 2-DE(이차원 겔 전기영동) 분석하

면 937개의 단백질 spot이 나타난다. 반면 본 연구에 의한 물추출법에 의하면 1240개의 단백질이 분리되었다. ESI Q-TOF 질량 스펙트로메트리 분석에 의한 경우에도 역시 본 연구에 의한 물추출법에서 비글루테린 단백질에 대한 선택성과 글루테린의 감소가 확인되었다. 따라서 본 연구에 의한 물추출법은 신속하고 효율적인 벼종자 프로테오믹스 분석 수단이 될 것으로 판단되었다.

본 연구에 의하면, 단순하고 신속하며 저렴하고 효율적으로 미량 비글루테린 단백질들을 용이하게 동정할 수 있게 한다. 본 연구에 의한 물추출법은 물을 용매로 사용하기 때문에 종래의 TCA 추출법에 비해 보다 온화한 추출조건을 요구한다. 따라서 벼종자 단백질의 고품량 비 단백질들인 글루테린 저장 단백질을 제거하는 방법에 관한 본 연구에 의한 물추출 방법은 벼종자 시료들의 프로테오믹스 분석에 매우 유용할 것이다.

적 요

본 연구는 벼종자 미량 단백질의 프로테오믹스 연구를 위하여 벼종자에 고품량으로 존재하는 벼종자 글루테린 저장 단백질을 제거하는 방법에 관한 것이다. 따라서 본 연구는, (A) 벼종자에 액체 질소를 가하고 분쇄하여 벼종자 가루를 만드는 분쇄단계; (B) 상기 분쇄된 벼종자 가루를 물에 현탁하여 현탁액을 만드는 현탁단계; (C) 상기 현탁액 중 미용해 물질을 제거하는 분리단계를 포함하는, 벼종자 미량 단백질의 프로테오믹스 연구를 위한 벼종자 글루테린 저장 단백질의 제거방법에 관하여 검토하였다. 본 연구의 결과, 단순하고 신속하며 저렴하고 효율적인 방법으로 미량 비글루테린 단백질들을 용이하게 동정할 수 있을 것으로 판단되었다.

인용문헌

Anderson, N. L. and N. G. Anderson. 2002. The human plasma proteome : history, character, and diagnostic prospects. *Molecular & Cellular Proteomics*. 1 : 845-67.

Sato, A. K., D. J. Sexton, L. A. Morganelli, E. H. Cohen, Q. L. Wu, G. P. Conley, Z. Streltsova, S. W. Lee, M. Devlin, D. B. DeOliveira, J. Enright, R. B. Kent, C. R. Wescott, T. C. Ransohoff, A. C. Ley, and R. C. Ladner. 2002. Development of mammalian serum albumin affinity purification media by peptide phage display. *Biotechnol. Prog.* 1 : 947-955.

Adkins, J. N., S. M. Varnum, K. J. Auberry, R. J. Moore, N. H. Angell, R. D. Smith, D. L. Springer, and J. G. Pounds.

2002. Toward a human blood serum proteome: Analysis by multidimensional separation coupled with mass spectrometry. *Molecular & Cellular Proteomics*. 1 : 947-955.
- Steel, L. F., M. G. Trotter, P. B. Nakajima, T. S. Mattu, G. Gonye, and T. Block. 2003. Efficient and specific removal of albumin from human serum samples. *Molecular & Cellular Proteomics*. 2 : 262-270.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of Protein utilizing the principle of Protein-Dye Binding. *Anal. Biochem.* 72 : 248-254.
- Finnie, C., S. Melchior, P. Roepstorff, and B. Svensson. 2002. Proteome analysis of grain filling and seed maturation in barley. *Plant Physiology*, 129, 1308-1319.
- Hirano, H., S. Komatsu, A. Nakamura, F. Kikuchi, H. Kajiwara, S. Tsunasawa, and F. Sakiyama. 1991. Structural homology between semi-dwarfism-related proteins and glutelin seed protein in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 83 : 153-158.
- Heukeshoven, J. and R. Dernick. 1985. Simplified method for silver staining of proteins in polyacrylamide gels and the mechanism of silver staining *Electrophoresis*. 6 : 103-112.
- Koller, A., M. P. Washburn, B. M. Lange, N. L. Andon, C. Deciu, P. A. Haynes, L. Hays, D. Schieltz, R. Ulaszek, J. Wei, D. Wolters, and J. R. Yates. 2002. Proteomic survey of metabolic pathways in rice. *Proc. Natl. Acad. Sci. USA*. 99 : 11969-11974.
- O'Farrell, P. H. 1975. High Resolution Two-Dimensional Electrophoresis of Proteins. *J. Biol. Chem.* 250 : 4007-4021.
- Perkins, D. N., D. J. Pappin, D. M. Creasy, and J. S. Cottrell. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 20 : 3551-3567.
- Pieper, R., C. L. Gatlin, A. J. Makusky, P. S. Russo, C. R. Schatz, S. S. Miller, Q. Su, A. M. McGrath, M. A. Estock, P. P. Parmar, M. Zhao, S. T. Huang, J. Zhou, F. Wang, R. Esquer-Blasco, N. L. Anderson, J. Taylor, and S. Steiner. 2003. The human serum proteome : Display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. *Proteomics*. 3 : 1345-1364.
- Ramani, S., and S. K. Apte. 1997. Transient expression of multiple genes in salinity-stressed young seedling of rice (*Oryza sativa* L.) CV. Bura. *Biochem. Biophys. Res. Commun.* 233 : 663-667.
- Takaiwa, F., K. Oono, D. Wing, and A. Kato. 1991. Sequences of three members and expression of a new major subfamily of glutelin gene from rice. *Plant Mol. Biol.* 17 : 875-885.
- Woo, S. H., M. Fukuda, N. Islam, M. Takaoka, H. Kawasaki, and H. Hirano. 2002. Efficient peptide mapping and its application to identify embryo proteins in rice proteome analysis. *Electrophoresis*. 23 : 647-654.
- Woo, S. H., A. Higa, M. Kimura, S. K. Jong, and I. Yamaguchi. 2003. Proteome Analysis of wheat Lemma. *Biosci. Biotechnol. Biochem.* 67 : 2486-2491.