

## 프로테오믹 해석에 의한 벼 게놈 기능해석과 응용

우선희\*, 김홍식<sup>1</sup>, 송범현<sup>1</sup>, 이철원<sup>1</sup>, 박영목, 정승근<sup>1</sup>, 조용구<sup>1</sup>  
한국기초과학지원연구원, 생명화학연구부, <sup>1</sup>충북대학교 농과대학 식물자원학과

## Rice Proteomics: A Functional Analysis of the Rice Genome and Applications

Sun Hee Woo\*, Hong Sig Kim<sup>1</sup>, Berm Heun Song<sup>1</sup>, Chul Won Lee<sup>1</sup>, Young Mok Park,  
Seung Keun Jong<sup>1</sup>, Yong Gu Cho<sup>1</sup>

*Division of Biological and Chemical Science, Korea Basic Science Institute, Oun-dong 52, Yuseong-gu,  
Daejeon 305-333, Korea*

<sup>1</sup>*Department of Plant Resources, College of Agriculture, Chungbuk National University, Gaesung-dong, San 48,  
Heungduck-gu, Cheongju 361-763, Korea*

**ABSTRACT** In this review, we described the catalogues of the rice proteome which were constructed in our program, and functional characterization of some of these proteins was discussed. Mass-spectrometry is the most prevalent technique to rapidly identify a large number of proteins in proteome analysis. However, the conventional Western blotting/sequencing technique has been used in many laboratories. As a first step to efficiently construct protein data-file in proteome analysis of major cereals, we have analyzed the *N*-terminal sequences of 100 rice embryo proteins and 70 wheat spike proteins separated by two-dimensional electrophoresis. Edman degradation revealed the *N*-terminal peptide sequences of only 31 rice proteins and 47 wheat proteins, suggesting that the rest of separated protein spots are *N*-terminally blocked. To efficiently determine the internal sequence of blocked proteins, we have developed a modified Cleveland peptide mapping method. Using this above method, the internal sequences of all blocked rice proteins (*i.e.*, 69 proteins) were determined. Among these 100 rice proteins, thirty were proteins for which homologous sequence in the rice genome database could be identified. However, the rest of the proteins lacked homologous proteins. This appears to be consistent with the fact that about 45% of total rice cDNA have been deposited in the EMBL database. Also, the major proteins involved in the growth and development of rice can be identified using the proteome approach. Some of these proteins, including a calcium-binding protein that turned out to be calreticulin, gibberellin-binding protein, which is ribulose-1,5-bisphosphate carboxylase/oxygenase activase in rice, and leginsulin-binding protein in soybean have functions in the signal transduction pathway. Proteomics is well suited not only to determine interaction between pairs of proteins, but also to identify multisubunit complexes. Currently, a protein-protein interaction database for plant proteins (<http://genome.c.kanazawa-u.ac.jp/Y2H>) could be a very useful tool for the plant research community. Also, the information thus obtained from the plant proteome would be helpful in predicting the function of the unknown proteins and would be useful in the plant molecular breeding.

**Key words:** gas-phase sequencing, peptide mass fingerprinting, proteomics, rice (*Oryza sativa* L.), two-dimensional electrophoresis

## 서 론

프로테오믹스 (proteomics)는 단백질 (protein)과 유전체 (게놈, genome)의 합성 용어이다. 기능성 단백질을 전체적으로 해석하고, 게놈의 기능, 더 나아가서는 생물기능을 해명하려고 하는 연구 분야이다. 프로테오믹스 연구에는 여러 생물종들에 있어서 조직 및 시기 특이적으로 발현하고 있는 단백질의 database을 만들어 생명현상을 이해하고 나아가서 병을 치료하거나 의약품의 개발, 환경문제와 식량문제를 극복하기 위한 유용한 식물을 만드는 데 목적이 있다. 미국이나 유럽에서는 생물의 유전자 정보를 해독하는 게놈연구가 활발하게 진행되어 왔으며, 1995년 이후 많은 생물에서 게놈의 전체 염기서열이 결정되어 유전자 구조와 유전자 발현에 대하여 더욱더 많은 이해가 가능하게 되었다. 그러나 게놈 해석에서 결정된 염기서열로부터 database를 이용한 상동단백질의 검색에 의하여 기능이 추정되는 단백질의 비율은 전 번역산물의 30%에서 50% 정도에 지나지 않는 것을 알 수 있었다. 그 기능이 알려지지 않은 유전자들을 동정하기 위하여 최근에는 포스트 게놈 연구로서 단백질의 생화학적 및 물리화학적 특성을 분석하여, 단백질과 그것들을 암호하는 게놈 DNA와의 관계를 밝히는 연구, 게놈 정보를 이용하면서 모든 유전자의 번역산물에 대하여 기능을 해명하는 연구, 즉 프로테오믹스 연구가 추진되어야 할 필요성이 대두되고 있다. 대다수의 단백질은 번역 중 또는 번역 후에 여러 가지 변형과정을 거쳐 특유의 입체구조를 갖고 성숙 단백질이 된다. 그리고, 상호적으로 작용하며 그 기능을 발휘한다. 따라서, 단백질의 기능을 밝히는 경우에는 반드시 기능을 하고 있는 단백질 그 자체를 분석하고, 번역 중 및 번역 후의 변형과정과 입체구조, 상호작용 등에 관한 정보를 수집할 필요가 있다. 따라서 게놈 해석에 대응하기 위하여 대규모 고효율적 (high-throughput) 방법인 단백질 해석, 프로테오믹스 해석이 필요할 것이며, 종들의 다양한 환경 하에서 또는 생육단계에서 또는 기관에서 발현하는 모든 단백질이 연구의 대상이 된다. 여기에서는 벼를 중심으로 프로테오믹스 연구의 현황과 그 응용에 대하여 소개하고자 한다.

## 식물의 프로테오믹스 해석의 역할, 방법 및 현황

게놈해석은 식물에 있어서 1990년대 후반부터 본격적으로 진행되었다. 그 결과 2000년에는 게놈 크기가 작은 애기장대 (약 150 Mbp)에서 전 염기배열이 결정되었다. 또한 약 430 Mbp의 게놈 크기를 갖는 벼는 일본을 중심으로 하는 국제 벼 게놈 프로젝트에 의하여 2002년 12월 벼 1번 염색체의 완전 염기서열이 밝혀졌다 (Sasaki et al. 2002). 아직까지 밝혀지지 않은 나머지 염색체의 완성을 위하여 연구가 수행 중에 있으며, 그 이외에도 콩, 담배, 밀 등 많은 식물에서 게놈해석이 현재 진행되고 있다. 주요 DNA database에 기능이 알려진 유전

자의 등록수는 여러 생물체에서 유전자체 연구가 급속히 진행됨에 따라 비약적으로 증대하고 있다. 그러나, 동물의 경우와 마찬가지로 식물에 있어서도 번역산물 중에서 기능이 알려져 있는 단백질의 비율은 전체 추정 유전자 수에 비하여 극히 낮다. 따라서 앞으로 식물에서는 게놈 염기서열 분석으로 얻어진 모든 염기서열 정보를 이용한 프로테오믹스 해석을 진행하여 아직 밝혀지지 않은 수많은 단백질의 기능을 해명할 필요가 있다.

## 전사 (transcription)와 프로테오믹스 (proteomics)

프로테오믹스 (proteomics)에 있어서는, 전 전사산물, 즉 mRNA의 총체를 트랜스크립토믹스 (transcriptomics)이라고 부르고 있다. 트랜스크립토믹스를 해석하면, 각각의 전사산물의 발현량을 고효율적으로 알아낼 수가 있다 (Veculescu et al. 1997; Zhang et al. 1997). 그러나, 이 결과에 의해서는 실질적으로 가능한 단백질의 실체를 직접 얻는 데에는 어려움이 있다. 단백질은 종류에 따라서, 또 존재하는 조건에 의해서 수명이 다르다. 수명이 긴 단백질은 다량으로 존재할 필요가 있어도 다량으로 번역될 필요가 없을 수 있다. 그 때문에 전사산물량과 번역산물의 존재량과의 사이에 상관관계는 매우 낮은 것으로 알려져 있다 (Anderson and Seilhamer 1997). 따라서, 기능을 갖고 있는 단백질의 존재량을 전사 산물로부터 정확히 추정할 수 없다. 한편, 전사물의 분석에 의하여 단백질의 번역 중 및 번역 후의 변형과정과 입체구조, 단백질 상호작용 등에 관한 정보를 얻는 데에는 어려움이 있다. 모든 단백질은 번역 중 또는 번역 후에 여러 가지 변형을 통하여 특유의 입체구조를 형성하여 성숙된 단백질이 된다. 그리고, 상호적으로 작용하여 그 기능을 발휘하게 된다. 따라서, 단백질의 기능을 알고자 하는 경우에는, 반드시 기능을 지니고 있는 단백질 그 자체를 분석하고, 번역 중 및 번역 후의 변형과 입체구조, 상호작용 등에 관한 정보를 수집할 필요가 있다. 따라서 미지의 단백질에 대한 구체적인 정보를 얻기 위하여 게놈 해석에서 얻어지는 정보를 활용하여 대규모적인 고효율적 단백질 해석이 필요하다.

## 식물육종에 있어서 프로테오믹스 해석의 역할

프로테오믹스 해석에 의하여 게놈의 염기서열과 그것들이 암호하고 있는 많은 단백질의 관계가 알려지고 있다. 이 해석에 의해서 식물의 기관분화, 성장, 노화 등에 관련된 새로운 단백질과 그와 관련된 유전자를 밝힐 수 있는 가능성은 매우 크다. 이미 앞에서 서술한 것처럼 지금까지 프로테오믹스 연구과정에서 분석된 전체 단백질 중에서 기능이 밝혀진 단백질은 30~50%로 추정되고 있지만, 대다수의 단백질은 그 기능이 알려져 있지 않은 상태이다. 프로테오믹스 해석에 의하여 이러한 미지의 단백질의 기능 및 발현양식에 대하여 보다 본질적인 정보를 얻게 된다면 이를 통하여 밝혀진 유전자 또는 단백질의

기능발현을 유전적으로 개량한 우량한 품종을 육성할 수 있을 것이다. 즉, 병충해 및 불량 환경 하에서 발현되는 단백질을 특이적으로 분석하거나 품질 또는 생산성과 관련된 단백질을 집중적으로 분석함으로써 얻어지게 되었고, 이렇게 기능이 밝혀진 유전자들을 이용하여 새로운 품종을 육성할 수 있는 길이 열릴 가능성이 있다.

### 프로테옴 해석의 방법

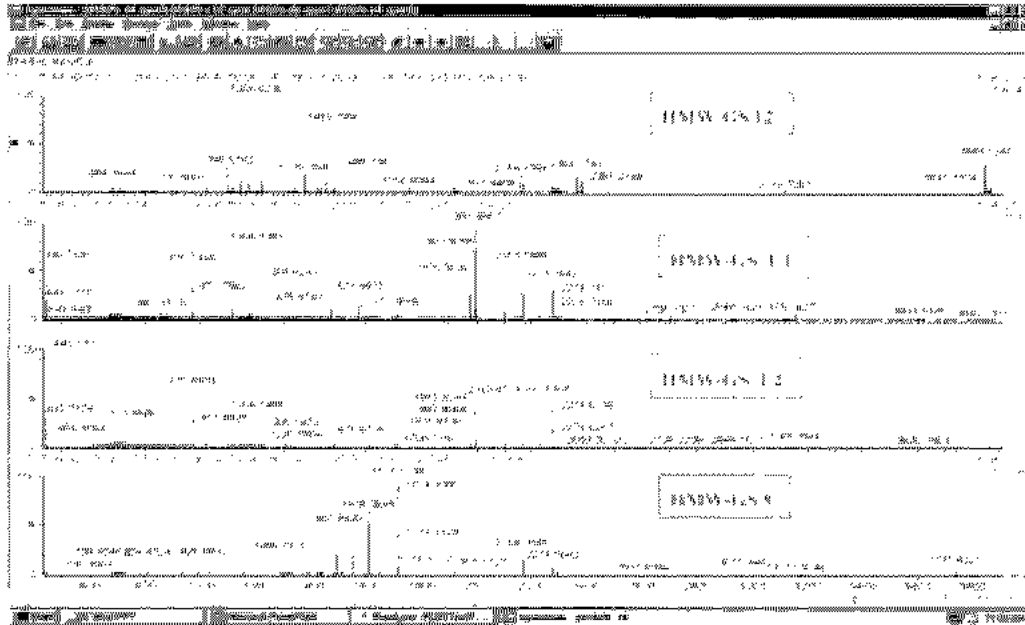
프로테옴 해석은 특정한 조건에서 발현된 전 단백질의 분리정제로부터 시작된다. 단백질의 분리정제에는 간단하고 신속하게 분리능력이 높은 이차원전기영동 (two-dimensional electrophoresis, 2-DE)이 일반적으로 이용되고 있다. 2-DE에서 분리된 단백질 각각에 대하여 질량분석을 이용하여 펩타이드를 만들기도 하고 (peptide mass finger print, PMF), 질량분석기 (mass spectrometry, MS)와 기상 아미노산분석기 (gas-phase amino acid sequencer)에 의하여 부분 아미노산 서열을 결정한다. 그리고, 이것들의 data를 기초로 하여 게놈 해석에서 만들어진 database를 검색하고 단백질과 유전자의 대응관계를 알아낸다. 유전자의 염기서열로부터 단백질의 완전한 아미노산 서열을 추정할 수 있기 때문에 이것을 이용하여 단백질 아미노산서열 database를 검색하면, 이미 기능이 알려져 있는 단백질이 있는지 없는지를 알 수 있다. 또 기능을 갖는 단백질과 아미노산서열에 유사성이 있는 경우에는 단백질의 기능을 추정할 수가 있다. 상동성 검색에 의하여 단백질 기능을 알 수 없으면 단백질의 각종 특성을 분석하여 단백질의 기능 해명을 밝힐 수 있다. 알려지지 않은 단백질의 기능을 알기 위해서는 단백질 번역 후의 변형과정, 입체구조, 발현량, 발현시기, 발현부위, 단백질 상호작용, 효소활성, 생리활성 등을 분석할 필요가 있다. 프로테옴 해석에서는 앞에서 밝혀진 분석결과와 게놈 염기서열을 이용하여 유전자의 실체를 알아낼 수 있다. 그것도 한, 두 종류의 단백질의 기능해명을 목표로 하는 것이 아니고 가능하면 모든 단백질의 기능을 해명하여 미지의 유전자를 밝혀낼 수 있다. 프로테옴 해석에 의해서 얻어진 2-DE 양상과 각각의 단백질의 서열과 구조에 관한 정보, 2-DE에서 분리된 단백질을 암호하는 유전자의 정보, 단백질의 기능에 관한 정보에 대하여는 database로 작성하여 많은 연구자들이 이용할 수 있게 하는 것이 중요하다.

### 프로테옴의 비교 분석

같은 식물 종의 동일 개체라고 하더라도 생육단계와 조직기관에 따라서 발현하는 단백질의 종류는 다르다. 또 환경에 의해서도 발현은 다르다. 프로테옴은 극히 다양한 변이를 가지고 있다. 이것은 게놈 DNA의 염기서열이 동일종 내에서는 같지만 생육조건에 따라서 단백질의 발현은 다양한 변이를 나타내기 때문이다. 2-DE를 이용하여 단백질을 비교 해석하

는 연구는 게놈 해석이 시작하기 전부터 연구가 이루어져 왔다 (Thiellement et al. 1999). 식물이 특정환경 하에서 생육할 경우 생육단계와 조직기관에 특이적으로 출현하거나 소실할 수 있으며 (Masson and Rossignol 1995), 호르몬 처리에 의해서 변동하는 단백질도 2-DE에서 분석되었다 (Santoni et al. 1997). 한편 벼의 반왜성 (semi-dwarf)과 같은 표현형을 갖는 준동질 유전자 계통 (Hirano et al. 1991)과 들연변이체 (Damerval and Le Guilloux 1998)에 출현 또는 소실하는 단백질이 검출되었다 (Santoni et al. 1997). Islam 등 (2002)은 밀의 프로테옴 연구에 있어서 컴퓨터 software 화상장치를 이용하여 39개의 ditelo centric (DT) lines과 정상적인 euploid의 계통 사이에 양적 변화를 탐색하였으며, 배우자 치사유전자를 이용하여 얻은 염색체 결실계통을 이용하여 염색체 결실부위에 위치하는 유전자가 암호하는 단백질을 동정하고 단백질과 유전자의 관계와 단백질 기능해명에 역점을 두었다 (Figure 1). 이와 같은 단백질의 해석은 식물의 분화, 성장 및 생식을 제어하는 단백질을 동정하는 데에 중요한 정보를 제공하고 있다.

식물에 있어서 병충해와 환경 스트레스에서 출현 또는 소실하는 단백질을 2-DE에 의하여 해석한 보고는 많이 있다 (Ramani and Apte 1997; Lund et al. 1998; Rey et al. 1998; Riccardi et al. 1998). 이러한 단백질 해석은 내병충성과 내염성 등의 스트레스 내성기구에 있어서 어떠한 단백질이 관련되어 있는가를 해명하는 데 중요한 정보를 제공한다. 또 식물이 갖는 알레르기 단백질 (Weiss et al. 1997), 맥주의 보리 발아 품질 관련 단백질 (Görg et al. 1992), 제빵성 관련 단백질 (Lafiandra and Kasarda 1985) 등이 이차원전기영동에서 동정이 되었는데, 이것들은 식물 단백질의 식품으로서의 기능을 밝히는데 중요성이 높다. 앞에서 열거한 단백질의 비교에 관한 연구는 프로테옴 연구로서 진행되어진 결과는 아니지만, 단백질의 기능해명에 관계되는 연구로서 프로테옴을 통한 연구보다 구체적인 정보를 제공할 수 있을 것이다. 최근 Gygi et al. (1999)와 Oda 등 (1999)은 처리를 달리하여 생육된 세포를 가지고 MS를 이용하여 단백질량의 차이를 정량적으로 측정하는 방법을 보고한 바 있다. 즉, 처리를 달리하여 생육된 세포의 단백질을 방사성 동위체 표식법 (isotope-coded affinity tag, ICAT) MS에 의한 분석에서 얻어지는 PMF를 비교하여 단백질 함량의 차이를 해명하였다. 이 방법은 단백질의 절대량을 정확히 파악하는 것은 어렵지만, 처리를 달리하여 생육된 세포조직의 단백질함량의 상대적인 차이가 근소한 경우에도 검출할 수 있다. 어떤 방법에서도 같은 단백질이 있으면 방사성 동위체 표식법에 의하여 물리화학적 성질에는 거의 차이가 없으므로 정확히 검출할 수 있다. 따라서, MS에서 미세분석이 필요한 경우에도 높은 정확도로 정량 분석을 수행할 수 있다. 이와 같은 기법을 이용해서 식물에서도 단백질을 비교 해석하는 연구가 수행되고 있다.



**Figure 1.** Peptide spots from 2-DE gel were digested according to the method described by Islam et al. (2002). The digests were desalted with ZipTip (Millipore, Boston) and subjected to analysis by MALDI-TOF/MS.

**단백질, 유전자의 동정**

1987년에 전기영동에서 분리된 단백질을 polyvinylidene difluoride membrane (PVDF)과 같은 막 필터에 전사하여 기상 아미노산분석기에서 부분 아미노산배열을 분석하는 방법이 개발되었지만 (Matsudaira 1987), 이 방법을 응용해서 1980년 후반부터 식물에 관해서도 2-DE에서 분리된 많은 단백질에 대한 부분 아미노산 배열을 분석하여 단백질을 동정하는 연구가 수행되었다 (Hirano 1989). 그 후 벼 (Hirano 1997; Komatsu et al. 1994, 1999; Tsugita et al. 1994; Woo et al. 2003), 옥수수 (Riccardi et al. 1998), 밀 (Gomez et al. 1991; Woo et al. 2003), 보리 (Flengsrud 1993), 애기장대 (Kamo et al. 1995; Santoni et al. 1998), 메밀 (Woo et al. 2001), 담배 (Rouquie et al. 1997) 등 많은 식물에서 유사한 연구가 이루어져 왔다.

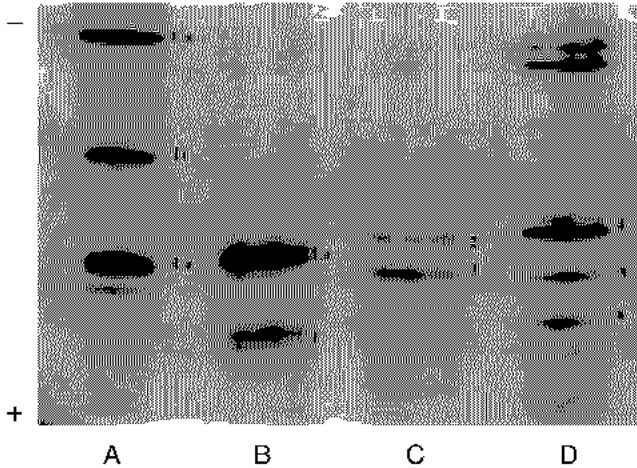
지금까지 벼의 배 단백질을 이차원전기영동에서 분리한 후, 겔내에서 trypsin 효소로 digestion을 하여 얻은 peptide를 matrix assistant lazer desorption ionization-time of flight mass spectrometry (MALDI-TOF/MS)를 이용하여 분석하고, PMF를 만듦에 따라서 단백질의 동정을 수행하여 왔다. 그러나, 게놈 DNA의 서열이 완전히 밝혀지지 않은 벼에서는 PMF만으로 동정할 수 있는 단백질과 유전자의 수는 제한되어 있다. 그 때문에 자주 다른 생물의 염기서열 database를 이용하여 단백질과 유전자를 동정하지 않으면 안되었다. 그러나, 다른 생물의 상동 또는 유사 단백질을 검색하기 위해서 PMF에 적어도 몇 잔기의 아미노산 배열 (아미노산 tag)을 밝힐 필요가 있다. 함유량이 비교적 많은 단백질에서는 보통 겔 전기영동에서 분리된 단백질을 PVDF막에 전사해서 기상아미노산 분석기를 이용하여 N말단 아미노산서열을 결정할 수가 있다. 그러나 지

금까지의 연구에서 벼 단백질의 다수는 N말단 아미노기가 block된 것을 증명해 왔다. block된 단백질은 양이 많다 하더라도 아미노산 분석기에서 N말단서열을 결정할 수 없다. 따라서 이러한 단백질의 서열을 기상 아미노산 분석기를 이용하여 결정하는 경우에는 Cleveland법을 이용하여 단백질 내부의 peptide를 단리해서 분석할 필요성이 있다. 실제 지금까지 이차원전기영동에서 분리한 단백질의 내부배열을 Cleveland법에서 결정하는 것은 시간이 많이 소요되었다. 따라서, 내부배열 결정에 아미노산서열 결정의 고효율 분석을 위하여 이차원전기영동에서 분리한 단백질의 peptide mapping을 작성하고 V8 protease를 이용하여 block된 부위의 아미노산 서열을 손쉽게 결정할 수 있는 방법을 개발하였다 (Woo et al. 2002). 우선, 단백질을 이차원전기영동에서 분리한 후, CBB로 염색한 후 겔을 건조시켰다. 건조겔의 스폿을 자르고 세분하게 절단했다. 이것을 SDS-PAGE겔의 시료구에 넣고, 100 μL의 10% 글리세롤을 포함한 전극액 (팽창액)을 시료구에 주입하여 약 1시간정도 방치했다. 건조겔은 전극액을 흡수하여 팽창했다. 겔을 팽창시킨 후 팽창액 위에 *Staphylococcus aureus* V8 protease (2 μg)를 포함하는 SDS-PAGE 시료완충액 20 μL을 증충했다. 그 위에 전극액을 주입하여 SDS-PAGE를 수행하였다. 단백질과 protease가 농축겔의 중앙부분까지 이동시켰을 때 전원을 끄고 이차원전기영동에서 분리된 단백질에 관하여 1일 정도에 20 시료의 peptide map를 작성할 수 있었다. 그 결과로서 N말단이 block되고 있다고 생각되는 단백질을 포함하여 100개의 단백질에 대하여 위의 방법을 이용하여 peptide mapping에서 얻어진 peptide의 서열을 기상 아미노산 분석기에 의하여 결정했다 (Figure 2, 3). 아미노산서열의 상동성 검색한 결과 100개의 단백질 중 30%은 이미 벼에서 해석된 상

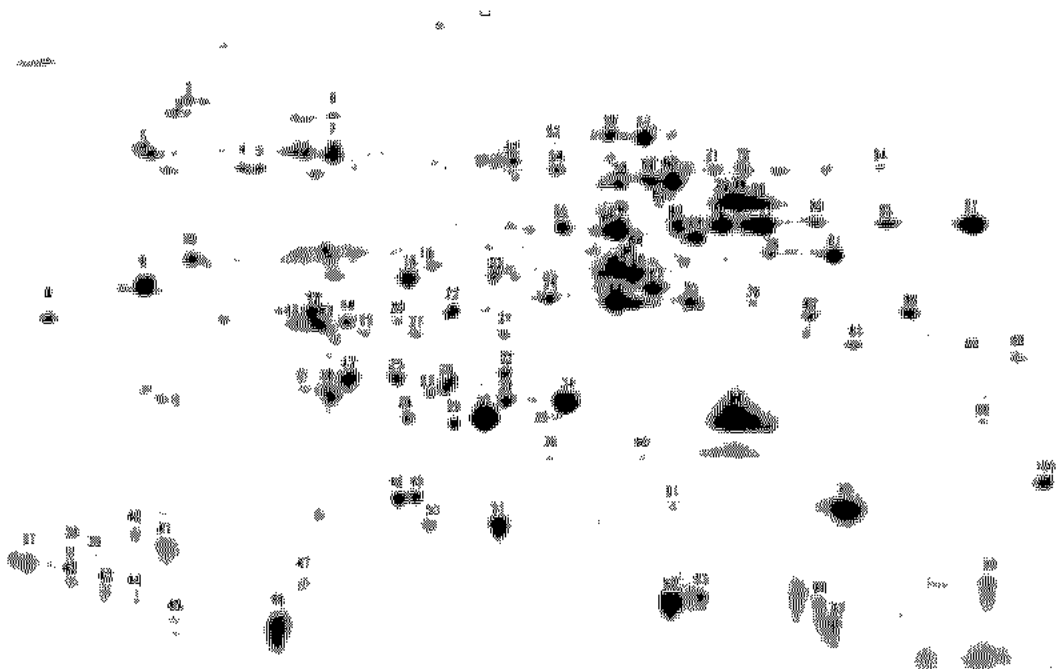
동성이 있는 단백질이며, 그리고 남은 70%은 지금까지 분석되고 있지 않았던 단백질이라는 것을 알아냈다. EMBL의 data base을 이용하여 분석한 단백질을 암호하는 EST를 검색한 결과, 45%의 단백질을 암호하는 EST를 검색할 수 있었다. 기능을 알 수 없는 55%의 단백질에 대하여는 앞으로 번역 후

변형, 단백질간의 상호작용, 입체구조, 비교 등의 해석에 의하여 기능을 밝히는 것이 중요하다고 생각된다. 또 단백질의 아미노산 서열을 electro spray ionization화 사중극 비행시간형 질량분석기와 같은 MS/MS에서 결정할 수 있다. 그러나 MS/MS에는 고효율 분석을 수행할 수 없는 단점이 있다. 그래서 본 연구개발에서 아미노산분석기에 의한 서열분석의 고효율화를 목표로 한 이유의 하나이다. 또한, 밀의 붉은 곰팡이 병의 감염부위인 이삭에서 발현하는 전 단백질을 전체적으로 해석하고 그것을 암호하는 유전자와 대응하여 단백질의 catalog화를 목표로 연구를 진행했다. 우선 함유량이 많은 70개의 단백질을 임의적으로 선발하여 아미노산 분석기를 이용하여 분석한 결과 47개는 N말단 서열을 결정할 수 있었지만, 나머지 23개는 결정할 수가 없었다. 즉 전체의 30% 정도가 N말단이 block라고 생각되어 Cleveland peptide mapping 방법과 역상 크로마토그래피 (RP-HPLC)의 방법을 이용하여 내부 아미노산 서열을 결정하여 상동성 단백질을 검색할 수 있었다. 또한 47개의 단백질에 대응하는 밀의 EST를 검색한 결과 100%의 단백질에 대응하는 EST가 검색되었다 (Woo et al. 2003).

최근, MALDI-TOF 질량분석기와 같은 MS가 발달해서 2-DE에서 분리된 단백질을 겔 속에서 트립신과 같은 protease에 의하여 절단하여 생기는 peptide를 MS에 의하여 고감도(高感度), 고정도(高精度) 및 고효율(高效率)로서 분석할 수 있게 되었다. 이때 얻어진 PMF를 database 내에서 단백질의 이론적인 PMF와 비교함에 의하여 식물에서도 단백질의 동정이 이루어지고 있다. 한편, MALDI-TOF/MS와 post-source-decay



**Figure 2.** Peptide map of the *S. aureus* V8 protease digests of three proteins. Low-molecular-weight standard marker proteins (100 pmol) were separated by SDS-PAGE and stained with Coomassie blue. After destaining, the gel was completely dried and then the gel pieces containing proteins were chopped into smaller pieces and inserted into the sample well for Cleveland peptide mapping. The resultant peptides on the SDS-PAGE gel were electroblotted on the PVDF membrane and subjected to gas-phase sequencing. La,  $\alpha$ -lactalbumin; Ti, trypsin inhibitor; Ca, carbonic anhydrase; A, intact proteins; B, digests of La; C, digests of Ti; D, digests of Ca.



**Figure 3.** Two-dimensional electrophoresis pattern of the rice embryo proteins. Proteins were separated by isoelectric focusing in the first dimension (right to left) and SDS-PAGE in the second dimension (top to bottom) and detected by Coomassie blue staining.

(PSD)을 조합한 MS와 electrospray ionization화한 사중극 TOF/MS와 같은 MS/MS를 이용함에 의하여 극미량의 peptide의 아미노산 서열을 결정할 수 있게 되었다. 그러나, 식물에 관해서는 MS를 이용하여 2-DE에서 분리한 단백질의 부분배열을 고효율로서 결정하고 단백질을 동정한 보고는 적다. Fukuda 등 (2003)은 전기영동에서 분리한 단백질의 동정법 개량에 초점을 맞추어 MALDI-TOF/MS에서 PSD에 의하여 생긴 단편을 분석에 의하여 MS/MS해석이 가능하지만, PSD spectro의 해석은 간단하지 않고 감도도 좋지 않다. peptide를 화학변형에 의하여 해석이 쉬운 PSD spectro를 얻은 보고가 있다. 몇 개 정도의 화학변형 중에는 N-Tris (2,4,6-trimethoxyphenyl) phosphine acetic acid N-hydroxysuccinimide ester bromide에 의한 화학변형에서는 peptide의 N말단만이 변형되어 정전하에 영향을 주어 단순한 PSD spectro를 얻을 수 있다는 보고가 있다. 이것을 겔의 내 trypsin소화물에 반응시켜 미량으로서 PSD spectro를 측정할 수 있는 방법을 확립하고 보다 정도가 높은 단백질의 database 검색을 가능하게 할 수 있도록 목표로 했다. N-Tris (2,4,6-trimethoxyphenyl) phosphine와 N-hydroxysuccinimide ester로부터 N-Tris (2,4,6-trimethoxyphenyl) phosphine acetic acid N-hydroxysuccinimide ester bromide를 합성했다. 수량은 얼마 없었지만, 질량분석의 결과로부터 N-Tris (2,4,6-trimethoxyphenyl)phosphine acetic acid N-hydroxysuccinimide ester bromide를 합성할 수 있었다. 몇 개 정도의 peptide에서 화학변형을 시험하였지만, 반응 수율이 낮기 때문에 peptide의 변형조건을 검토하고 있다.

프로테옴 해석은 2-DE에서 분리된 단백질과 게놈 염기서열분석에서 얻어진 유전자와의 대응관계를 밝혀낼 필요가 있다. 유전자의 동정에는 일반적으로 MS에 의하여 얻어진 PMF의 MS와 기상 아미노산 분석기에 의하여 결정된 부분 아미노산 배열을 이용하고 있다.

벼에서는 1996년 무렵에 게놈 프로젝트에 의하여 약 2만 개의 expressed sequence tag (EST)의 염기서열이 결정되었다. 이것들의 EST의 중복도는 약 50% 정도라고 추정되었다. 벼에는 약 4만 개의 유전자가 존재하리라 가정하면 cross-link된 EST는 전 유전자의 25~50%라고 추정된다. 2-DE에서 분리된 1,800개 이상의 배, 배유, 잎 및 뿌리의 단백질 중 126개의 아미노산서열이 기상 아미노산분석기에 의하여 결정되었는데, 아미노산서열을 기초로 하여 분석한 결과 약 30%의 단백질을 암호하는 EST가 검색되었다 (Hirano 1997). EST와 단백질간에 대응하는 단백질의 비율은 약 30%로서 상기의 클로닝된 EST가 2만 개 이상이 되지만, 아미노산서열을 이용하여 단백질을 검색한 결과, 60% 이상의 단백질을 암호하는 유전자 염기서열을 동정할 수 있었다.

애기장대에서는 잎의 원형질막 부분에 포함된 700개의 단백질이 2-DE에 의하여 분리되었다 (Santoni et al. 1998). 이 중 약 550개가 원형질막에 존재하고 있다고 추정되었다. 이 연구에서는 겔 속에 존재하는 단백질을 트립신에 의하여 절

단하여 얻어진 peptide를 고속액체크로마토그래프에서 정제하여 Edman 아미노산분석기를 이용하여 아미노산 서열을 분석하였다. 그 결과 82개 단백질 스왑의 부분 아미노산 서열이 결정되었다. 그 중에 기능을 알 수 없는 단백질은 65%이었으며, 기능을 알 수 없는 EST가 동정된 것은 약 20%이었다 (Santoni et al. 1998). 이 경우는 게놈 해석에서 얻어진 3만6천 개의 EST로부터 단백질에 대응하는 DNA 염기서열의 동정이 이루어졌다. 이상과 같이 벼 및 애기장대의 연구에 의하여 2-DE에서 분리된 단백질의 부분 아미노산 서열정보를 이용하여 게놈 염기서열 결과를 이용하여 유전자와 단백질의 대응관계를 밝혀낼 수 있었다.

## 벼 프로테옴의 기능해석

### 벼 프로테옴 해석의 현황

벼 게놈 프로젝트는 RFLP (제한효소단편장다형) 연쇄지도상의 DNA마커로서 이용하기 때문에 다수의 cDNA가 clone화되어 그 염기서열이 해석되었다. 그것들의 일부는 추정된 아미노산 서열의 상동성에 기본으로 단백질을 암호하는 DNA에 있어서 해명되어 왔지만, 대부분의 cDNA는 어떠한 기능을 갖고 단백질을 암호하는 것에 대하여는 밝혀지지 않았다. 한편, 기능해명에 관계하는 번역 후의 변형과정 등의 정보를 얻는 단백질의 구조해석기술은 많이 진보해 왔으며, 단백질의 이차원전기영동법을 이용하여 각각의 단백질을 해석할 수 있게 되었다. 프로테옴 해석방법을 벼의 단백질에 이용하여 각 조직별, 발육시기별 (배, 배유, 뿌리, 켈러스, 잎, 엽초, 개화기의 이삭 등)에 있어서 해석을 수행하여 왔다. 각 시기별, 조직 특이적으로 단백질을 추출하여 이차원전기영동을 수행하여 화상해석 후 각각의 단백질에 대하여 기상 아미노산 분석기와 질량분석기를 이용하여 아미노산서열을 결정했다. 그리고, 각각에 대하여 분자량, 등전점, 번역 후의 변형정보, 아미노산 서열정보, 상동성 검색 결과에 대한 data-file를 작성했다. 벼의 각시기별, 조직 특이적인 단백질의 이차원전기영동상에 확인할 수 있는 약 6500개의 단백질의 약 10% 정도 620개에 대하여 해석을 수행하고, 또, 반수의 322개의 단백질에 대하여 아미노산 서열정보를 얻었다. 그러나, 그 중에 약 70%의 단백질에 대하여는 N말단이 block되어 있기 때문에 내부 아미노산 서열정보를 얻었다. 얻어진 아미노산 서열정보를 기본으로 상동성 검색을 수행하여 40.4%의 단백질이 cDNA catalog중에서 검색할 수가 있었다 (Woo et al. in preparation).

또, 부분 아미노산서열이 결정된 322개의 단백질 중에 약 50종류에 대하여 겔로부터 정제한 단백질을 항원으로서 항체를 조정하여 그 기능해명에 이용되고 있다. 지금까지의 벼 켈러스의 재분화를 억제하는 인산화 단백질, 저온스트레스에 의한 인산화능을 향진하는 단백질, 병 감염에 의하여 유도되는 단백

질, 식물호르몬 자극에서 유도되는 단백질 등이 발견되어 항체를 조정하며 그 특이성을 해명하여 유전자를 단리하고 있다.

### 벼 프로테오믹 해석기술의 응용

#### 재분화에 관여하는 단백질의 기능해석

식물이 갖는 분화 전능성은 배양기술의 향상에 의하여 제어가 가능하지만, 전능성의 생리생화학적 기구는 불명한 점이 많다. 한편, 인산화에 의한 단백질의 활성조절기구는 넓은 생물계의 제어양식이며 세포재분화 과정에 있어서도 중요한 역할을 한다고 생각된다. 벼의 재분화에 관여하는 인산화 단백질을 프로테오믹 해석에서 발견하고 유전자 단리후 그 기능을 해석했다.

액체배양한 캘러스를 재분화 고형배지에 이식했을 때 재분화능이 있는 캘러스와 재분화능이 없는 캘러스 사이에 단백질의 수준에서의 차이를 해석하기 위하여 계대배양 2~6주 사이에서 재분화능이 있는 캘러스와 계대배양 55~59주 사이에 재분화능이 없는 캘러스를 이용했다. 단백질 추출 후 인산화 이차원전기영동을 수행하고 인산화 pattern을 비교함에 따라 캘러스 계대배양 중에는 계대배양 기간에 관계하지 않고 인산화 되었지만, 재분화능이 있는 캘러스에 있어서는 재분화 배지에 이식하면 인산화능이 감소하는 단백질을 발견했다. 이 단백질은 무기인산을 이용하여 인산화 반응에서 동등하게 인산화 되는 것도 확인되었다. 캘러스 단백질 data-file 중에는 분자량 56 kDa, 등전점 5.4의 단백질을 검색한 결과 SC-028이며, 별도로 정제하여 저장한 항체도 SC-028과 교차했다. 이 단백질의 N말단 및 내부 아미노산 서열은 옥수수의 calreticulin과 높은 상동성을 보였다 (Komatsu et al. 1996).

캘러스 단백질 data-file의 부분 아미노산 서열을 기초로 하여 probe를 조정하여 캘러스 cDNA library를 screening하여 full-length cDNA를 단리했다. ORF를 포함하는 1558-bp cDNA에서 424아미노산에 있었다. 상동검색의 결과, 동물의 calreticulin과 50~53%, 식물의 calreticulin과 70~93%의 상동성을 보였다. 또, binary vector pIG121-Hm의 CaMV35S promoter의 제어하에 calreticulin 유전자를 배치하여, agrobacterium EHA101을 이용하여 형질전환 벼를 작출했다. Calreticulin을 과잉발현시켜 형질전환벼에 있어서는 재분화율이 낮고 재분화한 후에도 성장속도가 지연 되었다. Calcium결합 단백질에서 calreticulin은 세포내 접착, 세포내 calcium의 homeostasis의 유지, 단백질의 folding, 환경 스트레스응답 등에 관여하는 다기능 조절인자로서 보고되고 있다 (Li and Komatsu 2000).

지베렐린은 발아, 신장, 화아형성 등 식물로서의 본질적인 생활환경 조절에 관여하고 있지만, 그 세포내 정보전달 기구는 불명한 점이 많다. 이것들의 응답에 대하여는 지베렐린 자극의 수용, 전달 후 발현되지만, 그 수용에 관한 기구는 거의 해명되어 있지 않다. 그래서 지베렐린과 결합하는 능력이 있는 벼 단백질을 프로테오믹 해석으로서 접근하고 유전자 단리

후 그 기능을 해석했다.

#### 지베렐린과 결합하는 단백질의 기능해석

벼 유묘기에 있어서 엽초, 엽신으로부터 단백질 추출 후 이차원전기영동을 하여 PVDF막에 전사하여 트리치움을 표식한 지베렐린을 이용하여 ligand binding assay을 했다. 방사선 patterns을 비교함에 의하여 활성형 지베렐린과 특이적으로 결합하는 능력이 있는 단백질을 발견했다. 이 단백질은 in vitro 에서 인산화 후 이차원전기영동을 하여 인산화된 것도 확인되었다. 엽초, 단백질 data-file중에는 분자량 47 kDa, 등전점 5.2의 단백질을 검색한 결과 LS-129이며, 별도로 정제하여 조정한 항체도 LS-120와 교차했다. 이 단백질의 N말단 및 내부 아미노산 서열은 보리 중의 ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) activase 와 높은 상동성을 보였다 (Komatsu et al. 1996).

엽초, 단백질 data-file의 부분 아미노산서열을 기초로하여 probe를 조정하고 엽초, 엽신 cDNA library를 screening하여 full-length cDNA를 단리했다. ORF를 포함하는 1591-bp cDNA에서 466아미노산으로부터 A1 type와 1676-bp cDNA에서 433아미노산으로부터 A2 type의 유전자가 단리되었다. 상동검색의 결과, 다른 고등식물의 ribulose-1,5-bisphosphate carboxylase/oxygenase activase와 73~89%의 상동성을 보였다. 또, binary vector pIG121-Hm의 CaMV35S promoter의 제어하에 ribulose-1,5-bisphosphate carboxylase/oxygenase activase 유전자를 배치하여 agrobacterium EHA101을 이용하여 형질전환체 벼를 작출하는 데 성공하였다. ribulose-1,5-bisphosphate carboxylase/oxygenase activaseA1을 과잉발현시켜 형질전환체 벼에 있어서는 생장이 촉진되었지만, ribulose-1,5-bisphosphate carboxylase/oxygenase activaseA2를 발현시킨 벼에서는 생장변동의 차이가 없었다. 지베렐린과 결합활성이 인정되어 ribulose-1,5-bisphosphate carboxylase/oxygenase activase은 본래 Rubisco를 활성화시키기 때문이라는 보고는 있지만 구체적인 것은 불명하다 (Zhang and Komatsu 2000).

### 프로테오믹 기능해석의 응용

#### 식물 호르몬상 peptide 구조와 기능

기능이 밝혀지지 않는 단백질의 기능을 번역 중 및 번역 후의 변형과정, 단백질의 상호작용, 입체구조 등을 고효율적인 해석으로 기능을 해명하는 것은 프로테오믹 해석의 중요한 목적이다. 2-DE에서 분리된 식물 단백질의 기능을 프로테오믹 해석의 방법으로서 분석한 예는 다음과 같다. 아미노산서열의 상동성 검색한 결과 100개의 단백질 중 30%은 이미 벼에서 해석된 상동성이 있는 단백질이며, 그리고 남은 70%은 지금까지 분석되고 있지 않았던 단백질이라는 것을 알아냈다.

EMBL의 data base를 이용하여 분석한 단백질을 암호하는 EST를 검색한 결과, 45%의 단백질을 암호하는 EST를 검색할 수 있었다. 기능을 알 수 없는 55%의 단백질에 대하여는 앞으로 번역 후 변형, 단백질간의 상호작용, 입체구조, 비교 등의 해석에 의하여 기능을 밝히는 것이 중요하다고 생각된다. 콩 종자에서 발현하는 단백질을 2-DE로 분석한 결과 27 kDa의  $\alpha$ -subunit와 16 kDa의  $\beta$ -subunit로 구성되는 염기성단백질 (Legumin-binding protein, LBP)이 검출되었다 (Hirano 1998). LBP는 콩 이외에 완두, 강낭콩, 팥 등의 콩과 식물과 당근 등에서도 발견되었다. LBP의 구조적 특징을 조사한 결과, LBP에는 동물의 인슐린 수용체와 구조상의 유사성이 있음을 알았다. 또 LBP는 protein kinase의 활성을 갖고 있으며, 인슐린 수용체와 같은 원형질막 일부 세포벽에 존재하는 것이 밝혀졌다. 그리고 LBP를 ligand와 affinity 크로마토그래피에 의하여 LBP에 결합하는 활성을 갖는 단백질을 정제한 결과, 분자량 3,900 Da의 peptide를 분리할 수 있었다. 레그인슐린이라 명명한 peptide는 37번째 아미노산부터 인슐린과 아미노산 서열에 상동성은 없지만, 인슐린과 같이 6번째에 시스테인을 갖고 있으며 이들 모두가 disulfide 결합에 관련하고 있다. 레그인슐린은 세포벽 부근에 위치하고 있고, LBP에 결합하고 있으며 그 protein kinase 활성을 촉진하였다. 이때 레그인슐린의 disulfide 결합을 절단하였지만 그 효과는 발견되지 않았다. 핵자기공명 스펙트로 (NMR)에 의하여 해석한 레그인슐린의 3차구조는 인슐린의 구조와 일부 유사한 점이 있었다. 또 레그인슐린을 당근 캘러스 배양 배지에 첨가하면 식물체 분화능과 배형 성능을 향상시키며 배양세포의 분열촉진에도 효과가 있는 것으로 나타났다. 이 결과로부터 레그인슐린은 레그인슐린 수용체라고 생각되며 LBP 인산화가 분화, 생장 및 세포분열의 제어와의 성장과 분화를 제어하고 있다. 그러나, 식물에서는 지금까지 이와 같은 호르몬과 그 정보 전달계의 존재는 밝혀지지 않았다. 상기의 연구에 의하여 2-DE에서 분리된 단백질을 프로테오믹스 해석의 기법을 이용하여, 호르몬과 유사한 기능을 갖는 peptide와 그것에 의한 정보 전달계의 존재를 시사할 수 있었다.

며 20S proteasome subunit의 N-말단수식이 효모의 N-말단과 차이가 있는가를 조사하기 위해 우선, 20S proteasome subunit을 2-DE로 분리하여 전 subunit의 부분 아미노산서열을 분석하였다. 그리고 20S proteasome subunit을 암호하는 EST를 동정하고 이것을 이용하여 대부분의 subunit을 암호하는 전체 DNA 단편을 클로닝하였다. 이들의 DNA 염기서열을 결정하고 모든 subunit의 아미노산서열을 추정하였다. 이 실험으로부터 벼에 있어서는 효모정상계통에서 아세틸화된 subunit 뿐만 아니라  $\beta$ -subunit의 N-말단이 아세틸화되어 있는 것을 알게 되었다. 이것은 벼와 효모의  $\beta$ -subunit에 기능적으로 차이가 존재할 가능성을 시사하고 있다. 이러한 N-말단수식과 더불어 인산화 등의 번역 후 수식, 단백질의 상호작용 등에 관한 정보가 프로테오믹스 해석에 의하여 얻어지면 이것들의 정

보를 기본으로 모든 subunit의 기능이 보다 명확하게 파악될 수 있을 것으로 생각된다. 프로테오믹스 해석에는 이런 종류들의 연구를 다수의 단백질을 대상으로 하여 고효율적으로 수행할 필요성이 있다. 그러나, 식물에서는 단백질의 기능을 고효율적으로 해석하는 연구는 초보적인 수준이다. 세포내의 모든 단백질에 대해서 기능을 설명하는 것은 현 시점에서는 용이하지 않지만, 현재의 분석 기술로 단백질 복합체 중의 단백질의 기능을 설명하는 것은 그렇게 어려운 일은 아니다.

#### 세포내 미소기관 단백질 복합체의 해석

세포 내에서 모든 단백질에 대한 기능을 설명하는 것은 용이하지 않지만, 단백질 복합체의 단백질의 기능을 설명하는 것도 가능하다. 단백질 복합체의 기능이 점차로 밝혀지게 되면 유전자에 의해 암호된 많은 단백질의 기능을 보다 효율적으로 해석할 수 있다. 단백질 복합체의 기능을 밝히려는 연구의 일환으로 벼에서 분리된 31종류의 subunit에서 거대한 단백질 복합체인 proteasome을 가지고 proteasome subunit의 모든 번역중 및 번역 후의 수식과 기능의 관계를 밝히기 위한 연구를 수행하였는데, 모든 subunit의 N-말단에 아세틸화 및 인산화에 있어서 번역후의 수식을 포괄적으로 해석하고 그들의 수식과 단백질의 기능이 관련이 있음을 시사하는 결과를 얻었다 (Ito et al. 2000; Kimura et al. 2000; Sassa et al. 2000). 이와 같이 프로테오믹스 해석에서 얻어지는 정보를 이용하면 단백질 복합체의 모든 subunit의 기능을 보다 명확하게 얻을 수 있을 것으로 생각된다.

#### 식물 프로테오믹스 분석 결과의 Database화

몇 개의 식물에서는 프로테오믹스 해석에 의하여 얻어진 2-

**Table 1.** Major analysis instruments by using proteome analysis.

Protein analysis and purification	Two-dimensional electrophoresis, protein electroblotting PVDF membranes, fluorescent computer analyzer
Two-dimensional electrophoresis and computer image analysis, Data-base	Computer image analyzer, software
Protein and identification of genes	MALDI-TOF/MS and ESI Q-TOF/MS, gas-phase amino acid sequencer
Protein post-translational modifications	ESI Q-TOF/MS
Crystal structure analysis of proteins	NMR, X-rays crystallography
Protein-protein interactions	NMR, protein chip, MALDI-TOF/MS and ESI Q-TOF/MS
Comparative protein analysis	Two-dimensional electrophoresis, MALDI-TOF/MS and ESI Q-TOF/MS



DE 양상과 분리된 단백질의 배열과 구조, 각 단백질을 암호하는 유전자, 단백질의 기능 등에 관한 정보가 database화되고 있는데, 애기장대 (<http://sphinx.rug.ac.be:8080/ppmdb/index.html>)와 옥수수 (<http://moulon.inra.fr/imgd>)의 경우 database화가 비교적 충실하게 진행되고 있다. 애기장대의 경우 database에 들어가면 하면 우선, 2-DE 양상을 볼 수 있어서 겔상의 spot을 클릭하면 단백질의 등전점, 분자량, MS data, 아미노산서열, 소수성, 세포내 위치, 비교, 다른 database와의 연결, 발표논문 등에 관한 정보를 입수할 수 있다. 그러나, 프로테옴 database는 DNA 염기서열과 단백질의 아미노산서열이 연결되어야 하는데 기존에 밝혀진 외부의 DNA 염기서열 database와 연결하여 분석 및 등록할 수 있는 체계가 없다. 따라서, 이러한 문제점은 앞으로 생물정보학 (bioinformatics)분야에서 개선되어야 할 과제이다.

식물 프로테옴 해석의 과제

이상의 결과는 프로테옴 해석기술이 기능성 단백질의 검출에 있어서 유용한 수단이 될 것이다. 또한, 프로테옴 해석에서는 어떤 조직에서, 어떤 시기에 발현하는 단백질의 해석이 필요하기 때문에 프로테옴 연구는 게놈프로젝트보다도 시간과 노력이 더 많이 요구될 것이라고 생각된다. 이런 연구를 신속히 추진하기 위해서는 대규모적인 프로테옴 해석을 진행하기 위한 연구조직의 구축이 필요할 것으로 생각된다. 게놈 크기가 작은 식물 프로테옴 해석에서는 소수의 연구 인력으로도 대응할 수 있을 가능성도 있다. 그러나, 게놈 크기가 큰 식물 프로테옴 해석을 수행할 경우에는 대규모 연구조직을 구축하는 것이 중요하다. 이렇게 되면, 생물기능을 인위적으로 조절하는 기술의 확립과 유용기능의 강화, 생산성 향상기술의 연구개발에 공헌할 수 있다. 특히 식물에서는 환경문제와 식량문제를 극복하기 위하여 유용식물을 작출할 수 있다고 생각된다. 한편, 프로테옴 해석에서는 2-DE, 2-DE 화상해석장치, MS, 핵자기공명장치, 프라즈몬 공명측정장치 등 많은 종류의 분석장치를 이용할 필요가 있다 (Table 1). 그러나 현재의 보유하고 있는 장치는 대규모적인 프로테옴 해석에 반드시 충분하지 않다. 더욱더 고감도화, 고정도화, 그리고 분석의 신속화, 자동화가 요망되고 있다. 또한, 얻어진 정보의 database화와 그 이용기술의 개발도 이 시점에서 절실하다.

적 요

프로테옴 연구는 여러 생물종에 있어서 조직별, 시기별 특이적으로 발현하고 있는 단백질의 database을 만들고, 병의 치료와 의약품의 개발, 환경문제와 식량문제를 극복하기 위하여 유용식물을 만드는데 목표로 하고 있다. 농림수산업 분야에서는 장래의 지구적 규모에서의 식량 및 환경문제의 해결을 위

하여 유전자 재조합 기술에 의한 신제품 개발이 기대되어 농업상의 중요한 유전자의 단리 및 그 기능해명에 있어서 국제적으로 연구경쟁이 수행되고 있다. 이와 같은 흐름 속에서 벼 게놈 2단계 프로젝트가 1998년에 시작되어 전염기서열의 해석 및 기능해명 프로젝트, 또 full-length cDNA library 프로젝트가 시작되었다. 그러한 상황 하에서 벼 프로테옴 연구에서는 조직, 시기 특이적인 변동과 기관형성과 분화제어 또는 환경응답기구등 식물특유의 생물기능을 동적으로 해명하고 있다. 벼의 프로테옴에는 벼의 각조직, 시기 특이적으로 단백질을 추출 또는 정제하고 이차원전기영동, 화상해석 후에 각각의 단백질에 대하여 기상 아미노산 분석기와 질량분석기를 이용하여 아미노산 서열을 결정하고 있다. 이차원 전기영동상에 확인되는 벼 embryo 단백질의 10%에 대하여 해석을 수행하고 현재 100개의 단백질에 대하여 아미노산 서열정보를 얻었다. 그 중 약 70%의 단백질에 대해서는 N말단이 block되어 있기 때문에 Cleveland peptide mapping법을 개발하여 내부 아미노산서열을 결정했다. 우선, 단백질을 이차원전기영동에서 분리한 후, CBB로 염색한 후 겔을 건조시켰다. 건조겔의 스폿을 자르고 세분하게 절단했다. 이것을 SDS-PAGE겔의 시료구에 넣고, 100 μL의 10% 글리세롤을 포함한 전극액 (팽창액)을 시료구에 주입하여 약 1시간 정도 방치했다. 건조겔은 전극액을 흡수하여 팽창했다. 겔을 팽창시킨 후 팽창액 위에 *Staphylococcus aureus* V8 protease (2 μg)를 포함하는 SDS-PAGE 시료원충액 20 μL을 증충했다. 그 위에 전극액을 주입하여 SDS-PAGE를 수행하였다. 단백질과 protease가 농축겔의 중앙부분까지 이동시켰을 때 전원을 끄고 이차원전기영동에서 분리된 단백질에 관하여 1일 정도에 20 시료의 peptide map를 작성할 수 있었다. 그 결과로서 N말단이 block되어 있다고 생각되는 단백질을 포함하여 100개의 단백질에 대하여 위의 방법을 이용하여 peptide mapping에서 얻어진 peptide의 서열을 기상 아미노산 분석기에 의하여 결정했다. 아미노산서열의 상동성 검색한 결과 100개의 단백질 중 30%은 이미 벼에서 해석된 상동성이 있는 단백질이며, 그리고 남은 70%은 지금까지 분석되고 있지 않았던 단백질이라는 것을 알아냈다. EMBL의 data base을 이용하여 분석한 단백질을 암호하는 EST를 검색한 결과, 45%의 단백질을 암호하는 EST를 검색할 수 있었다. 기능을 알 수 없는 55%의 단백질에 대하여는 앞으로 번역 후 변형, 단백질간의 상호작용, 입체구조, 비교 등의 해석에 의하여 기능을 밝히는 것이 중요하다고 생각된다. 또, 부분 아미노산서열이 결정된 단백질 중 약 60종류에 대하여 겔로부터 정제한 단백질을 항원으로서 항체를 조정하고 그 기능해명에 이용되고 있다.

이차원전기영동을 이용한 differential display법에 따라 벼의 재분화와 생장시 변동이 있는 인산화 단백질로서 calreticulin과 ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) activase을 검출하고 항체이용에 의한 단백질 상호작용해석과 그 유전자 이용에 의한 형질전환체작출에 의하여 기능을 해

석했다. 또, 콩 종자에서 발현하는 단백질을 2-DE로 분석한 결과 27 kDa의 subunit와 16 kDa의 subunit로 구성되는 염기성단백질 (Leginsulin-binding protein, LBP)이 검출되었다. LBP의 구조적 특징을 조사한 결과, LBP에는 동물의 인슐린 수용체와 구조상의 유사성이 있음을 알았다. DNA microarray를 이용하여 유전자발현조절 상황을 해석함에 의하여 단백질 수준에서의 전체적 해석 결과와 상호비교를 검토했다. 이상의 결과는 프로테오믹 해석 기술이 기능성 단백질의 검출에 대하여 유용한 수단이라고 생각된다. 뿐만 아니라 계놈 유전자의 기능발현양식에 대하여 이해를 깊게 파악할 수 있으므로 생물의 증식분화, 성장 그리고 노화 등의 기구를 인위적으로 제어하는 기술을 확립할 수 있는 가능성이 높다. 또한, 식물의 내병성등의 생산성이 높은 식물 분자유종도입 및 품종육성에 길이 열릴 가능성도 높다. 계놈연구와 더불어 프로테오믹 연구를 조직적으로 수행함에 따라 방대한 단백질정보를 얻을 수 있으며 그 생물기능정보와 함께 database화에 의하여 계놈기능의 해명에 박차를 가속화시킬 수 있다.

사사 - 본 연구는 과학기술부 21세기 프론티어 연구개발 사업인 작물유전체 기능연구 사업단인 연구비 지원 (CG 3112)에 의해 수행되었다.

## 인용문헌

- Anderson L, Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* 18: 533-537
- Damerval C, Le Guilloux C (1998) Characterization of novel proteins affected by the  $\alpha 2$  mutation and expression during maize endosperm development. *Mol Gen Genet* 257: 354-361
- Flengsrud R (1993) Separation of acidic barley endosperm proteins by two-dimensional electrophoresis. *Electrophoresis* 14: 1060-1066
- Fukuda M, Islam N, Woo SH, Yamagishi A, Takaoka M, Hirano H (2003) Assessing matrix assisted laser desorption/ionization-time of flight-mass spectrometry as a means of rapid embryo protein identification in rice. *Electrophoresis* 24: 1319-1329
- Gomez L, Sanchez-Monge R, Lopez-Otin C, Salcedo G (1991) Amino acid compositions and sequence analysis of the major low Mr globulins from *Triticum monococcum* L. endosperm. *J Cereal Sci* 14: 117-123
- Görg A, Postel W, Baumer M, Weiss W (1992) Two-dimensional polyacrylamide gel electrophoresis, with immobilized pH gradients in the first dimension, of barley seed proteins: Discrimination of cultivars with different malting grades. *Electrophoresis* 13: 192-203
- Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnol* 17: 994-999
- Hirano H (1989) Microsequence analysis of winged bean seed proteins electroblotted from two-dimensional gel. *J Protein Chem* 8: 115-130
- Hirano H (1997) Screening of rice genes from the cDNA catalog using the data obtained by protein sequencing. *J Protein Chem* 16: 533-536
- Hirano H (1998) Structure and function of hormone-like peptides in plants. *Danbaku, Kakusan and Kouso* (in Japanese) 43: 27-39
- Hirano H, Komatsu S, Nakamura A, Kikuchi F, Kajiwara H, Tsunawawa S, Sakiyama F (1991) Structural homology between semidwarfism-related proteins and glutelin seed protein in rice (*Oryza sativa* L.). *Theor Appl Genet* 83: 153-158
- Islam N, Woo SH, Tsujimoto H, Kawasaki H, Hirano H (2002) Proteome approaches to characterize seed storage proteins related to ditelocentric chromosomes in common wheat (*Triticum aestivum* L.). *Proteomics* 2: 1146-1155
- Ito K, Kimura Y, Sassa H, Hirano H (2000) Identification of the rice 20S proteasome subunits and analysis of their N-terminal modification. *Jpn J Electrophoresis* 44: 205-210
- Kamo M, Kawakami T, Miyatake M, Tsugita A (1995) Separation and characterization of *Arabidopsis thaliana* proteins by two-dimensional gel electrophoresis. *Electrophoresis* 16: 423-430
- Kimura Y, Takaoka M, Tanaka S, Sassa H, Tanaka K, Plevoda B, Sheman F, Hirano H (2000) N-acetylation and proteolytic activity of the yeast 20S proteasome. *J Biol Chem* 275: 4635-4639
- Komatsu S, Kajiwara H, Hirano H (1994) Rice protein library: A data file of rice proteins separated by two-dimensional electrophoresis. *Theor Appl Genet* 86: 935-942
- Komatsu S, Masuda T, Hirano H (1996) Rice gibberellin-binding phosphoprotein structurally related to ribulose-1,5-bisphosphate carboxylase/oxygenase activase. *FEBS Lett* 384: 167-171
- Komatsu S, Masuda T, Abe K (1996) Phosphorylation of a protein (pp56) is related to the regeneration of rice cultured suspension cells. *Plant Cell Physiol* 37: 748-753
- Komatsu S, Li Z, Rakwal R (1999) Separation and characterization of proteins in rice (*Oryza sativa* L.) suspension cultured cells. *Plant Cell Tiss Org Cult* 55: 183-192
- Lafandra D, Kasarda, DD (1985) One and two-dimensional polyacrylamide gel electrophoresis in a single gel: Separation of wheat proteins. *Cereal Chem* 62: 314-319
- Li Z, Komatsu S (2000) Molecular cloning and characterization of calreticulin, a calcium-binding protein involved in the regeneration of rice cultured suspension cells. *Eur J Biochem* 267: 737-745
- Lund AA, Blum PH, Bhatramakki D, Elthon TE (1998) Heat-stress response of maize mitochondria. *Plant Physiol* 116: 1097-1110
- Masson F, Rossignol M (1995) Basic plasticity of protein expression in tobacco leaf plasma membrane. *Plant J* 8: 77-85
- Matsudaira P (1987) Sequence from picomole quantities of proteins

- electroblotted onto polyvinylidene difluoride membranes. *J Biol Chem* 262: 10035-10038
- Oda Y, Huang K, Cross R, Cowburn D, Chait BT (1999) Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci USA* 96: 6591-6596
- Ramani S, Apte SK (1997) Transient expression of multiple genes in salinity stressed young seedlings of rice (*Oryza sativa* L.) cv. Bura Rata. *Biochem Biophys Res Commun* 233: 663-667
- Rey P, Pruvot G, Becuwe N, Eymery F, Rumeau D, Peltier G (1998) A novel thioredoxin-like protein located in the chloroplast is induced by water deficit in *Solanum tuberosum* L. plants. *Plant J* 13: 97-107
- Riccardi F, Gazeau P, de Vienne D, Zivy M (1998) Protein changes in response to progressive water deficit in maize. *Plant Physiol* 117: 1253-1263
- Rouquie D, Peltier JB, Marquis-Mansion M, Tounaire C, Dumas P, Rossignol M (1997) Construction of a directory of tobacco plasma membrane proteins by combined two-dimensional gel electrophoresis and protein sequencing. *Electrophoresis* 18: 654-660
- Santoni V, Rouquie D, Dumas P, Mansion M, Boutry M, Degand H, Dupree P, Packman L, Sherrif J, Prime T, Bauw G, Posada E, Rouze P, Dehais P, Sahnoun I, Barlier I, Rossignol M (1998) Use of a proteome strategy for tagging proteins present at the plasma membrane. *Plant J* 16: 633-641
- Santoni V, Delarue M, Caboche M, Bellini C (1997) A comparison of two-dimensional electrophoresis data with phenotypical traits in *Arabidopsis* leads to the identification of a mutant (*cri 1*) that accumulates cytokinins. *Planta* 202: 62-69
- Sassa H, Oguchi S, Inoue T, Hirano H (2000) Primary structural features of the 20S proteasome subunits of rice (*Oryza sativa* L.). *Gene* 250: 61-69
- Sasaki T, Matsumoto T, Yamamoto K, Sakata K, Baba T, Gojobori T (2002) The genome sequence structure of rice chromosome 1. *Nature* 420: 312-316
- Thiellement H, Bahman N, Damerival C, Plomion C, Rossignol M, Santoni V, de Vienne D, Zivy M (1999) Proteomics for genetic and physiological studies in plants. *Electrophoresis* 20: 2013-2026
- Tsugita A, Kawakami T, Uchiyama Y, Kamo M, Miyatake N, Nozu Y (1994) Separation and characterization of rice proteins. *Electrophoresis* 15: 708-720
- Vekulescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE, Hieter P, Vogelstein B, Kinzler KW (1997) Characterization of the yeast transcriptome. *Cell* 88: 243-251
- Weiss W, Huber G, Engel KH, Pethran A, Dunn M, Gooley AA, Gorg A (1997) Identification and characterization of wheat grain albumin/globulin allergens. *Electrophoresis* 18: 826-833
- Woo SH, Islam N, Takaoka M, Adachi T, Hirano H (2001) Preliminary proteome analysis of the buckwheat embryo-proteins. *Adv Buckwheat Res* 8: 544-548
- Woo SH, Fukuda M, Islam N, Takaoka M, Kawasaki H, Hirano H (2002) Efficient peptide mapping and its application to identification of embryo proteins in the rice (*Oryza sativa* L.) proteome analysis. *Electrophoresis* 23: 647-654
- Woo SH, Hirano H, Jong SK (2003) Separation and characterization of spikelet proteins expressed at the young microspore stage in rice (*Oryza sativa* L.). *Kor J Crop Sci* (revision)
- Woo SH, Higa A, Kimura M, Jong SK, Yamaguchi I (2003) Analysis of wheat lemma proteins separated by two-dimensional polyacrylamide gel electrophoresis. *Biosci Biotechnol Biochem* (in press)
- Zhang L, Zhou W, Vekulescu VE, Kem SE, Hruban RH, Hamilton SR, Vogelstein B, Kinzler KW (1997) Gene expression profiles in normal cancer cells. *Science* 276: 1268-1272
- Zhang Z, Komatsu S (2000) Molecular cloning and characterization of cDNA encoding two isoforms of ribulose-1,5-bisphosphate carboxylase/oxygenase activase in rice (*Oryza sativa* L.). *J Biochem* 128: 383-389

(접수일자 2003년 8월 6일, 수리일자 2003년 9월 4일)