Application of low molecular peptide-based biomarkers for the rapid identification of *Echinochloa* species

So Hyun PARK[1], Sung Yung YOO [1], Jae Ho AN [*2], Tae Wan KIM [*1]

[1] Institute of Ecological Phytochemistry/Department of Plant Life and Environmental Science, Hankyong National University, The Republic of Korea

[2] Department of Civil, Safety and Environmental Engineering, Hankyong National University, The Republic of Korea

[*]***Corresponding author: Prof. Dr. Tae Wan KIM,*** *E-mail: taewkim@hknu.ac.kr*

Prof. Dr. Jae Ho AN, *E-mail: jhan@hknu.ac.kr*

Abstract

The objective of the present study was to determine the low molecular peptides in Echinochloa seeds as biomarkers for classification. Some peptide-based biomarkers for the classification of Echinochloa species were identified using SELDI-TOF, a mass spectrometry (MS) technique. Proteomic profiling using SELDI-TOF MS techniques could be a useful and powerful tool to discover peptide biomarkers and to discriminate and classify Echinochloa species, especially under 20 kDa. In 12 Echinochloa species, a total of 72 peptides were significantly detected on strong anion exchanger (CM10) and weak cation exchanger (Q10) arrays. Sixteen peptides on CM10 and 12 peptides on Q10 were selected as peptide biomarkers. The hierarchical heat map analysis of the peptide-based biomarkers indicated that Echinochloa species were classified into two groups and then more precisely subdivided. One major cluster group included early barnyard grass (ECOR), large barnyard grass (ECREC), jungle ricegrass (ECOLON), meadow barnyard grass (EPRAT), and barnyard grass (ECHCG). Hairless barnyard grass (ECGL), cockspur (ECRCR), awned billion-dollar grass (ECFRAW), awnless billion-dollar grass (ECFRAWL), awnless barnyard grass (EMITIS), awnless Japanese barnyard millet (EESCAWL), and awned Japanese barnyard millet (EESCAW) belonged to the other group.

Keywords: *Barnyard grass, Echinochloa crus-galli, low molecular peptide biomarker, proteomic profiling, SELDI-TOF.*

Nomenclature: CHCA, a-cyano-4-hydroxycinnamic acid; jungle ricegrass, *Echinochloa colona* (L.) Link ECOLON; cockspur, *E. crus-galli* P. Beauv. var. *crus-galli* ECRCR; large barnyard grass, *E. crus-galli* var. *echinata* (Willd.) Honda ECREC; esculent barnyard grass, *E. esculenta* (A. Braun) H. Scholz, ECHCG; awnless Japanese barnyard millet, *E. esculenta* (A. Braun) H. Scholz EESCAWL; awned Japanese barnyard millet, *E. esculenta* (A. Braun) H. Scholz EESCAW; awnless barnyard grass, *E. crus-galli* (L.) P. Beauv. var. mitis (Pursh) Peterm EMITIS; meadow barnyard grass, *E. crus-galli* var. *praticola* Ohwi EPRAT; awned billion-dollar grass, *E. frumentacea* ECFRAW; awnless billion-dollar grass, *E. frumentacea* ECFRAWL; hairless barnyard grass, *E. glabrescens* Munro ex Hook. f. ECGL; early barnyard grass, *E. oryzoides* (Ard.) Fritsch ECOR; MS, mass spectrometry; *m/z*, mass-to-charge ratio; PCA, principal component analysis; SELDI, surface enhanced laser desorption/ionization; S:N, signal-to-noise; TOF, time-of-flight.

Introduction

Barnyard grass (Echinochloa crus-galli) is an annual grass weed included in the Global Compendium of Weeds (Randall, 2002), and is considered one of the world's most troublesome weeds. Barnyard grass is particularly abundant in flooded rice fields, where it reduces yields by up to 40% (Smith *et al.*, 1991). *Echinochloa* weed species are a major constraint to rice production worldwide. On the other hand, Japanese barnyard millet (*E. frumentacea*) is used as a food in Korea, India, and Japan (Sood *et al.*, 2015; Thathola and Srivastava, 2002). Since *Echinochloa* spp. cause serious problems, both in paddy areas and upland, they have been considered major weeds in many studies, and their appropriate

classification has been established on the basis of genetic and morphological variations (Damalas *et al.*, 2008; Im *et al*., 1989; Lopez-Martinez *et al.*, 1999; Moon *et al*., 2004; Nozawa *et al.*, 2006; Parani *et al.*, 2001; Ruiz-Santaella *et al.*, 2006; Tabacchi *et al.*, 2009). However, new emerging ecotypes, as well as hybrid species from crosses among the different species, make it more difficult to classify *Echinochloa* spp.

In addition to morphological characteristics, protein specificity, and processing suitability (Branlard *et al*., 2001), a peptide biomarker is necessary for the discrimination of wild barnyard grass species. Proteomics is an important technique for biological systems information, and broad instrument-intensive research in this area has rapidly advanced (Baginsky, 2009; Issaq *et al.*, 2002; van Wijk, 2001). The identification and informatics of low-molecular-weight peptides are important in the field of proteomics research. There have been no reports on peptide profiling and biomarker discovery in barnyard grass seeds using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS). This technique requires very small sample volumes. In addition, it may be the most effective at profiling low-molecular-weight proteins below 20 kDa and has a high throughput capacity for biomarker discovery (Issaq *et al.*, 2002; Mæland Nilsen *et al.*, 2011; Seibert *et al.*, 2004). Even in the field of seed science, very few reports have been published on this technique. However, it has been demonstrated that proteomic profiling using SELDI-TOF MS is useful and powerful for discovering biomarkers in humans and in cereal crops (Lee *et al.*, 2011; Park *et al.*, 2013).

The objective of this study was to identify peptide biomarkers and low molecular peptides in the seeds of barnyard grass species, using SELDI-TOF MS. In agriculture, these molecules may also be predictive of future phenotypic trait endpoints (Schudoma *et al.*, 2012). Low-molecular-weight protein profiling was conducted and small peptide biomarkers were selected for the classification of different barnyard grasses using SELDI-TOF MS.

Material and methods

Seed preparation

Seeds from 12 species of barnyard grass (*Echinochloa* spp.) were collected in the Korean Peninsula or obtained from the Seed Bank of Wild Plant Resources at Korea University, Korea.

Sample Preparation

For total soluble proteins, 100 mg of barnyard grass flour was suspended in 1 ml of extraction buffer [1.25 mM sodium borate, pH 10.0, 2% (v/v) 2-mercaptoethanol, 1% (v/v) Triton X-100] and mixed on a shaker inside an ice box for 1 h before being centrifuged at 12,000 x *g* for min. The total soluble protein content of the extract was quantified with the Bradford assay (Bradford, 1976; Seibert *et al.*, 2004; Zhong et al., 2010). The protein extract was used immediately for SELDI-TOF MS a5nalysis following the published protocol (Park *et al.*, 2013).

Surface-enhanced Laser Desorption/Ionization Time-of-flight Mass Spectrometry Peptide Profiling

Two kinds of ProteinChip arrays were conducted using Q10 (strong anion exchanger) and CM10 (weak cation exchanger) arrays (Bio-Rad Laboratories). Protein samples (1 mg/ml) were diluted by 1:10 with binding/washing buffer (#K20-00010-MSDS). Next, 100 μL of binding solution was added to each well and incubated at room temperature for 5 min with vigorous shaking (250 rpm). This procedure was repeated once. After the second incubation, the buffer was removed and 100 μL of each sample was added to individual wells. The samples were incubated at room temperature with vigorous shaking for 30 min, then removed, and the wells were washed three times with 200 μL of binding buffer for 5 min with agitation. Next, the wells were washed three times with 200 μL of distilled water. The arrays were air-dried for 20 min, and 1 μL of α-cyano-4-hydroxycinnamic acid (CHCA)

ProteinChip energy-absorbing molecule (EAM) solution was added to each well. An additional 1 μL of EAM was then added to each well. Finally, the array surfaces were allowed to dry completely before SELDI-TOF MS was performed (Agrawa *et al.*, 2013; Park *et al.*, 2013; Seibert *et al.*, 2004).

Data Collection

The two arrays (Q10 and CM10) were read in a ProteinChip SELDI System (Bio-Rad Laboratories). The arrays were analyzed with an ion acceleration potential of 25 kV. The mass range investigated was from 2,000 to 30,000 Da. Ten laser-shot pixels were averaged; two warning shots per pixel were taken before data collection, but the warning-shot data were not included in the data average. This set of acquisition protocols was performed on a reference sample and on the experimental samples. The laser energy for low calibration was 2,400 nJ. Before each SELDI-TOF MS analysis, a mass calibration was performed using the all-in-one peptide standard (#C10-00005-MSDS).

Data Analysis

All data were processed using ProteinChip Data Manager Software (Bio-Rad Laboratories). Baseline subtraction was performed with a smoothing of 25 points, auto-fitting width, filtering of 0.2 times expected, and noise range starting at 1,000 Da. The spectra were normalized by total ion current with a mass-to-charge ratio (m/z) range from 1,500 to 30,000 Da. To obtain repeatability and statistically significant differences at $p<0.01$ among the 12 barnyard grass species in the SELDI-TOF MS analysis. Peak cluster was performed according to the two-step parameter setting. For the first step, peaks were automatically detected according to the specified S:N (signal-to-noise) ratio of 5 and a minimum valley depth of 3, if they were found in at least 10% of all spectra, with a mass window of 5 peak width. The settings for the second step were an S:N of 2 and a minimum valley depth of 2.

Estimated peak was added using the "at cluster center" option. The m/z range was set between 2,000 and 30,000.

Statistical Analysis

ProteinChip Data Manager Software was used for statistical analysis (Bio-Rad Laboratories). Once the peaks were clustered, univariate analysis with a nonparametric test was performed, calculating the p-values associated with each cluster (and using the Kruskal–Wallis test for comparison of the 12 barnyard grass cultivars). Peaks with a p-value of <0.000001 were detected by multivariate analyses, with a supervised hierarchical heat map and principal component analysis (PCA) to evaluate their potential as biomarkers.

Results and Discussion

Peptide Profiling in Echinochloa spp. seeds

A total of 168 peak spectra from 12 barnyard grass species were detected. All peptide spectra on CM10 (weak cation) were practically equal to Q10 (strong anion) (Fig. 1 and 2). We obtained the peptide profiling data of barnyard grass seed in the molecular mass range of 1,500 to 30,000 Da. A total of 72 peaks, 39 on the CM10 array and 33 on the Q10 array, in the molecular weight range below 30,000 Da, were detected universally at $p<0.00001$ in the 12 species (Tables 1 and 2). This indicates that these peptides were universally released in the 12 *Echinochloa* species during protein extraction. This SELDI-TOF MS analysis has advantages in the profiling and discovery of low-molecular-weight peptides that could be used as biomarkers, at least in *Echinochloa* spp. Further research is needed to confirm peptide expression, the identification and repeatability of significantly different peptides released in *Echinochloa* seeds, and the discovery of their functions in seed metabolism. The advantage of SELDI-TOF MS technology is powerful in the discovery and analysis of low-molecular-weight protein below 20 kDa (Issaq *et al*., 2002).

Species Discrimination by Peptide Profiles

Significantly different peptide peaks ($p<0.000001$) among the 12 *Echinochloa* species were detected on CM10 and Q10 in the molecular mass range of 2000 to 30,000 Da (Tables 1 and 2). The numbers of significantly different peptide peaks ($p<0.000001$) among the 12 *Echinochloa* species were 16 on CM10 and 12 on Q10, respectively (bold cluster in Tables 1 and 2). Specifically, 16 peaks on CM10 and 10 peaks on Q10 were differentiated below 10 kDa. This technique using SELDI-TOF provides profiles for low-molecular-weight proteins in the seeds of *Echinochloa* species, as with earlier reports on other plant parts (Schudoma *et al.*, 2012) and rice seed (Park *et al.*, 2013).

Biomarker Selection by Up- and Down-regulated Peak

Based on the heat map analysis of significant difference at the $p<0.000001$ peak cluster, biomarkers (up- or down-regulated peptides) were determined in 16 peaks on CM10 and 12 peaks on Q10 (Table 3). On CM10, the up-regulated peak clusters were 2,358.3, 5,214.7, 7,398.9, 4,323.2, 3,883.4, 3,903.5, 4,113.3, 4,426.5, and 5,015.5 Da. On Q10, the up-regulated peak clusters were 3,509.4, 7,114.9, 7,174.4, 7,396.3, 7,471.9, 14,885.7, and 29,588.7 Da. On CM 10, the down-regulated peak clusters were 2,120.1, 2,130.8, 2,108.7, 6,095.1, 6,154.6, 6,341.6, and 6,408.3 Da. On Q 10, the down-regulated peak clusters were 2,599.2, 3,248.4, 3,322.5, 4,390.1, and 6,495.8 Da.

Many low-molecular-weight peptides in *Echinochloa* spp. were released universally at $p<0.000001$, and were significantly different at $p<0.000001$ among the 12 *Echinochloa* species. These results suggest a peptide profile of barnyard grass seed in the molecular weight range below 30 kDa.

Classification and Cluster Analysis of Echinochloa species

On the basis of significantly different peptide peaks, we attempted to discriminate among the 12 species through a principal component analysis and a heat map analysis via supervised hierarchical clustering at $p<0.01$ (Fig. 3 and 4). Twelve *Echinochloa* species were discriminated and separated into different two clusters. Group 1 included five *Echinochloa* species: ECOR, ECREC, ECOLON, EPRAT, and ECHCG. Group 2 included seven species: ECGL, ECRCR, ECFRAW, ECFRAWL, EMITIS, EESCAWL, and EESCAW.

In earlier studies, it was reported that several barnyard grass species could be morphologically identified as *E. crus-galli* P. Beauv. var. *crus-galli*, *E. crus-galli* var. *echinata* (Willd.) Honda, *E. crus-galli* var. *praticola* Ohwi, *E. oryzoides* (Ard.) Fritsch], *E. colona* (L.) Link, *E. glabrescens* Munro ex Hook. f., and *E. esculenta* (A. Braun) H. Scholz (Altop and Mennan, 2011; Ehara and Abe, 1950; Moon *et al.*, 2004; Tabacchi *et al.*, 2009). In a recent classification study, *Echinochloa* accessions were grouped into three different species (*E. crus-galli*, *E. erecta*, and *E. phyllopogon*) according to Pignatti's classification key, and into four different species according to Carretero's taxonomy (*E. crus-galli*, *E. hispidula*, *E. oryzicola*, and *E. oryzoides*) (Tabacchi *et al.*, 2009). The *E. crus-galli* accessions clustered as a specific group under both AFLP analysis and a morphological traits analysis carried out according to Pignatti's and Carretero's keys (Tabacchi *et al*., 2009). In Group 1, EPRAT (*E. crus-galli* var. *praticola*) and ECHCG *E. crus-galli* (L. Beauv) were closely clustered. The neighbouring species were ordered to ECREC (*E. crus-galli* var. *echinata* (Willd.) Honda), ECOR (*E. oryzoides* (Ard.) Fritsch), and then ECOLON (*E. colona* (L.) Link). In Group 2, however, awned billion-dollar grass (ECFRAW) was far from awnless billion-dollar grass (ECFRAWL). Awned Japanese barnyard millet was also far from awnless Japanese barnyard millet. Thus, the morphological characteristics of awn may be inconsistent as a critical key, at least for peptide-based classification. Nevertheless, in

Group 2, ECRCR, ECFRAW, EMITIS, and EESCAWL belonged to the same subdivided group, and ECGL neighbored them. Surprisingly, ECFRAWL was intimately clustered to EESCAWL. These subdivided clusters on CM10 were completely the same as on Q10, indicating that the repeatability and responsiveness of peptide-based classifications were significant. Overall, our result was closer to Carretero's taxonomy (Altop and Mennan, 2011; Tabacchi *et al.*, 2009) and other morphological classifications in Korean *Echinochloa* species (Moon *et al.*, 2004). In recent whole-genome genotyping of cultivated *Echinochloa* accessions, including jungle ricegrass (ECOLON) and barnyard grass (ECHCG), it was determined that there are probably four population clusters within the ECOLON accessions and three such clusters within ECHCG (Wallace *et al.*, 2015). As shown in Figures 3 and 4, there were two sub-clusters within Group 1, *i.e*. ECOLON and others, including ECHCG. These clusters match phylogenetic relationships, but by and large they do not correspond to classification into individual races or clusters based on morphology, as in Wallace at al.'s indication (Wallace *et al.*, 2015). These morphological mismatches were shown in group II, which were sub-clustered by ECFRAWL, including EESCAW and five other species: ECGL, ECRCR, ECFRAW, EMITIS, and EESCAWL. Two cultivated species, ECFRAWL and ECFRAW, were differently sub-clustered. EPRAT and ECHCG were almost the same clusters; their morphological traits were difficult to discriminate.

To classify potential biomarker peptides at a higher significance level of $p<0.000001$, we attempted a principal component analysis and heat map analysis. Three replicates of each *Echinochloa* species were well-grouped, and two groups were discriminated in three-dimensional spaces with three principal components (red and blue color in Fig. 5). A heat map analysis of supervised hierarchical clustering at $p<0.000001$ also showed that even three replicates of each species, grouped as a cluster, and two groups were well-separated into different clusters (Fig. 6). Interestingly, cluster distance among the *Echinochloa* species

revealed that two groups were equally divided on both CM10 and Q10. In other words, ECOR, ECREC, ECOLON, EPRAT, and ECHCG (Group 1) were distant from ECGL, ECRCR, ECFRAW, ECFRAWL, EMITIS, EESCAWL, and EESCAW (Group 2).

In conclusion, peptides of *Echinochloa* species, detected using SELDI-TOF MS, which has advantages in the profiling and discovery of low-molecular-weight peptides, could be used as biomarkers for *Echinochloa* species identification and discrimination. Although we obtained peptide profiles and biomarkers for *Echinochloa* species in this study, further research is needed to confirm peptide expression, the identification and repeatability of significantly different peptides released in rice seeds, and the discovery of their functions in seed metabolism.

Acknowledgements

We thank Professor Kang, B. H. of korea university for his morphological identification and the seed supply.

References

Agrawal G K, Jwa N S, Jung Y H, Kim S T, Kim D W, Cho K, Shibato J, Rakwal R (2013). Rice proteomic analysis: sample preparation for protein identification. In: *Rice Protocols*). Springer, 151-184.

Altop E K, Mennan H (2011). Genetic and morphologic diversity of Echinochloa crus-galli populations from different origins. *Phytoparasitica* **39**: 93-102.

Baginsky S (2009). Plant proteomics: concepts, applications, and novel strategies for data interpretation. *Mass spectrometry reviews* **28**: 93-120.

Bradford M M (1976). Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Analytical Biochemistry* **72**: 248-254.

Branlard G, Dardevet M, Saccomano R, Lagoutte F, Gourdon J (2001). Genetic diversity of wheat storage proteins and bread wheat quality. In: *Wheat in a global environment*). Springer, 157-169.

Damalas C A, Dhima K V, Eleftherohorinos I G (2008). Morphological and physiological variation among species of the genus Echinochloa in Northern Greece. *Weed science* **56**: 416-423.

Ehara K, Abe S (1950). Classification of the forms of Japanese barnyard millet. In: *Proceedings of the Crop Science Society of Japan*). 245-246.

Im I, Guh J, Lee Y (1989). Weed-Ecological Classification of the Collected Barnyardgrass [Echinochloa crus-galli (L.) Beauv.] in Korea-II. Classification of collected barnyardgrass in growth pattern by multivariate clustering. *Korean Journal of Weed Science* **9**: 1-15.

Issaq H J, Veenstra T D, Conrads T P, Felschow D (2002). The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochemical and biophysical research communications* **292**: 587-592.

Lee J Y, Kim J Y, Park G W, Cheon M H, Kwon K H, Ahn Y H, Moon M H, Lee H J, Paik Y K, Yoo J S (2011). Targeted Mass Spectrometric Approach for Biomarker Discovery and Validation with Nonglycosylated Tryptic Peptides from N-linked Glycoproteins in Human Plasma. *Molecular & Cellular Proteomics* **10**.

Lopez-Martinez N, Salva A P, Finch R P, Marshall G, De Prado R (1999). Molecular markers indicate intraspecific variation in the control of Echinochloa spp. with quinclorac. *Weed science*, 310-315.

Mæland Nilsen M, Uleberg K E, Janssen E A, Baak J, Andersen O K, Hjelle A (2011). From SELDI-TOF MS to protein identification by on-chip elution. *Journal of proteomics* **74**: 2995-2998.

Moon B C, Oh S M, DiTomaso J M, Fischer A, Kwon Y W (2004). Morphological classification of Echinochloa species collected in rice paddy fields of California, US. *Korean Journal of Weed Science* **24**: 213-220.

Nozawa S, Takahashi M, Nakai H, Sato Y I (2006). Difference in SSR variations between Japanese barnyard millet (Echinochloa esculenta) and its wild relative E. crus-galli. *Breeding Science* **56**: 335-340.

Parani M, Rajesh K, Lakshmi M, Parducci L, Szmidt A E, Parida A (2001). Species identification in seven small millet species using polymerase chain reaction-restriction fragment length polymorphism of trn S-psb C gene region. *Genome* **44**: 495-499.

Park S J, Lee Y H, Nam Y J, Baik B K, Kim T W (2013). Peptide-Based Biomarker Discovery for Identification of Rice Cultivars using Surface Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Crop Science* **53**: 987-995.

Randall R P (2002) A global compendium of weeds. R G and F J Richardson, Melbourne, Victoria. 906, abstr.

Ruiz-Santaella J, Bastida F, Franco A, De Prado R (2006). Morphological and molecular characterization of different Echinochloa spp. and Oryza sativa populations. *Journal of agricultural and food chemistry* **54**: 1166-1172.

Schudoma C, Steinfath M, Sprenger H, van Dongen J T, Hincha D, Zuther E, Geigenberger P, Kopka J, Köhl K, Walther D (2012). Conducting Molecular Biomarker Discovery Studies in Plants. In: *High-Throughput Phenotyping in Plants*). Springer, 127-150.

Seibert V, Wiesner A, Buschmann T, Meuer J (2004). Surface-enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI TOF-MS) and ProteinChip® technology in proteomics research. *Pathology-Research and Practice* **200**, 83-94.

Smith J, Seaton J, Osuji P, Daguiar P, Chesney P, Mcbean M, Haynes A, Harding C, Cumberbatch N (1991). Dry-Matter and Crude Protein Yields of Echinochloa-Pyramidalis on Coastal Clay Soil of Guyana. *Turrialba* **41**: 69-75.

Sood S, Khulbe R K, Gupta A K, Agrawal P K, Upadhyaya H D, Bhatt J C (2015). Barnyard millet–a potential food and feed crop of future. *Plant Breeding* **134**: 135-147.

Tabacchi M., Mantegazza R., Spada A. and Ferrero A. (2009). Morphological traits and molecular markers for classification of Echinochloa species from Italian rice fields.

Thathola A, Srivastava S (2002). Physicochemical properties and nutritional traits of millet-based weaning food suitable for infants of the Kumaon hills, Northern India. *Asia Pacific journal of clinical nutrition* **11**: 28-32.

van Wijk K J (2001). Challenges and prospects of plant proteomics. *Plant Physiol* **126**: 501-508.

Wallace J G, Upadhyaya H D, Vetriventhan M, Buckler E S, Tom Hash C, Ramu P (2015). The genetic makeup of a global barnyard millet germplasm collection. *The Plant Genome* **8**.

Zhong L, Taylor D L, Whittington R J (2010). Proteomic profiling of ovine serum by SELDI-TOF MS: optimisation, reproducibility and feasibility of biomarker discovery using routinely collected samples. *Comparative immunology, microbiology and infectious diseases* **33**: 47-63.

Table 1. Protein peak clusters and their average m/z and intensity by CM10 arrays of SELDI-TOF MS in *Echinochloa* spp.

		Peak *m/z*[†] (Da)			Intensity			Number of peaks	
No.	*p*-value	Mean	Median	%CV[‡]	Mean	Median	%CV	Total	Estimated[§]
1	**0.000001**	**2,108.7**	**2,108.7**	**0.0165**	**84.4**	**67.6**	**101.4**	**35**	**12**
2	**0.000001**	**2,120.1**	**2,120.1**	**0.0031**	**8.0**	**4.9**	**101.5**	**35**	**32**
3	**0.000002**	**2,131.0**	**2,130.8**	**0.0451**	**13.8**	**7.6**	**100.5**	**35**	**19**
4	**0.000001**	**2,358.3**	**2,358.5**	**0.0173**	**2.8**	**0.5**	**191.5**	**35**	**26**
5	0.000002	2,442.5	2,442.7	0.0170	49.2	23.8	115.1	35	12
6	0.000001	2,598.4	2,598.9	0.0639	23.0	11.2	103.2	35	14
7	0.000006	3,171.5	3,171.5	0.0146	17.0	3.5	145.8	35	24
8	0.000002	3,496.9	3,496.8	0.0126	6.5	5.0	62.2	35	28
9	**0.000001**	**3,883.4**	**3,883.4**	**0.0108**	**12.0**	**2.0**	**197.1**	**35**	**29**
10	**0.000001**	**3,903.3**	**3,903.8**	**0.0399**	**18.1**	**2.8**	**136.3**	**35**	**21**
11	0.000001	3,914.0	3,913.9	0.0089	10.2	2.5	220.9	35	30
12	**0.000001**	**4,113.5**	**4,113.4**	**0.0322**	**10.1**	**3.5**	**185.2**	**35**	**28**
13	0.000002	4,133.7	4,133.4	0.0286	12.5	5.3	101.3	35	22
14	0.000008	4,180.1	4,180.9	0.0772	12.8	5.4	134.8	35	26
15	0.000008	4,194.6	4,194.6	0.0183	12.0	5.7	107.4	35	25
16	0.000003	4,274.9	4,274.5	0.1389	18.0	18.6	81.3	35	11
17	**0.000001**	**4,323.2**	**4,323.3**	**0.0107**	**10.6**	**4.0**	**106.3**	**35**	**23**
18	0.000008	4,399.7	4,399.7	0.0512	11.1	11.8	96.9	35	22
19	**0.000001**	**4,429.2**	**4,426.6**	**0.1509**	**23.1**	**6.2**	**116.2**	**35**	**18**
20	0.000005	4,511.3	4,511.3	0.0107	7.5	2.6	118.0	35	24
21	0.000009	4,699.1	4,704.3	0.1798	22.1	12.9	92.1	35	8
22	0.000004	4,904.4	4,904.5	0.1351	18.4	18.2	77.2	35	9
23	**0.000001**	**5,014.7**	**5,015.5**	**0.1024**	**11.5**	**8.2**	**87.1**	**35**	**27**
24	**0.000001**	**5,215.7**	**5,214.6**	**0.2554**	**13.0**	**7.9**	**79.9**	**35**	**16**
25	0.000003	5,434.1	5,437.6	0.2455	9.4	5.0	102.0	35	21
26	**0.000001**	**6,095.1**	**6,095.5**	**0.0205**	**5.0**	**1.9**	**131.5**	**35**	**26**
27	**0.000001**	**6,154.6**	**6,154.6**	**0.0096**	**6.0**	**5.5**	**92.6**	**35**	**32**
28	0.000003	6,199.1	6,198.6	0.0578	9.8	8.6	88.7	35	21
29	0.000002	6,233.9	6,234.0	0.0717	11.2	11.2	74.5	35	8

30	0.000002	6,276.3	6,276.2	0.0189	8.1	7.9	79.7	35	30
31	**0.000001**	**6,341.6**	**6,341.6**	**0.0208**	**5.4**	**4.9**	**93.6**	**35**	**25**
32	**0.000001**	**6,407.9**	**6,408.7**	**0.0906**	**4.3**	**4.0**	**106.2**	**35**	**26**
33	0.000005	6,486.9	6,484.6	0.1269	3.6	2.4	99.1	35	17
34	**0.000001**	**7,396.3**	**7,399.0**	**0.2287**	**1.8**	**1.4**	**86.5**	**35**	**24**
35	0.000003	8,398.6	8,399.7	0.0831	5.8	3.2	103.5	35	15
36	0.000009	12,357.6	12,360.9	0.0785	11.1	9.3	76.4	35	23
37	0.000009	12,420.2	12,420.4	0.0442	11.7	12.4	65.2	35	8
38	0.000002	12,854.9	12,853.9	0.3175	2.7	2.9	73.3	35	27
39	0.000003	17,547.9	17,545.0	0.1281	0.2	0.1	78.8	35	15

†*m/z*, mass-to-charge ratio.

‡CV represents the coefficient of variance.

§Cluster in bold represents significantly different peak cluster ($p<0.00001$) among the *Echinochloa* spp.

Table 2. Protein peak clusters and their average m/z and intensity by Q10 arrays of SELDI-TOF MS in *Echinochloa* spp.

No.	*p*-value	Peak *m/z*† (Da)			Intensity			Number of peaks	
		Mean	Median	%CV‡	Mean	Median	%CV	Total	Estimated§
1	0.000009	2108.3	2107.8	0.0703	6.2	5.8	84.5	35	12
2	0.000002	2262.4	2262.2	0.0320	3.7	2.1	149.7	35	27
3	0.000008	2334.5	2334.5	0.0019	1.6	1.7	75.4	35	33
4	0.000005	2455.9	2455.3	0.0866	3.9	3.6	88.3	35	20
5	0.000008	2535.6	2535.9	0.0294	4.2	2.2	118.0	35	19
6	0.000001	2579.9	2580.0	0.0556	3.9	.2.5	119.3	35	18
7	**0.000001**	**2599.2**	**2599.3**	**0.0076**	**2.1**	**1.4**	**91.8**	**35**	**32**
8	**0.000001**	**2796.2**	**2796.1**	**0.0154**	**4.5**	**2.9**	**103.9**	**35**	**16**
9	0.000001	3248.4	3248.3	0.0146	9.8	4.2	144.2	35	15
10	**0.000001**	**3322.5**	**3322.6**	**0.0130**	**2.9**	**2.6**	**97..5**	**35**	**22**
11	**0.000001**	**3509.4**	**3510.7**	**0.0679**	**2.1**	**1.1**	**91.1**	**35**	**22**
12	0.000001	3560.8	3561.6	0.0647	4.9	3.5	79.1	35	18
13	0.000002	3662.8	3662.6	0.0171	2.0	2.0	71.6	35	25
14	**0.000001**	**4389.8**	**4390.5**	**0.0472**	**5.5**	**5.4**	**77.1**	**35**	**16**
15	0.000003	4686.4	4686.3	0.0125	3.3	2.8	60.3	35	24
16	0.000006	4899.7	4899.8	0.0119	31.7	18.5	98.3	35	11
17	0.000001	6377.5	6381.3	0.4253	3.4	3.7	76.9	35	16
18	**0.000001**	**6494.5**	**6496.2**	**0.0620**	**2.1**	**2.0**	**81.2**	**35**	**22**
19	0.000001	7034.8	7035.5	0.1285	10.2	8.3	58.3	35	3
20	**0.000001**	**7116.2**	**7115.3**	**0.0577**	**6.8**	**6.3**	**50.6**	**35**	**28**
21	**0.000001**	**7174.4**	**7174.7**	**0.0130**	**5.6**	**4.0**	**68.0**	**35**	**27**
22	**0.000001**	**7396.3**	**7396.2**	**0.0092**	**1.8**	**1.1**	**101.9**	**35**	**31**
23	**0.000001**	**7473.6**	**7472.2**	**0.2445**	**1.6**	**1.0**	**91.2**	**35**	**19**
24	0.000004	7694.8	7693.1	0.1195	1.3	1.3	78.1	35	14
25	0.000001	13840.7	13840.7	0.0179	6.1	5.1	45.7	35	33
26	0.000001	14.037.6	14024.4	0.3133	11.7	8.9	71.9	35	3
27	**0.000001**	**24890.1**	**14885.3**	**0.3055**	**1.2**	**0.8**	**71.0**	**35**	**29**
28	0.000002	17506.6	17508.0	0.1091	0.1	0.1	81.3	35	14
29	0.000001	27775.5	27854.7	0.6600	0.7	0.5	87.2	35	1

30	0.000002	28578.1	28608.9	0.3992	0.3	0.2	76.3	35	26
31	0.000002	29115.9	29127.6	0.3486	0.2	0.2	72.7	35	22
32	**0.000001**	**29588.5**	**29588.4**	**0.1384**	**0.1**	**0.1**	**68.9**	**35**	**27**
33	0.000002	29883.9	29889.2	0.1101	0.1	0.1	65.3	35	28

†*m/z*, mass-to-charge ratio.
‡CV represents the coefficient of variance.
§Cluster in bold represents significantly different peak cluster ($p<0.00001$) among the *Echinochloa* spp.

Table 3. Molecular weights (m/z) of selected small peptide biomarkers in barnyard grass species, separately fixed on different protein chips (CM10 and Q10)

Chip array	Regulation of peak intensity	m/z mean	Intensity average
CM10	Down-regulated	2,120.1	8
		2,130.8	13.6
		2,108.7	84.4
		6,095.1	5
		6,154.6	6
		6,341.6	5.4
		6,408.3	4.1
	Up-regulated	2,358.3	2.8
		5,214.7	10
		7,398.9	1.7
		4,323.2	10.6
		3,883.4	12
		3,903.5	18
		4,113.3	9.9
		4,426.5	8.8
		5,015.5	11
Q10	Down-regulated	2,599.2	2.1
		3,248.4	9.8
		3,322.5	2.9
		4,390.1	5.5
		6,495.8	2
	Up-regulated	3,509.4	2.1
		7,114.9	6.7
		7,174.4	5.6
		7,396.3	1.8
		7,471.9	1.6
		14,885.7	1.2
		29,588.7	0.1

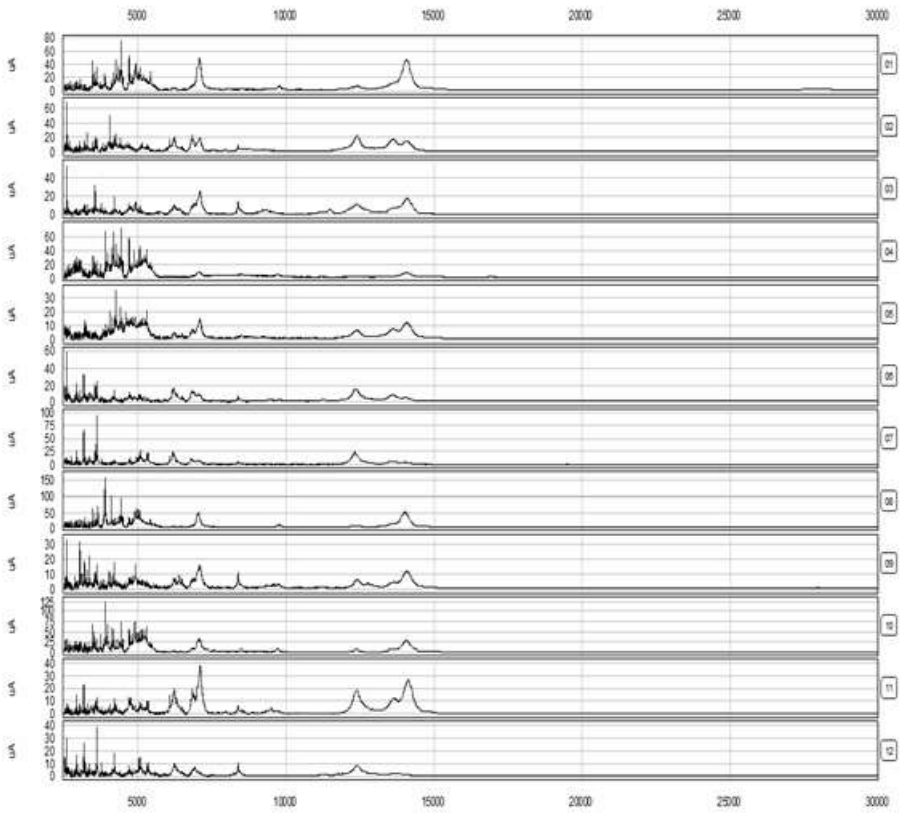

Fig. 1. SELDI-TOF MS spectra on CM10 protein chips. 1, ECOR; 2, ECGL; 3, ECRCR; 4, ECREC; 5, ECOLON; 6, ECFRAW; 7, ECFRAWL; 8, EPRAT; 9, EMITIS; 10, ECHCG; 11, EESCAWL; 12, EESCAW.

1: *E. oryzoides* (Ard.) Fritsch (early barnyard grass) ECOR
2: *E. glabrescens* Munro ex Hook. f. (hairless barnyard grass) ECGL
3: *E. crus-galli* (L.) P. Beauv. var. *crus-galli* (cockspur) ECRCR
4: *E. crus-galli* var. *echinata* (Willd.) Honda (large barnyard grass) ECREC
5: *E. colona* (L.) Link (jungle ricegrass) ECOLON
6: *E. frumentacea* (awned billion-dollar grass) ECFRAW
7: *E. frumentacea (*awnless billion-dollar grass) ECFRAWL
8: *E. crus-galli* var. *praticola* Ohwi (meadow barnyard grass) EPRAT
9: *E. crus-galli* (L.) P. Beauv. var. mitis (Pursh) Peterm (awnless barnyard grass) EMITIS
10: *E. crus-galli* (L.) P. Beauv. (barnyard grass) ECHCG
11: *E. esculenta* (A. Braun) H. Scholz (awnless Japanese barnyard millet) EESCAWL
12: *E. esculenta* (A. Braun) H. Scholz (awned Japanese barnyard millet) EESCAW

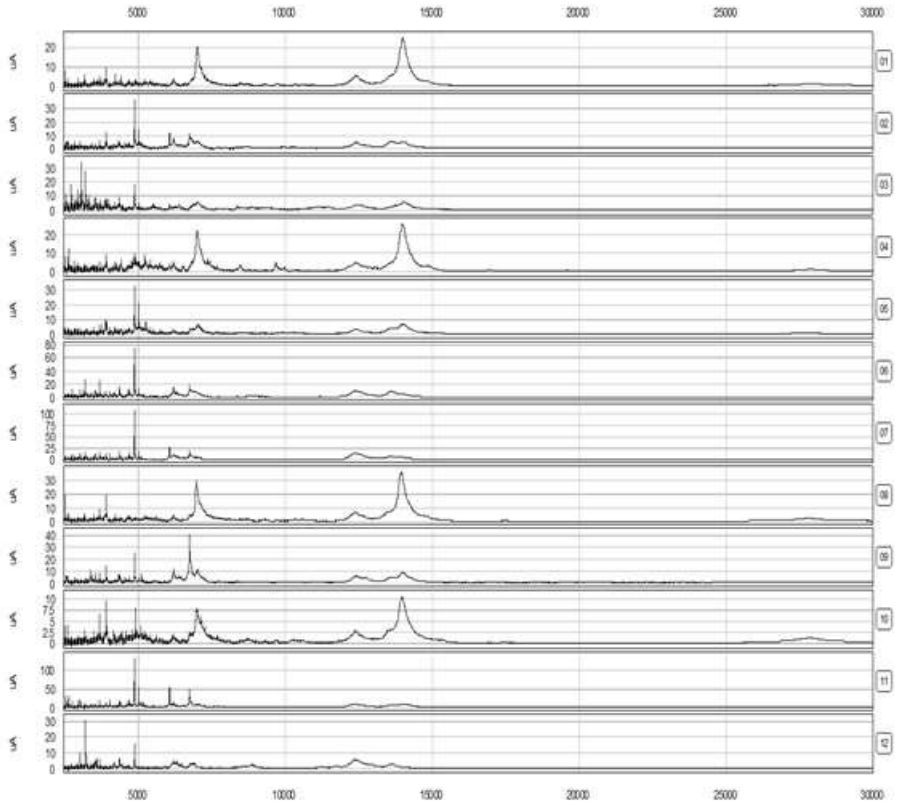

Fig. 2. SELDI-TOF MS spectra on Q10 protein. 1, ECOR; 2, ECGL; 3, ECRCR; 4, ECREC; 5, ECOLON; 6, ECFRAW; 7, ECFRAWL; 8, EPRAT; 9, EMITIS; 10, ECHCG; 11, EESCAWL; 12, EESCAW.

1: *E. oryzoides* (Ard.) Fritsch (early barnyard grass)
2: *E. glabrescens* Munro ex Hook. f. (hairless barnyard grass)
3: *E. crus-galli* P. Beauv. var. *crus-galli* (cockspur)
4: *E. crus-galli* var. *echinata* (Willd.) Honda (large barnyard grass)
5: *E. colona* (L.) Link (jungle ricegrass),
6: *E. frumentacea* (awned billion-dollar grass)
7: *E. frumentacea (*awnless billion-dollar grass)
8: *E. crus-galli* var. *praticola* Ohwi (meadow barnyard grass)
9: *E. crus-galli* (L.) P. Beauv. var. mitis (Pursh) Peterm (awnless barnyard grass)
10: *E. crus-galli* (L.) Beauv. (barnyard grass)
11: *E. esculenta* (A. Braun) H. Scholz (awnless Japanese barnyard millet)
12: *E. esculenta* (A. Braun) H. Scholz (awned Japanese barnyard millet)

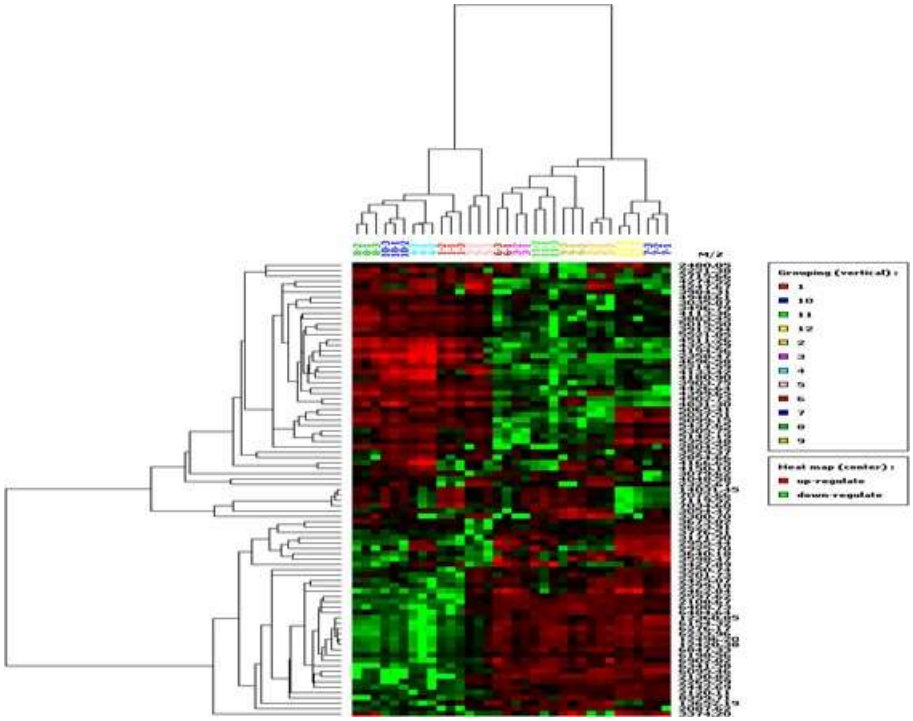

Fig. 3. Supervised hierarchical clustering and heat map using significantly different peaks on CM10 array of SELDI-TOF MS at $p<0.1$ in *Echinochloa* spp. 1, ECOR; 2, ECGL; 3, ECRCR; 4, ECREC; 5, ECOLON; 6, ECFRAW; 7, ECFRAWL; 8, EPRAT; 9, EMITIS; 10, ECHCG; 11, EESCAWL; 12, EESCAW.

1: *E. oryzoides* (Ard.) Fritsch (early barnyard grass)
2: *E. glabrescens* Munro ex Hook. f. (hairless barnyard grass)
3: *E. crus-galli* P. Beauv. var. *crus-galli* (cockspur)
4: *E. crus-galli* var. *echinata* (Willd.) Honda (large barnyard grass)
5: *E. colona* (L.) Link (jungle ricegrass)
6: *E. frumentacea* (awned billion-dollar grass)
7: *E. frumentacea (*awnless billion-dollar grass)
8: *E. crus-galli* var. *praticola* Ohwi (meadow barnyard grass)
9: *E. crus-galli* (L.) P. Beauv. var. mitis (Pursh) Peterm (awnless barnyard grass)
10: *E. crus-galli* (L.) Beauv. (barnyard grass)
11: *E. esculenta* (A. Braun) H. Scholz (awnless Japanese barnyard millet)
12: *E. esculenta* (A. Braun) H. Scholz (awned Japanese barnyard millet)

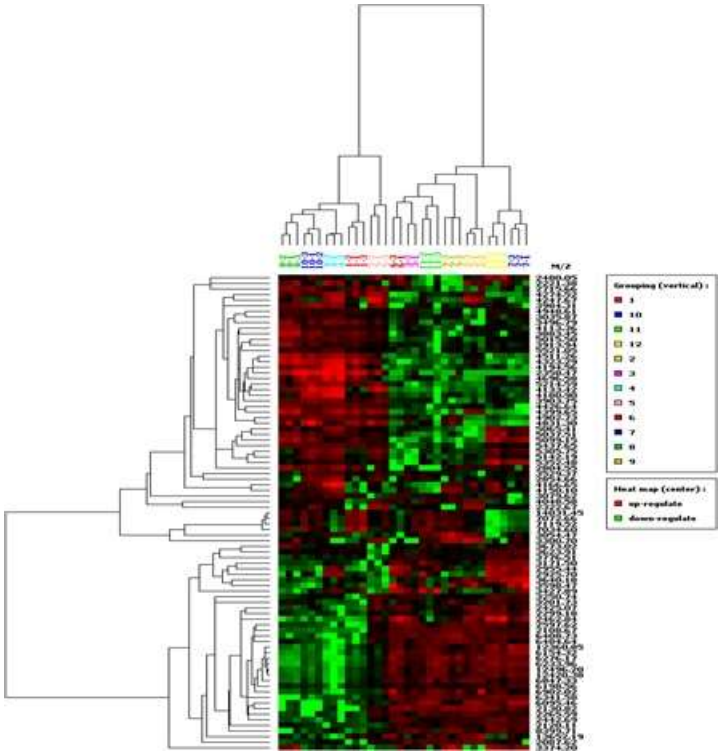

Fig. 4. Supervised hierarchical clustering and heat map using significantly different peaks on Q10 array of SELDI-TOF MS at p<0.01 in *Echinochloa* spp. 1. ECOR; 2, ECGL; 3, ECRCR; 4, ECREC; 5, ECOLON; 6, ECFRAW; 7, ECFRAWL; 8, EPRAT; 9, EMITIS; 10, ECHCG; 11, EESCAWL; 12, EESCAW.

1: *E. oryzoides* (Ard.) Fritsch (early barnyard grass)
2: *E. glabrescens* Munro ex Hook. f. (hairless barnyard grass)
3: *E. crus-galli* P. Beauv. var. *crus-galli* (cockspur)
4: *E. crus-galli* var. *echinata* (Willd.) Honda (large barnyard grass)
5: *E. colona* (L.) Link (jungle ricegrass)
6: *E. frumentacea* (awned billion-dollar grass)
7: *E. frumentacea (*awnless billion-dollar grass)
8: *E. crus-galli* var. *praticola* Ohwi (meadow barnyard grass)
9: *E. crus-galli* (L.) P. Beauv. var. mitis (Pursh) Peterm (awnless barnyard grass)
10: *E. crus-galli* (L.) Beauv. (barnyard grass)
11: *E. esculenta* (A. Braun) H. Scholz (awnless Japanese barnyard millet)
12: *E. esculenta* (A. Braun) H. Scholz (awned Japanese barnyard millet)

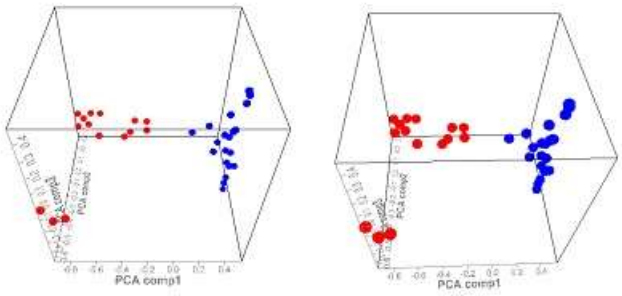

Fig. 5. Principal component analysis using significantly different peak clusters detected on CM10 (left) and Q10 (right) among barnyard grasses at p<0.000001 (red, Group 1; blue, Group 2).

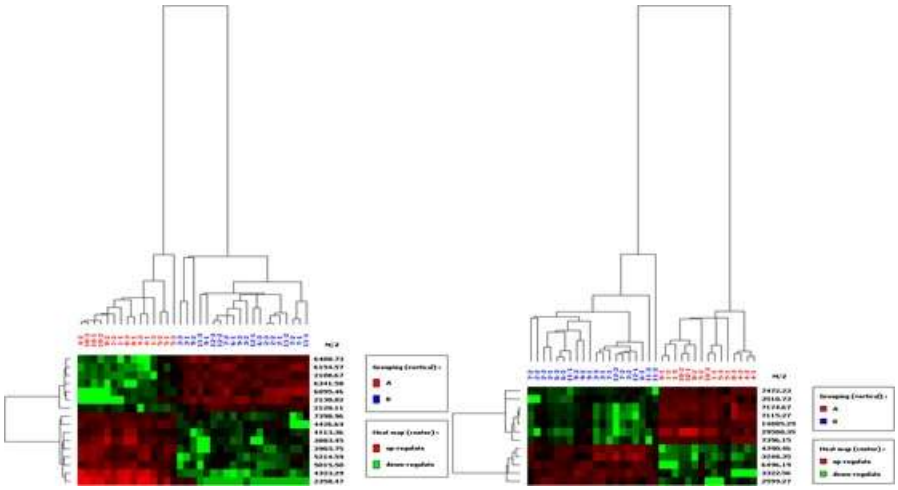

Fig. 6. Supervised hierarchical clustering and heat map using significantly different peaks on CM10 (left) and Q10 (right) array of SELDI-TOF MS at p<0.000001 in *Echinochloa* spp. 1, ECOR; 2, ECGL; 3, ECRCR; 4, ECREC; 5, ECOLON; 6, ECFRAW; 7, ECFRAWL; 8, EPRAT; 9, EMITIS; 10, ECHCG; 11, EESCAWL; 12, EESCAW.

1: *E. oryzoides* (Ard.) Fritsch (early barnyard grass)
2: *E. glabrescens* Munro ex Hook. f. (hairless barnyard grass)
3: *E. crus-galli* P. Beauv. var. *crus-galli* (cockspur)
4: *E. crus-galli* var. *echinata* (Willd.) Honda (large barnyard grass)
5: *E. colona* (L.) Link (jungle ricegrass)
6: *E. frumentacea* (awned billion-dollar grass)
7: *E. frumentacea (*awnless billion-dollar grass)
8: *E. crus-galli* var. *praticola* Ohwi (meadow barnyard grass)
9: *E. crus-galli* (L.) P. Beauv. var. mitis (Pursh) Peterm (awnless barnyard grass)
10: *E. crus-galli* (L.) Beauv. (barnyard grass)
11: *E. esculenta* (A. Braun) H. Scholz (awnless Japanese barnyard millet)
12: *E. esculenta* (A. Braun) H. Scholz (awned Japanese barnyard millet)