ISSN 0006-2979, Biochemistry (Moscow), 2009, Vol. 74, No. 5, pp. 480-490. © Pleiades Publishing, Ltd., 2009.
Published in Russian in Biokhimiya, 2009, Vol. 74, No. 5, pp. 593-606.
Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM08-158, April 12, 2009.

Proteome Analysis of the Moss *Physcomitrella patens* (Hedw.) B.S.G.

A. Yu. Skripnikov[1,2]*, N. B. Polyakov[1], E. V. Tolcheva[1], V. V. Velikodvorskaya[1#], S. V. Dolgov[3], I. A. Demina[4], M. A. Rogova[4], and V. M. Govorun[1,4]

[1]*Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, 117997 Moscow, Russia; fax: (495) 336-0777; E-mail: a.skripnikov@gmail.com*
[2]*Biology Faculty, Lomonosov Moscow State University, 119991 Moscow, Russia; fax: (495) 939-1268*
[3]*Branch of Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, pr. Nauki 6, 142290 Pushchino, Moscow Region, Russia; E-mail: dolgov@fibkh.serpukhov.su*
[4]*Scientific Research Institute of Physicochemical Medicine, ul. Malaya Pirogovskaya 1a, 119435 Moscow, Russia; fax: (495) 246-4401; E-mail: govorun@hotmail.ru*

Received April 30, 2008
Revision received July 3, 2008

Abstract—The sequencing of the moss *Physcomitrella patens* genome has facilitated studies of the plant proteome. To develop a proteome reference map based on the genome sequence, we conducted 2D electrophoreses of proteins extracted from moss protoplasts, protonemata, and gametophores grown under standard conditions on Petri dishes. On silver-stained gels, depending on the developmental stage of the moss, we resolved from 500 to 600 protein spots that were then excised and digested by trypsin, and 212 proteins were identified by PMF-MALDI-TOF. To enhance the proteome coverage, we performed 1D SDS-PAGE with subsequent separation of tryptic peptides derived from digested gel band slices by LC-ESI-MS/MS. The proposed approach allowed us to identify 186 proteins had not been determined by 2D PMF-MALDI-TOF. Proteins identified by both methods were categorized using a system of clusterization of orthologous genes as metabolism (26%), cellular processes and signaling (16%), and information storage and processing (7%). Proteome analysis by differential gel electrophoresis revealed moderate differences between filamentous protonemata and leafy shoots. Surprisingly, protoplasts isolated from protonema filaments displayed significant differences in protein composition compared with both protonemata and gametophores.

DOI: 10.1134/S0006297909050022

Key words: proteomics, 2D electrophoresis, MALDI-TOF, LC-ESI-MS/MS, *Physcomitrella patens*, protoplasts, DiGE

Recent publication of the complete nucleotide sequence of the genome of *Physcomitrella patens* (Hedw.) B.S.G. moss has considerably extended the capabilities of plant proteomics that have been used for proteome profiling of principal phases of the moss life cycle and protoplasts. In recent years proteomic technologies have been actively used in systems biology for identification of proteins and determination of their posttranslational modifications. Proteomic methodology is associated with identification, cataloging, qualitative, quantitative, and systems analysis of the whole set of proteins synthesized in an organism, tissue, or cell. The number of papers devoted to plant proteomics is significantly less than that devoted to proteomic investigations of prokaryotic, yeast, and mammalian objects [1]. This is very largely associated with difficulties in preparation of protein samples from plant tissues containing substances interfering with proteomic analysis, including polysaccharides of cellular walls, lipids, phenols, and chloroplast proteins dominating in

Abbreviations: DiGE, differential gel electrophoresis; DTT, dithiothreitol; KOG, EuKaryotic Orthologous Groups; LC-ESI-MS/MS, liquid chromatography-electrospray ionization-tandem mass spectrometry; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; PAGE, polyacrylamide gel electrophoresis; PMF, peptide mass fingerprinting; RuBisCo, ribulose bisphosphate carboxylase; SDS, sodium dodecyl sulfate.

* To whom correspondence should be addressed.
# Deceased.

green tissues, such as ribulose bisphosphate carboxylase (RuBisCo) [2]. Proteomic studies have been performed for more than 35 plant species. However, plant proteomic publications have been focused on two organisms [1]: the *Arabidopsis thaliana* L. model dicotyledonous plant and the *Oryza sativa* L. monocotyledonous rice. Evidently, these plants attract researchers by public availability of their genomic sequences, which are necessary for protein identification based on mass spectrometric analysis. Nevertheless, exploration of only two angiosperm plants (one of them is a cereal crop and another is a weed evolved as an ephemeral plant with a short life cycle) is obviously insufficient to solve a number of central problems of plant physiology and biochemistry. Moreover, cellular model systems are crucial for plant development study [3]. Traditional cell cultures of seed plants in many instances are inadequate as model systems for study molecular mechanisms of plant development [3, 4].

The genome sequence of *P. patens* was determined after publication of nucleotide sequences of *Arabidopsis*, rice, and poplar. The complete nucleotide sequence of the nuclear genome of *P. patens* was determined by an international consortium and assembled into 480 megabase pairs scaffold sequence [5]. From this study, 35,938 gene sequences of the moss became available in the online gene databases http://genome.jgi-psf.org/Phypa1_1/Phypa1_1.info.html (US DOI Joint Genome Institute (JGI)) and http://www.ncbi.nlm.nih.gov//genomes/geblast.cgi?bact=off&gi=5880 (National Center for Biotechnology (NCBI), USA).

Mosses are of fundamental interest for biology because they occupy an ideal phylogenetic position as representatives of the first and oldest (450 million years) branches of terrestrial plants. Molecular and biological study of mosses significantly contributes to reconstruction of evolutionary changes accompanying such important events in the Earth's history as conquest of land by plants [5]. Mosses are characterized by a large diversity and number of species; they occupy the second place after the angiosperms. Mosses are distinguished by large range of biochemical and physiological plasticity. They are capable of adaptation to extreme temperature, humidity, and light conditions in the terrestrial environment and form associations of symbiotic character with various organisms, especially with nitrogen-fixing cyanobacteria. Drought and cold resistance of mosses are of increased applied interest for studies of molecular mechanisms of stability of higher plants to extreme factors of the environment. Mosses have developed unique physiological mechanisms of gravireception and phototropism during adaptation to terrestrial habitat. These mechanisms are of interest for evolutionary, gravitational, and space biology.

The first studies of proteomic mapping of the moss were performed before the completion of the *P. patens* genome project using local bases of the EST fragments of the moss genome and bioinformatic algorithms [6, 7].

The goal of this study was proteomic mapping of the moss on the basis of the database of annotated genome sequences of *P. patens* using two-dimensional (2D) gel electrophoresis, MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) mass spectrometric analysis of peptides by peptide mass fingerprint (PMF), and one-dimensional (1D) gel electrophoresis under denaturing conditions with subsequent analysis of tryptic peptides by LC-ESI-MS/MS (liquid chromatography-electrospray ionization-tandem mass spectrometry). Basic phases of the moss life cycle (protonemata and gametophores) were used as objects of the proteomic analysis similarly to earlier moss proteomic studies [6, 7]. In addition, moss protoplasts that model early stages of development of an organism (primary chloronemata) from spores were analyzed for the first time.

MATERIALS AND METHODS

Growing of moss protonema and gametophores and preparation of protoplasts. Protonemata of *P. patens* of the Gransden strain was grown on Knop modified agarized medium with the addition of 500 mg/liter of ammonium tartrate (PPNH4 medium) [8, 9] under lighting with white light of luminescent lamps (F96T12/GRO/VHO/WS; Sylvania, USA) with photon flux of 61 $\mu M/m^2$ per sec under the conditions 16-h daylight at 26°C. The agar medium was poured into glass Petri dishes (9 cm in diameter) and allowed to form a gel. The dishes were covered with special cellophane disks to prevent growing protonema filaments into the medium. Water suspension (1 ml) of fragments of the protonema filaments (10-50 cells) prepared using a Virtis homogenizer (USA) was poured out on the surface of a cellophane disk of every Petri dish using a 2-ml glass pipette with a cut point. On the seventh day of the culture growth, the protonema filaments were collected using a spatula from the surface of a cellophane disk and placed in a 150-ml beaker with 100 ml of sterile distilled water. The sterile working nozzle of the homogenizer was immersed into the water, and the protonema filaments were fragmented for 1 min at a rotation rate of 2000 rpm. The suspension was taken with a sterile glass pipette with a cut point and poured on the fresh agarized medium. Passing was performed once every 7 days. Protonema at the age of 5 days was used for the proteomic analysis. The moss gametophores were grown on the Knop modified agarized medium (PPNO3 medium) [8, 9] in 9-cm Petri dishes under lighting with a white light from luminescent lamps (F96T12/GRO/VHO/WS) with a photon flux of 61 $\mu M/m^2$ per sec under the conditions of 16-h daylight at 26°C. Curtains of gametophores were divided using tweezers into fragments containing 4-5 shoots and transplanted onto fresh culture medium once a month. Gametophores at the age of one month were used for the proteomic analysis.

Protoplasts were isolated from the protonema filaments using 1% solution of Driselase (Fluka, Switzerland) in 0.48 M mannitol (Fluka) [10]. The protonema at the age of 5 days was collected using a spatula from a cellophane disk, placed in a sterile 5-cm Petri dish, supplemented with 10 ml of 1% solution of Driselase in 0.48 M mannitol, and incubated for 30 min under continuous careful stirring by manual circular motion (60 rpm) under dim illumination (5 $\mu M/m^2$ per sec) at 25°C. The suspension of isolated protoplasts was filtered trough 80 μm stainless steel sieve. The resulting protoplast suspension in the enzyme solution was incubated under the same conditions for 15 min. The protoplasts were precipitated by centrifugation in 10-ml glass tubes at 100*g* for 5 min using a bucket rotor, and the protoplast suspension was washed free of the enzyme with 0.48 M mannitol by centrifugation at the same rate two times. The freshly isolated protoplasts were immediately used for the protein isolation 1 h after the beginning of the enzymatic treatment of the protonema.

Isolation of proteins from protoplasts and tissues of the moss. Proteins were extracted from the moss protoplasts, protonema, and gametophores on the basis of the technique adapted for *P. patens* [7]. We proposed an additional preparative stage for the isolation of protein from the protoplasts. The protoplast suspension in 0.48 M mannitol was precipitated using a centrifuge with a bucket rotor at 100*g* for 5 min in 10-ml glass tubes. The supernatant was taken, and highly concentrated suspension remained at the bottom of a tube. It contained 4 million protoplasts in the volume of 200 μl. The protoplast suspension was immediately transferred into a porcelain mortar preliminarily cooled to −70°C and frozen in liquid nitrogen. Preparation of protein samples from the protonema filaments at the age of 5 days was started by the transfer of the moss filaments, which were collected from the surface of cellophane discs on the agar medium using a spatula, in a porcelain mortar preliminarily cooled to −70°C where they were immediately cooled in liquid nitrogen. For the preparation of protein samples from gametophores, moss shoots a the age of 1 month were cut off using a scalpel at a height of 1 mm from the surface of the agar medium and transferred by tweezers into a porcelain mortar preliminarily cooled to −70°C where they were immediately frozen in liquid nitrogen. Then, the frozen protoplasts, filaments of the protonema, and shoots of the moss were ground with a pestle preliminarily cooled to −70°C to a fine powder. The plant material was supplemented with 10% solution of trichloroacetic acid in acetone with 0.07% dithiothreitol (DTT) cooled to −20°C and incubated for 1 h at the same temperature for protein precipitation. The suspension with the protein precipitate was centrifuged at 30,000*g* for 15 min at 4°C. The protein precipitate was treated with acetone that was cooled to −20°C and contained 0.07% DTT, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 2 mM EDTA and vigorously shaken for removal of pigments and lipids from the preparation. Then, the protein suspension was centrifuged at 30,000*g* for 15 min at 4°C, and this procedure of washing of the protein preparation with acetone was repeated two times. The protein samples in polymeric centrifuge tubes (volume of 2 ml) were dried in a vacuum centrifuge and stored at −70°C no more than 10 days before 2D electrophoresis.

2D electrophoresis of proteins was carried out according to a published procedure [11] with insignificant modifications. After the completion of the electrophoresis, gels were marked and stained by treatment with silver with thiosulfate [12]. Images of the silver-stained gels were obtained using an Epson Perfection 4990 scanner and analyzed using the PDQuest 8.0 program (BioRad, USA).

1D electrophoresis of proteins under denaturing conditions in the presence of sodium dodecyl sulfate (SDS). One-dimensional fractionation of proteins in 5-20% gradient polyamide gel in the presence of SDS was performed according to the standard Laemmli procedure [13]. Coomassie G-250 was used for visualization of protein bands.

Proteins were hydrolyzed by trypsin in polyacrylamide gel by the method proposed by Shevchenko et al. [14] with insignificant changes.

MALDI mass spectrometry. The solution of tryptic peptides (2 μl) after extraction from the gel was mixed with 0.25 μl of 2,5-dihydrobenzoic acid (20 mg/ml in 30% solution of acetonitrile containing 0.5% acetic acid, v/v) on a steel target. The mixture was dried in air at 23°C for 30 min. The peptides prepared as described above were analyzed using a time-off-flight Ultraflex-TOF-TOF mass spectrometer (Bruker Daltonics, Germany) with a MALDI source equipped with a UV laser (337 nm) in a regime of detection of positive ions using a reflectron at the following settings of ionic source: voltages on IS1 and IS2 were 25 and 21.75 kV, respectively, voltage on lens(es) was 9.5 kV, and voltages on the reflectron were 26.43 (Ref1) and 13.80 kV (Ref2).

Ions were detected in the m/z range from 700 to 4000. Peaks of the trypsin autolysis (m/z 842.508, 1045.563, 2211.093), keratin (m/z 1475.780), and detected impurities that were specially determined by us and automatically removed from the final lists of detected masses were used as an internal standard.

Analysis of mass spectrometry results. The spectra were processed using the Flex Analysis 2.4 software (Bruker Daltonics). Smoothing according to the Savitzky and Golay algorithm [15] (width of 0.1 m/z, one cycle) and subtraction of a base line according to the Convex Hull algorithm [16] were applied to the spectra. The following parameters of peak detection were used: SNAP algorithm of the peak detection, the signal/noise ratio of 6, and the quality spectrum threshold of 100. Protein sequences of *P. patens* were searched in the database (http://genome.jgi-psf.org/Phypa1_1/Phypa1_1.home. html), v.1.1 using the MASCOT program complex (local version 2.1.03; Matrix

Science, Great Britain). The following parameters of the search were used: exactness of the mass determination was 100 ppm and methionine oxidation was considered as a possible posttranslational modification.

Proteins determined with the MOWSE score probability parameter being higher than the threshold parameter were considered to be significantly determined (significance of the determination was 95%, $p > 0.05$). This value was 58 for the database of protein sequences used in our study. Coincidence of molecular mass and/or isoelectric point of each particular protein was an additional control criterion.

1D liquid chromatography and electrospray ionization mass spectrometry (1D LC-ESI-MS/MS analysis). Bands of proteins from the protoplasts that were obtained using 1D gel electrophoresis and separate spots that were obtained by 2D gel electrophoresis but not determined by peptide mass fingerprinting were subjected to analysis by mass spectrometry on an Esquire 6000 plus (Bruker Daltonics) mass spectrometer equipped with a nanospray ion source with a quadrupole ionic trap as the mass analyzer. This device was combined with an LC Packings Ultimate nanochromatograph with an LC Packings Famos autosampler (Dionex, USA) in on-line regime. The chromatographic part of the installation consisted of two serially connected columns with an electromagnetic valve between them. The first column (100 μm × 3 cm) was filled with the Poros R2 polymeric phase (the hydrophobic polymeric phase with large pore diameter analogous to C_8) and used for preliminary concentration of the sample and its desalting. The second column (75 μm × 25 cm) was filled with Phenomenex C_{18} sorbent (particle size of 5 μm, pore diameter of 300 Å) and directly used for fractionation of the desalted mixture of tryptic peptides.

The chromatographic fractionation was performed at a flow rate of 200 μl/min (to splitter), and real flow rates for the desalting and the fractionation were 900 and 200 nl/min, respectively. Peptides were fractionated in a linear gradient of 75% acetonitrile and 25% isopropanol in 0.1% formic acid (from 5 to 60% within 48 min). Measurements were carried out within the m/z range from 300 to 2500 with the trap optimization mass of 700. Tandem experiments were performed only with ions whose charge number was equal or higher than 2 with intensity higher than the threshold. The resulting lists of determined masses were sent to the MASCOT search system. The search was performed both in the nuclear genome JGI database and in the NCBI database with the exception for the *P. patens* taxon. The search results were verified using the Scaffold program complex, version 01-07-00 (http://www.proteomesoftware.com) for approval of correctness of the determination of proteins and revelation of identical entries in the course of search in the two databases. Proteins with identification probability of 95% were saved in the final list. The number of determined peptides per protein was two or more.

2D electrophoresis of proteins with differential staining (DiGE). Total preparations of proteins isolated from protoplasts, protonemata, and gametophores were dissolved in 100 μl of buffer containing chromatographic water (Merck, USA), 8 M urea (Amersham, USA), 2 M thiourea (Amersham), 10 mM Tris-HCl, pH 8.0 (Amersham), 167 μl of solution of 30% CHAPS (Amersham), and 10% Nonidet P-40, pH 8.0 (Fluka). The samples were centrifuged at 14,000*g* for 15 min. The protein concentration was measured according to Bradford [17] using bovine serum albumin (BioRad) as a standard. The proteins were conjugated with the following cyanine fluorescent agents (Amersham): CyDye-DiGE Cy3 (proteins of protoplasts) and CyDye-DiGE Cy5 (proteins of protonemata), in accordance with the recommendations of the manufacturer (400 pmol per 50 mg of total protein). The mixture of the proteins with the cyanines was incubated for 30 min in the dark at 4°C. The reaction was stopped by addition of 10 mM solution of lysine. One-dimensional electrophoresis was performed for every protein preparation in 12% polyacrylamide gel for fluorimetry of cyanines. Before isoelectrofocusing, the protein preparations were mixed in the ratio providing equal total fluorescent signal from the proteins of every sample (protoplasts and protonema). Concentrations of DTT and Ampholine 3-10 were 80 mM and 0.2%, respectively. The isoelectrofocusing and the second direction of the electrophoresis were performed according to the standard procedure.

The resulting gels were scanned on a Typhoon device (Amersham) in the 600 PTM regime. After the fluorescent analysis, the gels were stained with silver by the procedure described above.

RESULTS

The total protein preparations obtained from protoplasts, protonemata, and gametophores of the moss were fractionated by 2D electrophoresis in a pH gradient within the 3-10 range. The fractionation was repeated no less than 6 times in order to obtain reproducible representative 2D map of the *P. patens* proteome. The program analysis of 2D electrophoregrams (PDQuest 8.0; BioRad) revealed high reproducibility of fractionation of proteins with molecular masses from 10 to 110 kDa for moss protoplasts, protonema, and gametophores. We resolved 504 (Fig. 1, see color insert), 596, and 585 protein spots on the electrophoregrams of the total protein extracts from the protoplasts, the protonema, and the gametophores, respectively (see Table 1 as Supplement for the on-line publication of this article at the site of *Biochemistry* (Moscow): http://protein.bio.msu.ru/biokhimiya). Proteins were identified in 212 spots cut from the gels by MALDI mass spectrometry using peptide mass fingerprinting and the MASCOT program (Table 1 from on-line publication).

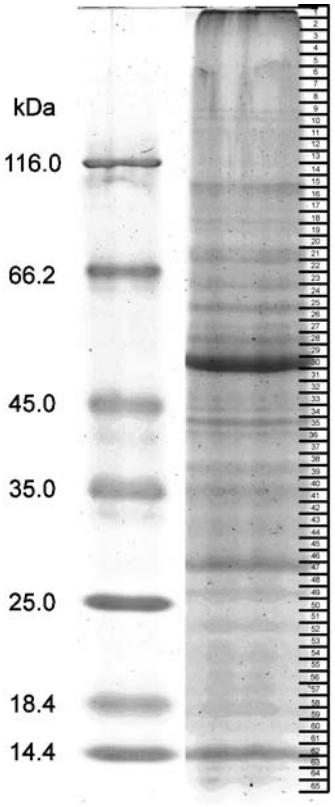

Fig. 2. 1D gel electrophoresis of proteins extracted from protoplasts in the presence of SDS.

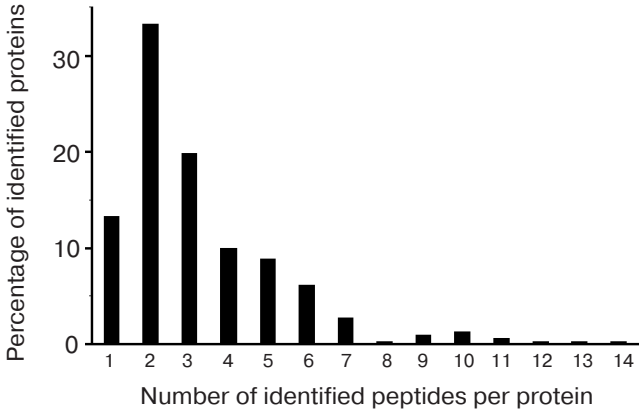

Fig. 3. Distribution of 228 proteins from protonema protoplasts determined by 1D PAGE LC-ESI-MS/MS according to the number of reliably determined peptides.

We fractionated proteins from the protoplast extracts by 1D electrophoresis under denaturing conditions in the presence of SDS in order to extend the spectrum of analyzed proteins (Fig. 2). After the protein fractionation, the gel was cut into separate bands and treated with trypsin. The peptide mixtures, which were prepared by the enzymatic cleavage by trypsin in the gels, were extracted and analyzed by liquid chromatography (LC) combined with the tandem mass spectrometer (MS/MS). We detected 228 proteins by this method, and 95% of them were determined according to two and more peptides (Fig. 3). The combination of liquid chromatography and tandem mass spectrometry revealed 186 proteins that were not found by 2D PAGE MALDI (Fig. 4). Proteins of secretion and intracellular transport, biosynthesis of secondary metabolites, transport, catabolism, biogenesis of membranes and cellular wall, metabolism and transport of coenzymes and nucleotides, control of cellular cycle, and division of cell were found among the determined proteins (Fig. 5). It should be noted that the methodology described above revealed a number of proteins with isoelectric points higher than 10, which were not identified by 2D PAGE MALDI (Fig. 6).

We classified the identified proteins based on isoelectric point and molecular mass (Figs. 6 and 7). These proteins had isoelectric points ranging from 4 to 12 and molecular masses from 7 to 210 kDa. It should be noted that fractionation of the proteins according to their isoelectric points had bimodal character (Fig. 6), which has recently been shown by Kiraga et al. to be universal for all living organisms, and the mode intensity correlated, in some cases, with both taxonomy and subcellular localization and with the econiche of an organism and the size of its proteome [18].

Direct comparison of the two proteomic methodologies (Fig. 4) demonstrated their good additivity: the number of proteins determined by both methods was 41 (10% of the total number of the determined proteins).

As mentioned above, proteins determined by LC-ESI-MS/MS were identified using two databases: database of annotated nuclear genome sequences (JGI) and NCBI database, because a number of proteins that were

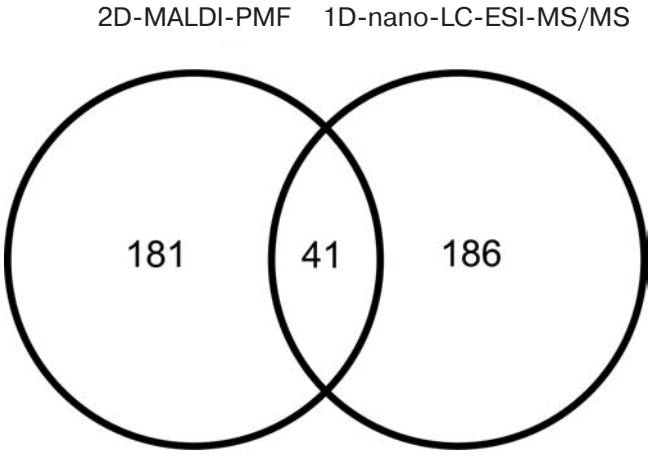

Fig. 4. Distribution of proteins determined by the two proteomic approaches (Venn diagram).

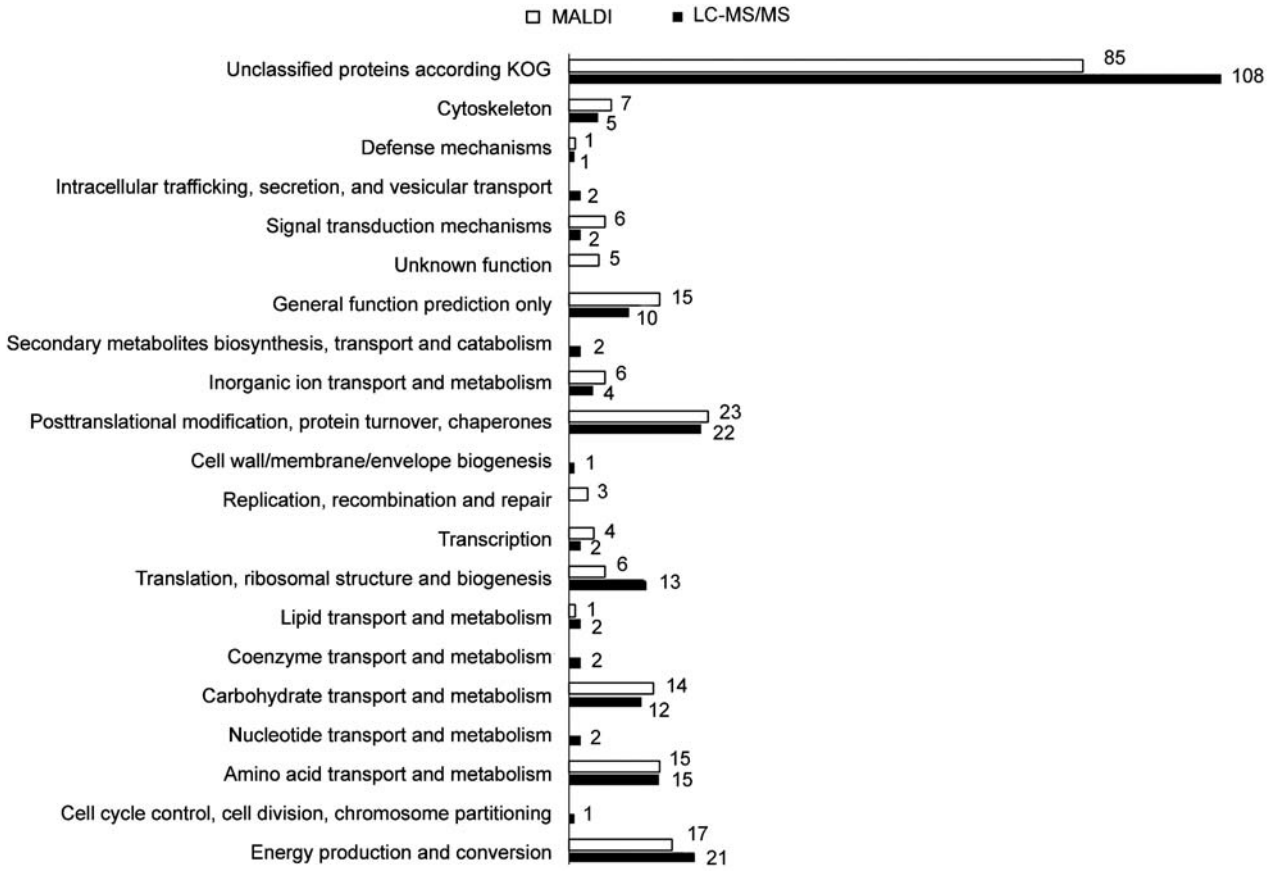

Fig. 5. Attribution of proteins from total extracts of protoplasts from *P. patens* to subclasses of the KOG system.

coded by the chloroplast or mitochondrial (but not by nuclear) genome would not be determined with the use of only one genome database. Indeed, comparison of the two data arrays revealed considerable overlapping between the identified proteins (181 proteins, 57%) (Fig. 8), but 12 proteins (3.8%) were determined only in the NCBI database.

We applied classification according to the KOG criteria (EuKaryotic Orthologous Groups, http://genome.jgi-psf.org/Chlre3/tutorial/kog.html) [19] to the proteins isolated from the protoplasts and identified by the two proteomic strategies. The identified proteins were divided into four classes and 25 subclasses (Figs. 9 and 10).

We identified 56% of the total number of the detected proteins in an automatic regime based on the JGI database of the moss using the MASCOT program complex and KOG classification. It was established that 16, 26, and 7% of the classified proteins participated in various cellular and transduction processes, processes of metabolism, and processing and storage of genetic information, respectively. Seven percent of the classified proteins were proteins with unknown function or proteins poorly characterized according to KOG.

A considerable group composed of 193 proteins (44% of the total number of determined proteins) is distinguished from the proteins identified by the two mass spectrometry approaches because it cannot be classified by the methodology of clusterization of the orthologous groups of proteins adapted for the eukaryotic genome. They can

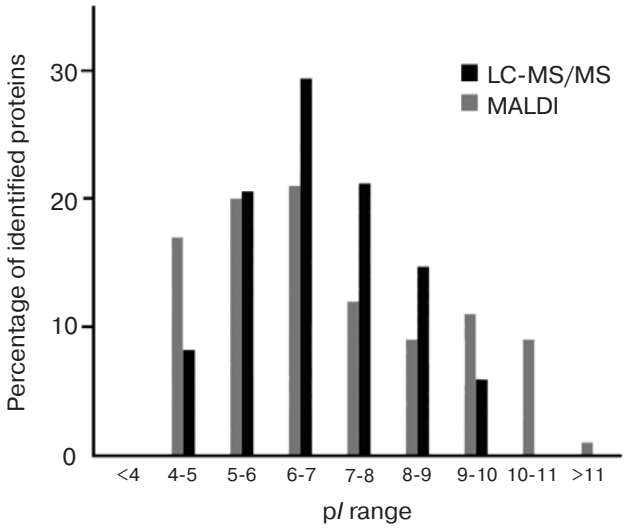

Fig. 6. Distribution of proteins from the moss protoplasts according to their isoelectric points.

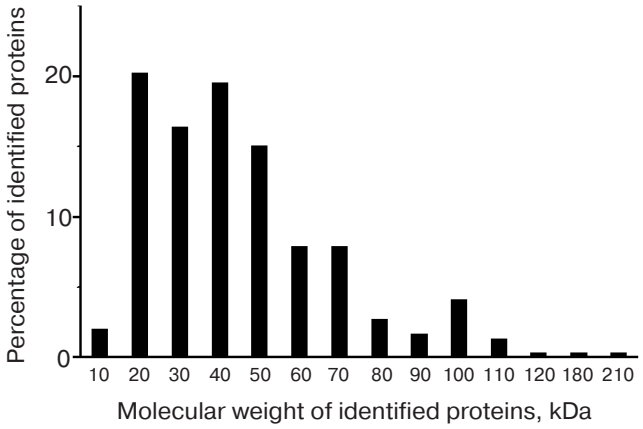

Fig. 7. Distribution of proteins from the moss protoplasts determined by 1D PAGE LC-ESI-MS/MS according to their molecular masses.

be attributed to none of the four KOG classes. This fact may be associated with both originality of these protein sequences of *P. patens* and with possible mistakes of the algorithms of automatic annotation of protein sequences.

Among the proteins annotated in accordance with KOG, 22% are associated with metabolism and post-translational modification. The proteins of cytoskeleton (9%) are mainly represented by actin and associated proteins, α- and β-tubulins, myosin, and others. Among the proteins identified by us, 5, 4, 3, and 1% are related to the systems of replication, reparation, and recombination, to the translation system, to RNA modification and processing, and to transcription, respectively. The identified proteins that participate in metabolism are divided into five classes: proteins of energy metabolism (12%), proteins associated with metabolism and transport of amino acids (13%), sugars (3%), coenzymes (3%), and inorganic ions (4%). The proteins with predicted function are separated as a subgroup (11%) and annotated on the basis of information about conservative functional domains with insufficient homology level that is established by the BLAST algorithm. The proteins with unknown function (3%) are isolated in a separate group. The remaining groups involve proteins associated with intracellular and vesicle transport (3%), transduction mechanisms, and possible protection mechanisms (1%).

We compared results of 2D electrophoregrams of proteins from protoplasts and protonema using specific fluorescent dyes for identification of proteins specific to different living forms of *P. patens* (Fig. 11, see color insert). The fluorescent electrophoregrams demonstrate difference in the protein compositions of protoplasts and protonema. Figure 11 demonstrates spots that are identified in the protein extracts from protoplasts and absent in those of protonema. On the other hand, spots of proteins that are specific only for protonema and absent in protoplast extracts are established.

Information on proteins fractionated by 2D gel electrophoresis and specific for protoplasts or protonema are given in the table. One can assume that appearance of new spots on the electrophoregrams of proteins from the protoplasts in comparison with those from the protonema extracts is associated with biochemical changes in the protonema cells in the course of isolation of the protoplasts from protonema filaments. The age of protoplasts by the moment of protein isolation from them in our study was 1 h. Thus, the changes in the moss protein spectrum during separation of the protoplasts occur rather fast. This suggests that changes that proceed during dozens of seconds in the protonema cells, which are converted into the protoplasts, are so significant that the protein spectrum detected on a 2D electrophoregram is changed. It is interesting that nine protein spots detected only on the protoplast protein electrophoregrams are identified by mass spectrometry as fragments of the large subunit of RuBisCo (P1, P2, P3, P7, P12, P13, P14, P15, and P62). Studies of RuBisCo in leaves of a number of angiosperm plants demonstrated that catabolism of this enzyme depended on endogenous factors and environmental influence that could cause stress reactions accompanied by protein fragmentation [20]. The treatment of protonema with Driselase (a cellulolytic agent) is apparently a stress process, which might initiate cellular reactions in the moss resulting in fragmentation of RuBisCo that is typical for stress reactions of green tissues of higher plants [20].

The spots corresponding to proteins that are not detected on the protoplast electrophoregrams were found on the electrophoregrams of proteins extracted from the protonema (Fig. 11, red fluorophore). This suggests that the protoplasts contain no proteins detected in the protonema, because there is a group of cells from which the protoplasts are predominantly isolated for the first time. It would be apical protonema cells according to our

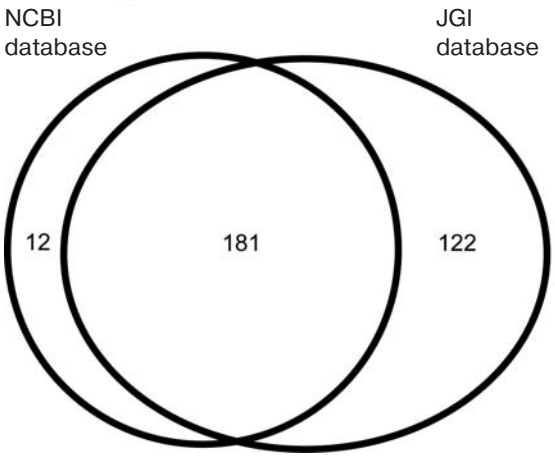

Fig. 8. Distribution of proteins determined in the two databases used for protein identification in this study (Venn diagram).

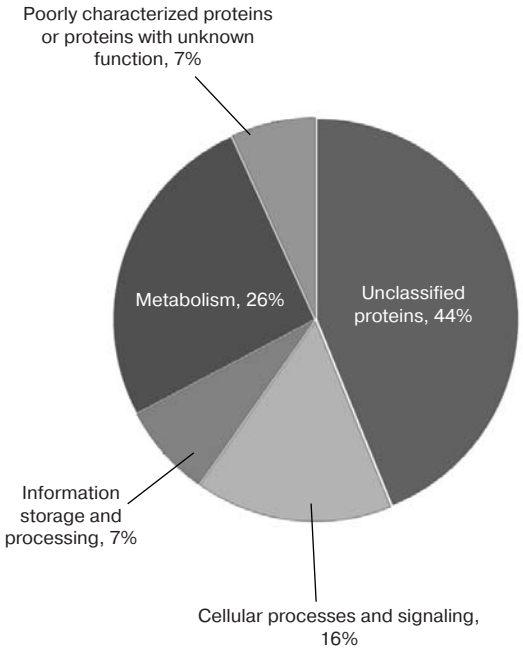

Fig. 9. Attribution of proteins from total extracts of protoplasts from *P. patens* moss to the classes of the KOG system.

microscopic observations during isolation of the protoplasts (the data is not given). It is possibly associated with the fact that they have less massive cellular walls on their tips [21], which might be degraded faster by Driselase. Thus, one can suppose that there is a number of basal protonema cells distant from growth points on tips of the filaments that are less subjected to the action of the enzyme for the isolation of protoplasts and have different protein composition in comparison with the apical cells and protoplasts. On the other hand, the absence of some proteins in the electrophoregrams of protoplasts in comparison with those of the protonema can be explained by fast elimination of these proteins during the process of protoplast isolation from the protonema cells.

Comparison of two main life forms of the moss (protonema and cormophyte gametophores) reveals significant similarity of their protein profiles (the data of fluorescent and electrophoretic studies are not given). Earlier, Cho et al. [7] found moderate differences in the protein profiles of protonema and gametophores of *P. patens*.

The proteins identified in protoplasts, protonema, and gametophores of the moss were additionally classified and attributed to groups and subgroups according to their functions (Table 1 from on-line publication and Figs. 9 and 10). Analysis of tables (Table 1 from on-line publica-

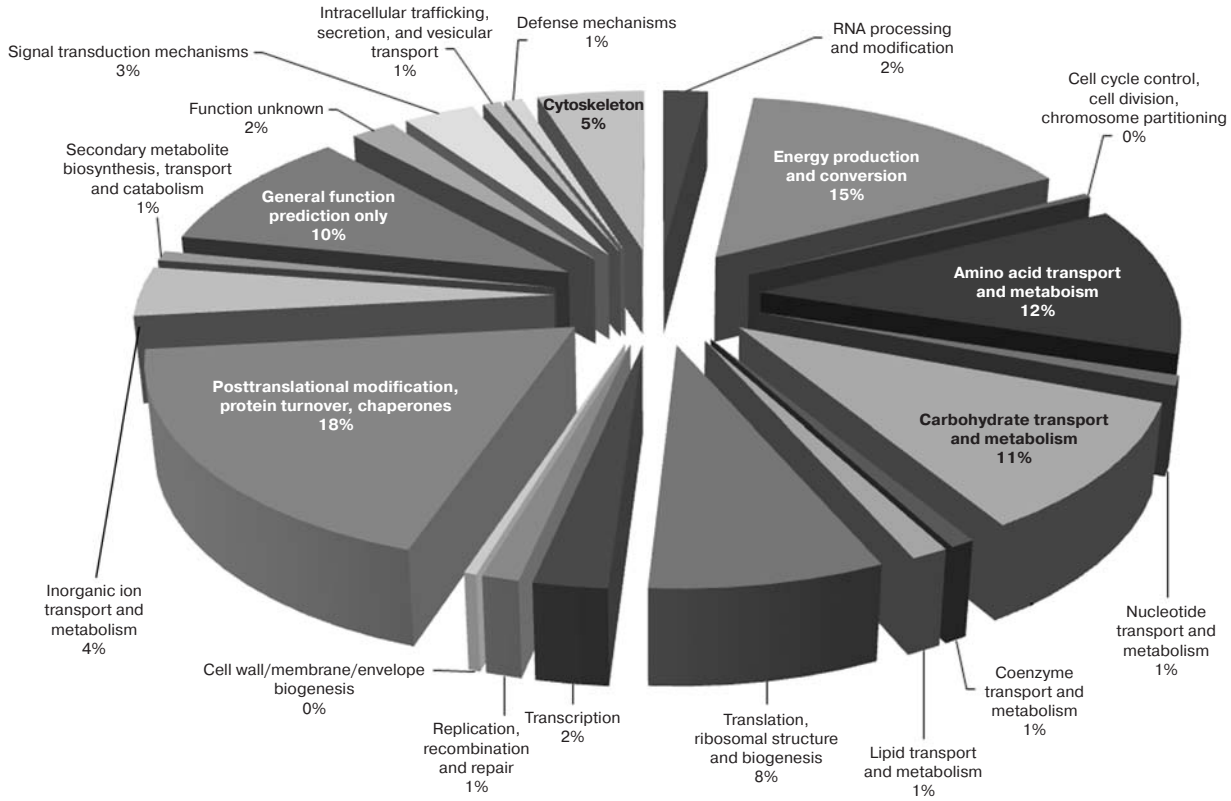

Fig. 10. Attribution of proteins from total extracts of protoplasts from *P. patens* moss to the subclasses of the KOG system.

Proteins unique for protoplasts (P) and protonemata (F)*

Spot	Protein ID	Protein details	MOWSE score	Molecular weight, Da	Database
P1	gi|34501407	RuBisCo, large subunit	60	52672	NCBI
P2	gi|34501408	–"–	62	52672	–"–
P3	gi|34501409	–"–	65	52672	–"–
P7	gi|34501409	–"–	86	52672	–"–
P10	95041 (gi|162668435) and 206166 (gi|162690347)	tyrosine kinase and ferric reductase	128	–	JGI NCBI
P11	191609	nucleoside diphosphate kinase 1	90	16851	JGI
P12	82369	hypothetical protein	63	80670	–"–
P12	gi|34501409	RuBisCo, large subunit	92	52672	NCBI
P13	gi|34501409	–"–	133	52672	–"–
P14	gi|34501409	–"–	126	52672	–"–
P15	gi|34501409	–"–	107	52672	–"–
P16	169825	protein containing bombesin-like peptide sequence	58	129075	JGI
P17	147634	cyclin H-1	66	35691	–"–
P20	162454	carbonic anhydrase	87	27590	–"–
P21	162316	kinesin-like protein	66	308692	–"–
F27	232987	hypothetical protein	59	21797	–"–
F28	35924	subunit of oxygen evolving system of photosystem II	74	22962	–"–
F29	144819	peroxidase	85	35199	–"–
F43	138701	aminopeptidase	184	99508	–"–
F45	108343	ATP-dependent Clp protease	234	101634	–"–
F46	108343	–"–	197	101634	–"–
F47	106487	–"–	98	92403	–"–
P50	160688	hypothetical protein	62	23265	–"–
P50	87196	–"–	59	24130	–"–
P53	111647	protein kinase	62	66507	–"–
P53	206554	glucosamine-phosphate acetyltransferase	58	17123	–"–
P54	63006	superoxide dismutase SOD1	83	17300	–"–
P55	117253	Argonaute family protein	58	100572	–"–
P56	100348	hypothetical protein	65	20382	–"–
P58	230273	HSP70 family chaperon	80	70931	–"–
P62	144819	peroxidase	128	35199	JGI
P62	gi|34501409	RuBisCo, large subunit	107	52672	NCBI
P64	85735	intracellular trafficking protein SNAP33	63	32359	JGI
P65	232922	HSP70 family chaperon	64	68262	–"–

* Proteins in spots P4-P6, P8, P9, P18, and P19 determined with insufficient MOWSE score (data not shown).

tion and the table from this article) and Figs. 1, 2, 5, 9-11 shows that the methods for proteomic analysis in our study allow fixation of only highly expressed proteins in the cells and tissues of the moss. These limitations are primarily associated with the dynamic range of the mass spectrometry equipment and instrumental limits of sensitivity of the used methods. Nevertheless, significant differences in the protein composition of various life forms of *P. patens* were found in this study even by the strategy of 2D electrophoresis with subsequent determination of a protein by peptide mass fingerprinting, although this strategy is seldom used in structural proteomics due to relatively small number of determined proteins. In addition, this methodology allows quick comparison and determination of changes in the level of expression and/or protein composition of various samples and conclusions about possible posttranslational changes of a protein.

DISCUSSION

In recent years the moss *P. patens* has become a leading model object for studies of molecular mechanisms of plant development. Increased interest in the ancient plant is due to the fact that *P. patens* is a single plant and one of three eukaryotic systems (along with yeasts and mouse stem cells) that can be highly effectively investigated by methods of reverse genetics and by the technology of gene targeting (substitution for the nucleotide sequences in target genes of an organism by homologous recombination of DNA) [9, 22, 23]. The relatively large genome of *P. patens* was sequenced and became available in open databases for comparative genome-evolutionary studies practically just after the determination of complete genome sequences of rice, poplar, and three species of green algae [5].

Moss *P. patens* is widely used as a model system for biochemistry, developmental biology, and genetics and has become an important fundamental object of proteomics. Now two independent projects of proteomic mapping of *P. patens* have been performed in Germany [6] and in Korea [7]. The total sequencing of the moss genome was not completed when these studies had been conducted, and limited data on the protein amino acid sequences of *P. patens* were used (dozens and hundreds of proteins) in the NCBI, TREMBL, and SWISSPROT databases. EST databases had been used for protein identification in the first two projects of proteomic mapping of *P. patens*. The first proteomic mapping of the moss were conducted using the clustered library including 110,087 EST fragments. It was the authors' opinion that this library comprised at least 95% of the moss genes [6]. In both papers, proteins were precipitated from 10% trichloroacetic acid in acetone, isolated, and subjected to 2D electrophoresis for the proteomic analysis. In the first paper [6], gels after fractionation of the proteins in two directions were treated with Coomassie dye, and 790 spots were detected. Then 276 spots were cut from the gel for the subsequent trypsin treatment and mass spectrometric analysis of peptides (MALDI-TOF and LC-ESI-MS/MS). As a result, 306 proteins isolated from the protonema were treated using the MASCOT search system and the BLASTIX comparison program [6]. In the second paper [7], gametophores, which were more differentiated and morphologically complex leafy shoots, were investigated along with the protonema. A number of methodological approaches to the proteomic analysis of *P. patens* described in the first paper [6] were critically re-examined in this article. Several significant modifications were proposed: homogenization of the tissues in liquid nitrogen in a conventional mortar with a pestle instead of a homogenizer with metallic beads; presence of thiourea in the buffer for protein solubilization and isoelectrofocusing (the authors believed that thiourea improved the quality of electrophoresis); application of silver for visualization of proteins instead of Coomassie dye [7].

Thus, methodological modifications of the protocols for preparation of protein samples and 2D electrophoresis played an important role in the project. The authors thought that this project should be a basis for establishment of a reference proteomic map of *P. patens* [7]. We found that the proposed protocol [7] is actually simple enough and reliable for proteomic mapping of the moss and used it as a methodical basis for this study. In general, the comparative proteomic analysis of the cormophyte gametophores and the filamentary protonema performed in our study and in paper [7] demonstrate considerable similarity of the protein composition of both life forms of the moss.

The new task of this study is the proteomic analysis of the protoplasts along with protonema and gametophores. The protoplasts are important objects for biochemistry, biophysics, and developmental biology and are used as a transformation system in the course of growing of transgenic plants, including investigation of physiological processes by the methods of reverse genetics. Protoplasts, like apical cells of a moss protonema, are simultaneously receptors and sites of fixation of light and gravitation stimuli. This property of protoplast cultures of mosses is promising for photobiological, gravitational, and space studies [24, 25]. New unicellular model systems based on moss protoplast culture are characterized by physiological and genetic homogeneity, distinct physiological reactions to external stimuli. The perception of biophysical stimuli and growth responses occur in one cell that is available for investigation and direct observation [26, 27]. Regeneration of the moss protoplasts is mainly similar to the germination of spores [28] and significantly differs from the regeneration of protoplasts of seed plants. The latter usually proceeds through the callus phase that is hardly achieved for many plants under experimental conditions [29]. On the contrary, this phase is absent in the moss protoplasts. The process of division of the moss protoplasts is considered to be "true" regen-

eration that gives rise to a new organism and starts from differentiation of the moss protonema [30].

Comparative proteomic analysis of protonema and gametophores in our experiments demonstrated insignificant differences in the protein spectra of the cells of two leading life forms of the moss similarly the studies of Cho et al. [7]. However, physiology of development of the filamentary protonema (it is attributed to a "two-dimensional" phase of the moss development [21]) is probably significantly different from that of the gametophores whose morphology has a profound three-dimensional character. Thus, elaboration of methodological approaches of comparative proteomics for determination of differences in the protein profiles of protonema and gametophores is an important problem of physiology of development of *P. patens*.

It is interesting that considerable differences in protein spectra are known to be found during analysis of electrophoregrams of protoplasts in comparison with those of protonema. Specific proteins are detected in protoplasts but are not found in the protonema. On the other hand, we found proteins that were present in the protonema and absent in the protoplasts. These results are in good agreement with those of studies of moss protoplasts as objects of developmental biology. It was found that protoplasts, like spores, gave rise to development of a new organism from the first division [26, 28, 31, 32]. A cell rapidly (within seconds) changes its form in the course of isolation of protoplasts from protonema and is converted into a geometrically ideal sphere. We propose that a cascade of stress signal reactions is generated when a protoplast is completely separated from a cell wall. Such reactions might be similar to those accompanied by plasmolysis of plant cells caused by drought or salt. Changes in the protein profile of the protoplasts can partially be associated with the process of their separation itself. This process has probably a stress character for a protonema cell, because a cell wall is degraded by enzyme, and the protoplast has to go out of its mechanically firm polysaccharide envelope-case into hypertonic solution.

Therefore, we can assume that both elimination of the proteins specific for the protonema cells and induction of synthesis of new regeneration proteins necessary for cell reprogramming and development of a new organism occur at the moment of isolation of intact protoplasts of *P. patens*. We observe change in the proteomic profile of a single plant cell that is associated with its transition into meristem state for the first time. Proteomic mapping of the moss protoplasts performed in this study opens up new approaches to studies of molecular basis of regeneration activity and totipotency of a plant cell.

REFERENCES

1. Rossignol, M., Peltier, J. B., Mock, H. P., Matros, A., Maldonado, A. M., and Jorrin, J. V. (2006) *Proteomics*, **6**, 5529-5548.
2. Espagne, C., Martinez, A., Valot, B., Meinnel, T., and Giglione, C. (2007) *Proteomics*, **7**, 3788-3799.
3. Lloyd, C. (1995) *Curr. Biol.*, **5**, 1085-1087.
4. Batygina, T. B. (1999) *Russ. J. Plant Physiol.*, **46**, 774-788.
5. Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P. F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., et al. (2008) *Science*, **319**, 64-69.
6. Sarnighausen, E., Wurtz, V., Heintz, D., van Dorsselaer, A., and Reski, R. (2004) *Phytochemistry*, **65**, 1589-1607.
7. Cho, S. H., Hoang, Q. T., Kim, Y. Y., Shin, H. Y., Ok, S. H., Bae, J. M., and Shin, J. S. (2006) *Plant Cell Rep.*, **25**, 475-488.
8. Ashton, N. W., and Cove, D. J. (1977) *Mol. Gen. Genet.*, **154**, 87-95.
9. Trouiller, B., Schaefer, D. G., Charlot, F., and Nogue, F. (2006) *Nucleic Acids Res.*, **34**, 232-242.
10. Grimsley, N. H., Ashton, N. W., and Cove, D. J. (1977) *Mol. Gen. Genet.*, **154**, 97-100.
11. O'Farrel, P. H. (1975) *J. Biol. Chem.*, **250**, 4007-4021.
12. Blum, H., Beier, H., and Gross, H. J. (1987) *Electrophoresis*, **8**, 93-99.
13. Laemmli, U. K. (1970) *Nature*, **227**, 680-685.
14. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) *Nature Protocols*, **1**, 2856-2860.
15. Savizky, A., and Golay, M. J. E. (1964) *Anal. Chem.*, **36**, 1627-1639.
16. Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2001) *Introduction to Algorithms*, McGraw-Hill, Columbus.
17. Bradford, M. M. (1976) *Anal. Biochem.*, **72**, 248-254.
18. Kiraga, J., Mackiewicz, K., Mackiewicz, D., Kowalczuk, M., Biecek, P., Polak, N., Smolarczyk, K., Dudek, M. R., and Cebrat, S. (2007) *BMC Genomics*, **8**, 163.
19. Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997) *Science*, **278**, 631-637.
20. Feller, U., Anders, I., and Mae, T. (2008) *J. Exp. Bot.*, **59**, 1615-1624.
21. Cove, D., Bezanilla, M., Harries, P., and Quatrano, R. (2006) *Annu. Rev. Plant Biol.*, **57**, 497-520.
22. Schaefer, D. G. (2001) *Curr. Opin. Plant Biol.*, **4**, 143-150.
23. Mittmann, F., Brucker, G., Zeidler, M., Repp, A., Abts, T., and Hartmann, E. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 13939-13944.
24. Sytnik, K. M., and Kordyum, E. L. (1994) in *Modern Problems of Space Sell Phytobiology* (Ugolev, A. M., ed.) [in Russian], Nauka, Moscow, pp. 243-249.
25. Wagner, T. A., and Sack, F. D. (1998) *Planta*, **205**, 352-358.
26. Cove, D. J., Quatrano, R. S., and Hartmann, E. (1996) *Development*, **122**, 371-379.
27. Kern, V. D., and Sack, F. D. (1999) *Planta*, **209**, 299-307.
28. Hartmann, E., and Jenkins, G. I. (1984) in *The Experimental Biology of Bryophytes* (Dyer, A. F., and Duckett, J. D., eds.) Academic Press, New York-London, pp. 203-228.
29. Siminis, C. I., Kanellis, A. K., and Roubelakis-Angelakis, K. A. (1994) *Plant Physiol.*, **105**, 1375-1383.
30. Hohe, A., and Reski, R. (2005) *Plant Cell Rep.*, **23**, 513-521.
31. Hartmann, E., and Weber, M. (1990) in *Bryophyte Development: Physiology and Biochemistry* (Chopra, R. N., and Bhatla, S. C., eds.) CRC Press, Boca Raton, Ann Arbor, Boston, pp. 33-54.
32. Schaefer, D., and Zryd, J.-P. (2001) *Plant Physiol.*, **127**, 1430-1438.