



Comprehensive nuclear proteome of *Arabidopsis* obtained by sequential extraction

Chieko Goto^a, Shoko Hashizume^b, Yoichiro Fukao^c, Ikuko Hara-Nishimura ^d, and Kentaro Tamura ^e

^aGraduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo, Japan; ^bDepartment of Botany, Graduate School of Science, Kyoto University, Kyoto, Japan; ^cDepartment of Bioinformatics, College of Life Sciences, Ritsumeikan University, Shiga, Japan; ^dFaculty of Science and Engineering, Konan University, Kobe, Japan; ^eDepartment of Environmental and Life Sciences, University of Shizuoka, Shizuoka, Japan

ABSTRACT

In eukaryotes, the nucleus plays key roles in fundamental cellular processes, including DNA replication, chromatin maintenance, transcription, and translation. To better understand the functional diversity of nuclei, we developed a method for the comprehensive extraction of the nuclear proteome from *Arabidopsis*. We used a buffer with a high sucrose concentration to purify nuclei and then conducted solubility-based fractionation to increase proteome coverage. We identified 1539 proteins and two novel nuclear envelope (NE) proteins in the nuclear fraction of *Arabidopsis* cultured cells. The localization of 25 proteins was determined by GFP fusion analyses; 23 of these proteins were localized either in the nucleus or the NE-associated endoplasmic reticulum. This result was indicative of the high quality of the proteome. These findings will be useful for clarifying novel nuclear functions in plants.

ARTICLE HISTORY

Received 30 November 2018
Revised 26 March 2019
Accepted 29 March 2019

KEYWORDS

Arabidopsis thaliana;
cultured cells; mass
spectrometry; nucleus;
nuclear bodies; nuclear
envelope; nucleolus;
proteome

Introduction

In plants and other organisms, the functions of the nucleus are crucial for cell proliferation and the regulation of gene expression during development and/or in response to biotic/abiotic stresses. Previous studies on vertebrates, yeasts, and protists have focused on different components of the nuclear proteome, including the whole nucleus [1–4], the nucleolus [5–7], the nuclear matrix [8,9], interchromatin granule clusters [10], and the nuclear envelope (NE) [11,12]. Studies on proteins and protein complexes in the nucleus have described the composition and roles of the nuclear pore complex (NPC) [13,14], the spliceosome complex, [15–18] the centromere complex[19], histones, [20] and intrinsically disordered proteins. [21] Other studies have focused on the nuclear proteome under specific conditions, for example, cell proliferation and differentiation[22], differentiation[23], embryonic development[24], organ development, [25] the DNA damage response[26], apoptosis[27], and viral infection. [28–30]

Several groups have conducted nuclear proteomics analyses for various plants [31–35]. In *Arabidopsis* cultured cells, Calikowski et al. identified 36 proteins

from nuclear matrix, demonstrating similarities in protein composition of the nuclear matrix across kingdoms[36]. Bigeard et al. identified a total of 879 proteins, of which 198 were phosphoproteins, in the chromatin-associated fraction from *Arabidopsis*[37]. To identify the stress-regulated nuclear proteins, it was reported the comparison of nuclear proteomes in response to cold stress [38] or MAMP-triggered immunity[39]. Whereas there are good methods for the extraction and analysis of proteomes from animals, there is still much room to improve methods to extract complete, high-quality proteomes from plants. The largest plant nuclear proteome to date is that of barley with more than 2,400 proteins (deposited in the database; <http://barley.gambrinus.ueb.cas.cz/>)[40]. Subsequently, more than 800 proteins was identified in rice [41,42] and *Arabidopsis*[37].

In this study, to improve the quality and quantity of the plant nuclear proteome, we used a buffer with a high concentration of sucrose to purify nuclei and then conducted solubility-based fractionation to increase proteome coverage [43,44]. We identified 1539 different proteins from the nuclei of *Arabidopsis* cultured cells and two novel NE proteins.

CONTACT Kentaro Tamura  tamura@u-shizuoka-ken.ac.jp  Department of Environmental and Life Sciences, University of Shizuoka, Shizuoka 422-8526, Japan
 Supplementary data for this article can be accessed [here](#).

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Our approach couples the isolation of pure organelles with solubility-based fractionation to allow comprehensive and efficient identification of the organelle proteome.

Materials and methods

Plant materials

Arabidopsis thaliana (Columbia-0) was used as the wild type. *Arabidopsis* seeds were germinated on Murashige and Skoog (MS) medium and grown at 22°C under continuous light (35 $\mu\text{mol m}^{-2} \text{s}^{-1}$). *Arabidopsis* cultured cells [45] were subcultured in the MS medium containing 2,4-dichlorophenoxyacetic acid at 23°C with continuous agitation in the dark. Tobacco (BY-2) cultured cells were maintained as described previously [46].

Nuclear isolation from *Arabidopsis* cultured cells

Arabidopsis cultured cells (4–5 day old) from 50–100 ml culture were collected and were suspended in 50 ml Cellulase enzyme solution (1% Cellulase ONOZUKA RS, 0.1% pectolyase Y-23, 0.4 M mannitol, 23.4 mM MES-KOH [pH5.7]). The cells were incubated in the Cellulase enzyme solution with shaking (55–65 rpm) at 28°C in dark for 3–4 hours to obtain protoplasts. The protoplasts were filtrated through a 125 μm nylon filter (NIPPON RIKAGAKU KIKAI) and then centrifuged at 1,000 $\times g$ for 3 minutes. The pellet of protoplasts was suspended in 15 ml Honda buffer (0.44 M sucrose, 2.5% ficoll, 5.0% dextran40, 25 mM Tris-HCl [pH8.0], 10 mM MgCl_2 , 3 mM CaCl_2 , 0.1% Triton X-100, 2.5 mM DTT, 1/50 ml complete EDTA-free) and homogenized with a glass homogenizer. The homogenate was centrifuged at 1,500 $\times g$ for 10 minutes at 4°C, and the pellet was homogenized in 15 ml Honda buffer. This step was repeated once. The homogenate was filtrated through 80 μm , 30 μm , and 20 μm nylon filters to remove cell debris. The solution passed through the filters was centrifuged at 1500 $\times g$ for 10 minutes at 4°C. The pellet was re-suspended in 10 ml 2.3 M sucrose buffer (2.3 M sucrose, 3 mM CaCl_2 , 10 mM

MgCl_2 , 25 mM Tris-HCl [pH8.0], 0.1% TritonX-100) and centrifuged at 65,000 $\times g$ for 45 minutes at 4°C. The pellet was re-suspended in Honda buffer and used as ‘isolated nuclei’. Immunostaining and an immunoblot analysis were performed as described previously (anti-histone 1:500, anti-BiP 1:10,000, anti-Nup43 1:1000, anti-PDI 1:2,000) [47]. The signal intensity of each band was semi-quantified with a Fiji software (<http://fiji.sc>).

Solubility-based fractionation of nuclear proteins

The isolated nuclei mentioned above were treated with DNase/RNase, and then suspended in a salt solution (1 M NaCl, 25 mM MES [pH5.6], 5 mM MgCl_2 , 10 mM KCl, 0.35 M sucrose, 30% glycerol), followed by incubation for 10 minutes on ice with sometimes stirring. After incubation, the nuclei were centrifuged at 3500 $\times g$ for 10 minutes. The supernatant was used as salt fraction. The pellet was suspended in an alkaline solution (0.1 M Na_2CO_3 , 25 mM MES [pH5.6], 5 mM MgCl_2 , 10 mM KCl, 0.35 M sucrose, 30% glycerol), followed by incubation for 10 minutes on ice with sometimes stirring. After incubation, the suspension was centrifuged at 3500 $\times g$ for 10 minutes. The supernatant is the Alkaline fraction. The pellet was suspended in Triton X-100 solution (1% Triton X-100, 25 mM MES [pH5.6], 5 mM MgCl_2 , 10 mM KCl, 0.35 M sucrose, 30% glycerol), followed by incubation for 10 minutes on ice with sometimes stirring. After incubation, the suspension was centrifuged at 3500 $\times g$ for 10 minutes. The supernatant is the Triton fraction. The pellet was suspended in Empigen BB solution (0.3% Empigen BB, 25 mM MES [pH5.6], 5 mM MgCl_2 , 10 mM KCl, 0.35 M sucrose, 30% glycerol), followed by incubation for 10 minutes on ice with sometimes stirring. After incubation, the suspension was centrifuged at 3500 $\times g$ for 10 minutes. The supernatant is the Empigen fraction. The pellet was suspended in SDS solution (4% SDS, 25 mM MES [pH5.6], 5 mM MgCl_2 , 10 mM KCl, 0.35 M sucrose, 30% glycerol), followed by incubation for 10 minutes at room temperature with sometimes stirring. After incubation, the suspension was centrifuged at 3500 $\times g$ for 10 minutes. The supernatant is the SDS fraction. The remained pellet,

which was not solubilized by the procedure mentioned above, was solubilized in 100% formic acid (the formic acid fraction).

Proteomics and bioinformatics

The proteins in each fraction were precipitated with cold acetone and subjected into in-solution digestion as described previously to obtain the peptides[48]. The peptides were analyzed by liquid chromatography–tandem mass spectrometry using LTQ-Orbitrap XL (Thermo Scientific) as described previously[48]. The Identified protein were analyzed with DAVID (<https://david.ncifcrf.gov>) to annotate GO terms (cellular components). The domain structures were predicted with ATTED-II (<http://atted.jp>), COILS (https://embnet.vital-it.ch/software/COILS_form.html), and SMART (<http://smart.embl-heidelberg.de>).

Plasmid construction

Plasmid DNA construction was performed as described previously[48]. Briefly, to construct the genes for localization analysis of GFP/YFP-tagged proteins by the transient expression, a cDNA fragment encompassing the entire coding sequence of each gene except for a stop codon was amplified from cDNA from flowers of wild-type plants by PCR and cloned into pENTR1A (Invitrogen), and then fused downstream/upstream of sGFP or eYFP tag and downstream of the constitutive cauliflower mosaic virus promoter 35S in plant transformation vectors (pUGW6 or pUGW5). To construct Pro35S:GFP-At1g07970, Pro35S:At1g07970-GFP, and Pro35S:At3g08870-GFP, a genomic fragment containing the entire coding sequence of At1g07970 or At3g08870 without a stop codon was cloned into pENTR1A (Invitrogen) and then fused downstream/upstream of sGFP tag and downstream of the promoter 35S in a plant transformation vector (pGWB406 or pGWB405). The primers used are shown in Supplemental Table S4.

Transient and stable expression of fusion proteins

For transient expression, protoplasts of *Arabidopsis* cultured cells were transformed with PEG-mediated

method[49]. For stable expression, *Arabidopsis* wild-type plants and tobacco BY-2 cells were transformed as described previously [46,50] with *Agrobacterium tumefaciens* (GV3101).

Results and discussion

Isolation of nuclei from *Arabidopsis* cultured cells

To isolate sufficient quality of nuclei, *Arabidopsis* cultured cells were used as the starting materials. The nuclear isolation method was based on a solution

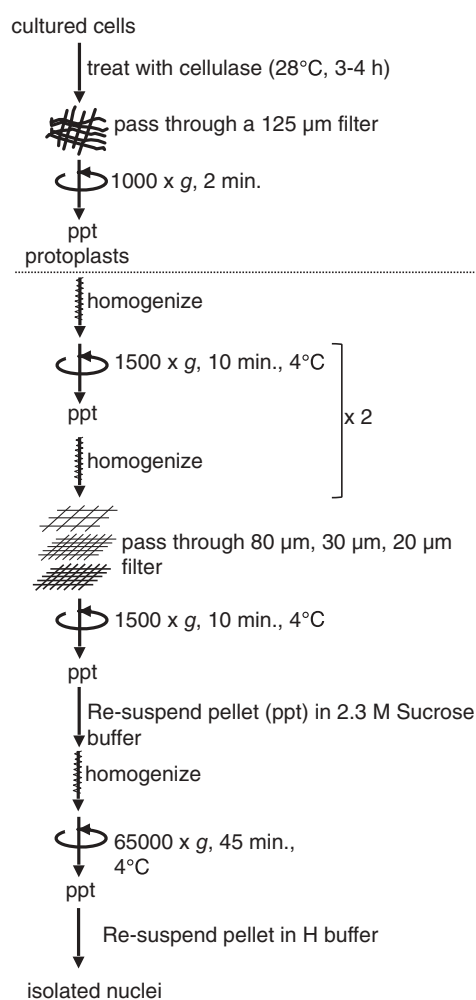


Figure 1. Nuclear isolation procedure used in this study.

Arabidopsis cultured cells were treated with cellulase solution to obtain protoplasts by degrading cell wall. Protoplasts were passed through a filter, collected by centrifugation, homogenized, and centrifuged again, and then passed through filters. Obtained cell lysate was re-suspended in sucrose buffer, homogenized, and centrifuged to concentrate nuclei. Pellet containing nuclei was re-suspended in Honda (H) buffer. See Materials and methods for details.

with a high concentration of sucrose (Figure 1) instead of the Percoll gradient commonly used in previous nuclear proteomics studies [36,38,51]. After isolating cultured protoplasts, a rough nuclear fraction was obtained by a combination of filtration and low-speed centrifugation. Next, we resuspended the fraction in 2.3 M sucrose buffer, and then subjected the mixture to ultracentrifugation to purify the nuclei. The purified nuclei were characterized by fluorescence microscopy after staining with 4',6-diamidino-2-phenylindole (DAPI), which indicated successful isolation of nuclei (Figure 2A). Staining with the lipophilic dye FM4-64 confirmed the intact membrane structure of the isolated nuclei (Figure 2B). Immunofluorescent staining detected both Nup43 and RAE1, which are components of the NPC [48,52], in the nuclear rim (Figure 2C). These results confirmed that the isolated nuclei had intact NEs with NPCs.

We then evaluated the purity of the isolated nuclei by immunoblot analyses with antibodies against histone (a nuclear marker), Nup43 (an NE marker), BiP (an ER luminal marker) [53], and AtALEU (a soluble vacuolar marker) [54]. The histone and Nup43 proteins were significantly concentrated in the isolated nuclear fraction (N) compared with the whole cell lysate fraction (W) (Figure 2D and Figure S1). The ER luminal proteins, BiP and PDI, were detected in both fractions, indicative of direct continuity between the lumen of the NE and the ER. The relative signal intensities of BiP and PDI in N to W are 0.77 and 0.57, respectively, suggesting that ER proteins were not concentrated in the nuclear fraction. These results confirmed that nuclei were highly purified by our isolation method.

Mass spectrometric analyses of nuclear proteome

To increase nuclear proteome coverage, we conducted solubility-based fractionation using sequential extraction with salt, alkaline, detergent, and a strong acid buffer (Figure 3A). The first step was digestion with DNase and RNase to solubilize intranuclear nucleic acids. The salt and alkaline buffers were used to extract the most soluble and peripheral proteins. To extract membrane proteins effectively, we used three types of detergent; Triton X-100, Empigen BB, and sodium dodecyl sulfate (SDS). We attempted to extract insoluble nuclear proteins, which potentially include those required for the formation of functional

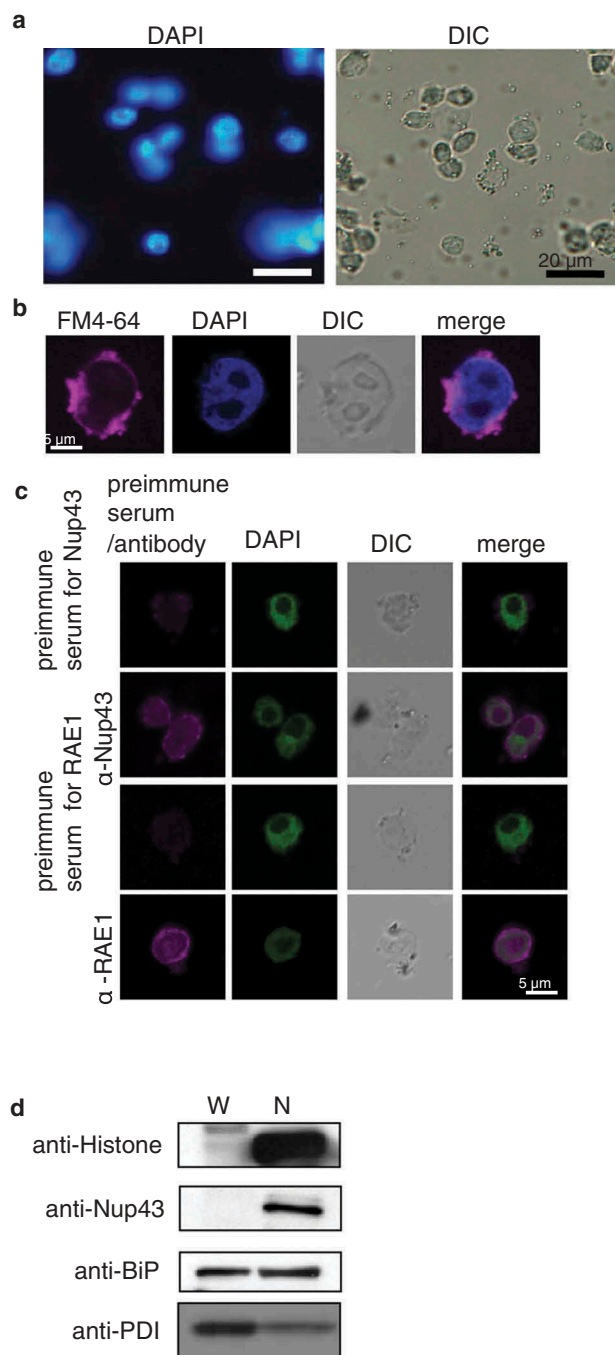


Figure 2. Isolated nuclei.

(A) Nuclei isolated by procedure shown in Figure 1 and stained with 4',6-diamidino-2-phenylindole (DAPI). (B) Isolated nuclei co-stained with DAPI and FM4-64. (C) Immunostaining of isolated nuclei with preimmune serum (negative control) or antibodies against the nucleoporins Nup43 and RAE1. (D) Immunoblot analysis of whole cell lysate (W) and isolated nuclei (N) with anti-histone antibody (nuclear marker), anti-Nup43 antibody (nuclear envelope marker), anti-BiP antibody (ER luminal marker), and anti-AtALEU antibody (soluble vacuolar marker). DIC, differential interference contrast. Uncropped images of blots are shown in Figure S1.

intranuclear compartments [55], with a formic acid buffer. Each extract was characterized by shotgun

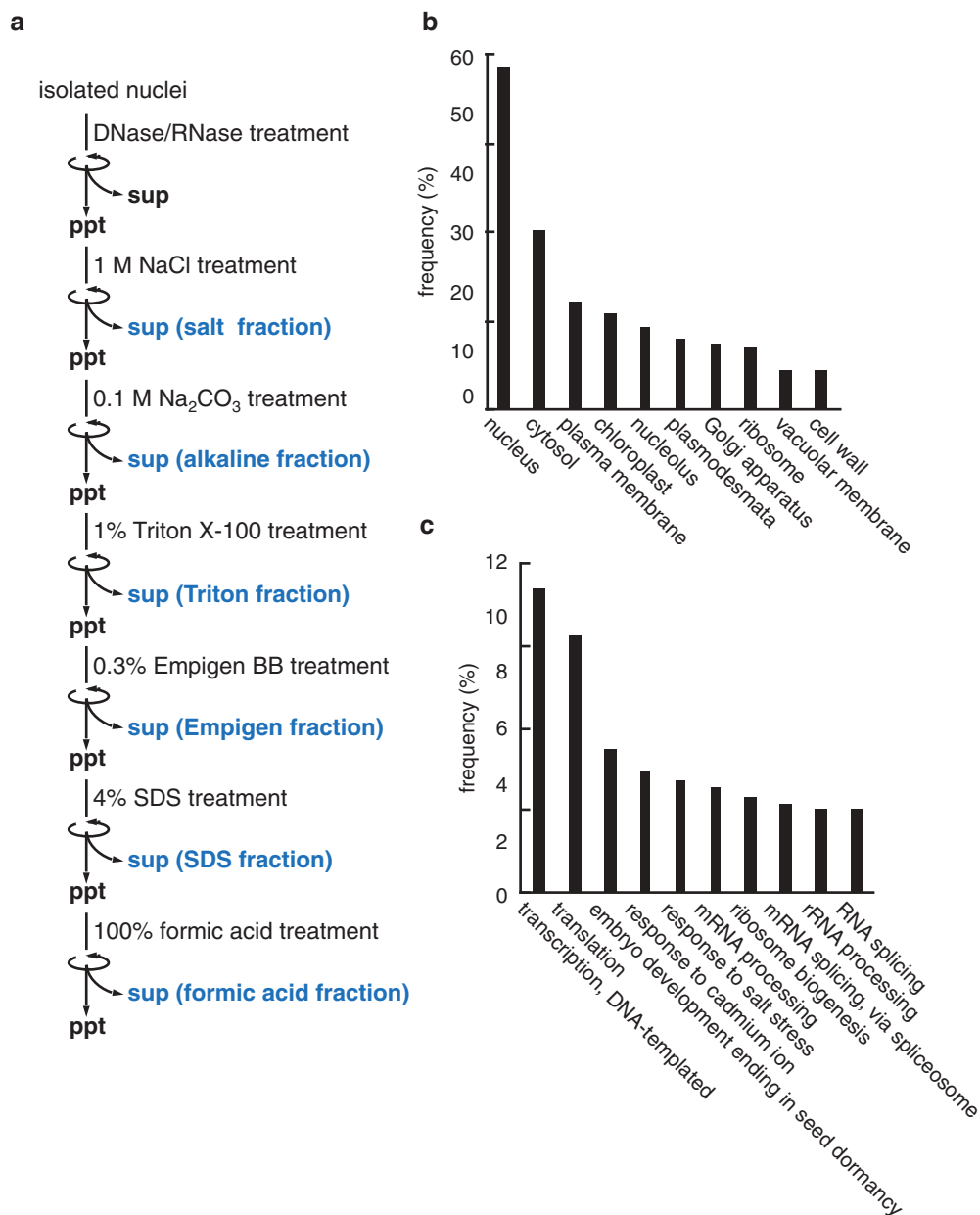


Figure 3. Analysis of nuclear proteome extracted from *Arabidopsis* cultured cells.

(a) Solubility-based fractionation of nuclear proteins. Proteins from isolated nuclei were sequentially fractionated, and proteins in each fraction were identified by mass spectrometry. (b) Distribution of 10 most abundant GO terms in cellular component category. (c) Distribution of 10 most abundant GO terms in biological process category.

proteomics using an LC-MS/MS (LTQ-Orbitrap) system. The mass spectrometry analyses identified 1539 different proteins (Table S1) (189 proteins in the salt fraction, 229 proteins in the alkaline fraction, 1271 proteins in the Triton X-100 fraction, 79 proteins in the Empigen fraction, 345 proteins in the SDS fraction, and 186 proteins in the formic acid fraction) (Figure 3A).

In gene ontology (GO) enrichment analyses, 1450 proteins were annotated in the cellular component

category. Among them, 58% were further classified into 'nucleus', 14.1% into 'nucleolus', and 6.6% into 'ribosomes' (Figure 3B). This result indicated that nuclear proteins were efficiently concentrated and identified using our preparation and analysis methods. In addition, 162 proteins were annotated in the biological process category (Table S3), suggesting a variety of nuclear functions in plants. Most of them were involved in RNA metabolism ('transcription', 10.2%; 'translation', 8.6%; 'mRNA processing', 3.5%), while

others were classified into ‘embryo development (4.7%)’, ‘response to cadmium ion (4.1%)’, and ‘response to salt stress (3.8%)’ (Figure 3C). These results suggested functional differentiation of the plant nucleus between the developmental and stress response pathways.

Nuclear envelope proteins identified in this study

The NE contains three distinct functional domains; the outer nuclear membrane (ONM), the inner nuclear membrane (INM), and the nuclear lamina. Proteins associated with the NE in plants have been characterized by several groups [56]. *Arabidopsis* has more than 30 nucleoporins, which are components of the NPC associated with the nuclear membrane [48,57]. We identified 14 nucleoporins in our proteomics analysis (Table 1). MAD1, which interacts with the NUA nucleoporin in interphase cells (Ding et al., 2012), was also detected in our proteomics analysis. In addition to nucleoporins, nine known NE proteins were identified (Table 2). The most

highly identified NE protein was RANGAP1. The NE localization of RANGAP1 depends on its interaction with the linker of nucleoskeleton and cytoskeleton (LINC) complex composed of WIPs, WITs, and SUNs in *Arabidopsis* [58,59]. In mammalian cells, the LINC complex spans both ONM and INM and associates with the lamina structure [60]. Our proteomics analysis identified four LINC complex constituents (WIT1, WIP3, SUN1, and SUN2) and three putative lamina components (CRWN1, CRWN4, and KAKU4) as NE proteins. These results indicated that our preparation method produced a nuclear proteome with comprehensive coverage of NE proteins.

In our nuclear proteome, we identified two novel NE proteins, At1g07970 and At3g08870, with putative transmembrane domains (Figure 4A). At1g07970 had two transmembrane domains at the N-terminal and a short coiled-coil region in the center of the protein. At3g08870 had a signal peptide, a lectin domain, a transmembrane domain, and a kinase domain. Stable expression of GFP-fused proteins revealed that they exclusively localized on the NE in tobacco BY-2 and *Arabidopsis* root cells (Figure 4B). Searches of publicly available microarray data revealed that At1g07970 is broadly expressed in vegetative and reproductive tissues while At3g08870 is highly expressed in senescent leaves. This result suggested that At3g08870 is involved in a specific function of the NE during the senescence process. This is the first report of a kinase-domain-containing protein localized on the NE in plants. Based on its primary structure, At3g08879 was predicted to be a type-I integral membrane protein with its kinase domain located at the cytoplasmic side. It would be interesting to explore

Table 1. nucleoporins identified by MS analysis in this study.

AGI code	Name	Reference
AT5G40480.1	gp210	Tamura et al. (2010)
AT2G41620.1	Nup93a	Tamura et al. (2010)
AT1G79280.1	NUA/TPR	Xu et al. (2007) Plant Cell; Tamura et al. (2010)
AT1G33410.1	Nup160	Muthuswamy and Meier (2011)
AT1G73240.1	NDC1	Boruc et al. (2012)
AT5G51200.1	Nup205	Tamura et al. (2010)
AT1G64350.1	Seh1	Tamura et al. (2010)
AT3G10650.1	Nup136	Tamura et al. (2010)
AT3G14120.1	Nup107	Tamura et al. (2010)
AT4G32910.1	Nup75	Tamura et al. (2010)
AT1G24310.1	Nup54	Tamura et al. (2010)
AT2G30050.1	Sec13	Tamura et al. (2010)
AT4G37130.1	Nup58	Tamura et al. (2010); Ferrández-Ayala et al. (2013)

Table 2. NE proteins identified by MS analysis in this study (except for nucleoporins).

AGI code	Name	Reference
AT3G63130.1	RANGAP	Rose and Meier (2001); Xu et al. (2007) Current Biology; Xu et al. (2008); Zhou et al. (2012)
AT1G67230.1	CRWN1	Dittmer et al. (2007); Sakamoto and Takagi (2013) (;Ciska et al. [2013]); Graumann (2014)
AT5G04990.1	SUN1	Graumann et al. (2010); Oda and Fukuda (2011); Graumann (2014); Graumann et al. (2014); Tatout et al. (2014); Zhou et al. (2015)
AT5G65770.1	CRWN4	Sakamoto and Takagi (2013)
AT4G31430.1	KAKU4	Goto et al. (2014)
AT5G11390.1	WIT1	Zhao et al. (2008); Zhou and Meier (2014); Zhou et al. (2015)
AT5G49880.1	MAD1/NES1	Ding et al. (2012)
AT3G10730.1	SUN2	Graumann et al. (2010); Oda and Fukuda (2011); Zhou et al. (2015)
AT3G13360.1	WIP3	Xu et al. (2007) Current Biology

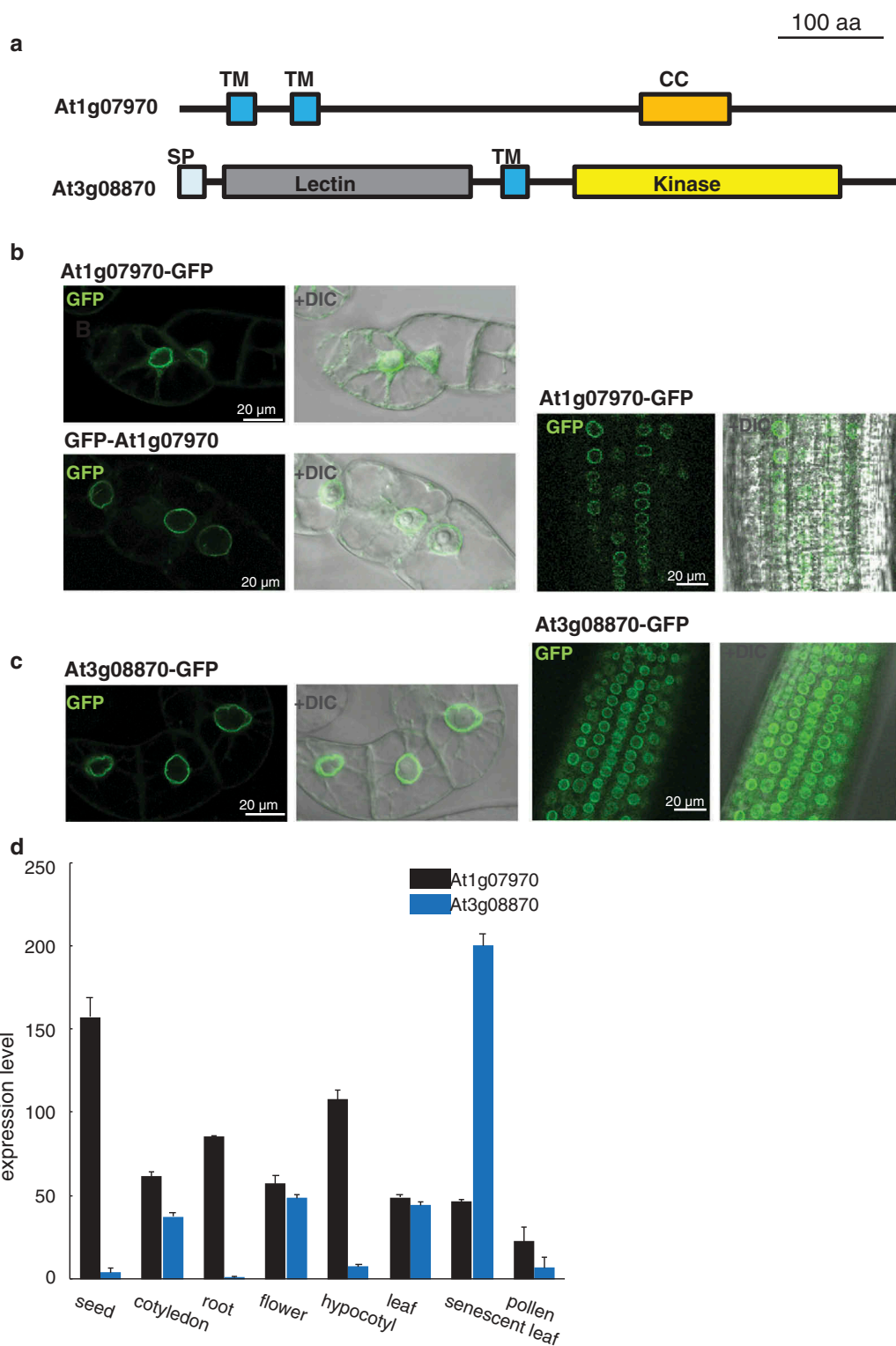


Figure 4. Identification of two novel nuclear envelope proteins.

(a) Schematic representation of At1g07970 and At3g08870 proteins. At1g07970 has putative two transmembrane domains (TM) and a coiled-coil domain (CC). At3g08870 protein has putative signal peptide (SP) followed by lectin domain, transmembrane domain (TM), and kinase domain. (b and c) Fluorescence images of tobacco BY-2 cells and *Arabidopsis* root cells stably expressing GFP-tagged At1g07970 (b) or At3g08870 (c). DIC, differential interference contrast. (d) Transcript levels of At1g07970 and At3g08870 in various tissues. Data were obtained from resource page of AtGenExpress project (<http://jsp.weigelworld.org/AtGenExpress/resources/>) [73]. Error bars indicate standard deviation.


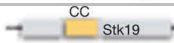

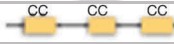
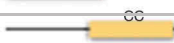






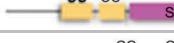









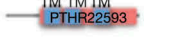

AGI code	Gene Name	Protein Structure	Localization*	Panel*	Ref.
At5g26210	Alfin-like 4 (AL4)		nucleus	a	70,62
At2g20495	-		nucleus	b	-
At5g53800	-		nucleus	c	-
At1g19980	-		nucleus	d	-
At1g61000	-		nucleus	e	-
At5g60030	-		nucleolus	f	-
At5g57120	-		nucleolus	g	-
At1g10510	EMBRYO DEFECTIVE 2004 (EMB2004)		nucleolus	h	-
At3g07050	DAR GTPase 4 (DGP4)/ NUCLEOSTEMIN-LIKE1 (NSN1)		nucleolus	i	71, 63, 72, 73, 74
At5g05210	POT1A		nucleolus	j	64
At5g10060	-		nucleus and cytoplasm	k	-
At5g10710	-		nucleus and cytoplasm	l	-
At5g65180	-		nucleus, cytoplasm and nucleolus	m	-
At1g61150	-		nucleus, cytoplasm and nucleolus	n	-
At1g22200	-		ER	o	-
At1g30130	-		ER	p	-
At1g34350	Arabidopsis Transmembrane Protein 18 (AtTMEM18)		ER	q	75
At1g56270	-		ER	r	-
At2g35760	-		ER	s	-
At5g50550	-		ER	t	-
At3g44330	-		ER	u	-
At4g03470	Ankyrin 4 (ANK4)		ER	v	76
At4g13270	-		ER	w	-
At5g60210	ROP INTERACTIVE PARTNER 5 (RIP5)		PM	x	77
At1g54200	-		PM	y	-

Figure 5. Schematic representation of predicted domain features of proteins characterized in this study. Localization analysis is shown in Figure S2. References are follows: 1[74]; 2[64]; 3[75]; 4[65]; 5[76]; 6[77], 7[78]; 8[66]; 9[79]; 10[80]; 11[81].

the role of kinase signaling on the NE in the senescence signaling pathway.

Localization analyses of proteins with coiled-coil or transmembrane domains

Nuclear coiled-coil proteins are involved in providing the nucleoskeleton and molecular scaffolds that organize membrane systems [61], while integral membrane proteins on the NE play multiple roles in shaping the nuclear membrane [62,63]. To identify novel proteins involved in the functional organization of the nucleus, we selected 25 proteins with either a coiled-coil domain (14 proteins) or a transmembrane domain (11 proteins; Figure 5) and analyzed their subcellular localization by transient expression in *Arabidopsis* cultured cells. Five proteins (At5g26210, At2g20495, At5g53800, At1g19980, and At1g61000) fused with GFP were localized in the nucleus. Among them, At5g26210 (AL4) has been reported to localize to the nucleus previously [64], while the other four proteins (At2g20495, At5g53800, At1g19980, and At1g61000) were newly characterized in this study (Figure S2). Five proteins (At5g60030, At5g57120, At1g10510, At3g07050, and At5g50210) fused with GFP were localized in the nucleolus exclusively. At3g07050 (Nucleostemin-like1: NSN1) has been reported to localize in the nucleolus and regulate the cell cycle [65], while At5g05210 (Protection of telomerase 1: POT1A) has been reported to localize both to the cytoplasm and to the nucleus [66]. POT1A (At5g05210) interacts with the nucleolar protein TERT-V (I8) [66], indicating that POT1A is able to localize the nucleolus under certain conditions. Four proteins (At5g10060, At5g10710, At5g65180, and At1g61150) showed diverse patterns of localization, including in the nucleus, nucleolus, and cytoplasm. It is possible that these proteins are trafficked between the nucleus and cytoplasm. Indeed, both At5g10710 and At5g65180 had an Epsin N-terminal homology (ENTH/VHS) domain, whose structure is similar to those of karyopherin and beta-katenin [67], which are known to be shuttled between the nucleus and the cytoplasm [68]. The GFP fusions of nine proteins with a transmembrane domain showed a typical NE and ER-localization pattern. Two coiled-coil proteins (At5g60210 and At1g54200) fused to GFP localized in dot-like structures at the PM. Consistently,

At5g60210 (RIP5) has been identified in another PM proteomics analysis [69]. Together, the results of these GFP-fusion localization analyses revealed that 23 out of 25 analyzed proteins were localized in either the nucleus or NE-associated ER. In the transient expression system, fusion proteins can be highly accumulated in the cell. Therefore, we could not exclude the possibility that the expressed fusion proteins were leaked from nuclear membrane to ER membrane. Stable expression system under control of own promoter will be further required for assessing exact protein localisation.

A successful proteomic analysis of complex biological samples is often hindered by high-abundance proteins [70]. Therefore, the ability to selectively deplete high-abundance proteins for efficient detection of minor proteins is increasingly important in proteomic studies [44]. Our sequential fractionation based on protein solubility is a powerful approach to reduce proteome redundancy after enrichment of not only nucleus but also other organelles. Indeed, for protein preparation in 2 dimensional electrophoresis, sequential extraction strategies are universally exploited, most with alternations in pH and ionic strength of extraction buffers [71]. Recently, Blavet et al. created a barley (*Hordeum vulgare* L., cv. Morex) nuclear protein database [40]. This database contains large numbers of proteins identified from nuclei at the G1, S, and G2 cell cycle phases, and follows on from their previous study reporting 803 proteins [72]. Their database and the results obtained in our study provide comprehensive information about plant nuclear proteins.

Abbreviations

ER	endoplasmic reticulum
GO	gene ontology
INM	inner nuclear membrane
ONM	outer nuclear membrane
NE	nuclear envelope
NPC	nuclear pore complex(es)
PM	plasma membrane

Acknowledgments

We thank Tsuyoshi Nakagawa (Shimane University, Japan) for donating the Gateway vectors. We also thank Jennifer Smith, PhD, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by a 'Specially Promoted Research' Grant-in-Aid for Scientific Research to I.H.-N. (22000014), and by a Grant-in-Aid for Scientific Research to K.T. (26711017, 25650096, and 18K06283).

ORCID

Ikuko Hara-Nishimura  <http://orcid.org/0000-0001-8814-1593>

Kentaro Tamura  <http://orcid.org/0000-0001-7387-2211>

References

- [1] Zhang Y, Fang C, Bao H, et al. Nuclear proteome profile of C57BL/6J mouse liver. *Sci China Life Sci.* 2013;56:513–523.
- [2] Jankowska U, Latosinska A, Skupien-Rabian B, et al. Optimized procedure of extraction, purification and proteomic analysis of nuclear proteins from mouse brain. *J Neurosci Methods.* 2016;261:1–9.
- [3] Henrich S, Cordwell SJ, Crossett B, et al. The nuclear proteome and DNA-binding fraction of human Raji lymphoma cells. *Biochim Biophys Acta.* 2007;1774:413–432.
- [4] Fernandez-Irigoyen J, Zelaya MV, Tunon T, et al. Anatomic-proteomic characterization of human basal ganglia: focus on striatum and globus pallidus. *Mol Brain.* 2014;7:83.
- [5] Scherl A, Coute Y, Deon C, et al. Functional proteomic analysis of human nucleolus. *Mol Biol Cell.* 2002;13:4100–4109.
- [6] Andersen JS, Lyon CE, Fox AH, et al. Directed proteomic analysis of the human nucleolus. *Curr Biol.* 2002;12:1–11.
- [7] Andersen JS, Lam YW, Leung AK, et al. Nucleolar proteome dynamics. *Nature.* 2005;433:77–83.
- [8] Engelke R, Riede J, Hegermann J, et al. The quantitative nuclear matrix proteome as a biochemical snapshot of nuclear organization. *J Proteome Res.* 2014;13:3940–3956.
- [9] Albrethsen J, Knol JC, Jimenez CR. Unravelling the nuclear matrix proteome. *J Proteomics.* 2009;72:71–81.
- [10] Saitoh N, Spahr CS, Patterson SD, et al. Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell.* 2004;15:3876–3890.
- [11] Schirmer EC, Florens L, Guan T, et al. Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science.* 2003;301:1380–1382.
- [12] Field MC, Adung'A V, Obado S, et al. Proteomics on the rims: insights into the biology of the nuclear envelope and flagellar pocket of trypanosomes. *Parasitology.* 2012;139:1158–1167.
- [13] Rout MP, Aitchison JD, Suprpto A, et al. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol.* 2000;148:635–651.
- [14] Cronshaw JM. Proteomic analysis of the mammalian nuclear pore complex. *J Cell Biol.* 2002;158:915–927.
- [15] Neubauer G, Gottschalk A, Fabrizio P, et al. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proc Natl Acad Sci U S A.* 1997;94:385–390.
- [16] Neubauer G, King A, Rappsilber J, et al. Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat Genet.* 1998;20:46–50.
- [17] Gottschalk A, Neubauer G, Banroques J, et al. Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snRNP. *Embo J.* 1999;18:4535–4548.
- [18] Rappsilber J, Ryder U, Lamond AI, et al. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* 2002;12:1231–1245.
- [19] Obuse C, Iwasaki O, Kiyomitsu T, et al. A conserved Mis12 centromere complex is linked to heterochromatic HP1 and outer kinetochore protein Zwint-1. *Nat Cell Biol.* 2004;6:1135–1141.
- [20] Shechter D, Dormann HL, Allis CD, et al. Extraction, purification and analysis of histones. *Nat Protoc.* 2007;2:1445–1457.
- [21] Skupien-Rabian B, Jankowska U, Swiderska B, et al. Proteomic and bioinformatic analysis of a nuclear intrinsically disordered proteome. *J Proteomics.* 2016;130:76–84.
- [22] Turck N, Richert S, Gendry P, et al. Proteomic analysis of nuclear proteins from proliferative and differentiated human colonic intestinal epithelial cells. *Proteomics.* 2004;4:93–105.
- [23] Nasrabadi D, Larijani MR, Fathi A, et al. Nuclear proteome analysis of monkey embryonic stem cells during differentiation. *Stem Cell Rev.* 2010;6:50–61.
- [24] Teperek M, Miyamoto K, Simeone A, et al. Sperm and spermatids contain different proteins and bind distinct egg factors. *Int J Mol Sci.* 2014;15:16719–16740.
- [25] Amin B, Maurer A, Voelter W, et al. New potential serum biomarkers in multiple sclerosis identified by proteomic strategies. *Curr Med Chem.* 2014;21:1544–1556.
- [26] Bennetzen MV, Larsen DH, Bunkenborg J, et al. Site-specific phosphorylation dynamics of the nuclear proteome during the DNA damage response. *Mol Cell Proteomics.* 2010;9:1314–1323.
- [27] Hwang SI, Lundgren DH, Mayya V, et al. Systematic characterization of nuclear proteome during apoptosis: a quantitative proteomic study by differential extraction and stable isotope labeling. *Mol Cell Proteomics.* 2006;5:1131–1145.
- [28] Lam YW, Evans VC, Heesom KJ, et al. Proteomics analysis of the nucleolus in adenovirus-infected cells. *Mol Cell Proteomics.* 2010;9:117–130.

- [29] Kroeker AL, Ezzati P, Coombs KM, et al. Influenza A infection of primary human airway epithelial cells up-regulates proteins related to purine metabolism and ubiquitin-related signaling. *J Proteome Res.* **2013**;12:3139–3151.
- [30] DeBoer J, Jagadish T, Haverland NA, et al. Alterations in the nuclear proteome of HIV-1 infected T-cells. *Virology.* **2014**;468-470:409–420.
- [31] Agrawal GK, Rakwal R. Rice proteomics: A move toward expanded proteome coverage to comparative and functional proteomics uncovers the mysteries of rice and plant biology. *Proteomics.* **2011**;11:1630–1649.
- [32] Erhardt M, Adamska I, Franco OL. Plant nuclear proteomics—inside the cell maestro. *Febs J.* **2010**;277:3295–3307.
- [33] Casati P. Recent advances in maize nuclear proteomic studies reveal histone modifications. *Front Plant Sci.* **2012**;3:278.
- [34] Petrovska B, Sebela M, Dolezel J. Inside a plant nucleus: discovering the proteins. *J Exp Bot.* **2015**;66:1627–1640.
- [35] Yin X, Komatsu S. Nuclear proteomics reveals the role of protein synthesis and chromatin structure in root tip of soybean during the initial stage of flooding stress. *J Proteome Res.* **2016**;15:2283–2298.
- [36] Calikowski TT, Meulia T, Meier I. A proteomic study of the arabidopsis nuclear matrix. *J Cell Biochem.* **2003**;90:361–378.
- [37] Bigeard J, Rayapuram N, Bonhomme L, et al. Proteomic and phosphoproteomic analyses of chromatin-associated proteins from *Arabidopsis thaliana*. *Proteomics.* **2014**;14:2141–2155.
- [38] Bae MS, Cho EJ, Choi EY, et al. Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress. *Plant J.* **2003**;36:652–663.
- [39] Fakh Z, Ahmed MB, Letanneur C, et al. An unbiased nuclear proteomics approach reveals novel nuclear protein components that participates in MAMP-triggered immunity. *Plant Signal Behav.* **2016**;11:e1183087.
- [40] Blavet N, Urinovska J, Jerabkova H, et al. UNcleProt (Universal Nuclear Protein database of barley): the first nuclear protein database that distinguishes proteins from different phases of the cell cycle. *Nucleus.* **2017**;8:70–80.
- [41] Tan F, Li G, Chitteti BR, et al. Proteome and phosphoproteome analysis of chromatin associated proteins in rice (*Oryza sativa*). *Proteomics.* **2007**;7:4511–4527.
- [42] Aki T, Yanagisawa S. Application of rice nuclear proteome analysis to the identification of evolutionarily conserved and glucose-responsive nuclear proteins. *J Proteome Res.* **2009**;8:3912–3924.
- [43] Nuez-Ortin WG, Carter CG, Nichols PD, et al. Sequential protein extraction as an efficient method for improved proteome coverage in larvae of Atlantic salmon (*Salmo salar*). *Proteomics.* **2016**;16:2043–2047.
- [44] Ronci M, Pieroni L, Greco V, et al. Sequential fractionation strategy identifies three missing proteins in the mitochondrial proteome of commonly used cell lines. *J Proteome Res.* **2018**.
- [45] Glab N, Labidi B, Qin LX, et al. Olomoucine, an inhibitor of the cdc2/cdk2 kinases activity, blocks plant cells at the G1 to S and G2 to M cell cycle transitions. *FEBS Lett.* **1994**;353:207–211.
- [46] Tamura K, Yamada K, Shimada T, et al. Endoplasmic reticulum-resident proteins are constitutively transported to vacuoles for degradation. *Plant J.* **2004**;39:393–402.
- [47] Tamura K, Takahashi H, Kunieda T, et al. *Arabidopsis* KAM2/GRV2 is required for proper endosome formation and functions in vacuolar sorting and determination of the embryo growth axis. *Plant Cell.* **2007**;19:320–332.
- [48] Tamura K, Fukao Y, Iwamoto M, et al. Identification and characterization of nuclear pore complex components in *Arabidopsis thaliana*. *Plant Cell.* **2010**;22:4084–4097.
- [49] Tamura K, Shimada T, Ono E, et al. Why green fluorescent fusion proteins have not been observed in the vacuoles of higher plants. *Plant J.* **2003**;35:545–555.
- [50] Clough SJ, Bent AF. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **1998**;16:735–743.
- [51] Pendle AF, Clark GP, Boon R, et al. Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. *Mol Biol Cell.* **2005**;16:260–269.
- [52] Tamura K, Hara-Nishimura I. Involvement of the nuclear pore complex in morphology of the plant nucleus. *Nucleus.* **2011**;2:168–172.
- [53] Hatano K, Shimada T, Hiraiwa N, et al. A rapid increase in the level of binding protein (BiP) is accompanied by synthesis and degradation of storage proteins in pumpkin cotyledons. *Plant Cell Physiol.* **1997**;38:344–351.
- [54] Rogers SW, Burks M, Rogers JC. Monoclonal antibodies to barley aleurain and homologs from other plants. *Plant J.* **1997**;11:1359–1368.
- [55] Takata H, Nishijima H, Ogura S, et al. Proteome analysis of human nuclear insoluble fractions. *Genes Cells.* **2009**;14:975–990.
- [56] Tamura K, Goto C, Hara-Nishimura I. Recent advances in understanding plant nuclear envelope proteins involved in nuclear morphology. *J Exp Bot.* **2015**;66:1641–1647.
- [57] Meier I, Richards EJ, Evans DE. Cell biology of the plant nucleus. *Annu Rev Plant Biol.* **2017**;68:139–172.
- [58] Zhao Q, Brkljacic J, Meier I. Two distinct interacting classes of nuclear envelope-associated coiled-coil proteins are required for the tissue-specific nuclear envelope targeting of *Arabidopsis* RanGAP. *Plant Cell.* **2008**;20:1639–1651.
- [59] Zhou X, Graumann K, Evans DE, et al. Novel plant SUN-KASH bridges are involved in RanGAP anchoring and nuclear shape determination. *J Cell Biol.* **2012**;196:203–211.
- [60] Meier I. LINCing the eukaryotic tree of life - towards a broad evolutionary comparison of nucleocytoplasmic bridging complexes. *J Cell Sci.* **2016**;129:3523–3531.
- [61] Rose A, Schraegle SJ, Stahlberg EA, et al. Coiled-coil protein composition of 22 proteomes—differences and

- common themes in subcellular infrastructure and traffic control. *BMC Evol Biol.* **2005**;5:66.
- [62] Starr DA, Fischer JA. KASH 'n Karry: the KASH domain family of cargo-specific cytoskeletal adaptor proteins. *Bioessays.* **2005**;27:1136–1146.
- [63] Wilhelmsen K, Ketema M, Truong H, et al. KASH-domain proteins in nuclear migration, anchorage and other processes. *J Cell Sci.* **2006**;119:5021–5029.
- [64] Lee WY, Lee D, Chung WI, et al. Arabidopsis ING and Alfin1-like protein families localize to the nucleus and bind to H3K4me3/2 via plant homeodomain fingers. *Plant J.* **2009**;58:511–524.
- [65] Wang X, Xie B, Zhu M, et al. Nucleostemin-like 1 is required for embryogenesis and leaf development in Arabidopsis. *Plant Mol Biol.* **2012**;78:31–44.
- [66] Rossignol P, Collier S, Bush M, et al. Arabidopsis POT1A interacts with TERT-V(I8), an N-terminal splicing variant of telomerase. *J Cell Sci.* **2007**;120:3678–3687.
- [67] De Camilli P, Chen H, Hyman J, et al. The ENTH domain. *FEBS Lett.* **2002**;513:11–18.
- [68] Hyman J, Chen H, Di Fiore PP, et al. Epsin 1 undergoes nucleocytoplasmic shuttling and its eps15 interactor NH (2)-terminal homology (ENTH) domain, structurally similar to Armadillo and HEAT repeats, interacts with the transcription factor promyelocytic leukemia Zn(2)+ finger protein (PLZF). *J Cell Biol.* **2000**;149:537–546.
- [69] Benschop JJ, Mohammed S, O'Flaherty M, et al. Quantitative phosphoproteomics of early elicitor signaling in Arabidopsis. *Mol Cell Proteomics.* **2007**;6:1198–1214.
- [70] Wilson R, Bateman JF. A robust method for proteomic characterization of mouse cartilage using solubility-based sequential fractionation and two-dimensional gel electrophoresis. *Matrix Biol.* **2008**;27:709–712.
- [71] Wu X, Xiong E, An S, et al. Sequential extraction results in improved proteome profiling of medicinal plant *Pinellia ternata* tubers, which contain large amounts of high-abundance proteins. *PLoS One.* **2012**;7:e50497.
- [72] Petrovska B, Jerabkova H, Chamrad I, et al. Proteomic analysis of barley cell nuclei purified by flow sorting. *Cytogenet Genome Res.* **2014**;143:78–86.
- [73] Schmid M, Davison TS, Henz SR, et al. A gene expression map of Arabidopsis thaliana development. *Nat Genet.* **2005**;37:501–506.
- [74] Liang X, Lei M, Li F, et al. Family-wide characterization of histone binding abilities of PHD domains of AL proteins in Arabidopsis thaliana. *Protein J.* **2018**;37:531–538.
- [75] Hill TA, Broadhvest J, Kuzoff RK, et al. Arabidopsis SHORT INTEGUMENTS 2 is a mitochondrial DAR GTPase. *Genetics.* **2006**;174:707–718.
- [76] Wang X, Gingrich DK, Deng Y, et al. A nucleostemin-like GTPase required for normal apical and floral meristem development in Arabidopsis. *Mol Biol Cell.* **2012**;23:1446–1456.
- [77] Jeon Y, Park YJ, Cho HK, et al. The nucleolar GTPase nucleostemin-like 1 plays a role in plant growth and senescence by modulating ribosome biogenesis. *J Exp Bot.* **2015**;66:6297–6310.
- [78] Wang Z, Wang X, Xie B, et al. Arabidopsis NUCLEOSTEMIN-LIKE 1 (NSN1) regulates cell cycling potentially by cooperating with nucleosome assembly protein AtNAP1;1. *BMC Plant Biol.* **2018**;18:99.
- [79] Dou XY, Yang KZ, Ma ZX, et al. AtTMEM18 plays important roles in pollen tube and vegetative growth in Arabidopsis. *J Integr Plant Biol.* **2016**;58:679–692.
- [80] Du J, Wang X, Zhang M, et al. Unique nucleotide polymorphism of ankyrin gene cluster in Arabidopsis. *J Genet.* **2007**;86:27–35.
- [81] Li S, Gu Y, Yan A, et al. RIP1 (ROP Interactive Partner 1)/ICR1 marks pollen germination sites and may act in the ROP1 pathway in the control of polarized pollen growth. *Mol Plant.* **2008**;1:1021–1035.