Molecular Systems Biology 4; Article number 193; doi:10.1038/msb.2008.32
Citation: ***Molecular Systems Biology*** **4:193**

www.molecularsystemsbiology.com

molecular systems biology

REPORT

Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in *Arabidopsis*

Naoyuki Sugiyama[1,2,5], Hirofumi Nakagami[3,5], Keiichi Mochida[3], Arsalan Daudi[3], Masaru Tomita[1,2], Ken Shirasu[3,*] and Yasushi Ishihama[1,4,*]

[1] Institute for Advanced Biosciences, Keio University, Tsuruoka, Japan, [2] Human Metabolome Technologies, Tsuruoka, Japan, [3] RIKEN Plant Science Center, Yokohama, Japan and [4] PRESTO, Japan Science and Technology Agency, Tokyo, Japan
[5] These authors contributed equally to this work
* Corresponding authors. Y Ishihama, Institute for Advanced Biosciences, Keio University, 403-1, Daihoji, Tsuruoka, Yamagata 997-0017, Japan. Tel.: +81 235 29 0571; Fax: +81 235 29 0536; E-mail: y-ishi@ttck.keio.ac.jp or K Shirasu, RIKEN Plant Science Center, Yokohama, Japan. Tel.: +81 45 503 9574; Fax: +81 45 503 9573; E-mail: ken.shirasu@psc.riken.jp

Received 6.2.08; accepted 4.4.08

Protein phosphorylation regulates a wide range of cellular processes. Here, we report the proteome-wide mapping of *in vivo* phosphorylation sites in *Arabidopsis* by using complementary phosphopeptide enrichment techniques coupled with high-accuracy mass spectrometry. Using unfractionated whole cell lysates of *Arabidopsis*, we identified 2597 phosphopeptides with 2172 high-confidence, unique phosphorylation sites from 1346 proteins. The distribution of phosphoserine, phosphothreonine, and phosphotyrosine sites was 85.0, 10.7, and 4.3%. Although typical tyrosine-specific protein kinases are absent in *Arabidopsis*, the proportion of phosphotyrosines among the phospho-residues in *Arabidopsis* is similar to that in humans, where over 90 tyrosine-specific protein kinases have been identified. In addition, the tyrosine phosphoproteome shows features distinct from those of the serine and threonine phosphoproteomes. Taken together, we highlight the extent and contribution of tyrosine phosphorylation in plants.
Molecular Systems Biology 6 May 2008; doi:10.1038/msb.2008.32
Subject Categories: proteomics; plant biology
Keywords: *Arabidopsis*; phosphoproteome; tyrosine kinase; tyrosine phosphorylation

Introduction

Protein phosphorylation is a critical regulatory step in signaling networks and is arguably the most widespread protein modification affecting almost all basic cellular processes in various organisms (Hunter, 2000; Manning *et al*, 2002).

Advances in mass spectrometry (MS)-based technologies accompanied with phosphopeptide enrichment methods paved the way for high-throughput, large-scale *in vivo* phosphorylation site mapping, and indeed, several pioneering plant phosphoproteome studies have been reported in the past 5 years (Nuhse *et al*, 2003, 2004, 2007; de la Fuente van Bentem *et al*, 2006; Benschop *et al*, 2007). Although these studies provided new insights into phosphorylation events in plants, the analyses were restricted to subfractionated samples, such as plasma membrane proteins, containing a few hundred phosphoproteins. No plant study has yet been reported using unfractionated whole cells to provide a wide-ranging view of cellular phosphorylation events.

More than 1000 phosphorylation sites have recently been identified in animal and yeast cells, using a combination of two or more methods for phosphopeptide enrichment coupled with mass spectrometric phosphopeptide-oriented techniques, such as neutral loss-triggered MS^3 to generate fragment ions after elimination of labile phosphate groups, multistage activation, and electron transfer dissociation (Olsen *et al*, 2006; Bodenmiller *et al*, 2007a; Chi *et al*, 2007; Molina *et al*, 2007; Villen *et al*, 2007). We also reported the identification of more than 2000 *in vivo* phosphorylated sites in unstimulated HeLa cells employing an aliphatic hydroxy acid-modified metal oxide chromatography (HAMMOC) as a phosphopeptide enrichment method (Sugiyama *et al*, 2007). Since different phosphopeptide enrichment methods are likely to have

distinct preferences for particular properties of phosphopeptides (Bodenmiller *et al*, 2007b), it is reasonable to use two or more phosphopeptide enrichment methods for evaluation of proteome-wide phosphorylation.

Comparative genome analyses revealed substantial differences in the ensembles of kinases (kinomes) in eukaryotes (Diks *et al*, 2007). The *Arabidopsis* genome encodes at least two times more protein kinases than the human genome (Manning *et al*, 2002; Champion *et al*, 2004). Importantly, the *Arabidopsis* genome (Initiative, 2000) does not contain any predicted human-type TKs (Rudrabhatla *et al*, 2006). However, plants are likely to utilize tyrosine phosphorylation signaling, as *bona fide* tyrosine-specific protein phosphatases do exist in *Arabidopsis* (Xu *et al*, 1998; Luan, 2003), and a few early studies detected tyrosine phosphorylation by using pY antibodies (Barizza *et al*, 1999; Kameyama *et al*, 2000; Luan, 2002). In addition, a previous *Arabidopsis* phosphoproteome study identified a small number of phosphorylated tyrosine residues, although the actual data sets were missing in the report (Benschop *et al*, 2007). Thus, evidence for tyrosine phosphorylation in plants is limited so far.

Here, we present a large-scale phosphoproteome analysis of *Arabidopsis*, providing an overview of *in vivo* phosphorylation events in *Arabidopsis* at the cellular level. Importantly, we show the extent of tyrosine phosphorylation in plants, which has been largely underestimated to date.

Results and discussion

Large-scale *in vivo* phosphorylation site mapping in *Arabidopsis*

To collect a comprehensive data set of *Arabidopsis* phosphorylation sites, we employed six distinct methods for phosphopeptide enrichment (Supplementary information). Our approach identified 2172 unique phosphorylation sites with very high confidence on 1346 proteins from unfractionated *Arabidopsis* cell lysates; this is one of the largest data sets available for a plant to date (Supplementary Table I and Supplementary information). A large majority (1155; 85.8%) of the identified phosphoproteins are novel, while 191 (14.2%) were reported in the previous phosphoproteome studies that focused on plasma membrane proteins (Nuhse *et al*, 2004; Benschop *et al*, 2007) (Supplementary Figure 2).

Arabidopsis phosphoproteome

To obtain an overview of phosphorylation events in *Arabidopsis*, protein abundance distribution, cellular localization, molecular function, and biological processes of identified phosphoproteins were analyzed and compared with those of all proteins encoded by the *Arabidopsis* genome (Supplementary Figures 3 and 4). Phosphoproteins were generally less abundant, as expected, even when we did not take account of the degree of phosphorylation (Supplementary information). Proteins from all subcellular compartments were found to be targets for phosphorylation. However, approximately 40% of phosphorylation occurred on predicted nuclear proteins. Since nuclear proteins account for only approximately 20% of all genome-encoded proteins and 15% of the experimentally identified proteins in this study, phosphorylation is likely to target nuclear proteins preferentially (Supplementary Figures 4A and 5). The distributions of the molecular function and biological processes of phosphoproteins and that of all genome-encoded proteins were relatively similar (Supplementary Figures 4B and C). This indicates that most cellular processes in *Arabidopsis* are likely to be regulated at least in part by various phosphorylation events.

Table I Numbers of identified phosphopeptides, phosphoproteins, phosphorylation sites, and the content of phosphorylated residues

Items	Number
Number of phosphopeptides[a]	2597
Number of phosphoproteins[b]	1346
Number of unique phosphorylation sites	2172
Phosphorylated residues (Ser:Thr:Tyr)	1847:231: 94 (85.0%) (10.6%) (4.3%)

[a]The number of phosphopeptides is based on unique sequences containing missed cleavage products, oxidization of methionine, and phosphorylation of different sites.
[b]Multiple distinct proteins except splicing variants matched against a single peptide are also counted.

To our surprise, of the 2172 identified phosphorylation sites, we found 94 sites to be tyrosine residues (Table I). The kinome of *Arabidopsis* does not contain any of the typical TKs found in humans, suggesting that plants and humans do not share mechanistic features of tyrosine phosphorylation. Nevertheless, the relative abundances of pS, pT, and pY in *Arabidopsis* were estimated to be 85.0, 10.7, and 4.3%, which are strikingly close to the human phosphoproteome profile recently reported. The proportion of pY among phospho-residues in human cells is estimated to be between 1.8 and 6.0%, depending on the analyzed samples (Olsen *et al*, 2006; Molina *et al*, 2007; Sugiyama *et al*, 2007). These data clearly indicate that the importance of tyrosine phosphorylation in plants has been greatly underestimated.

Arabidopsis tyrosine phosphoproteome

The 94 identified pY residues were mapped on 95 proteins (Supplementary Table II). The difference in the number of pY residues and corresponding proteins is due to matching of single phosphopeptides to several different proteins. Since the sequences surrounding tyrosine phosphorylation sites on the listed protein kinases are often well conserved, the number of protein kinases is over-represented. To investigate whether tyrosine phosphorylation is targeted to a specific subset of proteins, gene ontology analyses of serine-, threonine-, or tyrosine-phosphorylated proteins were performed as described (Figure 1). Tyrosine phosphorylation preferentially occurs on proteins that possess kinase activity or transferase activity (Figure 1B). Otherwise, no outstanding differences were found in the distributions.

Location of phosphorylation sites on characterized protein domains

To assess whether trends or patterns exist in the position of tyrosine phosphorylation sites, we investigated whether these

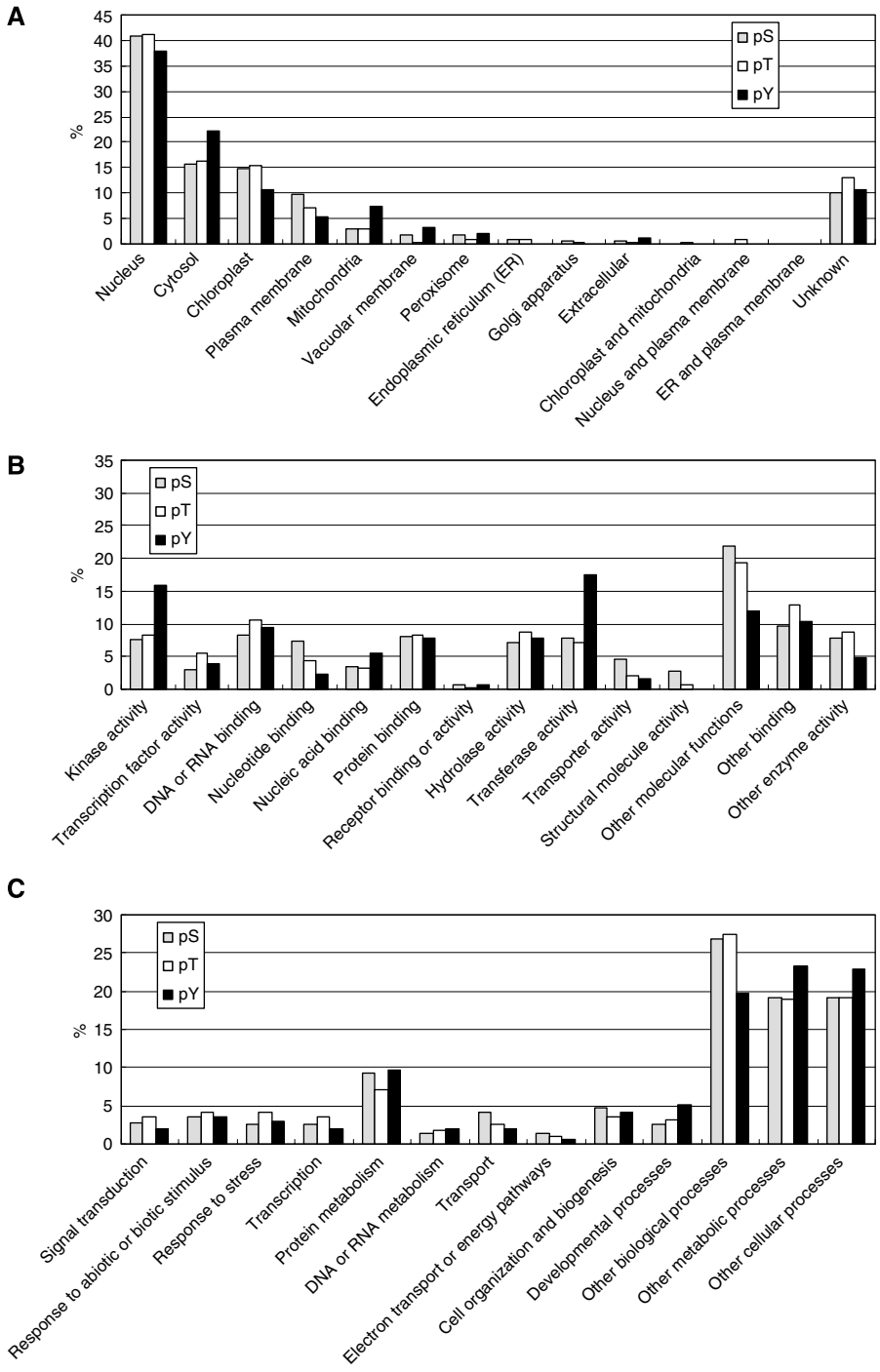

Figure 1 Gene ontology analysis of the serine-, threonine-, or tyrosine-phosphorylated proteins. (**A**) Cellular localization, (**B**) molecular function, and (**C**) biological process.

sites are located in conserved domains. Pfam search (Bateman *et al*, 2004) was used to extract domain information of the identified phosphoproteins. Of the 1346 proteins, we obtained domain information for the 1118 proteins. In these proteins, 77.95% of phosphorylation sites (1548 sites) were located outside of conserved domains (Table II). The tendency that the majority of phosphorylation occurs outside of conserved domains is consistent with the observations from the phosphoproteome study of plasma membrane proteins (Nuhse *et al*, 2004). Interestingly, however, nearly half (48.5%) of pYs were found to be located on conserved domains (Table II). These data indicate that tyrosine phosphorylation may have more impact on domain-associated function compared to serine and threonine phosphorylation.

Table II Location of phosphorylation sites on characterized protein domains

	Number of proteins possessing Pfam domain	Number of phosphorylation sites		
		Pfam domain[a]		Total (%)
		ON (%)	OUT[b] (%)	
pS	1014	317 (19.1)	1340 (80.9)	1657 (100)
pT	195	74 (32.2)	156 (67.8)	230 (100)
pY	87	49 (48.5)	52 (51.5)	101 (100)
All	1118	440 (22.1)	1548 (77.9)	1988 (100)

[a]Whether the phosphorylation sites are located on the conserved domains annotated in the Pfam database was analyzed.
[b]A single phosphorylation site was counted only once, regardless of how many domains are found in a single phosphoprotein.

Conservation of tyrosine phosphorylation sites in plant homologs

Conservation of the tyrosine phosphorylation sites between homologous proteins in *Arabidopsis*, rice (*Oryza sativa*), and poplar (*Populus trichocarpa*) was investigated to get an overview of tyrosine phosphorylation events in other plant species. Of the 95 tyrosine-phosphorylated proteins (109 tyrosine phosphorylation sites), 84 proteins (97 sites) were validated to possess homologs in *Arabidopsis* (paralogs), while 89 (103 sites) and 92 (106 sites) proteins had corresponding homologs (orthologs) in rice and poplar, respectively (Supplementary Table II). Multiple sequence alignments of the homologous proteins were created using ClustalW (Thompson *et al*, 1994), and the conservation of the phosphorylatable tyrosine residues was verified manually (Figure 2). In total, 72 sites are conserved within *Arabidopsis* paralogs, while 72 and 79 sites are conserved in rice and poplar orthologs, respectively. Most of these sites (61 sites) are conserved in all three plant species, indicating that the tyrosine phosphorylation sites are nearly equally conserved in paralogs and orthologs. This observation is in clear contrast to the case of serine phosphorylation sites, which are less conserved in paralogs compared to orthologs (Nuhse *et al*, 2004).

Distribution of the phosphorylation sites

We found that most (76.3%) of the pY-containing phosphopeptides are multiply phosphorylated, while the majority (80.9%) of phosphopeptides are singly phosphorylated (Table III). In other words, tyrosine phosphorylation seems to occur near other phospho-residues (Supplementary Table III). Since the amino acids surrounding the phosphorylation sites often contribute substantially to recognition by protein kinases, the phosphorylation status of neighboring residues is an essential factor in determining whether the phosphorylation site is targeted by a particular protein kinase. It would be very interesting to investigate whether and how the phosphorylation state of the neighboring residues affects tyrosine phosphorylation events in *Arabidopsis*.

Tyrosine phosphorylation motifs

An obvious question arising from our finding is, which kinases carry out tyrosine phosphorylation in plants? To address this question, we attempted to extract significant patterns surrounding the pY residues from our data set, assuming that conserved phosphorylation sites within functionally related proteins tend to be well targeted by structurally similar protein kinases. We have extracted 20 pY motifs through the substrate-driven approach (Schwartz and Gygi, 2005) (Supplementary information). Most of the identified pY motifs are novel and distinct from the human pS, and pT and pY motifs in the human protein reference database (Amanchy *et al*, 2007). These results indicate that tyrosine phosphorylation in plants is carried out by a novel class(es) of plant kinases. One candidate might be dual-specific serine/threonine/tyrosine protein kinases (Rudrabhatla *et al*, 2006). Other possible candidates would be tyrosine-specific protein kinase-like kinases (TKLs), which are especially abundant in plants: 776 in *Arabidopsis* and nearly 1000 in rice, compared to 55 in humans (Miranda-Saavedra and Barton, 2007). Tyrosine phosphorylation by human TKLs has not been reported. Functions of plant TKLs remain also unknown, but the large number of TKLs in plants may suggest that they carry out important and diverse plant-specific functions. In this sense, it is of particular interest to investigate if any of TKL possesses tyrosine phosphorylation activity.

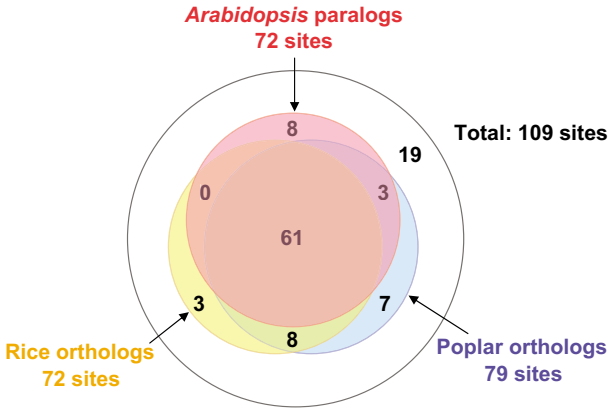

Figure 2 Venn diagram showing the number of conserved tyrosine phosphorylation sites in homologs. Blue, yellow, and red circles indicate the conserved sites in *Arabidopsis* paralogs, rice orthologs, and poplar orthologs, respectively.

Materials and methods

Plant material

Arabidopsis cell suspension line (ecotype Landsberg erecta) (Maor *et al*, 2007) was grown in Murashige and Skoog medium (pH 5.7) containing 3% sucrose, 0.59 g/l MES, 100 mg/l myo-inositol, 10 mg/l thiamine-HCl, 1 mg/l pyridoxine-HCl, 1 mg/l nicotinic acid, 0.5 mg/l 1-naphthaleneacetic acid, and 0.05 mg/l 6-benzylaminopurine under a 16-h light/8-h dark cycle at 22°C. Seven-day-old *Arabidopsis* suspension cultures were harvested by vacuum filtration, frozen immediately in liquid nitrogen, and kept at −80°C until the analysis.

Digestion of *Arabidopsis* cell cytoplasmic fraction

Arabidopsis cells (0.2 g, wet) were frozen in liquid nitrogen and then disrupted with a Multi-beads shocker (MB400U; Yasui Kikai). The disrupted cells were suspended in 0.1 M Tris–HCl (pH 8.0), containing

Table III Comparison of singly and multiply phosphorylated peptides

All		pS		pT		pY	
Single	Multi	Single	Multi	Single	Multi	Single	Multi
1888 (80.9%)	445 (19.1%)	1712 (80.3%)	419 (19.7%)	153 (61.0%)	98 (39.0%)	23 (23.7%)	74 (76.3%)

Whether pS-, pT-, or pY-containing phosphopeptides are singly or multiply phosphorylated was analyzed.

protein phosphatase inhibitor cocktails 1 and 2 (Sigma) and protease inhibitors (Sigma). The homogenate was centrifuged at 1500 *g* for 10 min and the supernatant was reduced with dithiothreitol, alkylated with iodoacetamide, and digested with Lys-C, followed by dilution and trypsin digestion as described (Saito *et al*, 2006). These digested samples were desalted using StageTips with C18 Empore disk membranes (3 M) (Rappsilber *et al*, 2003). The peptide concentration of the eluates was adjusted to 1.0 mg/ml with 0.1% TFA and 80% acetonitrile.

Enrichment of phosphopeptides

HAMMOC using titania and zirconia was performed as described previously (Sugiyama *et al*, 2007) with some modifications. Custom-made MOC tips were prepared using C8-StageTips and metal oxide bulk beads (0.5 mg beads per 10 µl pipette tip), as described for SCX(beads)-C18 tips (Ishihama *et al*, 2006). Prior to loading samples, the MOC tips were equilibrated with 0.1% TFA, 80% acetonitrile, containing a hydroxy acid as a selectivity enhancer (solution A). As an enhancer, lactic acid was used at a concentration of 300 mg/ml for titania MOC tips and β-hydroxypropanoic acid at 100 mg/ml for zirconia MOC tips. The digested sample from 100 µg of *Arabidopsis* total proteins was diluted with 100 µl of solution and loaded onto the MOC tips. After successive washing with solution A and solution B (0.1% TFA and 80% acetonitrile), 0.5% ammonium hydroxide or 1.0% disodium hydrogen phosphate was used for elution. The eluted fraction was acidified with TFA, desalted using C18-StageTips as described above, and concentrated in a Tony CC-105 vacuum evaporator (Tokyo, Japan), followed by the addition of solution A for subsequent nanoLC-MS/MS analysis.

Fe-IMAC was conducted using Phos-Select (Sigma) as described previously (Kokubu *et al*, 2005; Ishihama *et al*, 2007), except for the use of C8-StageTips instead of C18-StageTips for packing Phos-Select beads. Briefly, after loading the sample solutions, the tips were rinsed with 0.5 ml of 50% ACN in 0.3% TFA. Then, 0.5% ammonium hydroxide or 1.0% disodium hydrogen phosphate was used for elution. The eluted fraction was acidified with TFA and desalted using C18-StageTips using HAMMOC methods. The eluted phosphopeptide fraction was concentrated in the vacuum evaporator and resuspended in solution A for nanoLC-MS/MS analysis.

NanoLC-MS system

A Finnigan LTQ-Orbitrap (Thermo Fisher Scientific, Bremen, Germany) coupled with a Dionex Ultimate3000 pump (Germering, Germany) and an HTC-PAL autosampler (CTC Analytics AG, Zwingen, Switzerland) was used for nanoLC-MS/MS analyses throughout this study. ReproSil C18 materials (3 µm; Dr Maisch, Ammerbuch, Germany) were packed into a self-pulled needle (150 mm length × 100 µm i.d., 6 µm opening) to prepare an analytical column needle with 'stone-arch' frit (Ishihama *et al*, 2002). An *x*–*y*–*z* nanospray interface (Nikkyo Technos, Tokyo, Japan) was used to hold the column needle and to set the appropriate spray position. A spray voltage of 2400 V was applied. The injection volume was 5 µl and the flow rate was 500 nl/min. The mobile phases consisted of (A) 0.5% acetic acid and (B) 0.5% acetic acid and 80% acetonitrile. A three-step linear gradient of 5–10% B in 5 min, 10–40% B in 60 min, 40–100% B in 5 min and 100% B for 10 min was employed throughout this study. The MS scan range was m/z 300–1400, and the top 10 precursor ions were selected for subsequent MS/MS scans. A lock mass function was used for the LTQ-Orbitrap to obtain constant mass accuracy during gradient analysis (Olsen *et al*, 2005).

Database searching

Mass Navigator v1.2 (Mitsui Knowledge Industry, Tokyo, Japan) was used to create peak lists on the basis of the recorded fragmentation spectra. Peptides and proteins were identified by means of automated database search using Mascot v2.1 (Matrix Science, London) against TAIR7_pep_20070425 (ftp://ftp.arabidopsis.org/home/tair/Sequences/blast_datasets/TAIR7_blastsets/) with a precursor mass tolerance of 3 p.p.m., a fragment ion mass tolerance of 0.8 Da and strict trypsin specificity (Olsen *et al*, 2004), allowing for up to two missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification, and oxidation of methionines and phosphorylation of serine, threonine, and tyrosine were allowed as variable modifications. Peptides were considered identified if the Mascot score was over the 95% confidence limit based on the 'identity' score of each peptide and at least three successive y- or b-ions with a further two and more y-, b-, and/or precursor-origin neutral loss ions were observed, based on the error-tolerant peptide sequence tag concept (Mann and Wilm, 1994). A randomized decoy database created by a Mascot Perl program estimated a 2.1% false-positive rate for identified peptides within these criteria. Note that most sulfated peptides can be discriminated from phosphopeptides because of the ultrahigh accuracy of the Orbitrap instrument that we used.

Phosphorylated sites were unambiguously determined when y- or b-ions between which the phosphorylated residue exists were observed in the peak lists of the fragment ions.

Bioinformatics

We used the KAGIANA tool (http://pmnedo.kazusa. or.jp/kagiana/index.html) to extract cellular localization information of *Arabidopsis* proteins predicted by the WoLF PSORT program (Horton *et al*, 2007). For molecular function and biological process annotations extraction, the TAIR gene ontology annotation search tool (Berardini *et al*, 2004) was used.

For the homologs search, BlastP searches (Altschul *et al*, 1997) were performed against the protein databases, TAIR7_pep_2007425, rap1_all_orf_amino, and proteins.Poptr1_1.JamboreeModels for *Arabidopsis*, rice, and poplar, respectively (Ohyanagi *et al*, 2006; Tuskan *et al*, 2006). The *E*-value cutoff of 10^{-3} was used for the initial search and if there were no protein hits, the cutoff value was lowered stepwise to 10^{-2} and 10^{-1}. In some cases, *E*-value cutoff of 10^{-6} was used for AT1G70520.1, *E*-value cutoff of 10^{-5} was used for AT3G05140.1, and *E*-value cutoff of 10^{-4} was used for AT2G30940.1. For multiple sequence alignment, ClustalW (Thompson *et al*, 1994) was performed with default parameter settings. The aligned sequences were further manually analyzed using the MEGA4 program (Tamura *et al*, 2007).

Pfam domain information was extracted from the database, TAIR7_all.domains (ftp://ftp.arabidopsis.org/home/tair/Proteins/Domains/).

Supplementary information

Supplementary information is available at the *Molecular Systems Biology* website (www.nature.com/msb) and at the RIKEN website (OmicBrowse, http://omicspace.riken.jp/gps) (Toyoda *et al*, 2007) connected to the Keio Univ PepBase (http://pepbase.iab.keio.ac.jp).

Acknowledgements

We thank Sumiko Ohnuma (Keio University) for technical assistance, Alex Jones (The Sainsbury Laboratory, UK) for initial introduction of mass spectrometry at RIKEN, and Tetsuro Toyoda and Yoshiki Mochizuki (RIKEN) for kindly uploading the phosphorylation site data to a RIKEN OmicBrowse database. This study was supported by research funds from Yamagata Prefecture and Tsuruoka City to Keio University, MEXT Grants-in-Aid for Scientific Research (nos. 19678001 and 19039034 to KS and no. 19710169 to HN) and a grant from the Japan Society for the Promotion of Science to AD.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402

Amanchy R, Periaswamy B, Mathivanan S, Reddy R, Tattikota SG, Pandey A (2007) A curated compendium of phosphorylation motifs. *Nat Biotechnol* **25:** 285–286

Barizza E, Lo Schiavo F, Terzi M, Filippini F (1999) Evidence suggesting protein tyrosine phosphorylation in plants depends on the developmental conditions. *FEBS Lett* **447:** 191–194

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR (2004) The Pfam protein families database. *Nucleic Acids Res* **32:** D138–D141

Benschop JJ, Mohammed S, O'Flaherty M, Heck AJ, Slijper M, Menke FL (2007) Quantitative phosphoproteomics of early elicitor signaling in *Arabidopsis*. *Mol Cell Proteomics* **6:** 1198–1214

Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang P, Mueller LA, Yoon J, Doyle A, Lander G, Moseyko N, Yoo D, Xu I, Zoeckler B, Montoya M, Miller N, Weems D, Rhee SY (2004) Functional annotation of the *Arabidopsis* genome using controlled vocabularies. *Plant Physiol* **135:** 745–755

Bodenmiller B, Malmstrom J, Gerrits B, Campbell D, Lam H, Schmidt A, Rinner O, Mueller LN, Shannon PT, Pedrioli PG, Panse C, Lee HK, Schlapbach R, Aebersold R (2007a) PhosphoPep—a phosphoproteome resource for systems biology research in *Drosophila* Kc167 cells. *Mol Syst Biol* **3:** 139

Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R (2007b) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat Methods* **4:** 231–237

Champion A, Kreis M, Mockaitis K, Picaud A, Henry Y (2004) *Arabidopsis* kinome: after the casting. *Funct Integr Genomics* **4:** 163–187

Chi A, Huttenhower C, Geer LY, Coon JJ, Syka JE, Bai DL, Shabanowitz J, Burke DJ, Troyanskaya OG, Hunt DF (2007) Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry. *Proc Natl Acad Sci USA* **104:** 2193–2198

de la Fuente van Bentem S, Anrather D, Roitinger E, Djamei A, Hufnagl T, Barta A, Csaszar E, Dohnal I, Lecourieux D, Hirt H (2006) Phosphoproteomics reveals extensive *in vivo* phosphorylation of *Arabidopsis* proteins involved in RNA metabolism. *Nucleic Acids Res* **34:** 3267–3278

Diks SH, Parikh K, van der Sijde M, Joore J, Ritsema T, Peppelenbosch MP (2007) Evidence for a minimal eukaryotic phosphoproteome? *PLoS ONE* **2:** e777

Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res* **35:** W585–W587

Hunter T (2000) Signaling—2000 and beyond. *Cell* **100:** 113–127

Initiative TAG (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408:** 796–815

Ishihama Y, Rappsilber J, Andersen JS, Mann M (2002) Microcolumns with self-assembled particle frits for proteomics. *J Chromatogr A* **979:** 233–239

Ishihama Y, Rappsilber J, Mann M (2006) Modular stop and go extraction tips with stacked disks for parallel and multi-dimensional peptide fractionation in proteomics. *J Proteome Res* **5:** 988–994

Ishihama Y, Wei FY, Aoshima K, Sato T, Kuromitsu J, Oda Y (2007) Enhancement of the efficiency of phosphoproteomic identification by removing phosphates after phosphopeptide enrichment. *J Proteome Res* **6:** 1139–1144

Kameyama K, Kishi Y, Yoshimura M, Kanzawa N, Sameshima M, Tsuchiya T (2000) Tyrosine phosphorylation in plant bending. *Nature* **407:** 37

Kokubu M, Ishihama Y, Sato T, Nagasu T, Oda Y (2005) Specificity of immobilized metal affinity-based IMAC/C18 tip enrichment of phosphopeptides for protein phosphorylation analysis. *Anal Chem* **77:** 5144–5154

Luan S (2002) Tyrosine phosphorylation in plant cell signaling. *Proc Natl Acad Sci USA* **99:** 11567–11569

Luan S (2003) Protein phosphatases in plants. *Annu Rev Plant Biol* **54:** 63–92

Mann M, Wilm M (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem* **66:** 4390–4399

Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. *Science* **298:** 1912–1934

Maor R, Jones A, Nuhse TS, Studholme DJ, Peck SC, Shirasu K (2007) Multidimensional protein identification technology (MudPIT) analysis of ubiquitinated proteins in plants. *Mol Cell Proteomics* **6:** 601–610

Miranda-Saavedra D, Barton GJ (2007) Classification and functional annotation of eukaryotic protein kinases. *Proteins* **68:** 893–914

Molina H, Horn DM, Tang N, Mathivanan S, Pandey A (2007) Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc Natl Acad Sci USA* **104:** 2199–2204

Nuhse TS, Bottrill AR, Jones AM, Peck SC (2007) Quantitative phosphoproteomic analysis of plasma membrane proteins reveals regulatory mechanisms of plant innate immune responses. *Plant J* **51:** 931–940

Nuhse TS, Stensballe A, Jensen ON, Peck SC (2003) Large-scale analysis of *in vivo* phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Mol Cell Proteomics* **2:** 1234–1243

Nuhse TS, Stensballe A, Jensen ON, Peck SC (2004) Phosphoproteomics of the *Arabidopsis* plasma membrane and a new phosphorylation site database. *Plant Cell* **16:** 2394–2405

Ohyanagi H, Tanaka T, Sakai H, Shigemoto Y, Yamaguchi K, Habara T, Fujii Y, Antonio BA, Nagamura Y, Imanishi T, Ikeo K, Itoh T, Gojobori T, Sasaki T (2006) The Rice Annotation Project Database (RAP-DB): hub for *Oryza sativa* ssp. *japonica* genome information. *Nucleic Acids Res* **34:** D741–D744

Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M (2006) Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127:** 635–648

Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, Makarov A, Lange O, Horning S, Mann M (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics* **4:** 2010–2021

Olsen JV, Ong SE, Mann M (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* **3:** 608–614

Rappsilber J, Ishihama Y, Mann M (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **75:** 663–670

Rudrabhatla P, Reddy MM, Rajasekharan R (2006) Genome-wide analysis and experimentation of plant serine/threonine/tyrosine-specific protein kinases. *Plant Mol Biol* **60:** 293–319

Saito H, Oda Y, Sato T, Kuromitsu J, Ishihama Y (2006) Multiplexed two-dimensional liquid chromatography for MALDI and nanoelectrospray ionization mass spectrometry in proteomics. *J Proteome Res* **5:** 1803–1807

Schwartz D, Gygi SP (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol* **23:** 1391–1398

Sugiyama N, Masuda T, Shinoda K, Nakamura A, Tomita M, Ishihama Y (2007) Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol Cell Proteomics* **6:** 1103–1109

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24:** 1596–1599

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22:** 4673–4680

Toyoda T, Mochizuki Y, Player K, Heida N, Kobayashi N, Sakaki Y (2007) OmicBrowse: a browser of multidimensional omics annotations. *Bioinformatics* **23:** 524–526

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V *et al* (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313:** 1596–1604

Villen J, Beausoleil SA, Gerber SA, Gygi SP (2007) Large-scale phosphorylation analysis of mouse liver. *Proc Natl Acad Sci USA* **104:** 1488–1493

Xu Q, Fu HH, Gupta R, Luan S (1998) Molecular characterization of a tyrosine-specific protein phosphatase encoded by a stress-responsive gene in *Arabidopsis*. *Plant Cell* **10:** 849–857