J. Plant Biol. (2016) 59:467-477
DOI 10.1007/s12374-016-0197-y

ORIGINAL ARTICLE

Molecular Evolution of ACTIN RELATED PROTEIN 6, a Component of SWR1 Complex in Arabidopsis

Eunjoo Seo[1,†], Chulmin Park[1,5,†], Kyuha Choi[1,6], Dongseon Lee[2], Chaok Seok[2,3] and Ilha Lee[1,4,*]

[1]*National Research Laboratory of Plant Developmental Genetics, School of Biological Sciences, Seoul National University, Seoul 151-742, Korea*
[2]*School of Chemistry and Molecular Engineering, Seoul National University, Seoul 151-747, Korea*
[3]*School of Computational Sciences, Korea Institute for Advanced Study, Seoul 130-722, Korea*
[4]*Plant Genomics and Breeding Institute, Seoul National University, Seoul 151-742, Korea*
[5]*Gregor Mendel Institute of Molecular Plant Biology, Austrian Academy of Sciences, 1030 Vienna, Austria*
[6]*Division of Applied Life Sciences, BK21[+], Gyeongsang National University, Jinju 600-701, Korea*

Received: April 17, 2016 / Accepted: June 8, 2016

Abstract To date, it has been assumed that the evolution of a protein complex is different from that of other proteins. However, there have been few evidences to support this assumption. To understand how protein complexes evolve, we analyzed the evolutionary constraints on ACTIN RELATED PROTEIN 6 (ARP6), a component of the SWR1 complex. Interspecies complementation experiments using transgenic plants that ectopically express trans*ARP6s* (ARP6s from other organisms) showed that the function of ARP6s is conserved in plants. In addition, a yeast two-hybrid analysis revealed that this functional conservation depends on its ability to bind with both PIE1 and AtSWC6. ARP6 consists of 4 domains similar to actin. Functional analysis of chimeric*ARP6s* (domain-swapped ARP6s between Arabidopsis and mouse) demonstrated that each domain of ARP6s imposes differential evolutionary constraints. Domains 1 and 3 of ARP6 were found to interact with SWC6 and PIE1, respectively, and domain 4 provides a nuclear localization signal. Moreover, domains 1 and 3 showed a slower evolution rate than domain 4, indicating that the interacting domains have higher evolutionary constraints than non-interacting domains do. These findings suggest that the components of this protein complex have evolved coordinately to preserve their interactions.

Keywords: Actin fold, ARP6, Molecular evolution, Protein complex, SWR1 complex

[†]These authors contributed equally to this article.

*Corresponding author; Ilha Lee
Tel : +82-2-872-6950
E-mail : ilhalee@snu.ac.kr

Introduction

Evolutionary concepts have been recognized since at least the time of Aristotle (384-322 BC); however, the mechanism of evolution by natural selection was proposed by Charles Darwin in 1859 (M.A. 1859). Since the advance in molecular biology from the 1950s, evolutionary studies have expanded from species to the DNA and protein levels. Increased accumulation of the information of DNA and protein sequences has revealed that the rate of evolution is not homogeneous among proteins (Uzzell and Corbin 1971; Yang 1993). Previous studies have suggested that the underlying evolutionary features of proteins with slow evolution rates impose geometrical constraints on three-dimensional (3D) structures and biochemical constraints on enzyme and binding activity preservation (Branden and Tooze 1999). Additionally, the evolutionary properties of protein-protein interactions have been intensively analyzed by computational approaches by sequence comparisons of orthologous genes using high-throughput protein interaction data (Papp et al. 2003; Mintseris and Weng 2005). However, the evolutionary features of protein complexes, which accompany many obligate protein-protein interactions, have not yet been examined. Therefore, the evolutionary properties of protein complexes should be understood in the context of interactions among protein components. To elucidate the molecular evolution mechanisms of protein complexes, we examined the SWR1 complex (SWR1C), a chromatin remodeling complex (CRC).

SWR1C, which catalyzes ATP-dependent replacement of histone H2A with histone variant H2A.Z, has been intensively studied in yeast (Mizuguchi et al. 2004; Zlatanova and Thakar 2008). SWR1C is composed of 14 subunits, and

SWR1, ARP6, SWC6 and SWC2 are core subunits forming the sub-complex. In SWR1C, ARP6 coupled with SWC6 links SWR1 and SWC2, which are involved in histone exchange by directly binding to H2A.Z (Wu et al. 2005; Nguyen et al. 2013). H2A.Z deposition by SWR1C is required for transcriptional regulation, heterochromatic barrier formation, and genome stability in yeast (Kamakaka and Biggins 2005). Interestingly, during the evolution of eukaryotes, histone variant H2A.Z is more highly conserved than H2A; H2A.Zs found in yeast, humans, and Arabidopsis exhibit ~90% sequence similarity (Iouzalen et al. 1996; Deal et al. 2007). Components of SWR1C, especially core subunits (SWR1, ARP6, SWC6, and SWC2), have also been highly conserved during the evolution of eukaryotes, including humans and Arabidopsis. Moreover, the interaction patterns among components of core subunit have been sustained in eukaryotes (Choi et al. 2007; Nguyen et al. 2013).

Until now, developmental roles of SWR1C in higher eukaryotes have mainly been studied in Arabidopsis because mutations in animal SWR1C components cause embryonic lethality (Zlatanova and Thakar 2008). In Arabidopsis, all four homologs of SWR1C core subunits have been reported-these include *PHOTOPERIOD-INDEPENDENT EARLY FLOWERING 1* (*PIE1*), a homolog of *SWR1*; *ACTIN RELATED PROTEIN 6* (*AtARP6*)/*SUPPRESSOR OF FRIGIDA 3* (*SUF3*)/*EARLY IN SHORT DAYS 1* (*ESD1*); *AtSWC6/SERRATED LEAVES AND EARLY FLOWERING* (*SEF*); and *AtSWC2* (Noh and Amasino 2003; Deal et al. 2005; Kandasamy et al. 2005; Martin-Trillo et al. 2006; Choi et al. 2007; Lazaro et al. 2008; March-Diaz and Reyes 2009). Mutations in any of the core subunits cause similar developmental defects such as early flowering, leaf serration, and production of extra petals, short siliques, and weak apical dominance. Functional analyses in Arabidopsis have revealed that PIE1, AtARP6, AtSWC6, and AtSWC2 in SWR1C are required for the deposition of H2A.Z (Choi et al. 2007; Deal et al. 2007).

SWR1C provides an attractive system to study the molecular evolution of protein complexes because the core components are conserved in all eukaryotes and the interactions among components are already known. In particular, ARP6 has several useful features for studying molecular evolution. Actin-related proteins (ARPs, ARP1~ARP10) are a branch of the actin superfamily that share the same structural architecture, known as the "actin fold" (Kabsch and Holmes 1995). In addition, ARP4~ARP10 have specific insertions or deletions, whereas ARP1~ARP3 do not (Muller et al. 2005). The actin fold in ARP6 is organized into four domains (domain 1, 2, 3 and 4) and two insertion regions (insertion region I and II) (Choi et al. 2005; Muller et al. 2005).

In the present study, we focused on the evolutionary constraints of ARP6 proteins that bind to both SWR1s and SWC6s. Our results showed that domains 1 and 3 of AtARP6 interacted with AtSWC6 and PIE1 respectively, and insertion region II in domain 4 provided a nuclear localization signal. Moreover, we proposed that stronger evolutionary constraints were imposed on domains 1 and 3 because they require coevolution with the interaction partners in SWR1C; however, domains 2 and 4 are more flexible to evolutionary changes. The present study provides insight into the molecular evolution of protein complexes.

Results

ARP6s are Evolutionarily Conserved Among Eukaryotes

To provide a framework for ARP6 evolution, we first analyzed the phylogeny based on the amino acid sequences from 13 species embracing three eukaryotic kingdoms: Fungi (budding yeast *Saccharomyces cerevisiae* and fission yeast *Schizosaccharomyces pombe*), Animalia (*Caenorhabditis elegans*, *Drosophila melanogaster, Xenopus laevis, Gallus gallus*, *Mus musculus*, and *Homo sapiens*), and Planta (moss *Physcomitrella patens*, monocots, rice *Oryza sativa,* and maize *Zea mays*, and dicots, *Populus tremula,* and *A. thaliana*). As a result, it was revealed that amino acid sequences among eukaryotic ARP6s were highly conserved (Fig. 1A). Interestingly, the maximum parsimonious phylogenetic analysis of ARP6s from diverse eukaryotes exhibited three distinct branches in the neighbor-joining tree, which exactly recapitulated the phylogenetic tree of the three eukaryotic kingdoms as Fungi, Animalia, and Planta (Fig. 1B). In addition, all aligned ARP6s contained two major insertion regions when they were compared to the relevant actins from the same organisms (Fig. S1). Although the sequences and the length of insertions were variable, the positions of two insertion regions were relatively conserved and the second insertion region was longer than the first insertion region (Fig. 1C; Fig. S1). Therefore, our results showed that ARP6 proteins are conserved during eukaryote evolution, and the two insertion regions in ARP6s are conserved features of ARP6 distinguishable from actin and other ARPs.

ARP6s are Functionally Conserved in Kingdom Planta

To determine the extent of functional conservations of ARP6s, we examined whether *ARP6* genes from diverse organisms (*transARP6s*) could complement *suf3*, a mutant in *AtARP6*. We generated transgenic plants overexpressing *transARP6s* (*ScARP6, SpARP6, CeARP6, DmARP6, MmARP6, PpARP6,* and *OsARP6*) in the *suf3* mutant using a strong constitutive promoter, cauliflower mosaic virus (CaMV) 35S promoter. *35S-AtARP6* in *suf3* was used as a positive control. Since

A

	Identity	Similarity
PtARP6	73%	85%
OsARP6	63%	76%
ZmARP6	61%	78%
PpARP6	50%	68%
GgARP6	35%	56%
MmARP6	34%	55%
XtARP6	34%	54%
HsARP6	33%	55%
DmARP6	33%	52%
CeARP6	31%	52%
SpARP6	31%	53%
ScARP6	26%	43%

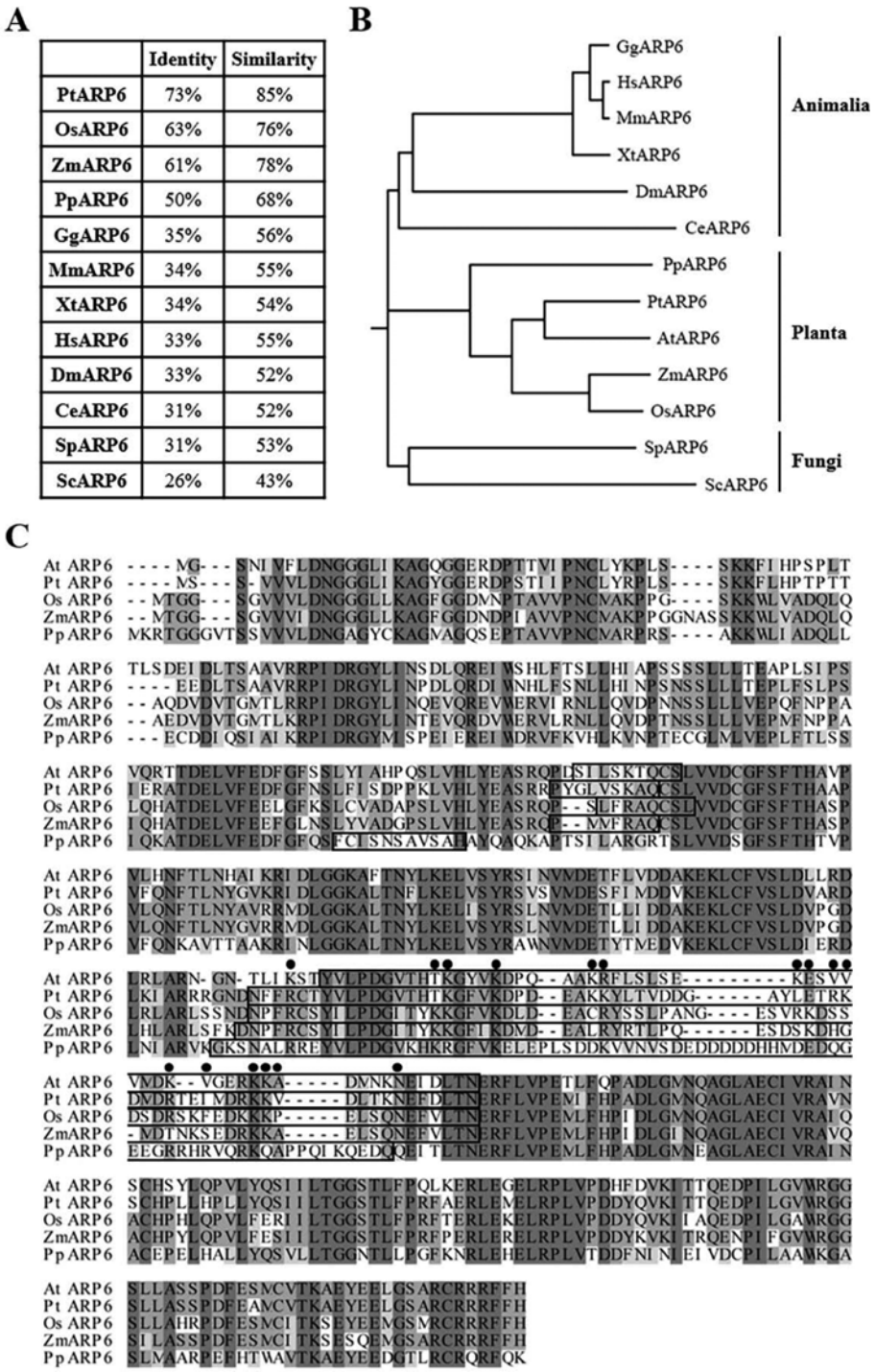
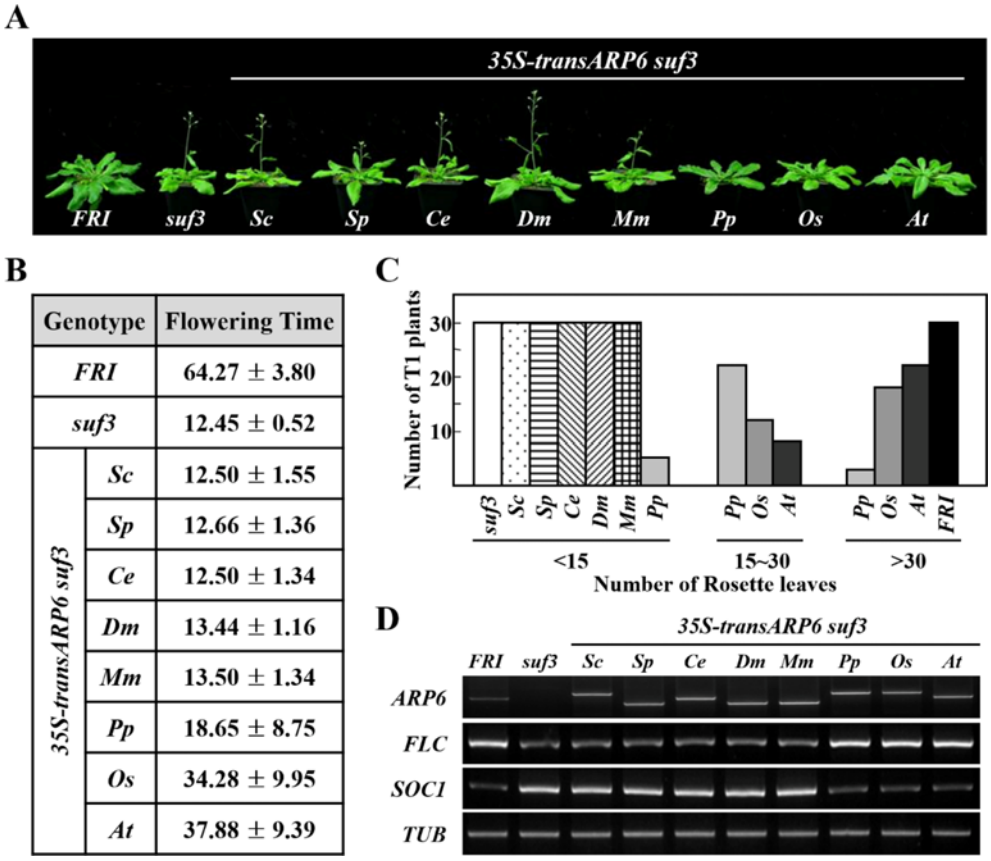


Fig. 1. Phylogenetic relationships among eukaryotic ARP6s. (A) Identities and similarities of amino acid sequences between AtARP6 and other ARP6 homologs; dicots poplar (*Populus tremula, Pt*), monocots rice (*Oryza sativa, Os*) and corn (*Zea mays, Zm*), moss (*Physcomitrella patens, Pp*), Animalia (*Gallus gallus, Gg; Mus musculus, Mm; Xenopus laevis, Xt; Homo sapiens, Hs; Drosophila melanogaster, Dm; Caenorhabditis elegans*, *Ce*), fission yeast (*Schizosaccharomyces pombe, Sp*), and budding yeast (*Saccharomyces cerevisiae, Sc*), (B) Phylogenic tree of ARP6 proteins, (C) Aligned amino acid sequences of planta ARP6s using the CLUSTAL W. The boxed regions represent two insertion regions. Dotted residues represent the putative nuclear localization signals.

suf3 mutations cause pleiotropic phenotypes, we firstly determined whether the early flowering phenotype of *suf3* is rescued by *35S-transARP6s* (Fig. 2). When *35S-AtARP6* was introduced into *suf3*, all the transgenic plants flowered later than *suf3*; the number of rosette leaves at flowering was 64.3 for the wild type (*FRI*), 12.5 for *suf3*, and 37.9 for *35S-*



Genotype		Flowering Time
FRI		64.27 ± 3.80
suf3		12.45 ± 0.52
35S-transARP6 suf3	*Sc*	12.50 ± 1.55
	Sp	12.66 ± 1.36
	Ce	12.50 ± 1.34
	Dm	13.44 ± 1.16
	Mm	13.50 ± 1.34
	Pp	18.65 ± 8.75
	Os	34.28 ± 9.95
	At	37.88 ± 9.39

Fig. 2. Complementation experiments using *transARP6s.* (A) Morphology of wild type (*FRI*), *suf3*, and T1 plants of *35S-transARP6s suf3* grown for 30 d under long day conditions, (B) Flowering time of wild type (*FRI*), *suf3*, and *35S-transARP6s suf3* T1 plants grown under long day conditions. Flowering time was measured by counting the rosette leaf number at the time of bolting (n=30), (C) Distribution of flowering time among *35S-transARP6s suf3* T1 plants grown under long day conditions, (D) The expression of flowering time genes, *FLC* and *SOC1*, in *35S-transARP6s suf3* transgenic plants. RNA was isolated from young leaves of 32-day-old plants. For transgenic lines, RNA was extracted from leaves of at least 10 independent T1 plants. The expression of *TUB* was used as a quantitative control. The expression of each *TransARP6s* (*ARP6*) was checked in its own transgenic plants.

AtARP6 (Fig. 2B). As a result of possible positional effect of the transgene insertions, *35S-AtARP6* showed broad distribution of flowering time, but 76% flowered with more than 30 leaves, 23% flowered with 15–30 leaves, and none of them flowered with less than 15 leaves (Fig. 2C). Similarly, *35S-OsARP6* showed a broad distribution of flowering time but all flowered later than *suf3*, indicating almost complete complementation. For *35S-PpARP6*, most of the transgenics flowered with 15–30 leaves (76%) and a few of them flowered with less than 15 leaves (17%), thus, *PpARP6* also rescued *suf3,* although not completely. In contrast, the overexpression of *ScARP6, SpARP6, CeARP6, DmARP6*, and *MmARP6* could not rescue the early flowering phenotype of *suf3* (Fig. 2A-C). Then, we analyzed the expression levels of *FLOWERING LOCUS C* (*FLC*), a strong repressor of flowering which is a direct target of Arabidopsis SWR1C, and *SUPPRESSOR OF OVEREXPRESSION OF CO1* (*SOC1*), a flowering activator that is negatively regulated by FLC (Choi et al. 2007; Deal et al. 2007). Consistent with the flowering phenotype, *35S-PpARP6, 35S-OsARP6,* and *35S-AtARP6* plants showed increased expression of *FLC* but decreased expression of *SOC1* similar to the wild type, whereas *35S-ScARP6, 35S-SpARP6, 35S-CeARP6, 35S-DmARP6*, and *35S-MmARP6* showed similar expressions of *FLC* and *SOC1* with *suf3* (Fig. 2D). The other phenotypes such as leaf serration, extra petal production, and weak apical dominance were also rescued completely by *35S-AtARP6,* and *35S-OsARP6,* and partially by *35S-PpARP6,* whereas they were not rescued at all by *35S-ScARP6, 35S-SpARP6, 35S-CeARP6, 35S-DmARP6*, and *35S-MmARP6* (Table S1). Taken together, our results showed that moss and rice ARP6s are capable of substituting Arabidopsis ARP6. Considering that moss is the basal taxon in kingdom Planta, it suggests that ARP6s are functionally conserved among plants.

Interaction with Both PIE1 and AtSWC6 is Critical for AtARP6 Function

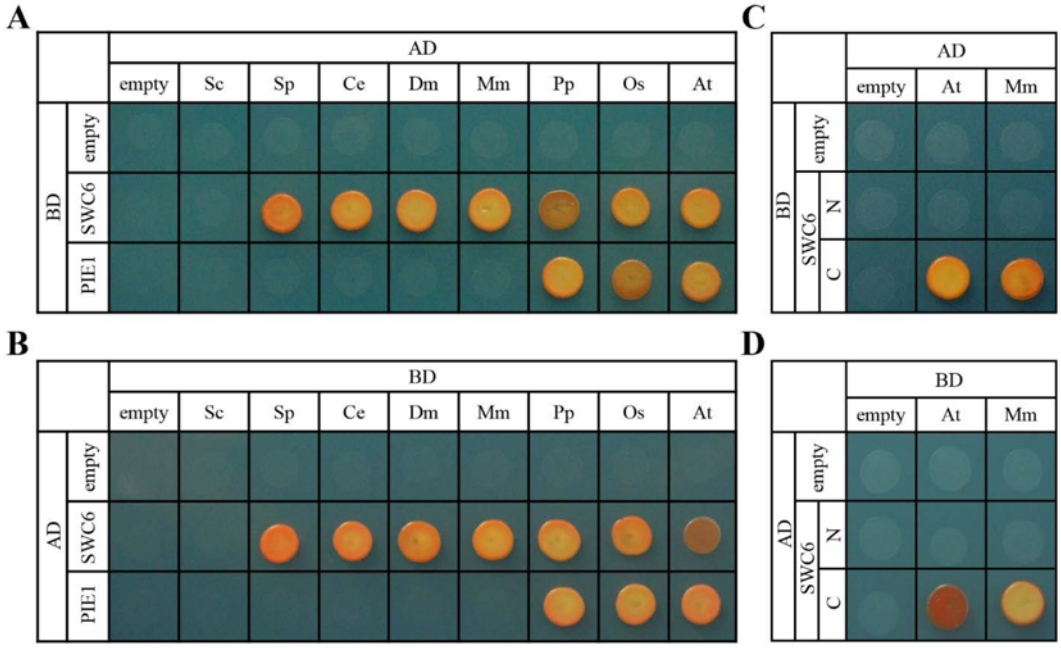


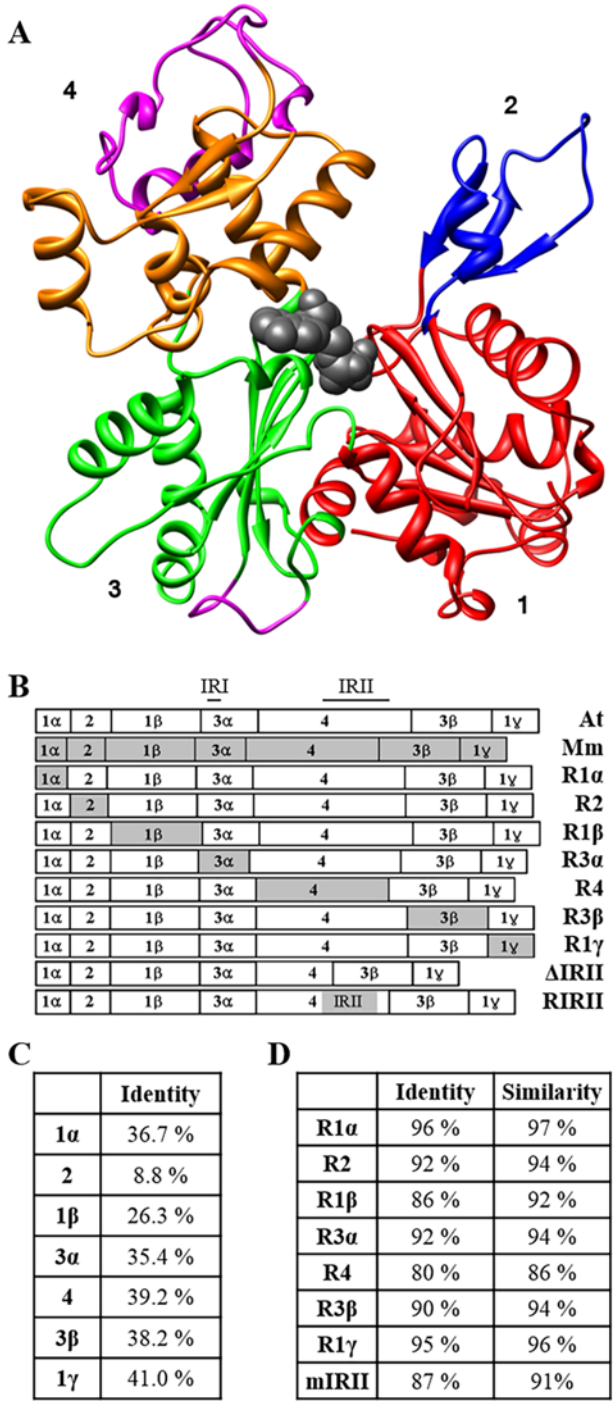
Fig. 3. Interactions between transARP6s, PIE1, and AtSWC6 by yeast two-hybrid. (A) Yeast two-hybrid analysis of protein interaction between transARP6s and AtSWC6 (upper) or PIE1 (bottom). Clones containing each combination of bait and prey were grown on selective medium (-His). Empty vectors, pGBKT7 for BD and pGADT7 for AD, were used as a negative control, (B) The same interaction analysis with (A) but bait and prey were exchanged, (C) Interactions between AtARP6s or MmARP6 and AtSWC6 N-terminus or C-terminus, (D) The same interaction analysis with (C) but bait and prey were exchanged.

The components of Arabidopsis SWR1 complex are interconnected similarly to the yeast SWR1 complex; the C-terminal region of PIE1, a homolog of SWR1, binds to both AtARP6 and AtSWC6, and AtSWC6 interacts with AtARP6 and AtSWC2 (Choi et al. 2007). We reasoned that the functional conservation of transARP6 reflects the conserved interaction with PIE1 and AtSWC6. To confirm this, we performed yeast two-hybrid analyses. The full length cDNA sequences of transARP6s, PIE1, and AtSWC6 proteins were fused to the GAL4 DNA-binding domain (BD) or GAL4 transcriptional activation domain (AD). As expected, only PpARP6 and OsARP6 interacted with both PIE1 and SWC6 (Fig. 3A, B). In contrast, the other transARP6s failed to interact with PIE1, although they interacted with AtSWC6 except ScARP6 (Fig. 3A, B). Considering the functional conservation of moss and rice ARP6s, it supports that the interactions with both AtSWC6 and PIE1 are critical for the function of transARP6s in Arabidopsis.

Next, we addressed the general binding characteristics of AtSWC6 to transARP6s. Similar to ARP6, SWC6, especially the Zn-finger domain at C-terminus, is highly conserved among eukaryotes (Fig. S2; Choi et al. 2007). To address if this highly conserved Zn-finger domain is responsible for the general binding characteristics of SWC6, we divided AtSWC6 to AtSWC6N (1 to 123 aa) and AtSWC6C (124 to 171 aa), and performed yeast two-hybrid analyses. The results showed that C-terminus of AtSWC6 binds to AtARP6 (Fig. 3C, D), indicating that the general binding of AtSWC6 to transARP6s is possibly due to the highly conserved C-terminal Zn-finger domain. This suggests that during the evolution of SWR1 complex, SWC6 has coevolved with ARP6 and the C-terminus of AtSWC6 has been conserved by geometrical constraint to maintain the Zn-finger domain and biochemical constraint to maintain the binding activities to ARP6 (Fig. 7C). Although AtSWC6 successfully binds to transARP6, allowing PIE1, AtSWC6, and transARP6 to come together, the complex does not work functionally because there is no interaction between PIE1 and ARP6. Therefore, the interaction of PIE1 and ARP6 is critical for the biochemical activity of the SWR1 complex.

Four Domains of ARP6 Show Differential Evolution

To gain further insight into the molecular evolution of ARP6 in the SWR1 complex, we analyzed the domain-specific functions of AtARP6 based on the 3D structure. Since ARP6 shares the actin fold and has two insertion regions compared with actin, the 3D structure of AtARP6 was predicted with a homology modeling method using multiple templates (Fig. 4A). Similar to actin, ARP6 has four domains (domains 1, 2, 3 and 4) and insertion regions I and II (IRI and IRII) are



C

	Identity
1α	36.7 %
2	8.8 %
1β	26.3 %
3α	35.4 %
4	39.2 %
3β	38.2 %
1γ	41.0 %

D

	Identity	Similarity
R1α	96 %	97 %
R2	92 %	94 %
R1β	86 %	92 %
R3α	92 %	94 %
R4	80 %	86 %
R3β	90 %	94 %
R1γ	95 %	96 %
mIRII	87 %	91%

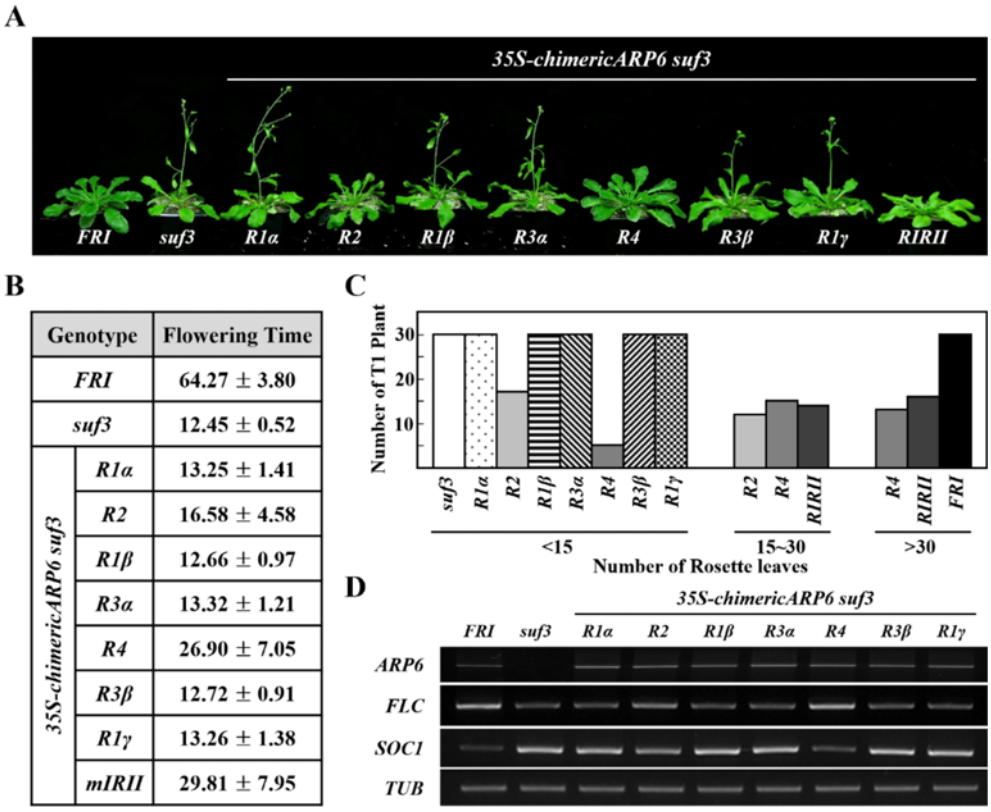
Fig. 4. Three-dimensional structure of ARP6 and constructions for chimericARP6s. (A) The predicted structure of AtARP6 by homology modeling. The four domains of the predicted structure are colored in red (domain 1), blue (domain 2), green (domain 3), and orange (domain 4). The two insertion regions (loop region I in domain 3 and loop region II in domain 4) are colored in magenta, and ATP is represented in gray spheres in its putative binding site, (B) Schematic representations of chimericARP6s swapped between AtARP6 and MmARP6, (C) Amino acid sequence identity of subdomains between Arabidopsis and mouse ARP6s, (D) Identities and similarities of amino acid sequences between AtARP6 and chimericARP6s.

located in domains 3 and 4, respectively (Fig. 4A, B). In the primary structure, ARP6 can be divided into 7 subdomains. Domain 1 is organized with 3 separate subdomains (subdomains 1α, 1β, and 1γ), domain 3 is organized with 2 separate subdomains (subdomains 3α, 3β), and domains 2 and 4 are organized with single subdomains (subdomains 2 and 4, respectively) (Fig. 4B).

To investigate if any domains in ARP6 are functionally conserved during eukaryotic evolution, we swapped each subdomain of *AtARP6* with the corresponding subdomain of *MmARP6* for the complementation test. We constructed a series of *chimericARP6s* (Fig. 4B) and produced transgenic plants overexpressing *chimericARP6s* in *suf3* background. Then, we determined whether the various mutant phenotypes of *suf3* were rescued by *35S-chimericARP6s*. The results showed that *R4 chimericARP6* had the strongest rescue effect, although it had the lowest identity (Fig. 4D; Fig. 5). Most *35S-R4* transgenic plants flowered with more than 15 leaves and only 7% flowered with less than 15 leaves. Consistently, the representative lines showed increased *FLC* expression and decreased *SOC1* expression (Fig. 5D). Similarly, *35S-R4* showed a reduced number of extra petals and increased silique length (Table S2), indicating the other phenotypes of *suf3* were also rescued. *R2 chimericARP6,* which shows 92% identity with *AtARP6*, also showed a strong rescue effect in both flowering time and other phenotypes of *suf3* (Fig. 4D; Fig. 5; Table S2). The phenotypic observation showed that *R2* partially rescues and *R4* completely rescues various mutant phenotypes of *suf3,* which suggests that the R2 domain of mouse partially substitutes that of Arabidopsis and the R4 domain of mouse fully substitutes that of Arabidopsis. In contrast, *R1α, R1β, R1γ, R3α,* and *R3β chimericARP6s,* which have more than 90% identity with *AtARP6*, failed to rescue *suf3* mutant (Fig. 4D; Fig. 5; Table S2). Especially, since *R1α* and *R1γ chimericARP6s* showing more than 95% identity failed to rescue *suf3* suggests that domain-specific identity, and not overall identity, is more critical for ARP6 function. Together, our results demonstrate the importance of domains 1 and 3 for kingdom-specific functions and domains 2 and 4 for generic functions.

Domains 1 and 3 of AtARP6 Interact with AtSWC6 and PIE1, Respectively

Having established that AtARP6 binds to both PIE1 and AtSWC6, we determined which domain is responsible for each binding. We performed the yeast two-hybrid analysis using the chimericARP6s and PIE1, AtSWC6. If a certain domain is necessary for binding, the swapped chimericARP6 for that domain should fail to interact with the binding partner. The result showed that R1α and R1β failed to bind



Genotype		Flowering Time
FRI		64.27 ± 3.80
suf3		12.45 ± 0.52
35S-chimericARP6 suf3	*R1α*	13.25 ± 1.41
	R2	16.58 ± 4.58
	R1β	12.66 ± 0.97
	R3α	13.32 ± 1.21
	R4	26.90 ± 7.05
	R3β	12.72 ± 0.91
	R1γ	13.26 ± 1.38
	mIRII	29.81 ± 7.95

Fig. 5. Complementation experiments using *chimericARP6s.* (A) Morphology of wild type (*FRI*), *suf3*, and *35S-chimericARP6s suf3* T1 plants grown for 30 d under long day conditions, (B) Flowering time of wild type (*FRI*), *suf3*, and *35S-chimericARP6s suf3* T1 plants grown under long day conditions. Flowering time was measured by counting the rosette leaf number at the time of bolting (n=30), (C) Distribution of flowering time among *35S-chimericARP6s suf3* T1 plants grown under long day conditions, (D) The expression of flowering time genes, *FLC* and *SOC1*, in *35S-chimericARP6s suf3* transgenic plants. RNA was isolated from young leaves of 32-day-old plants. For transgenic lines, RNA was extracted from leaves of at least 10 independent T1 plants. The expression of *TUB* is used as a quantitative control. The expression of each *chimericARP6s* (*ARP6*) was checked in its own transgenic plants.

to AtSWC6 although they interacted with PIE1 (Fig. 6A, B). In contrast, R3α and R3β failed to bind to PIE1 but successfully interacted with AtSWC6. Our results suggest that domain 1 is responsible for binding to AtSWC6 and domain 3 is responsible for binding to PIE1 (Fig. 7B). It is noteworthy that the domains with kingdom-specific functions play a critical role in the interaction with binding partners. It also indicates that ARP6, SWC6, and PIE1 have coevolved to maintain their interactions due to biochemical constraints (Fig. 7C).

It was unexpected that chimeric ARP6s failed to interact with AtSWC6, because the full length protein of MmARP6 interacts with AtSWC6 (Fig. 3; Fig. 6A, B). In addition, although R1γ successfully bound to both PIE1 and AtSWC6 the transgene failed to rescue the *suf3* mutation (Fig. 5; Fig. 6A, B). Thus, it is possible that the replacement of subdomain 1 causes a malfunction of ARP6 through the alteration of the overall 3D structure.

ARP6 Domain 4 has a Nuclear Localization Signal

The domain swapping analysis indicated that domain 4 has a generic function. Since it has insertion region II (IRII), we determined if IRII is responsible for the generic function. Compared with actin, ARP6 localizes in the nucleus (Choi et al. 2005) and the IRII sequence is its most distinct feature (Fig. S1). From the visual analysis, we recognized putative nuclear localization signals in IRII (Fig. 1C). To determine if IRII is required for nuclear localization, we generated a fusion construct of GFP and AtARP6 with a deletion in IRII (AtARP6ΔIRII-GFP). The transient expression of this construct in Arabidopsis protoplasts showed that AtARP6ΔIRII-GFP proteins were retained in cytoplasm (Fig. 6E), which supports that IRII is necessary for nuclear localization. Then, we checked if the replacement of Arabidopsis IRII with mouse IRII effected the nuclear localization of AtARP6. The protoplast transient expression of AtARP6:RIRII-GFP, which has a replacement of IRII, showed that this protein localizes in the nucleus similar to AtARP6-GFP (Fig. 4B; Fig. 6F). Therefore, our results suggest that IRII provides a nuclear localization signal and this function is conserved among plant and animal kingdoms.

Finally, we checked if *AtARP6:RIRII*, a *chimericARP6*,

could complement the *suf3* mutant (Fig. 4; Fig. 5). We postulated that if the function of IRII is restricted to nuclear localization, *AtARP6:RIRII,* whose protein localizes in the nucleus, could functionally substitute *AtARP6* despite its low identity. Indeed, the result showed that *35S-AtARP6:RIRII* completely rescued the *suf3* mutation (Fig. 5). Therefore, this result indicates that IRII is replaceable as long as it retains a nuclear localization signal, thus is under low evolutionary constraint compared to the other regions. This also explains why *35S-R4* rescues *suf3* so completely although the amino acid sequence identity is relatively low.

Evolution Rates of Domains 1 and 3 are Lower than That of Domain 4

Our results suggested that the four domains of ARP6 are under different evolutionary constraints, thus the evolution rate for each domain is expected to be different. Since domains 1 and 3 participate in protein-protein interactions, whereas domain 4 provides a nuclear localization signal, we hypothesized that domains 1 and 3 evolve more slowly than domain 4. To test this hypothesis, we calculated the ratio of the number of nonsynonymous substitutions (Ka) to that of synonymous substitutions (Ks) as an indication of the evolution rate in pairwise comparisons between Arabidopsis and poplar ARP6s (Nei and Gojobori 1986). As shown in Fig. 7A, domain 4 showed the highest Ka/Ks value, indicating the greatest degree of amino acid diversification and faster evolution than the other domains. The subdomains of 1γ, 3α, and part of 1β showed the lowest Ka/Ks value, indicating that these regions are highly conserved (Fig. 7A). These results strongly support our hypothesis that the four domains of ARP6 evolve at different rates; domains 1 and 3 evolve more slowly than domain 4.

Discussion

The molecular evolution of the protein complex must be understood in the context of interaction among its components. To gain insight into the molecular evolution of protein complex, we investigated the evolution of ARP6, a component of SWR1 complex, in Arabidopsis. We experimentally examined the functionalities of transARP6s and chimericARP6s *in vivo*. Our results revealed the precise domain-specific functions of ARP6 and the evolutionary features of ARP6.

Functional Conservation of ARP6 in the Plant Kingdom

Interspecies complementation experiments showed that moss and rice *ARP6s* could rescue the phenotype of *suf3* mutations (Fig. 2). Moss is the basal taxon of kingdom Planta, and rice and Arabidopsis belong to different taxa in angiosperms, that is, monocot and eudicot, respectively. It suggests that ARP6 and SWR1C have been functionally conserved during the past 450~700 million years of Kingdom Planta evolution (Hedges 2002). Similar experiments were performed using orthologs of *LEAFY* (*LFY*), a plant specific transcription factor (Maizel et al. 2005). However, moss LFY was not functional in Arabidopsis although PpLFY and PpARP6 showed similar amino acid sequence conservation, 49% and 50% identity, respectively, compared to Arabidopsis orthologs. This suggests that amino acid sequence homology does not directly reflect functional conservation.

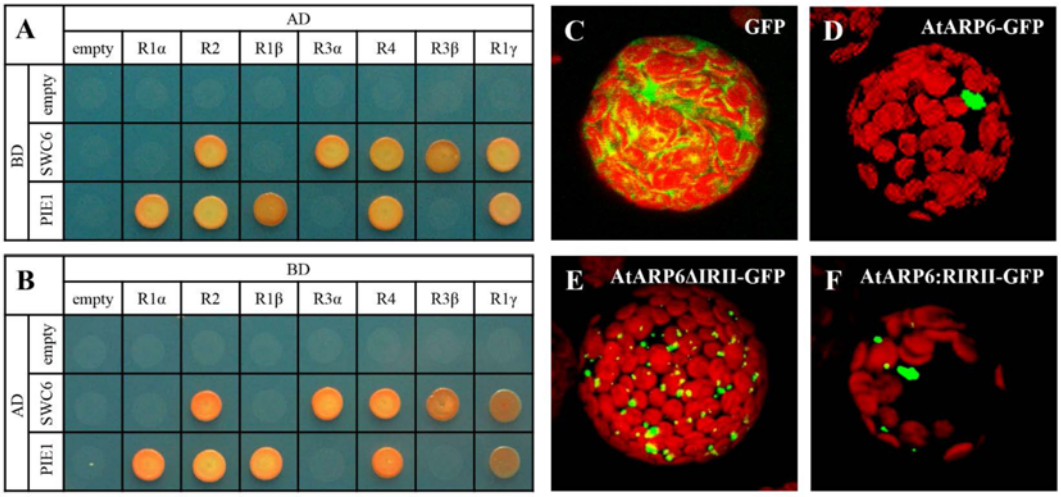


Fig. 6. Domain-specific functions of ARP6. (A and B) Interactions between chimericARP6s and AtSWC6 or PIE1 by yeast two-hybrid assays, (C-F) Subcellular localizations of GFP (C), AtARP6-GFP (D), AtARP6ΔIRII-GFP (E), and AtARP6:RIRII-GFP (F). ΔIRII means the deletions of insertion region II (IRII), and RIRII represents the substitution of the mouse version IRII. GFP fluorescence is shown in green, and chloroplasts are in red.

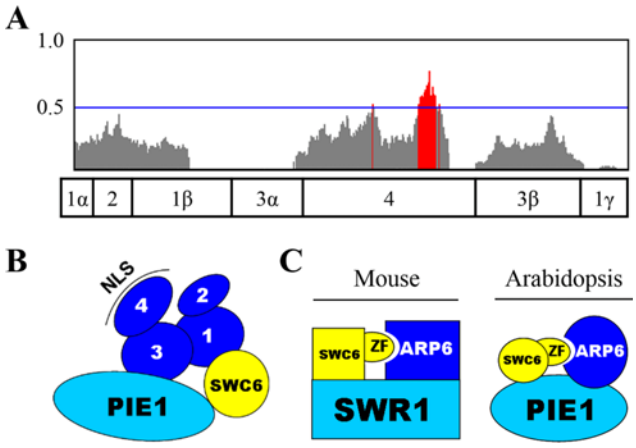


Fig. 7. Molecular evolution of ARP6 in SWR1 complex. (A) Ka/Ks ratio of ARP6 comparing Arabidopsis and popular proteins. Ka/Ks ratios were calculated using a sliding window analysis with a window size of 90 bp and sliding size of 3 bp, (B) Model of interaction of 4 domains of ARP6 with PIE1 and SWC6, (C) Model of interaction of SWC6 and ARP6 through the zinc finger region.

Consistent with complementation results, only plant ARP6s could bind to PIE1 (Fig. 3), suggesting that the interaction capability with PIE1 is critical for ARP6 function. In addition, chimericARP6s not interacting with AtSWC6 failed to rescue *suf3* mutants (Fig. 5; Fig. 6A, B). These results suggest that ARP6 has evolved to maintain interactions with both PIE1 and SWC6, which possibly provides a strong evolutionary constraint during the evolution of Kingdom Planta. Thus, our study experimentally supports the concept that coevolution between two interacting proteins occurs to maintain binding specificity (Moyle et al. 1994).

Domain-specific Roles of ARP6 in the SWR1 Complex

In this study, the domain-specific functions of ARP6 were revealed through swapping analysis (Fig. 4; Fig. 5; Fig. 6). Domains 1 and 3 are required for interaction with AtSWC6 and PIE1 respectively, and domain 4 provides a nuclear localization signal (Fig. 6). However, the function of domain 2 was not revealed. Since AtARP6 interacts only with PIE1 and AtSWC6 in Arabidopsis, domain 2 is not likely involved in protein-protein interaction.

Domains 1 and 3 are composed of 3 and 2 segments, respectively, which are dispersed in the primary structure of ARP6. For domain 3, the substitution of either subdomain (3α or 3β) disrupted interaction with PIE1. However, the substitutions of three subdomains of domain 1 (1α, 1β, and 1γ) caused unexpected results. Although full length proteins of MmARP6 successfully bind to AtSWC6, the substitutions of 1α and 1β failed to interact with AtSWC6 (Fig. 3; Fig. 6A, B). This indicates that an intradomain interaction also imposes evolutionary constraint so that the proper interaction of 1α and 1β is maintained to sustain the correct structure for the interaction with SWC6.

Interestingly, although R1γ has the capacity to bind to both AtSWC6 and PIE1, it could not rescue *suf3* mutations (Fig. 5; Fig. 6A, B). The replacement of 1γ possibly affects the biochemical function of ARP6, such as ATP incorporation. This is consistent with the fact that the 1γ region is highly conserved based on the Ka/Ks ratio (Fig. 7A), thus a minor change can cause a dramatic effect.

We observed that each domain of ARP6s has significantly different Ka/Ks ratios, and domains involved in protein interaction showed low Ka/Ks ratios (Fig. 7A). Thus, interacting domains in ARPs are more evolutionarily conserved than non-interacting domains, indicating that the binding property of a protein can act as an evolutionary constraint. In addition, this study provided evidence that ARP6, SWC6, and PIE1 have coevolved to maintain their interactions. For example, fungi and animal ARP6s are incapable of binding to PIE1, an Arabidopsis ortholog of SWR1, although ARP6 interacts with SWR1 in their own species (Fig. 3A, B).

All nuclear ARPs (ARP4~9) have specific insertions or deletions unlike cytoplasmic ARPs (Blessing et al. 2004; Muller et al. 2005). Although nuclear ARPs function as the components of chromatin remodeling complexes (CRCs) or histone acetyltransferase (Olave et al. 2002; Shen et al. 2003; Mizuguchi et al. 2004), the function of insertion regions has never been examined or proposed so far. In this study, the insertion region II of ARP6 placed in domain 4 is required for nuclear localization. Similar putative nuclear localization signals are also detected in domain 4 of AtARP4 and AtARP5 (data not shown). Based on these observations, it is proposed that the insertion of the nuclear localization signal in domain 4 provides ARPs new functions, such as a component of CRCs.

In conclusion, some principles for the molecular evolution of protein complexes have been disclosed in this study: 1) the functional conservation of the protein complex is determined not by sequence homology, but by their roles in the complex; 2) the interacting domains have higher evolutionary constraints than the non-interacting domains; and 3) the interacting domains coevolve with the interacting partners to preserve their interactions.

Materials and Methods

Plant Materials and Growth Conditions

The wild type, named *FRI,* was the *A. thaliana* Col:*FRI*^{SF2} strain, and the *suf3* mutant used in this study was *suf3-1* (Choi et al. 2005). Plants were grown in long day conditions (16 h light/8 h dark) at 22°C with 60% relative humidity. Flowering time was determined by counting

the number of primary rosette leaves when plants bolted.

Plasmid Construction

To construct *35S-transARP6s*, full length *transARP6* cDNAs were amplified by PCR, and were introduced into a binary vector myc-pBA. The *chimericARP6s* between *AtARP6* and *MmARP6* were generated by overlapping PCR, and they were cloned into a binary vector myc-pBA to produce *35S-chimericARP6s*. For yeast two-hybrid assays, *transARP6s*, *chimericARP6s*, *PIE1*, *AtSWC6*, *AtSWC6-N,* and *AtSWC6-C* were cloned into pGADT7 and pGBKT7 vectors. For cellular localization experiments, *AtARP6ΔIRII* and *AtARP6:RIRII* were amplified except stop codon by overlapping PCR and cloned into p326-GFP. Primer sequences and details of cloning procedures are available as Supporting information (Supplementary Experimental Procedures).

Prediction of 3D Structure of AtARP6

The 3D structure of AtARP6 was predicted with a template-based modeling method using MODELLER 9v2 (Sali 1995). The experimental structures of five homologous proteins (PDB ID: 1qz5A, 1c0gA, 1yagA, 1k8kB, and 1dx4A) were used as templates in the homology modeling with MODELLER. The structures of the two insertion regions (loop region I: 145P-153Q and loop region II: 246L-290N) were generated by further loop modeling using a method that employs fragment assembly and analytical loop closure (Coutsias et al. 2004; Lee et al. 2008). The model with the best DOPE score was used as the template for loop modeling (Shen and Sali 2006). The candidate loop structures were generated by a fragment assembly method satisfying the connectivity requirement of the loop regions with the rest of the structure (Coutsias et al. 2004; Lee et al. 2008). Among candidates, the loop structures that did not interfere with the interface between the loop region I by domain 1 and 3 and the loop region II by domain 2 and 4 were firstly chosen, and the loop structure with the best DFIRE-B score was selected (Zhou and Zhou 2002). The putative ATP binding site was predicted by a structure alignment of template protein structures with a bound ATP, 1qz5, using the MatchMaker option in CHIMERA (Pettersen et al. 2004).

Analysis of Gene Expression

Total RNA was prepared using TRIZOL (Sigma-Aldrich, St Louis, USA). Primers for *FLC, SOC1,* and *TUB2* and the RT-PCR conditions have been described previously (Choi et al. 2007).

Yeast Two-Hybrid Assay

The yeast strains and vectors were obtained from Clontech, and the yeast two-hybrid assay was performed as previously described (Choi et al. 2011).

Protoplast Transient Expression Assay

For the isolation and transformation of protoplasts, rosette leaves from wild type plants were used as described previously (Choi et al. 2007). GFP-fusion constructs were prepared by the QIAGEN Plasmid Midi Kit, and about 30 μg plasmid DNAs were transformed to protoplasts. Protoplasts were observed 12 hours after incubation at 22°C in the dark.

Calculating the Ka/Ks Value

Ka/Ks ratios were calculated using a sliding window analysis with window size of 90 bp and sliding size of 3 bp by a WSPMaker (http://wspmaker.kobic.kr:8080/wspmaker/).

Acknowledgements

This work was supported by National Research Foundation of Korea (NRF) Grant from the Korean Government (No. 2014023132).

Author's Contribution

IL supervised the research; CP and KC designed and performed the experiments; DL and CS contributed to predict 3D structure; Manuscript was prepared by IL, ES, and CP. All the authors agreed on the contents of the paper and post no conflicting interest.

Supporting Information

Fig. S1. Alignments of ARP6 and actin from each organism.
Fig. S2. Alignments of SWC6 between Arabidopsis and mouse.
Table S1. Morphological phenotype analysis of *35S-transARP6s suf3.*
Table S2. Morphological phenotype analysis of *35S-chimericARP6s suf3.*
Table S3. Primers used for plasmid construction in this study.
Supplementary Experimental Procedures

References

Blessing CA, Ugrinova GT, Goodson HV (2004) Actin and ARPs: action in the nucleus. Trends Cell Biol 14:435–442

Branden C, Tooze J (1999) Introduction to Protein Structure Garland Publishing, New York

Choi K, Kim J, Hwang HJ, Kim S, Park C, Kim SY, Lee I (2011) The FRIGIDA Complex Activates Transcription of *FLC*, a Strong Flowering Repressor in *Arabidopsis*, by Recruiting Chromatin Modification Factors. Plant Cell 23:289–303

Choi K, Kim S, Kim SY, Kim M, Hyun Y, Lee H, Choe S, Kim SG, Michaels S, Lee I (2005) *SUPPRESSOR OF FRIGIDA3* Encodes a Nuclear ACTIN-RELATED PROTEIN6 Required for Floral Repression in Arabidopsis. Plant Cell 17:2647–2660

Choi K, Park C, Lee J, Oh M, Noh B, Lee I (2007) *Arabidopsis* homologs of components of the SWR1 complex regulate flowering and plant development. Development 134:1931–1941

Coutsias EA, Seok C, Dill KA (2004) A Kinematic View of Loop Closure. J Comput Chem 25:510–528

Deal RB, Kandasamy MK, McKinney EC, Meagher RB (2005) The Nuclear Actin-Related Protein ARP6 Is a Pleiotropic Developmental Regulator Required for the Maintenance of *FLOWERING LOCUS C* Expression and Repression of Flowering in *Arabidopsis*. Plant Cell 17: 2633–2646

Deal RB, Topp CN, McKinney EC, Meagher RB (2007) Repression of Flowering in Arabidopsis Requires Activation of *FLOWERING LOCUS C* Expression by the Histone Variant H2A.Z. Plant Cell 19:74–83

Hedges SB (2002) THE ORIGIN AND EVOLUTION OF MODEL ORGANISMS. Nat Rev Genet 3:838–849

Iouzalen N, Moreau J, Mechali M (1996) H2A.ZI, a new variant histone expressed during Xenopus early development exhibits several distinct features from the core histone H2A. Nucleic Acids Res 24:3947–3952

Kabsch W, Holmes KC (1995) PROTEIN MOTIFS 2. The actin fold.

FASEB J 9:167–174
Kamakaka RT, Biggins S (2005) Histone variants: deviants? Genes Dev 19:295–310
Kandasamy MK, Deal RB, McKinney EC, Meagher RB (2005) Silencing the nuclear actin-related protein AtARP4 in Arabidopsis has multiple effects on plant development, including early flowering and delayed floral senescence. Plant J 41:845–858
Lazaro A, Gomez-Zambrano A, Lopez-Gonzalez L, Pineiro M, Jarillo JA (2008) Mutations in the Arabidopsis SWC6 gene, encoding a component of the SWR1 chromatin remodelling complex, accelerate flowering time and alter leaf and flower development. J Exp Bot 59:653–666
Lee D-S, Seok C, Lee J (2008) Protein Loop Modeling using Fragment Assembly. J Korean Phys Soc 52:1137–1142
MA CD (1859) On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life, London
Maizel A, Busch MA, Tanahashi T, Perkovic J, Kato M, Hasebe M, Weigel D (2005) The Floral Regulator LEAFY Evolves by Substitutions in the DNA Binding Domain. Science 308:260–263
March-Diaz R, Reyes JC (2009) The Beauty of Being a Variant: H2A.Z and the SWR1 Complex in Plants. Molecular plant 2:565–577
Martin-Trillo M, Larazo A, Poethig RS, Gomez-Mena C, Pineiro MA, Martinez-Zapater JM, Jarillo JA (2006) *EARLY IN SHORT DAYS 1* (*ESD1*) encodes *ACTIN-RELATED PROTEIN 6* (*AtARP6*), a putative component of chromatin remodelling complexes that positively regulates *FLC* accumulation in *Arabidopsis* Development 133:1241–1252
Mintseris J, Weng ZP (2005) Structure, function, and evolution of transient and obligate protein-protein interactions. Proc Natl Acad Sci USA 102:10930–10935
Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C (2004) ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. Science 303:343–348
Moyle WR, Campbell RK, Myers RV, Bernard MP, Han Y, Wang X (1994) Co-Evolution of Ligand-Receptor Pairs. Nature 368: 251–255
Muller J, Oma Y, Vallar L, Friederich E, Poch O, Winsor B (2005) Sequence and Comparative Genomic Analysis of Actin-related Proteins. Mol Biol Cell 16:5736–5748
Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426
Nguyen VQ, Ranjan A, Stengel F, Wei D, Aebersold R, Wu C, Leschziner AE (2013) Molecular Architecture of the ATP-Dependent Chromatin-Remodeling Complex SWR1. Cell 154: 1220–1231
Noh YS, Amasino RM (2003) *PIE1*, an ISWI family gene, is required for *FLC* activation and floral repression in Arabidopsis. Plant Cell 15:1671–1682
Olave IA, Reck-Peterson SL, Crabtree GR (2002) NUCLEAR ACTIN AND ACTIN-RELATED PROTEINS IN CHROMATIN REMODELING. Annu Rev Biochem 71:755–781
Papp B, Pal C, Hurst LD (2003) Dosage sensitivity and the evolution of gene families in yeast. Nature 424:194–197
Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—A Visualization System for Exploratory Research and Analysis. J Comput Chem 25:1605–1612
Sali A (1995) MODELER: Implementing 3D protein modeling, mc^2 Molecular Simulations Inc., Burlington, MA
Shen M-Y, Sali A (2006) Statistical potential for assessment and prediction of protein structures. Protein Sci 15:2507–2524
Shen X, Ranallo R, Choi E, Wu C (2003) Involvement of Actin-Related Proteins in ATP-Dependent Chromatin Remodeling. Mol Cell 12:147–155
Uzzell T, Corbin KW (1971) Fitting Discrete Probabillity Distributions to Evolutionary Evcints. Science 172:1089–1096
Wu WH, Alami S, Luk E, Wu CH, Sen S, Mizuguchi G, Wei D, Wu C (2005) Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. Nat Struct Mol Biol 12:1064–1071
Yang ZH (1993) MAXIMUM-LIKELIHOOD-ESTIMATION OF PHYLOGENY FROM DNA-SEQUENCES WHEN SUBSTITUTION RATES DIFFER OVER SITES. Molecular Biology and Evolution 10:1396–1401
Zhou HY, Zhou YQ (2002) Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. Protein Sci 11:2714–2726
Zlatanova J, Thakar A (2008) H2A.Z: View from the top. Structure 16:166–179