SOG1: a master regulator of the DNA damage response in plants

Kaoru Okamoto Yoshiyama

1Life Sciences, Kyoto Sangyo University, Kamigamo-Motoyama, Kitaku, Kyoto 603-8555, Japan
2Japan Society for the Promotion of Science, 5-3-1 Koujimachi, Chiyodaku, Tokyo 102-0083, Japan

(Received 8 February 2015, accepted 26 March 2015; J-STAGE Advance published date: 26 November, 2015)

The DNA damage response (DDR) is a critical mechanism to maintain the genome stability of an organism upon exposure to endogenous and exogenous DNA-damaging factors. The DDR system is particularly important for plants as these organisms, owing to their intrinsic immobility, are inevitably exposed to environmental stress factors, some of which induce DNA damage. Arabidopsis thaliana has orthologs of several DDR factors that are present in animals; however, some of the important animal regulators, such as the tumor suppressor p53 and the DDR kinases CHK1 and CHK2, have not been found in plants. These observations imply a unique DDR system in plants. The present review focuses on recent advances in our understanding of the DDR in A. thaliana and, in particular, on the function and role of SUPPRESSOR OF GAMMA RESPONSE 1 (SOG1), a plant-specific transcription factor that regulates the DDR. The most obvious response to DNA damage in A. thaliana is a rapid and robust change in the transcriptional regulation of numerous genes, in which SOG1 is an essential regulatory factor. Mutation of SOG1 causes various defects in the activation of cell cycle arrest, programmed cell death, and endoreduplication in response to DNA damage. These observations indicate that SOG1 is a master regulator of the DDR. Phylogenetic analyses of SOG1 reveal that orthologs of this crucial transcription factor are present not only in angiosperms but also in gymnosperms, suggesting that the SOG1 system is conserved across spermatophytes. Finally, future prospects for SOG1 research are also discussed.

Key words: ATM, DNA damage, genome stability, p53, plants

INTRODUCTION

The nuclear genome in each of our cells accumulates thousands of DNA lesions every day (Sancar et al., 2004). Although most of these lesions occur as byproducts of normal cell metabolism, they are also induced by radiation and toxic environmental chemicals (Jackson and Bartek, 2009; Ciccia and Elledge, 2010). DNA damage interferes with DNA replication and transcription, resulting in mutations and chromosomal aberrations. Genome integrity is constantly monitored by a sophisticated mechanism called the DNA damage response (DDR) (Ciccia and Elledge, 2010). The DDR is a signal transduction pathway that senses DNA damage and transduces this signal to downstream factors to activate various pathways, including cell cycle arrest (checkpoint), DNA repair and apoptosis. The kinases ataxia telangiectasia mutated (ATM) and Rad3-related (ATR) are the major regulators of the DDR (Cimprich and Cortez, 2008). ATM responds to DNA double-strand breaks (DSBs), while ATR responds to stalled replication forks and single-stranded DNA structures. The DDR in plants is less well studied than it is in animals and yeast. This may be because mutations of DDR factors cause various genetic diseases in humans, and understanding the causes of these diseases is important for advances in their prevention and treatment (Jackson and Bartek, 2009). This is not the case in plants, and the importance of the DDR in plants has therefore been overlooked. While animals and land plants are both multicellular eukaryotes, plants are sessile and so cannot escape from environmental stresses such as drought, high and low temperatures, and high light intensity. These stresses induce the production of reactive oxygen species (ROS), which lead to DNA damage (Baxter et al., 2014). Furthermore, while sunlight is required for photosynthesis to convert solar energy to
chemical energy by the plant’s chloroplasts, the ultraviolet component of sunlight damages the genome, and photosynthesis itself generates ROS (Foyer and Shigeoka, 2011). Aluminum and boron taken up from the soil also act as DNA-damaging agents in plants (Sakamoto et al., 2011; Nezames et al., 2012). In addition to abiotic stresses, plants are constantly exposed to a wide range of pathogens in nature, and biotic stress can also trigger the release of ROS (Laloi et al., 2004). Given these differences between animals and plants, it is important to understand the DDR in plants, as plants have probably evolved different strategies to minimize the effects of harmful genotoxic agents and infections (Jones and Dangl, 2006).

Orthologs of animal ATM and ATR have been found in the Arabidopsis thaliana genome. The A. thaliana atm mutant is sensitive to DSBI-inducing agents (e.g., ionizing radiation and methyl methanesulfonate), and the atr mutant is sensitive to replication stress-inducing agents (e.g., aphidicolin and hydroxyurea) (Garcia et al., 2003; Culligan et al., 2004). These phenotypes are similar to those of the corresponding animal mutants, indicating that the role of these proteins is conserved in plants. Under non-stress conditions, disruption of ATM in humans leads to a genomic instability disorder, whereas disruption of ATR is lethal (Xu et al., 1996; Brown and Baltimore, 2000). In contrast, the A. thaliana atm mutant is phenotypically identical to the wild type except for partial sterility, and the atr mutant is viable and fertile (Garcia et al., 2003; Culligan et al., 2004). These phenotypic differences between animals and plants imply that ATM and ATR in plants have distinct functions from their animal counterparts.

Intriguingly, a relationship between the plant immune response and DNA damage has been reported recently. Yan et al. (2013) showed that the plant immune hormone salicylic acid triggers DNA damage in the absence of a genotoxic agent, and the DDR factors ATR and RAD17 are required for effective immune responses. Furthermore, microbial pathogens reportedly induce DSBs in A. thaliana genomic DNA (Song and Bent, 2014). These observations indicate that DNA damage serves as a signal to induce immune responses and that the DDR and defense mechanisms are closely related. Therefore, the plant DDR system is crucial not only for maintaining genome stability, but also for coping with microbial pathogens.

THE DNA DAMAGE RESPONSE IN A. THALIANA

A) Induction of transcriptional response Wild-type A. thaliana activates various responses upon DNA damage. The most obvious response is that the expression of more than 100 genes changes, in five-day-old whole seedlings, within 1.5 h after gamma irradiation (fold change cutoff ≥ 4, q-value < 0.05) (Culligan et al., 2006). Upregulated genes include DNA repair genes such as BRCA1, RAD51 and PARP-1/2, indicating that the mechanism for maintaining genome stability is activated. Downregulated genes include cell cycle-related genes such as CDKB2;1 (required for progression from G2 to M) and KNOLLE (required for cytokinesis), suggesting that cell cycle arrest is induced in response to irradiation. This immediate transcriptional response to gamma irradiation is governed by ATM, but not by ATR (Culligan et al., 2006). In contrast to the robust transcriptional response in A. thaliana, microarray studies in human cells show that the magnitude of the change in the expression of irradiation-responsive genes is not as great (Rieger and Chu, 2004). Furthermore, few DNA repair genes are transcriptionally induced by gamma irradiation in human cells. For example, BRCA1, RAD51 and PARP-1/2 are strongly induced in plants by gamma irradiation, but not in human cells (Rieger and Chu, 2004). These differences show that the regulation mechanisms of these DNA repair genes are different in plants and animals. As human ATM also plays an important role in regulating gene expression in response to DSBs (Heinloth et al., 2003; Elkon et al., 2005), the transcriptional regulation mediated by ATM seems to be conserved between animals and plants. Based on the transcriptional response in A. thaliana, it is clear that both DNA repair and cell cycle arrest are immediately activated in response to DNA damage.

B) Programmed cell death in stem cells Apoptosis induced by DNA damage is an important mechanism of tumor suppression, and is a common DNA damage response in animals (Nobury and Zhivotovsky, 2004). In A. thaliana, cell death was observed in root and shoot stem cells upon treatment with radiomimetic drugs and X-rays (Fulcher and Sablowski, 2009). The pattern of cell death as observed by electron microscopy did not show apoptotic features (cell shrinkage, peripheral chromatin condensation and nuclear fragmentation), but instead showed autolytic features (numerous vesicles, intact nuclei and lost internal organization), similar to those observed in plant programmed cell death during development. ATM and ATR kinases are required for the death of root stem cells (Fulcher and Sablowski, 2009), suggesting that stem cell death is a genetically programmed response to DNA damage mediated by ATM and ATR rather than a consequence of the DNA damage itself. Genes involved in animal apoptosis, including BAX, PUMA and NOXA, have not been identified in A. thaliana. However, when animal BAX is expressed in plants, it triggers cell death (Lacomme and Santa Cruz, 1999; Yoshinaga et al., 2005). In addition, the cell death suppressor BAX inhibitor-1 (BI-1) is conserved in both animals and plants, and A. thaliana BI-1 functions as a cell death attenuator for biotic and abiotic types of cell death (Watanabe and Lam, 2006). Furthermore, meta-
caspases (MCs), which are structurally related to caspases, are found in plants, although the key initiator and effector caspases that are expressed in animals are missing in plants (Uren et al., 2000). *A. thaliana MC8* (*AtMC8*) is specifically induced by treatment with ultraviolet C light and ROS-inducing agents (He et al., 2008). Transgenic overexpression of *AtMC4* and *AtMC8* increases the level of cell death induction upon treatment with ROS-inducing agents, while the loss of *AtMC4* and *AtMC8* functions results in a delay or decrease in cell death (Lam and Zhang, 2012). These results demonstrate that *AtMC4* and *AtMC8* are positive regulators for ROS activation of cell death. Further studies are required to identify factors involved in stem cell death induced by DNA damage.

C) Induction of endoreduplication Endoreduplication is a common process in plants in which DNA is replicated without cell division, which leads to elevated genomic DNA content and polyploidy (Lee et al., 2009). Polyploid cells are often observed in developing tissues, such as root hairs, xylem and endosperm; therefore, the regulation of endoreduplication plays an important role in normal differentiation (De Veylder et al., 2001). In *A. thaliana*, endoreduplication is one of the common responses to DNA damage in addition to cell cycle arrest and cell death. DSBs induced by the lack of chromatin assembly factor 1 activity accelerate endoreduplication in seedlings and leaves (Endo et al., 2006; Schonrock et al., 2006; Ramirez-Parra and Gutierrez, 2007). Additionally, wild-type *A. thaliana* treated with zeocin, which is known to produce DSBs, show an increase in DNA ploidy (Ramirez-Parra and Gutierrez, 2007; Adachi et al., 2011). These results indicate that in contrast to stem cells, damaged proliferating cells exit the cell cycle by endoreduplication in response to DNA damage. The advantage of this may be that entry into a non-dividing state by endoreduplication prevents DNA-damaged cells from proliferating, yet also prevents them from dying. A positive correlation between DNA content and cell size has been recognized in many plant cell types (Sugimoto-Shirasu and Roberts, 2003). To compensate for a reduction in cell number in the damaged tissue because of cell cycle arrest and to maintain growth and tissue structure, cells may enter into endoreduplication cycles. In animals, many different tissues including the brain, skin and placenta have polyploid cells, indicating that endoreduplication is an essential part of normal development (Sherman, 1972; Corash et al., 1989; Zanet et al., 2010; Unhavaithaya and Orr-Weaver, 2012). On the other hand, telomere shortening and defects in chromatin assembly, which activate the DDR, appear to induce endoreduplication; however, the function of this response is not yet clear. Davoli and de Lange (2011) observed that polyploidy is associated with mitotic progression of several types of cancers. Therefore, unscheduled endoreduplication induced by the DDR may promote tumorigenesis in animals.

ISOLATION OF SOG1: A MASTER REGULATOR OF THE DNA DAMAGE RESPONSE

The *xpf-2* mutant of *A. thaliana* is defective in an ortholog of the human repair endonuclease XPF, which is involved in nucleotide excision repair, and the *A. thaliana* *xpf-2* mutant is sensitive to gamma irradiation (Fig. 1). Although irradiated *xpf-2* mutant seeds exhibit normal germination, the production of true leaves is delayed (Preuss and Britt, 2003): true leaves appear 8–9 days after irradiation in wild-type, but not until 17 days after irradiation in *xpf-2* mutants. As the cotyledon is a part of the embryo within the seed, cell expansion alone is required to open the cotyledon. However, cell division is required for formation of true leaves; therefore, it is hypothesized that the delay in true leaf formation in the irradiated *xpf-2* mutant is due to the arrest of the cell cycle. A possible explanation for such cell cycle arrest is that irradiated plants activate the checkpoint system and enforce an arrest of cell division, providing time for the cell to repair the damage before the chromosomes segregate into separate daughter cells. If this hypothesis is correct, it should be possible to isolate checkpoint mutants as suppressor mutants of *xpf-2*. Preuss and

![Wild-type](image1.png) ![xpf-2](image2.png) ![xpf-2 sog1-1](image3.png)

**Fig. 1.** Morphology of plants grown from irradiated seeds. Seeds were irradiated with 150 Gy, sown on MS plates, and photographed nine days later. The *xpf-2* mutant germinates normally but true leaves do not form until 17 days after irradiation, whereas *xpf-2 sog1-1* displays true leaves by nine days after irradiation. In contrast, no difference was observed among the three genotypes in plants grown from unirradiated seeds. (This figure is cited from Yoshiyama et al. (2009), reproduced with permission.)
Britt (2003) succeeded in isolating several such suppressor mutants termed sog (suppressor of gamma response) (Fig. 1). SOG1 encodes a NAC (NAM, ATAF1/2, and CUC2) domain protein (Yoshiyama et al., 2009). The genome of A. thaliana contains more than 100 genes encoding NAC transcription factors, making this one of the largest plant-specific families of transcription factors (Ooka et al., 2003). The A. thaliana NAC family is divided into ten major groups (Jensen et al., 2010), and SOG1 belongs in subgroup IX-1, which is a distinct clade. Members of this subgroup deviate from the characteristic NAC structure in having N-terminal extensions. Although some NAC transcription factors have been shown to play critical roles in processes as diverse as the establishment of the shoot apical meristem, lateral root formation, environmental stress responses and xylem cell specification, the function of most NAC proteins is still largely unknown (He et al., 2005; Olsen et al., 2005; Yamaguchi et al., 2010). SOG1 was the first transcription factor identified as being involved in the DDR in plants. The N terminus of SOG1 includes the NAC domain, which is important for DNA binding, and the C terminus of the molecule includes the transcription regulatory region (Fig. 2). SOG1 also has an N-terminal extension of approximately 40 amino acid residues. Since the function of the N-terminal extension is unknown, it is to be hoped that future research will clarify its function. The sog1-1 mutant carries a missense mutation affecting a highly conserved amino acid in the NAC domain, implying that the mutation eliminates the DNA binding activity of SOG1.

THE FUNCTION AND ROLE OF SOG

As mentioned above, within 1.5 h after gamma irradiation of A. thaliana seedlings, the expression of more than 100 genes changes. Intriguingly, irradiated sog1-1 seedlings are almost completely defective in the induction of gene expression after gamma irradiation (Yoshiyama et al., 2009). This phenotype is similar to that of the atm mutant, which is also required for the transcriptional response to gamma irradiation (Culligan et al., 2006). These results indicate that the great majority of the transcriptional response to gamma irradiation is regulated through SOG1, and raised the possibility that ATM and SOG1 work in the same pathway. The cell cycle-related transcripts CDKB2;1 and KNOLLE are downregulated by gamma irradiation (Culligan et al., 2006). Although this suppression lasts for 24 h after irradiation in wild-type plants, it lasts for only 5–10 h in the sog1-1 mutants (Yoshiyama et al., 2009), suggesting that SOG1 plays an important role in the maintenance of cell cycle arrest. In fact, root growth of wild-type seedlings is inhibited from two days after transfer to zeocin-containing plates, whereas no such inhibition is observed in the sog1-1 mutants, which thus display longer roots than those of wild-type plants (Adachi et al., 2011; Yoshiyama et al., 2014). Chromatin immunoprecipitation (ChIP)-PCR analysis showed that SOG1 binds directly to the promoter regions of SMR5 and SMR7, two plant-specific cyclin-dependent kinase inhibitors, in a DNA damage-dependent manner (Yi et al., 2014). These data suggest that SOG1 induces cell cycle arrest through SMR5 and SMR7 in response to DNA damage. Furthermore, neither cell death of root stem cells nor endoreduplication of root epidermal and cortex cells is observed in the sog1-1 mutant upon treatment with DNA-damaging agents (Furukawa et al., 2010; Adachi et al., 2011). Thus, SOG1 contributes to the maintenance of genome stability after DNA damage by inducing several pathways, including cell cycle arrest,
SOG1 is a master regulator of the DDR in plants. DNA repair, programmed cell death and endoreduplication, through the regulation of many genes. Recently, Hu et al. (2015) reported that SOG1 is involved in the replication checkpoint activated by deficiency of RTEL1 (an ortholog of human Regulator of Telomere Length 1). This result indicates that SOG1 plays an important role not only in the DNA damage response but also in the replication checkpoint. As it regulates many pathways, SOG1 is clearly a master regulator of DNA damage and replication stress responses.

**LOCALIZATION AND REGULATION MECHANISM OF SOG1**

Transgenic *A. thaliana* plants carrying SOG1 fused to β-glucuronidase (SOG1-GUS) were used to determine the tissue expression of SOG1. SOG1-GUS was observed as blue staining in meristematic tissues, such as the shoot and root apical meristems, lateral root primordium, root stele and young flowers (Yoshiyama et al., 2013) (Fig. 3). All of these tissues contain dividing cells, indicating that SOG1 function is required in actively proliferating cells. In experiments investigating the intracellular localization of SOG1, SOG1-GFP protein localized to the nucleus, and neither the intensity nor the localization of the signal was affected by zeocin treatment (Yoshiyama et al., 2013). Additionally, DNA-damaging stress does not induce SOG1 transcription. These results show that SOG1 is not regulated by protein accumulation, subcellular localization or the level of its mRNA, suggesting that SOG1 is activated by posttranscriptional modification. In response to DSB-inducing treatment, part of the SOG1 population is hyperphosphorylated in an ATM-dependent and ATR-independent manner, again suggesting that SOG1 activity is controlled at the posttranscriptional level (Yoshiyama et al., 2013). However, SOG1 is not hyperphosphorylated after treatment with the replication inhibitors aphidicolin and hydroxyurea (Yoshiyama et al., 2013). The C terminus of SOG1 has five copies of the SQ (serine followed by glutamine) motif (Fig. 2), which is the preferred site for phosphorylation by ATM and ATR. Since the DNA damage-dependent SOG1 hyperphosphorylation was not observed with the SOG1(5A) construct, which possesses serine-to-alanine substitutions at all five SQ motifs, one or more of the SQ motifs must be targets for the hyperphosphorylation in response to DNA damage. Complementation analysis showed that the SOG1(5A) construct cannot rescue various *sog1-1* phenotypes, although SOG1 can do so. Taken together, these data suggest that ATM-dependent hyperphosphorylation of SQ motif(s) is essential for SOG1 functions.

**COMPARISON BETWEEN PLANT SOG1 AND ANIMAL p53**

The functions and regulatory mechanisms of SOG1 and animal p53 show many similarities (Yoshiyama et al., 2014). Like SOG1, p53 is a transcription factor that regulates many genes involved in the cell cycle, DNA repair and apoptosis (Rozan and El-Deiry, 2007). p53 is phosphorylated by ATM shortly after DNA damage, resulting in enhanced activity of p53 (Banin et al., 1998; Canman et al., 1998). Although plants apparently lack a p53 ortholog, the discovery and analysis of SOG1 indicates that SOG1 is a functional homolog of animal p53 in plants. p53 is normally maintained at low levels in unstressed cells by degradation through continuous ubiquitination; however, when cells suffer DNA damage, p53 ubiquitination is suppressed and the polypeptide is stabilized (Lavin and Gueven, 2006). In contrast, the level of SOG1 protein is not affected by DNA damage, indicating that the regulatory mechanisms of SOG1 and p53 differ. Apart from protein stability, these two proteins share common features; however, their amino acid sequences are unrelated to each other (Fig. 2). Thus, it is likely that a NAC protein, not a p53 ancestor, evolved to acquire a function in the DDR during plant evolution.
The evolution of plants has led to widely varying levels of complexity, from the earliest green algae through bryophytes, lycophytes, and ferns to the complex gymnosperms and angiosperms. NAC proteins are present in mosses, lycophytes, gymnosperms and most of the angiosperms, but not in the unicellular green alga *Chlamydomonas reinhardtii* or in the multicellular green alga *Klebsormidium flaccidum* (Zhu et al., 2012). It is important to determine when plants acquired SOG1, a member of the NAC protein family, during their evolution. When the entire SOG1 amino acid sequence was used to search for SOG1-like proteins in other plants, predicted SOG1-like proteins were identified in most land plants, from moss to angiosperms (Yoshiyama et al., 2014). *A. thaliana* SOG1 has two characteristic features: an N-terminal extension, and five SQ motifs (Yoshiyama et al., 2014). Therefore, SOG1-like proteins found in other species were examined for these two features (Fig. 4). Eudicot (soybean [*Glycine max*], barrel medic [*Medicago truncatula*], grape [*Vitis vinifera*], black cottonwood [*Populus trichocarpa*]) and monocot (rice [*Oryza sativa*], maize [*Zea mays*], sorghum [*Sorghum bicolor*]) SOG1-like sequences have five SQ motifs at conserved positions in their C-terminal regions, as well as conserved N-terminal extensions. A SOG1-like protein found in an ancestor of the angiosperms, *Amborella trichopoda*, has four of the five conserved SQ motifs. The gymnosperm *Picea glauca* has three, and, interestingly, it also has an N-terminal extension, which *A. trichopoda* lacks. The SOG1-like protein found in the bryophyte *Physcomitrella patens*, a member of the most primitive plant division, possesses two SQ motifs; however, their positions differ from those found in *A. thaliana* SOG1. In addition, the moss SOG1-like protein lacks an N-terminal extension. Therefore, it is not clear whether the moss SOG1-like protein is an ortholog of *A. thaliana* SOG1. We have identified a candidate SOG1-like gene in the fern *Selaginella moellendorffii*, but since the sequence includes only the NAC domain and is C-terminally truncated it is currently unclear whether lycophytes possess a functional SOG1 ortholog. From these observations, it is apparent that SOG1 had already been acquired in the gymnosperms at the latest.

**CONCLUSION AND OUTLOOK**

*A. thaliana* activates several pathways in response to DNA damage, cell cycle arrest, DNA repair, programmed cell death and endoreduplication to maintain its genome stability. The transcription factor SOG1 plays an important role in regulating more than 100 genes to activate these responses. When exposed to DSB inducers, SOG1 is activated through phosphorylation by the DNA damage response kinase ATM. These SOG1 functions and regulatory mechanisms are similar to those of animal p53; however, the amino acid sequences of the two proteins are unrelated, implying that plants and animals independently acquired these regulators during evolution. According to an alignment of SOG1-like proteins found in

![Fig. 4. Structural features of SOG1-like proteins found in diverse plant species. One plant species is selected as a representative for each division of plants. The N-terminal extension, NAC domain and transcription regulatory domain are shown. Conserved SQ motifs found in *A. thaliana* SOG1 are represented by red boxes. Blue boxes represent SQ motifs found in the transcription regulatory domain of other species’ SOG1-like protein. *Physco: Physcomitrella.*](image-url)
SOG1 is a master regulator of the DDR in plants

other plants, it appears that algae do not possess SOG1, but gymnosperms do.

Further studies on SOG1 are required to better understand the plant’s response to DNA damage. Specifically, it is necessary to identify genes that are directly regulated by SOG1, using a combination of ChIP-seq and microarray analyses, to determine how signal transduction occurs and what pathways are activated. It is also necessary to identify conserved motifs as SOG1 target sequences. Identification of factors that interact with SOG1 is necessary to understand how SOG1 is regulated. These studies will facilitate the identification of novel DDR factors, which in turn will provide insights into the molecular mechanism of the plant DDR. Furthermore, the relationship between SOG1 structure and function should be analyzed. This analysis is important for understanding SOG1 function, and the resultant information will be critical for comparisons between A. thaliana SOG1 and SOG1-like proteins found in other plant species. It will also be interesting to determine more precisely when SOG1 was acquired during the evolution of plants, and whether the functions of SOG1 vary in different plant species. Finally, it will be intriguing to study the role of SOG1 in response to pathogens. Recent studies have revealed that the DDR and immune responses are closely related in plants (Yan et al., 2013; Song and Bent, 2014). Since SOG1 may also play a crucial role in the plant immune system, it is imperative to examine whether the sog1-1 mutation affects sensitivity to various pathogens. To understand how plants adapt to stressful environments, further analysis of the DDR in plants is eagerly anticipated in the future.

I thank Seisuke Kimura for helpful suggestions. This work was supported by a research fellowship from the Japan Society for the Promotion of Science to K. O. Y. (13J40017).

REFERENCES


Hu, Z., Cools, T., Kalhorzadeh, P., Heyman, J., and De Veylder,


