HHS Public Access

Author manuscript

Nat Plants. Author manuscript; available in PMC 2020 October 01.

Published in final edited form as:

Nat Plants. 2019 October ; 5(10): 1057–1065. doi:10.1038/s41477-019-0511-z.

CYP72A Enzymes Catalyze 13-Hydrolyzation of Gibberellins

Juan He[a,d], **Qingwen Chen**[a], **Peiyong Xin**[c], **Jia Yuan**[a,c], **Yihua Ma**[a,c], **Xuemei Wang**[a,c], **Meimei Xu**[b], **Jinfang Chu**[c], **Reuben J Peters**[b], **Guodong Wang**[a,d,*]

[a]State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, INASEED, Chinese Academy of Sciences, Beijing 100101, China

[b]Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, IA, 50011, USA

[c]National Centre for Plant Gene Research (Beijing), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing100101, China

[d]University of Chinese Academy of Sciences, Beijing 100039, China

Abstract

Bioactive gibberellins (GAs, diterpenes) are essential hormones in land plants, controlling many aspects of plant growth and developments. In flowering plants, 13-OH (low bioactivity; such as GA_1) and 13-H GAs (high bioactivity; such as GA_4) frequently coexist. However, the *bona fide* GA 13-hydroxylase and its physiological functions in Arabidopsis remain unknown. Here, we report that novel cytochrome P450 genes (*CYP72A9* and its homologs) encode active GA 13-hydroxylases in Brassicaceae plants. *CYP72A9*-overexpressing plants exhibited semi-dwarfism, which was caused by significant reduction in GA_4 levels. Biochemical assays revealed that recombinant CYP72A9 protein catalyzed the conversion from 13-H GAs to the corresponding 13-OH GAs. *CYP72A9* was expressed predominantly in developing seeds in Arabidopsis. Freshly harvested seeds of *cyp72a9* mutants germinated more quickly than wild-type, while long-term storage and stratification-treated seeds did not. The evolutionary origin of GA 13-oxidases from the CYP72A subfamily also was investigated and discussed here.

INTRODUCTION

Gibberellins (GAs), biosynthesized from geranylgeranyl diphosphate (GGDP) via multiple enzymatic steps, are essential phytohormones for plant growth and development. Among

*To whom correspondence should be addressed: Tel: 86-10-64860462, Fax: 86-10-64873428, gdwang@genetics.ac.cn.

Author contributions

G.W. designed the research. J.H. performed the majority of experiments. Q.C. and Y.M. generated the Arabidopsis transgenic plants and analyzed the phenotypes of these plants. J.Y., X.W. and M.X. performed part of biochemical assays of CYP72A members from soybean and rice. P.X. and J.C. measured the gibberellins in plants by UPLC-QQQ-MS/MS. J.H. and G.W. analyzed the data. R.J.P. and G.W. wrote the manuscript with input from all authors.

Data availability

All data and materials generated during this study are available from the corresponding author upon request.

Competing interests

The authors declare no competing interests.

more than 130 discovered GAs to date, GA_4, GA_1 (also known as 13-OH GA_4), GA_7 and GA_3 (also known as 13-OH GA_7) are common bioactive GAs in flowering plants. The bioactivity of GA_1 is approximately 1000-fold lower than GA_4 in plants[1–5]. The plant gibberellin biosynthetic pathway and its regulation have been well elucidated, mainly in Arabidopsis and rice[6] (Fig. 1). Recent studies have shown that not only GA biosynthesis, but also GA deactivation is important for GA homeostasis *in planta*. Thus far, three types of GA deactivation enzymes and their encoding genes have been identified, including GA 2-oxidase (soluble 2-oxoglutarate-dependent dixoygenase; seven *GA2ox* genes in Arabidopsis and ten *GA2ox* genes in rice), GA methyltransferase (*GAMT1/2* from Arabidopsis), and GA 16,17-oxidase (*CYP714D1/EUI* from rice)[7–9]. More recently, reverse genetic studies in rice have suggested a predominant role for CYP714B1 and CYP714B2 (CYP714D1 homologs) in reducing GA activity through 13-hydroxylation of GA_{12} to form GA_{53}[3]. Interestingly, both *cyp714d1* mutant and *cyp714b1/cyp714b2* double mutant rice display an elongated uppermost internode, in where these three *P450*s are highly expressed, at the heading stage.

In Arabidopsis, GA_4 is the predominant bioactive GA in most tissues while GA_1 accumulates at relatively high levels in developing seeds/siliques[8,10,11]. These results suggest that 13-hydroxylation of GA_4 to form GA_1 plays an important role in seed maturation and germination, at least in Arabidopsis. Two Arabidopsis EUI homologs (CYP714A1 and CYP714A2) had been shown to catalyze distinct oxidations events in GA_{12} *in vitro* and *in planta*: CYP714A1 converts GA_{12} to 16-carboxylated GA_{12}, while CYP714A2 catalyzes the hydroxylation at the C12 or C13 (only one twentieth of C12 hydroxylase activity) position of GA_{12} and its upstream precursor, *ent*-kaurenoic acid[12]. However, the C_{19}-GA profiles, including bioactive GA_4 and GA_1, of both *cyp714a1* and *cyp714a2* single mutants are comparable to those in wild-type[13]. These results suggest that other genes should encode *bona fide* GA 13-hydroxylase in Arabidopsis.

In this study, we report that one member of the CYP72A subfamily, *CYP72A9*, encodes gibberellin 13-oxidase, which catalyzes the conversion of 13-H GAs (GA_{12}, GA_9 and GA_4) to the corresponding13-OH GAs (GA_{53}, GA_{20} and GA_1). *CYP72A9* is predominantly expressed in developing seeds. *cyp72a9* mutants show a deficiency in GA_1 and an increase in the concentration of GA_4, suggesting that CYP72A9 plays a key role in the 13-hydroxylation of bioactive GAs in *Arabidopsis thaliana*. We further demonstrated that the conversion of GA_4 to GA_1 is an indispensable factor for primary seed dormancy in Brassicaceae plants.

RESULTS

Overexpression of *CYP72A9* in Arabidopsis Results in Dwarf Phenotypes Caused by GA_4 Deficiency

The *Arabidopsis thaliana* genome contains one geranylfarnesyl pyrophosphate synthase (C25)-sesterterpene synthase-P450 (*GFPPS-sesterTPS-P450*) gene cluster, in which eight tandem duplicated *CYP72A*s are functionally unknown[14–16]. To elucidate the biochemical and biological functions of each member of the *CYP72A* subfamily (CYP72A7, At3g14610; A8, At3g14620; A9, At3g14630; A10, At3g14640; A11, At3g14650; A13, At3g14660; A14, At3g14680; A15, At3g14690) in Arabidopsis, we firstly generated transgenic

Arabidopsis overexpressing each *CYP72A* gene (Supplementary Fig. 1). It is noteworthy that we failed to generate *CYP72A15*-overexpressing plants, in which *CYP72A15* was upregulated only by 2–3-fold (Supplementary Fig. 1). Among these *CYP72A*-overexpressing Arabidopsis lines, *CYP72A9*-overexpressing plants exhibited a semi-dwarf and late-flowering phenotype at the mature stage (Fig. 2A and Supplementary Fig. 2), which phenocopied the gibberellin deficiency mutants (like *ga1-t* (SALK_023192), Col-0 ecotype) or plants overexpressing GA deactivation genes[3,8,12,17]. To determine whether the bioactive GA deficiency caused the phenotypes of *CYP72A9*-overexpressing plants, we treated *CYP72A9*-overexpressing seedlings with 2 μM bioactive GA_3, which was able to reverse this semi-dwarf phenotype (Fig. 2B), indicating a defect in GA metabolism.

We further applied UPLC-QQQ-MS/MS (ultra-high-performance liquid chromatography coupled to a triple quadrupole mass spectrometer; MRM (multiple reaction monitoring) mode, for detailed parameters see Supplementary Table 1) to profile the endogenous GAs in wild-type (WT, Col-0 ecotype) and two independent *CYP72A9*-overexpressing lines (*72A9* OE-1 and *72A9* OE-2; Supplementary Fig. 1). The results showed that 13-H GAs (GA_{12}, GA_{15}, GA_{24}, GA_9, GA_4, and GA_{34}) were greatly reduced in both *CYP72A9*-overexpressing lines, while 13-OH GAs (GA_{44}, GA_{19}, GA_{20}, GA_1, and GA_8) were greatly increased compared with those in WT (Fig. 2c). GA_3 and GA_7 were not detected in any of the tested plants. The expression levels of many genes involved in GA metabolism and signaling also were significantly altered in the *CYP72A9*-overexpressing lines, with those involved in biosynthesis (*GA20ox1–3* and *GA3ox1*) significantly upregulated, while those involved in catabolism (*GA2ox1*, *2*, *6*, *CYP714A1*, *CYP714A2* and *GAMT2*) or signaling (*RGA* and *GID1c*) were significantly downregulated (Supplementary Fig. 3). Altogether, these data suggested that *CYP72A9* played a role in the production of 13-OH GA in Arabidopsis.

Biochemical Characterization of CYP72A9

To clarify the biochemical functions of CYP72A9 in GA metabolism, we expressed each *CYP72A9* gene (pESC-Leu vector) in the yeast strain WAT11, which has the Arabidopsis *cytochrome P450 reductase 1* (AtCPR1, At4G24520) gene integrated into its chromosome. We incubated 13-H GAs (bioactive GA_4, GA_9, and GA_{12}) in the heterologous culture and checked the resulting products by GC-MS (gas chromatography-mass spectrometry, and selected ion mode) after extraction and derivation. The results clearly showed that CYP72A9 was promiscuous; hydroxylating GA_4, GA_9, and GA_{12} at C13 position to produce GA_1, GA_{20}, and GA_{53}, respectively (Fig. 3 and Supplementary Fig. 4a, c). Additionally, CYP72A9 converted *ent*-kaurenoic acid to produce both steviol (*ent*-13-hydroxy kaurenoic acid, minor product) and *ent*-16β,17-dihydroxy kaurenoic acid (main product; Supplementary Fig. 4d and Supplementary Fig. 5).

*CYP72A*9 Predominantly Expressed in Developing Seeds and Involved in Primary Seed Dormancy

To determine the more specific physiological functions of *CYP72A*9 in Arabidopsis, we checked the tissue-specificity of *CYP72A*9 using qRT-PCR. *CYP72A*9 was predominantly expressed in developing seeds/siliques and was barely detected in other tested tissues (Fig. 4a). Consistent with the RT-PCR results, upon transformation with a P_{72A9}*::GUS* construct,

strong GUS activity was mainly detected in the developing seeds (Fig. 4b). Subcellular localization of 72A9:GFP fusion protein in protoplasts indicated that CYP72A9 is a prototypical CYP, which usually is an endoplasmic reticulum (ER)-bound protein (Fig. 4c). Two mutant alleles were obtained, the first, designated *cyp72a9–1*, was a T-DNA insertion (SALK_130811C; Supplementary Fig. 6), while the second, designated *cyp72a9–2*, was generated using the CRISPR/Cas9 technique (Supplementary Fig. 7). Note that *CYP72A*9 transcripts were still present at a low level (approximately 5% of WT) in the *cyp72a9–1* mutant, which likely caused by the T-DNA insertion at the fourth intron of *CYP72A9* genomic DNA (Supplementary Fig. 6a, c). GA profiling of developing seeds/siliques of WT and the two independent *cyp72a9* mutants indicated that bioactive GA_4 (0.84 ± 0.016 ng/g fresh weight in WT vs 1.46 ± 0.032 ng/g fresh weight in *cyp72a9–1* and 2.20 ± 0.26 ng/g fresh weight in *cyp72a9–2*, $n = 3$) was significantly increased in both *cyp72a9* mutants, while 13-OH GAs were decreased compared with WT (Fig. 4d). Indeed, GA_{20} and GA_1 were present at levels below the detection limits in the *cyp72a9* mutants (Fig. 4d). These results strongly indicate that *CYP72A*9 plays a regulatory role in the homeostasis of bioactive GA_4 in the developing seeds/siliques of *Arabidopsis thaliana*.

The expression levels of a number of genes involved in GA biosynthesis, including *GA20ox1*, *GA20ox2* and *GA3ox1*, were significantly downregulated in the developing seeds/siliques of *cyp72a9–1* mutants. However, the expression levels of the seed-predominant GA deactivation genes *CYP714A1/2* and *GAMT1/2* are not significantly changed (Supplementary Fig. 8b). Moreover, the GA signaling genes *GID1a* and *GID1c* were significantly downregulated (Supplementary Fig. 8c). Overall, the pattern of changes in expression for genes involved in GA metabolism from the *cyp72a9–1* mutant were opposite of those observed in the *CYP72A9*-overexpressing lines.

Previous studies have demonstrated that bioactive GA_4 plays a key role in primary seed dormancy and germination[18,19]. Enzymatic deactivation of GA_4 by CYP72A9 and predominant expression of *CYP72A*9 in the developing seeds inspired us to ascertain the involvement of *CYP72A*9 in primary seed dormancy and germination. Freshly harvested seeds from WT and the two *cyp72a9* mutants were used to investigate the effect of *cyp72a9* on seed dormancy. The germination ratio of *cyp72a9* seeds was significantly higher than that of WT in the absence of stratification (exposure to 4 °C for three days), even at 5 days after sowing (55% for WT versus 78% for *cyp72a9–1* and 80% for *cyp72a9–2*; Fig. 5a). Previous studies have demonstrated that primary seed dormancy could be released by a period of dry storage and stratification, which induced the *de novo* biosynthesis of GA[19]. We did not observe any difference when all seeds were subjected to stratification (4°C for 3 days) or dry storage for six months prior to germination with the germination ratio of all three tested lines reaching almost 100% at 3 days after sowing (Fig. 5b, c). Moreover, *cyp72a9* seeds in developing siliques also germinated more quickly than WT (Fig. 5d). These results suggested that the disruption of *CYP72A*9 reduced the primary seed dormancy in Arabidopsis. Thus, the conversion of GA_4 to GA_1 during seed maturation appears to be necessary for primary seed dormancy.

Biochemical Screening Other Plant CYP72As

The CYP72A subfamily was expanded with a lineage-specific pattern, mainly due to independent tandem gene-duplication events. Thus, it is very difficult to predict the biochemical activities of CYP72As based only on comparisons of primary protein sequences[20]. The presence of both the CYP72A subfamily and 13-OH GAs in flowering plants inspired us to test whether CYP72A from other plant species also displayed GA 13-oxidase activity. Using the abovementioned yeast system, we further tested all CYP72As from five representative plant species, including three *Brassicales* plants (8 CYP72As from *Arabidopsis thaliana*, 5 CYP72As from *Capsella rubella* and 3 CYP72As from *Brassica rapa*), one *Fabales* plant (12 CYP72As from *Glycine max*), and one *Poales* plant (13 CYP72As from *Oryza sativa*; Fig. 6). A total twelve CYP72As from these 41 candidates, including CYP72A9, were found to utilize at least one of the tested substrates (*ent*-kaurenoic acid, GA_{12}, GA_9, and GA_4). All twelve active CYP72As had the ability to convert *ent*-kaurenoic acid to *ent*-16β,17-dihydroxy kaurenoic acid as a main product (Fig. 6 and Supplementary Fig. 9). Moreover, CYP72A262 from *Brassica rapa* had a similar activity profile to CYP72A9, including hydroxylating GA_4 at C13 position to produce GA_1 (Supplementary Fig. 10), while CYP72A15 from Arabidopsis, CYP72A272 from *Brassica rapa*, CYP72A484 from *Capsella rubella*, and CYP72A135 from *Glycine max* utilized *ent*-kaurenoic acid, GA_{12} and GA_9 as substrates rather than GA_4 (Supplementary Figs.11–14). Interestingly, both *CYP72A262* and *CYP72A484* showed a silique/seed-predominant expression pattern, like *CYP72A9*, in *Brassica rapa* and *Capsella rubella* (Supplementary Fig. 15). AtCYP72B1 (also named BAS1), which is the closest relative to CYP72As (around 40% identity at protein level), inactivates castasterone and brassinolide (both are triterpenoids) via C26 hydroxylation[21]. However, CYP72B1 did not show activity toward four tested diterpenoids in this study (Fig. 6 and Supplementary Fig. 16).

DISCUSSION

GA_1 Formation via a Distinct Biosynthetic Pathway in Arabidopsis Seeds

Both GA_1 (weak bioactivity) and GA_4 (strong bioactivity) coexist in flowering plants. It is well-established that GA_{12} and GA_{53} are converted to GA_4 and GA_1 via two parallel pathways, which are catalyzed by GA 20-oxidases (GA20ox) and 3-oxidases (GA3ox; Fig. 1). The parallel formation of GA_1 and GA_4 was also supported by GA profiling in the *cyp714b1/cyp714b2* rice mutant (both CYP714B1 and CYP714B2 catalyze 13-hydroxylation of GA_{12}, rather than GA_9 and GA_4): all examined 13-OH GAs were significantly decreased[3]. In this study, we discovered a GA_4 13-hydroxylase, encoded by *CYP72A9*, in *Arabidopsis thaliana*. GA_{20} and GA_1 were not detected in the developing seeds/siliques of *cyp72a9* mutants, while their upstream intermediates (GA_{19} and GA_{44}) were almost unchanged in *cyp72a9* mutants. These results suggested that the efficiency of converting GA_{19} to GA_{20} by GA20ox was very low, at least in developing silique/seeds tissue of Arabidopsis. This proposal is consistent with the previous biochemical assays of GA20ox from pumpkin (*Cucurbita maxima* L.). The pumpkin GA20ox catalyzed the three-step conversions of GA_{53} to GA_{44} to GA_{19} to GA_{17} (a tricarboxylic acid GA) and GA_{20}; however, the conversion of GA_{19} to GA_{20} was 20 times lower than that of GA_{19} to GA_{17}[22]. Thus far, only GA20ox1 from Arabidopsis has been partially characterized using GA_{53} and

GA_{19} as substrates[23]. The detailed catalytic efficiencies of all GA20oxs toward various GA substrates (five *GA20ox* genes in Arabidopsis and four *GA20ox* genes in rice) are needed to explain the difference in GA metabolism between Arabidopsis and rice[24]. Nevertheless, the increased GA_4 observed in the *cyp72a9* mutants indicates that CYP72A9 is mainly responsible for the formation of GA_1 (from GA_4) and GA_{20} (from GA_9) in *Arabidopsis thaliana*. It is noteworthy that, as shown in Fig. 4d, GA_8 was detected in the *cyp72a9* mutants with a comparable level to wild-type although GA_1 was almost undetectable. One possible explanation is that there are one or several uncharacterized oxidases (e.g. P450 or 2-oxoglutarate-dependent dioxygenaes etc), which catalyzed the 13-hydroxylation of GA_{34} to from GA_8 in Arabidopsis. Thus, we have updated the GA metabolic network in Arabidopsis, as summarized in Fig. 7.

GA Deactivation and Primary Seed Dormancy

It is well-known that the ABA level is increases and GA level decreases during seed maturation to maintain primary seed dormancy. Although there are many lines of evidence supporting the importance of ABA biogenesis in primary seed dormancy, the mechanism by which GA biogenesis is regulated during seed maturation has been largely unknown[19,25]. Recent GA profiling across seed development stages has clearly revealed that the presence of strong bioactive GA_4 in the early stages of seed development, which is dramatically decreased in later stages. In contrast, 13-OH GAs, such as GA_1, GA_3, GA_8, GA_{20}, and GA_{29} (2-hydroxylation of GA_{20}) are highly accumulated in mid-development during seed maturation[10], which is tightly correlated with the *CYP72A9* expression pattern (Fig. 4a, b). Disruption of *CYP72A9* significantly reduced the primary seed dormancy, most likely by increasing GA_4 level (*cyp72a-1* had 1.74 folds and *cyp72a-2* had 2.62 folds higher than WT; Fig. 4c). Interestingly, all GA deactivation genes identified in *Arabidopsis thaliana* to date, including *CYP714A1*, *CYP714A2*, two *GAMT*s, and *CYP72A9* (as shown here), are highly expressed in developing seeds[8,13], suggesting that GA_4 deactivation plays an important role in seed development and germination. Consistent with this hypothesis, *CYP714A1*, *GAMT1*, and *GAMT2* null mutants showed higher seed germination ratio than WT, although only in the presence of inhibitors of GA *de novo* biosynthesis (paclobutrazol or ancymidol)[8,13]. It is notable that dry stored seeds, rather than fresh harvested seeds, were used in these experiments. Further germination assays with fresh harvested seeds are needed to ascertain whether *CYP714A1*, *GAMT1*, and *GAMT2*, like *CYP72A9*, are also involved in primary seed dormancy in *Arabidopsis thaliana*. Furthermore, the conservation of GA-13 hydroxylase activity and seed-predominant expression pattern of *CYP72A262* (from *Brassica rapa*) and *CYP72A484* (from *Capsella rubella*) suggests to us that GA deactivation is indispensable for primary seed dormancy in *Brassicaceae* plants.

Divergence of Terpene Oxidase among the CYP72A Subfamily

The CYP72A subfamily has expanded dramatically in some plant species during plant evolution[20,26]. Previously characterized CYP72A subfamily members utilized monoterpene (CYP72A1 from *Catharanthus roseus*) or triterpenes as substrates[27–30]. Here, we found at least one diterpene oxidase (*ent*-kaurenoic acid or GA) from the CYP72A subfamily in the examined flowing plants, in which this subfamily was expanded mainly due to tandem gene duplication. These results suggested that the common ancestor of the CYP72A subfamily

might have encoded a diterpene oxidase, as *ent*-kaurene and *ent*-kaurenoic acid are widely detected in flowering plants[31]. The coexistence of diterpene oxidase (CYP72A135 and CYP72A151) and triterpene oxidase (CYP72A61 and CYP72A69) from this subfamily in soybean further suggests that CYP72A subfamily underwent rapid functional divergence in specific plant lineages. Parallel gain of GA_9 13-hydroxylase activity by CYP72A9/CYP72A15/CYP72A262/CYP72A272/CYP72A484 in *Brassicales* and CYP72A135 in *Fabales* also would support this conclusion. It should be noted that the tandem CYP72A members are linked to *GFPPS-sesterTPS* to form a gene cluster in *Brassicaceae* plants[16] (Fig. 2). It seems plausible to suggest that one or more of these AtCYP72As are involved in the oxidation of the sesterterpene backbones[14,15], and metabolomic analysis of the transgenic plants generated in this study will shed light on the further functional assignation of the CYP72A subfamily in Arabidopsis.

In summary, we conclude that GA deactivation (GA_4 to GA_1) by CYP72As has an important molecular function in fine-tuning GA homeostasis, which is particularly important for primary seed dormancy in *Brassicaceae* plants. The identification and functional characterization of the various *CYP72A* genes encoding diterpene (GA) oxidases in this study pave the way for further understanding of the functional diversity of the CYP72A subfamily in flowering plants.

METHODS

Plant Materials and Chemicals

The *Arabidopsis thaliana* lines (both WT and transgenic plants are Col-0 ecotype) and other plant species (*Brassica rapa* and *Capsella rubella*) mentioned in this study were grown on soil at 22 °C under a 16-h light/8-h dark cycle. The *ga1-t* mutant (SALK_023192) was a gift from Dr. Xiangdong Fu (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing). The T-DNA insertion line *cyp72a9–1*(SALK_130811C) was ordered from ABRC (http://abrc.osu.edu/). To construct transgenic plants overexpressing *AtCYP7A7-AtCYP7A15*, coding regions of *AtCYP7A7-AtCYP7A15* were amplified from cDNA clones using sets of primers (Supplementary Table 1). The PCR products were first ligated into pENTR/D-TOPO vector (Life Technologies Corporation, USA) and then ligated into the binary vector pCHF3 by the LR reaction. The *cyp72a9–2* mutant was generated using the CRISPR/Cas9 technique. Cas9 editing target site primers were designed (Supplementary Table 1) and ligated into pCAMBIA 1300-pYAO-cas9.

All chemicals used in this study were purchased from Sigma-Aldrich (St. Louis, MO), except gibberellins (GA_1, GA_4, GA_9 and GA_{20}), which were purchased from OlChemIm company (http://www.olchemim.cz/) and *ent*-16β,17-dihydroxy kaurenoic acid, which was purchased from BOS Science (http://www.bocsci.com/).

Quantitative RT-PCR Analysis and Promoter: GUS Staining

RNA extraction, reverse transcription reaction and quantitative RT-PCR were performed as described previously[32,33]. Two sets of reference genes (three for each set) were selected for different qRT-PCR assay by following previous studies[9,34]. The annealing temperature,

specificity and amplification efficiency of all primers used in the study were tested and are listed in Supplementary Table 1.

The promoter fragments for *CYP72A9* (644 bp upstream of the ATG codon) were subcloned into the pMDC162 GATEWAY binary vector using a PCR-mediated technique (see Supplementary Table 1 for primer information). Arabidopsis transformation, screening for homozygous plants and GUS staining were performed as described previously[35].

Subcellular Localization of CYP72A9

Constructs of the CYP72A9-GFP fusion protein (pJIT163-hGFP vector), Arabidopsis leaf protoplast preparations, transformation, and image collection using a laser scanning confocal microscope were all performed as described previously[36]. The plasmid harboring the gene encoding ER-localized mCherry marker protein (plasmid# CD3–959)[37] was ordered from ABRC (http://abrc.osu.edu/). For detailed primer information, see Supplementary Table 1.

Heterologous Expression in Yeast and Identification of CYP72As Products

All *CYP72A* genes tested in this study were subcloned into the pESC-Leu vector for expression of Myc-tagged CYP72A in the WAT11 yeast strain[38]. Related primers used in the constructs are listed in Supplementary Table 1. The resulting positive yeast clones were cultured in 5 mL of SD dropout medium (-Ura, -Leu), and then harvested and induced with SG dropout medium (-Ura, -Leu) at 30°C for one day in a shaking incubator (200 rpm). The medium was diluted with fresh SG medium to OD_{595} = 0.4, 5 mL of the diluted medium was incubated with 1.0 μg of the different GA substrates for an additional 16 h. Ethyl acetate extracts of these recombinant cultures were evaporated to dryness by nitrogen flux. The dried chemicals were then derivatized with *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA) before analysis by GC-MS. One microliter of analytical sample was loaded onto the GC-MS for analysis. The initial oven temperature was held at 60°C for 1 min and then ramped at 10°C/min to 325°C and maintained at 325°C for 10 min. The inlet temperature was 270°C. Positive mode EI ionization was used. The temperatures of the ion source and quadrupole were set at 230°C and 150°C, respectively. Mass spectra were acquired within a scanning range of *m/z* 50–600.

Western Blot Assay

Yeast microsome was prepared as previously described[39]. Briefly, collected yeast cells (30 mL culture) were suspended in 4 mL of TE buffer (50 mM Tris-HCL, pH 7.4, 1 mM EDTA with 1 mM protease inhibitor Cocktail) and were disrupted with glass beads by vortexing. The broken cells were centrifuged at 10,000 *g* for 15 min, and the resulting supernatant was further centrifuged at 100,000 *g* for 2 h. The microsomal proteins were re-suspended in 200 μl TE buffer adding 1% Triton X-100 and quantified using Bradford assay. Twenty μg microsomal proteins were separated by 12% SDS-PAGE, transferred to a polyvinylidene difluoride membrane, and blotted with anti-Myc tag primary antibody (1:5,000; ZSGB-BIO, China) and goat anti-mouse igG (H+L)-HRP as the second antibody (1:10,000).

Germination Assay and Vivipary Testing

For the Arabidopsis germination test, freshly harvested seeds were sown on water-saturated filter paper and then germinated under a 16-h light/8-h dark photoperiod at 22 °C. The number of germinated seeds was counted every 12 h until approximately 100% of the seeds had germinated. Radicle protrusion was used to indicate germination and was scored at the indicated time points. More than 3 biological replicates consisting of 50 seeds per replicate were used for each experiment. The vivipary assay was performed as previously described[40]. Briefly, developing siliques at the long-green stage were collected, sterilized with 70% ethanol for 1 min and 25% bleach for 10 min and then plated on ½ MS medium for germination.

Quantification of Endogenous Gibberellins in Arabidopsis

The quantification of endogenous GAs levels was performed as reported previously with modifications of the sample pretreatment[41]. Briefly, 200 mg of the ground plant material powder was extracted with 5 mL of 90% aqueous methanol. Simultaneously, 2 ng of each D-labeled GA compound was added to the extracting solvents as internal standards for GA content measurement. The MAX cartridge (Waters Corporation, Milford, USA) was activated and equilibrated with MeOH, water, 5% NH_4OH, and 90% MeOH in turn, while MCX (Waters Corporation, Milford, USA) was equilibrated with MeOH, water and 90% MeOH. Subsequently, the crude extracts were loaded onto the tandem cartridges connected to an adapter. The MAX cartridge was then disconnected and rinsed with 5% NH_4OH in 5% MeOH, MeOH in turn. Finally, GA compounds were eluted with 90% MeOH containing 2% FA. The eluent was dried under a N_2 stream and redissolved in 150 μL of 40% MeOH prior to UPLC-MS/MS analysis. GA analysis was performed on a quadrupole linear ion trap hybrid mass spectrometer (QTRAP 6500, AB SCIEX, Foster City, CA) equipped with an electrospray ionization (ESI) source and coupled to a UPLC (Waters, Milford, MA, USA). The UPLC inlet method and ESI source parameters were set as previously reported[41]. GAs were detected in negative multiple reaction monitoring (MRM) mode. Two transitions were monitored for each GA compound, including one as a quantifier and the other as a qualifier. The MRM transitions and the corresponding collision energies were listed in Supplementary Table 2.

Molecular Phylogenetic Analysis

The protein sequences of the CYP72A subfamily from three Brassicaceae plants (*Arabidopsis thaliana*, *Capsella rubella*, and *Brassica rapa*), rice (*Oryza sativa*) and soybean (*Glycine max* L.) were extracted from a previous report[20]. A maximum likelihood tree was constructed using MEGA6.0 software with 1,000 boot-strap replicates[42].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We cordially thank Dr. Xiangdong Fu (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences) for providing the *ga1-t* mutant seeds and Dr. Qi Xie (Institute of Genetics and Developmental Biology,

Chinese Academy of Sciences) for providing the vector pCAMBIA 1300-pYAO-cas9. We would also like to thank Dr. Jianqiang Wu (Kunming Institute of Botany, Chinese Academy of Sciences) for giving us the GA standard. This work was financially supported by the National Key R&D Program of China (Grant No.2018YFA0900600) to G.W., "Priority Research Program" of the Chinese Academy of Science (Grant No. ZDRW-ZS-2019-2) to G.W., and the State Key Laboratory of Plant Genomics of China (SKLPG2016A-15) to G.W., and National Natural Science Foundation of China (Grant No. 31770398) to J.C.

References

1. Cowling RJ, Kamiya Y, Seto H & Harberd NP Gibberellin dose-response regulation of GA_4 gene transcript levels in Arabidopsis. Plant Physiol. 117, 1195–1203 (1998). [PubMed: 9701576]
2. Yang YY et al. Effects of gibberellins on seed germination of phytochrome deficient mutants of *Arabidopsis thaliana*. Plant Cell Physiol. 36, 1205–1211 (1995). [PubMed: 8564296]
3. Magome H et al. CYP714B1 and CYP714B2 encode gibberellin 13-oxidases that reduce gibberellin activity in rice. Proc. Natl. Acad. Sci. USA 110, 1947–1952 (2013). [PubMed: 23319637]
4. Blazquez MA, Green R, Nilsson O, Sussman MR & Weigel D Gibberellins promote flowering of Arabidopsis by activating the LEAFY promoter. Plant Cell 10, 791–800 (1998). [PubMed: 9596637]
5. Eriksson S, Bohlenius H, Moritz T & Nilsson O GA_4 is the active gibberellin in the regulation of LEAFY transcription and Arabidopsis floral initiation. Plant Cell 18, 2172–2181 (2006). [PubMed: 16920780]
6. Hedden P & Thomas SG Gibberellin biosynthesis and its regulation. Biochem. J 444, 11–25 (2012). [PubMed: 22533671]
7. Zhu YY et al. ELONGATED UPPERMOST INTERNODE encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice. Plant Cell 18, 442–456 (2006). [PubMed: 16399803]
8. Varbanova M et al. Methylation of gibberellins by Arabidopsis GAMT1 and GAMT2. Plant Cell 19, 32–45 (2007). [PubMed: 17220201]
9. Rieu I et al. Genetic analysis reveals that C19-GA 2-oxidation is a major gibberellin inactivation pathway in Arabidopsis. Plant Cell 20, 2420–2436 (2008). [PubMed: 18805991]
10. Hu YL et al. Gibberellins play an essential role in late embryogenesis of Arabidopsis. Nature Plants 4, 289–298 (2018). [PubMed: 29725104]
11. Kanno Y et al. Comprehensive hormone profiling in developing Arabidopsis seeds: examination of the site of ABA biosynthesis, ABA transport and hormone interactions. Plant Cell Physiol. 51, 1988–2001 (2010). [PubMed: 20959378]
12. Zhang YY et al. Two Arabidopsis cytochrome P450 monooxygenases, CYP714A1 and CYP714A2, function redundantly in plant development through gibberellin deactivation. Plant J. 67, 342–353 (2011). [PubMed: 21457373]
13. Nomura T et al. Functional analysis of arabidopsis CYP714A1 and CYP714A2 reveals that they are distinct gibberellin modification enzymes. Plant Cell Physiol. 54, 1837–1851 (2013). [PubMed: 24009336]
14. Wang C et al. Structural analyses of short-chain prenyltransferases identify an evolutionarily conserved GFPPS clade in Brassicaceae plants. Mol. Plant 9, 195–204 (2016). [PubMed: 26537048]
15. Shao J et al. (+)-Thalianatriene and (−)-retigeranin B catalyzed by sesterterpene synthases from *Arabidopsis thaliana*. Org. Lett 19, 1816–1819 (2017). [PubMed: 28350168]
16. Chen Q et al. Recently duplicated sesterterpene (C25) gene clusters in *Arabidopsis thaliana* modulate root microbiota. Sci. China Life Sci doi:10.1007/s11427-019-9521-2 (2019, in press).
17. Jiang CF, Gao XH, Liao L, Harberd NP & Fu XD Phosphate starvation root architecture and anthocyanin accumulation responses are modulated by the gibberellin-DELLA signaling pathway in Arabidopsis. Plant Physiol. 145, 1460–1470 (2007). [PubMed: 17932308]
18. Seo M et al. Regulation of hormone metabolism in Arabidopsis seeds: phytochrome regulation of abscisic acid metabolism and abscisic acid regulation of gibberellin metabolism. Plant J. 48, 354–366 (2006). [PubMed: 17010113]

19. Finkelstein R, Reeves W, Ariizumi T & Steber C Molecular aspects of seed dormancy. Annu. Rev. Plant Biol 59, 387–415 (2008). [PubMed: 18257711]
20. Prall W, Hendy O & Thornton LE Utility of a phylogenetic perspective in structural analysis of CYP72A enzymes from flowering plants. PLoS One 11, e0163024 (2016). [PubMed: 27669508]
21. Turk EM et al. CYP72B1 inactivates brassinosteroid hormones: An intersection between photomorphogenesis and plant steroid signal transduction. Plant Physiol. 133, 1643–1653 (2003). [PubMed: 14605216]
22. Lange T, Hedden P & Graebe JE Expression cloning of a gibberellin 20-oxidase, a multifunctional enzyme involved in gibberellin biosynthesis. Proc. Natl. Acad. Sci. USA 91, 8552–8556 (1994). [PubMed: 8078921]
23. Xu YL et al. The *ga5* locus of *Arabidopsis thaliana* encodes a multifunctional gibberellin 20-oxidase – molecular cloning and functional expression. Proc. Natl. Acad. Sci. USA 92, 6640–6644 (1995). [PubMed: 7604047]
24. Yamaguchi S Gibberellin metabolism and its regulation. Annu. Rev. Plant Biol 59, 225–251 (2008). [PubMed: 18173378]
25. Shu K, Liu XD, Xie Q & He ZH Two faces of one seed: hormonal regulation of dormancy and germination. Mol. Plant 9, 34–45 (2016). [PubMed: 26343970]
26. Nelson D & Werck-Reichhart D A P450-centric view of plant evolution. Plant J. 66, 194–211 (2011). [PubMed: 21443632]
27. Yano R et al. Metabolic switching of astringent and beneficial triterpenoid saponins in soybean is achieved by a loss-of-function mutation in cytochrome P450 72A69. Plant J. 89, 527–539 (2017). [PubMed: 27775214]
28. Irmler S et al. Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450CYP72A1 as secologanin synthase. Plant J. 24, 797–804 (2000). [PubMed: 11135113]
29. Itkin M et al. Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. Science 341, 175–179 (2013). [PubMed: 23788733]
30. Umemoto N et al. Two cytochrome P450 monooxygenases catalyze early hydroxylation steps in the potato steroid glycoalkaloid biosynthetic pathway. Plant Physiol. 171 (2016).
31. Zi JC, Mafu S & Peters RJ To gibberellins and beyond! Surveying the evolution of (Di)terpenoid metabolism. Annu. Rev. Plant Biol 65, 259–286 (2014). [PubMed: 24471837]
32. Wang G et al. Terpene biosynthesis in glandular trichomes of hop. Plant Physiol. 148, 1254–1266 (2008). [PubMed: 18775972]
33. Li W et al. Nicotinate *O*-glucosylation is an evolutionarily metabolic trait important for seed germination under stress conditions in *Arabidopsis thaliana*. Plant Cell 27, 1907–1924, (2015). [PubMed: 26116607]
34. Czechowski T, Stitt M, Altmann T, Udvardi MK & Scheible WR Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. 139, 5–17 (2005). [PubMed: 16166256]
35. Wang G & Pichersky E Nicotinamidase participates in the salvage pathway of NAD biosynthesis in Arabidopsis. Plant J. 49, 1020–1029 (2007). [PubMed: 17335512]
36. Xu H et al. Characterization of the formation of branched short-chain fatty acid:CoAs for bitter acid biosynthesis in hop glandular trichomes. Mol. Plant 6, 1301–1317 (2013). [PubMed: 23300257]
37. Nelson BK, Cai X & Nebenfuhr A A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. Plant J. 51, 1126–1136 (2007). [PubMed: 17666025]
38. Urban P, Mignotte C, Kazmaier M, Delorme F & Pompon D Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. J. Biol. Chem 272, 19176–19186 (1997). [PubMed: 9235908]
39. Pompon D, Louerat B, Bronine A & Urban P Yeast expression of animal and plant P450s in optimized redox environments. Methods Enzymol. 272, 51–64 (1996). [PubMed: 8791762]

40. Martinez-Andujar C et al. Induction of 9-cis-epoxycarotenoid dioxygenase in *Arabidopsis thaliana* seeds enhances seed dormancy. Proc. Natl. Acad. Sci. USA 108, 17225–17229 (2011). [PubMed: 21969557]

41. Ma XD et al. CHR729 is a CHD3 protein that controls seedling development in rice. PLoS One 10, e0138934 (2015). [PubMed: 26398683]

42. Tamura K, Stecher G, Peterson D, Filipski A & Kumar S MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol 30, 2725–2729 (2013). [PubMed: 24132122]

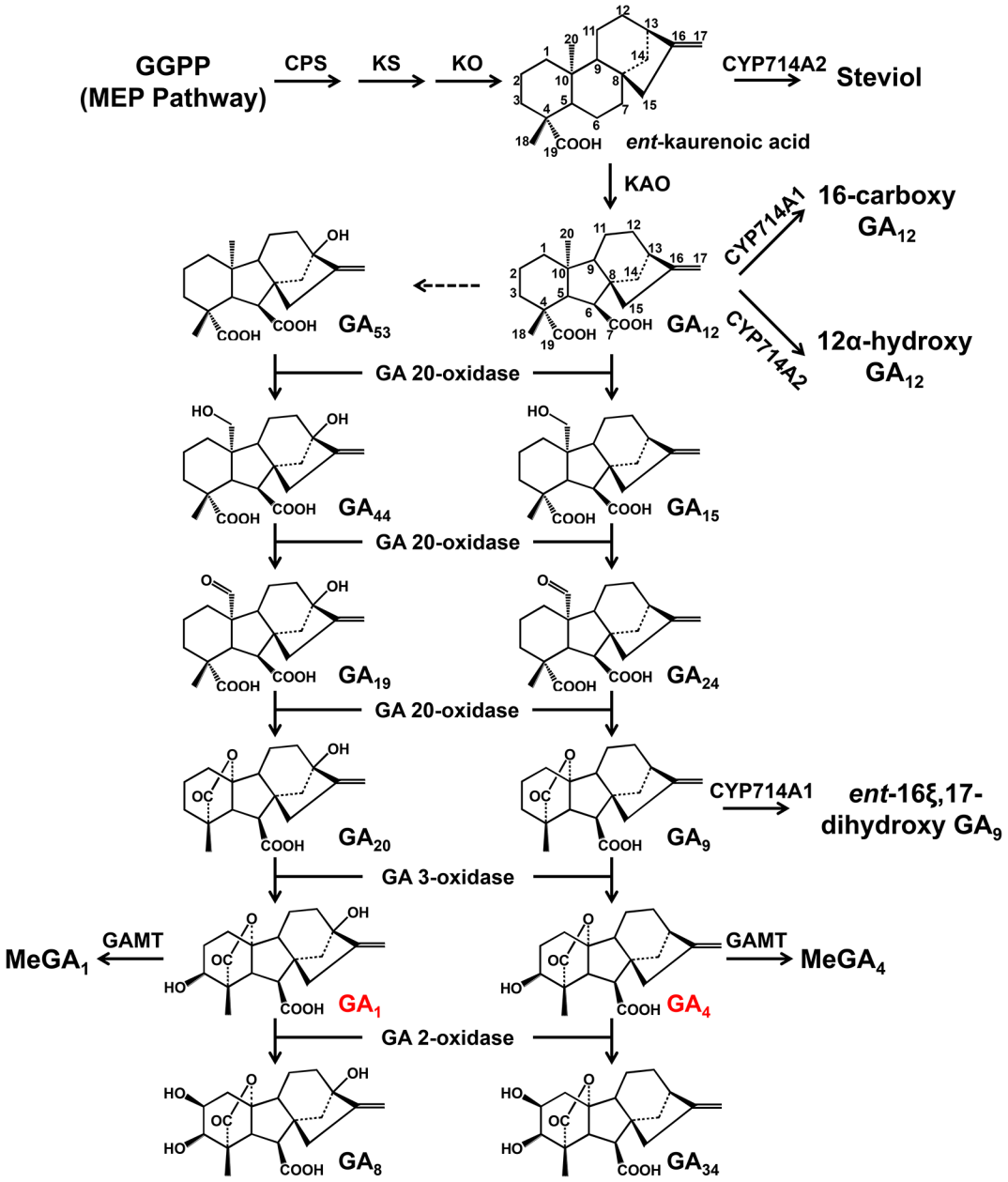

Fig. 1. The GA biosynthesis pathway in Arabidopsis.

All enzymes mapped to the GA biosynthesis pathway were verified with enzymatic assays and chemical profiling of loss-of-function mutants. The dashed line indicates uncharacterized enzymatic step in Arabidopsis. The carbon backbone of GA_{12} is labeled with numbers, and the bioactive GAs (GA_1 and GA_4) are marked in red. CPS, *ent*-copalyl diphosphate synthase; GAMT, GA methyltransferase; GGPP, geranylgeranyl diphosphate; KAO, *ent*-kaurenoic acid oxidase; KO, *ent*-kaurene oxidase; KS, *ent*-kaurenoic acid synthase; MEP, 2-*C*-methyl-*D*-erythritol 4-phosphate.

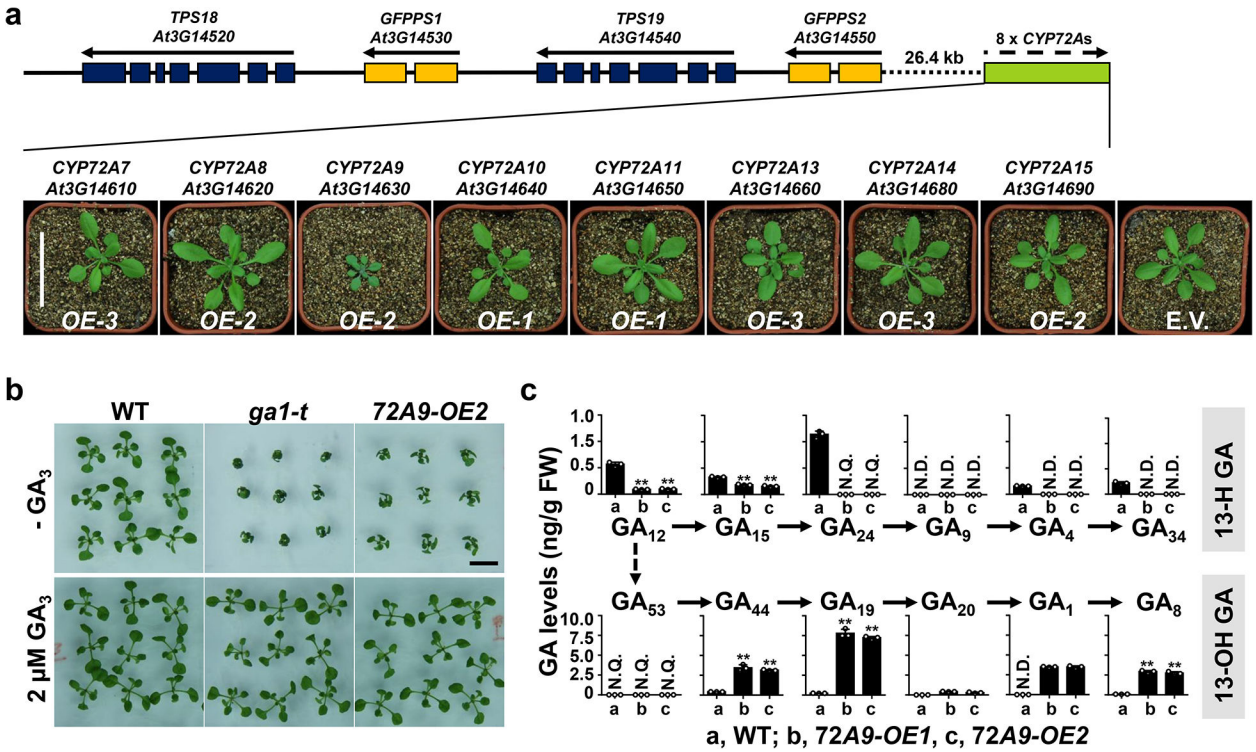

Fig. 2. Overexpression of *CYP72A9* results in a dwarf phenotype and decreases endogenous GA_4 levels.

a. GFPPS-sesterTPS-P450 gene cluster and phenotypes of Arabidopsis with increased expression level of each *AtCYP72* (*CYP72A7, A8, A9, A10, A11, A13, A14*, and *A15*). All plants were grown under the same growth conditions and images were taken at 24 days after germination. This experiment was repeated at least three times with similar results. E.V., empty vector; Scale bar = 5 cm.

b. Exogenous applied GA_3 rescues the growth of *ga1-t* mutant and *CYP72A9*-overexpressing Arabidopsis. All one-week-old seedlings, which grown on ½ MS agar medium, were transferred onto ½ MS agar medium and grown for another week with or without 2 μM bioactive GA_3. This experiment was repeated three times with similar results. WT, wild type. Scale bar = 1 cm.

c. Profile of endogenous GAs in WT and two independent *CYP72A9*-overexpressing lines. The GA levels in the rosette leaves of 4-week-old Arabidopsis are presented as the means ± SDs (n = 3 biologically independent samples). **, significant difference from WT ($P < 0.01$; two-tailed Student's t-test). N.D., not detectable; N.Q., detected, but not quantifiable due to low abundance.

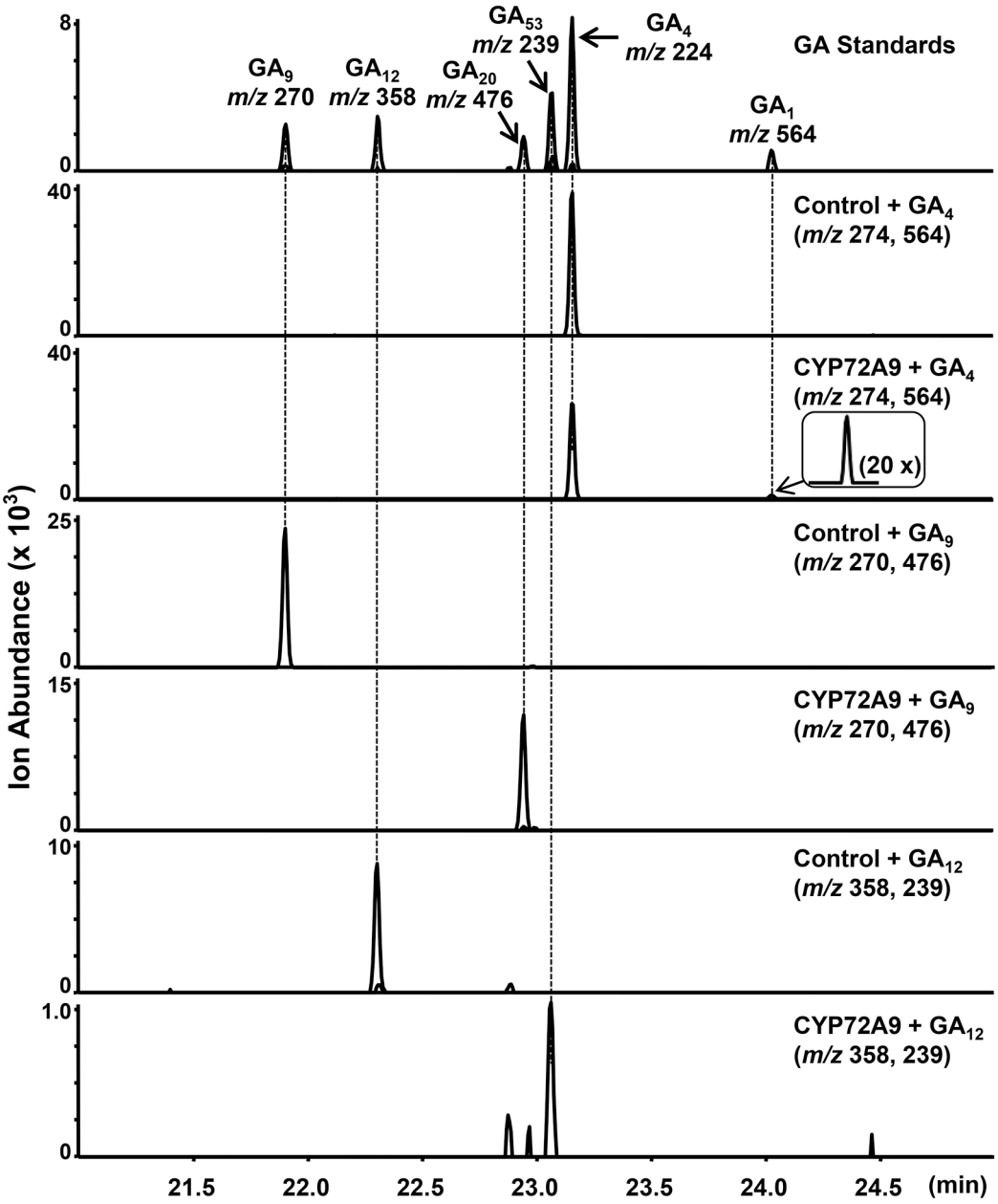

Fig. 3. CYP72A9 is a GA 13-hydroxylase.

CYP72A9 converted GA_4, GA_9, and GA_{12} to GA_1, GA_{20}, and GA_{53}, as verified by comparison to authentic standards. Chromatogram of selected ions of *m/z* 546 for GA1, *m/z* 224 for GA_4, *m/z* 270 for GA_9, *m/z* 358 for GA_{12}, *m/z* 476 for GA_{20}, and *m/z* 239 for GA_{53}. It is noteworthy that the *y* axis scale for each reaction is arbitrary for clarity, and the GA_1 region has been amplified 20 times to allow easier visualization (inlet window). Control, yeast strain harboring pESC-Leu empty vector. This experiment was repeated at least three times with similar results.

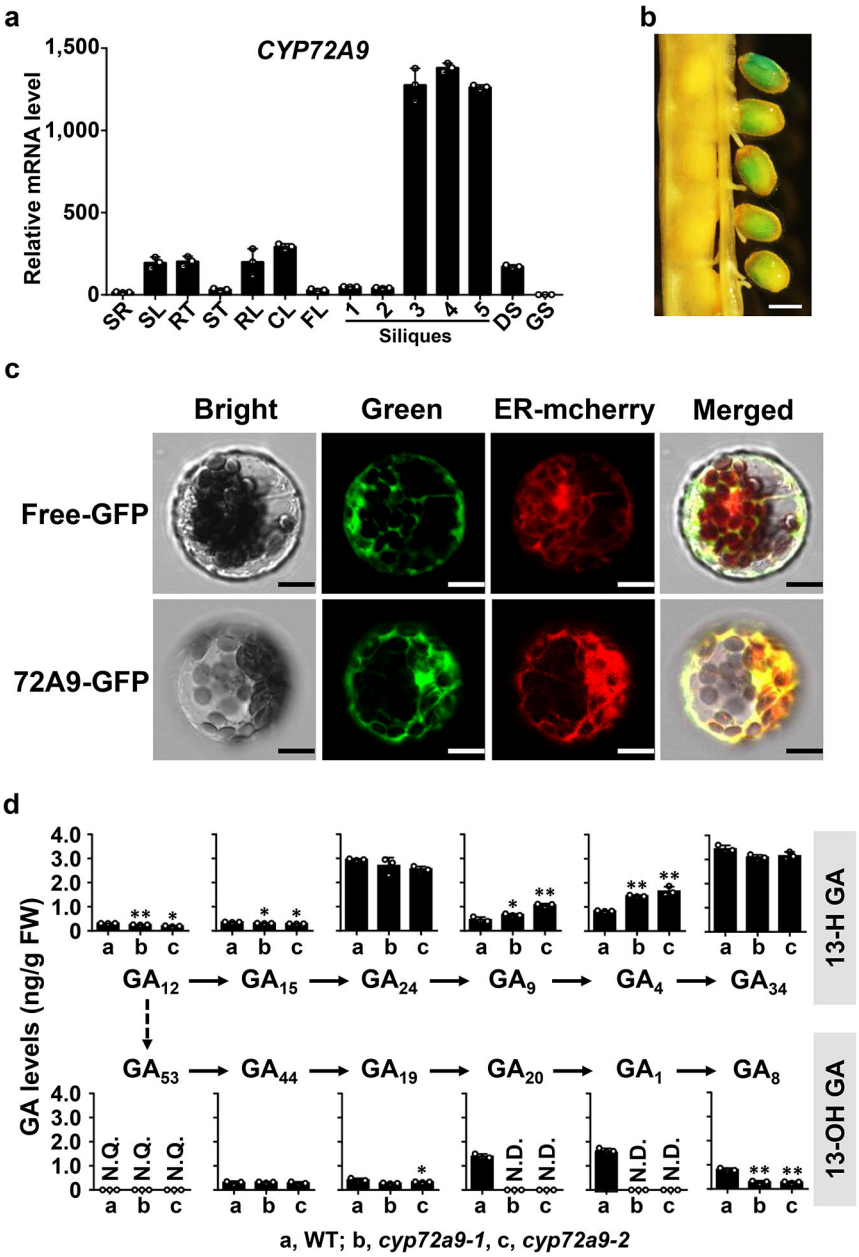

Fig. 4. Tissue-specific expression and subcellular localization of CYP72A9.

a. qRT-PCR analysis of *CYP72A9* transcript levels in different tissues. Error bars represent the SD of three independent experiments. SR, root of 10-day-old seedlings; SL, leaf of 10-day-old seedlings; RO, root of mature plants; RL, rosette leaf; CL, cauline leaf; FL, flowers; GS, germinating seeds; DS, dry seeds; Silique samples were prepared followed the reference by Varbanova et al[8]. *At1g13320*, *At2g28390*, and At4g34270 were used as reference genes in this analysis. The lowest level of *CYP72A9* transcript in germinating seeds was set as 1.0. **b**. Histochemical GUS staining of siliques (stages 9 and 10) from *Pro72A9:GUS* transgenic plants. Scale bar = 0.2 mm. Silique samples had been stained for 24 h before imaging. This experiment was repeated two times with similar results.

c. Subcellular localization of CYP72A9 in Arabidopsis leaf-mesophyll protoplasts. The ER was revealed by mCherry marker protein (Nelson et al., 2007). Scale bar = 5 μm. This experiment was repeated two times with similar results.

d. Profiles of endogenous GAs in WT and two independent *cyp72a9* mutants. The GA levels in the developing seeds/siliques of mature Arabidopsis are presented as the means ± SDs (n = 3 biologically independent samples). **, significant difference from WT ($P < 0.01$; two-tailed Student's t-test); *, ($P < 0.05$; two-tailed Student's t-test). N.D., not detectable; N.Q., detected, but not quantifiable due to low abundance.

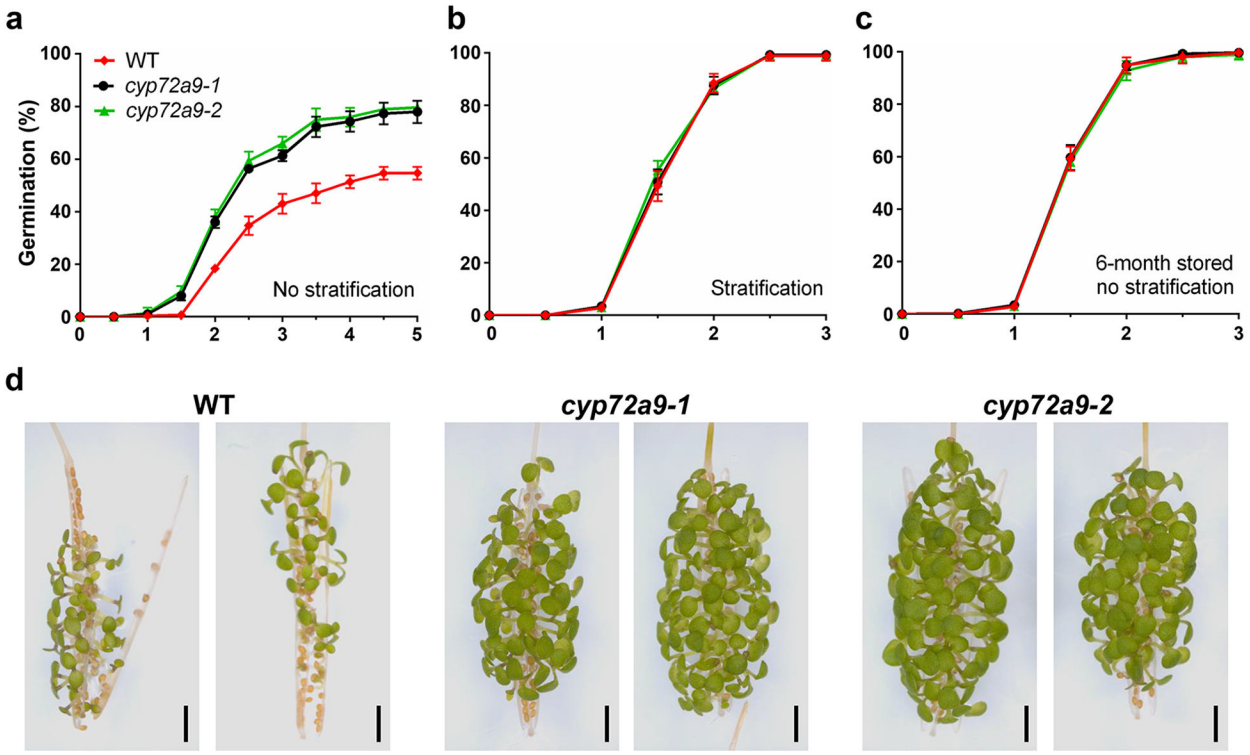

Fig. 5. Decreased primary seed dormancy of *cyp72a9* mutants.

a. Germination of fresh harvested WT and *cyp72a9* seeds on water-saturated filter paper without stratification treatment (4 °C for 3 days). Each data point represents the means ± SDs (n = 5 biologically independent experiments).

b. Germination of fresh harvested WT and *cyp72a9* seeds on water-saturated filter paper with stratification treatment (4 °C for 3 days). Each data point represents the means ± SDs (n = 5 biologically independent experiments).

c. Germination of 6-month dry storage WT and *cyp72a9* seeds on water-saturated filter paper without stratification treatment. Each data point represents the means ± SDs (n = 5 biologically independent experiments).

d. Representative images of precocious germination tests of WT and *cyp72a9* siliques. Pictures were taken 14 days after plating on ½MS medium.

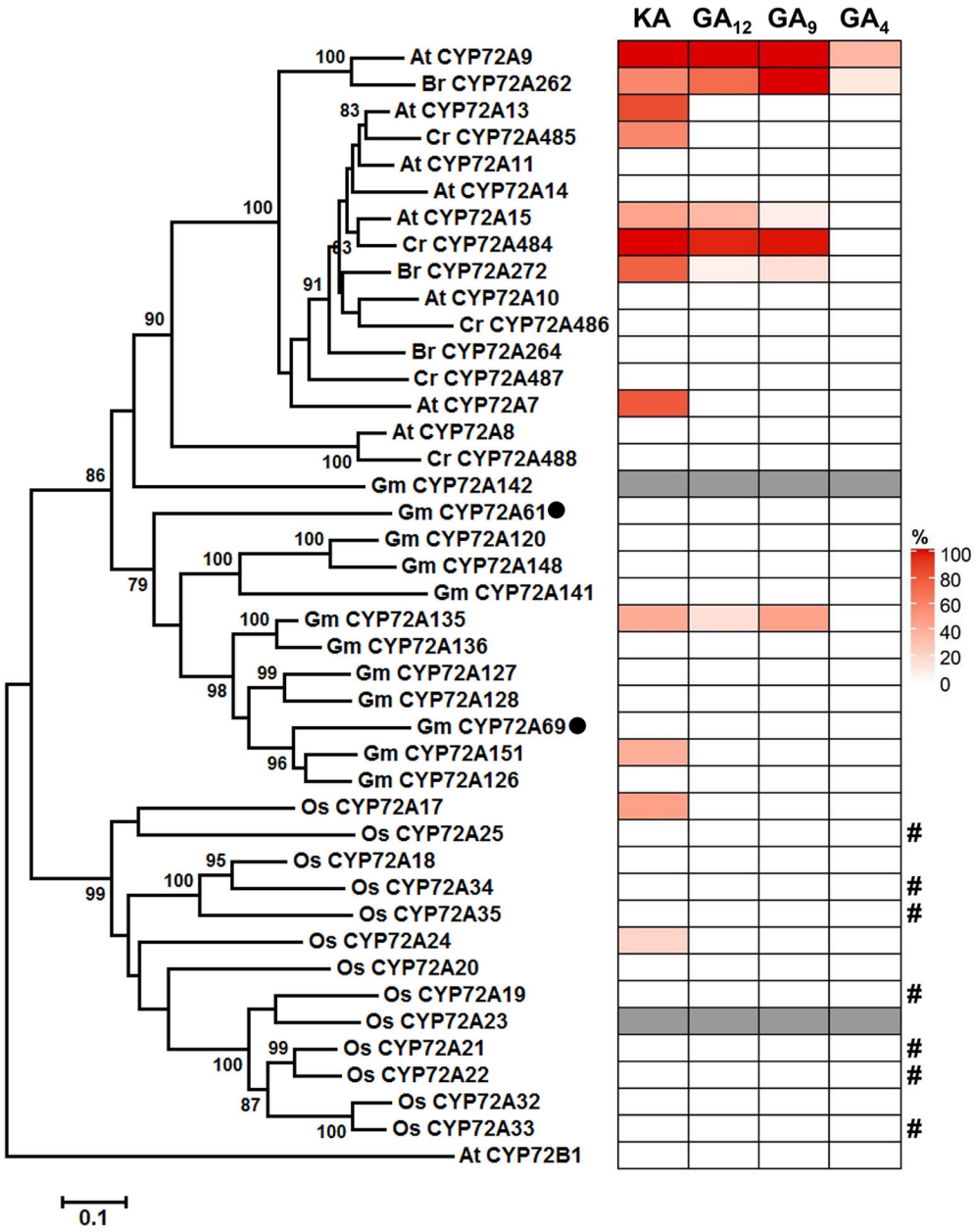

Fig. 6. Phylogenetic and biochemical analysis of CYP72A proteins from three Brassicaceae plants, rice, and soybean.

Gray color means no gene cloned in this study. GmCYP72A61 and GmCYP72A69, previously identified as triterpene oxidase, are indicated by black circles. Relative activity of various CYP72A proteins is expressed as the substrate conversion ratio (%). Values represent means from two independent experiments. It is noteworthy that the product of AtCYP72A15 and GmCYP72A135 using GA_9 as the substrate is a mixture of GA_{20} and one unidentified hydroxylated GA_9 (Supplementary Figs. 12 and 13). #, Inactivity with tested GA substrates might just reflect unsuccessful P450 protein expression in WAT11 yeast strain, which not detected by western blot (Supplementary Fig. 16).

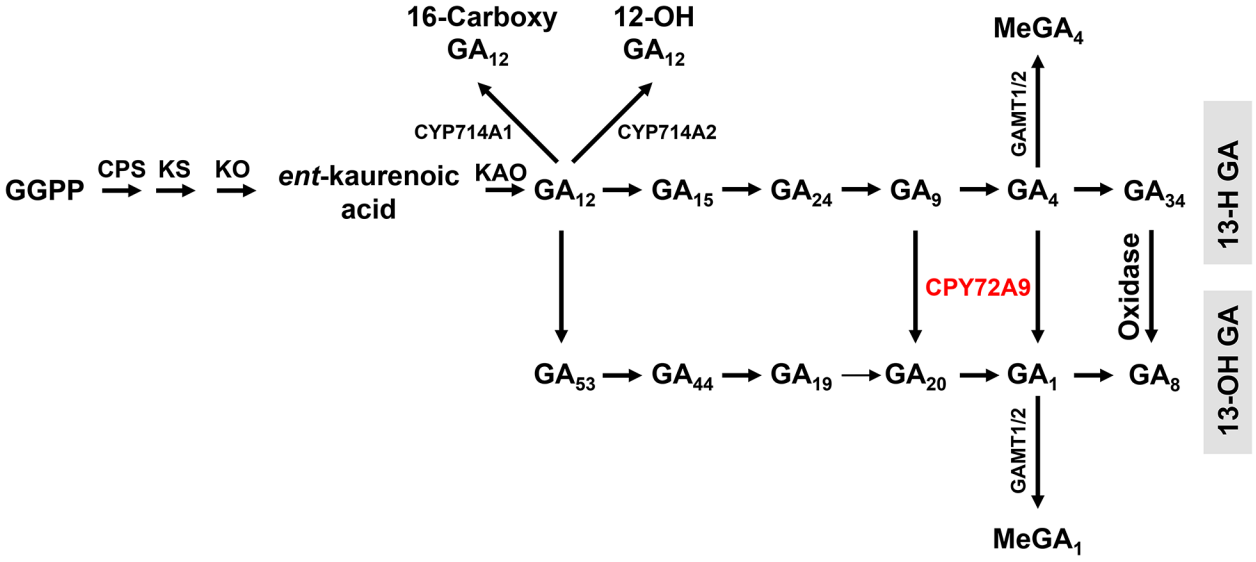

Fig. 7. Updated GA metabolism in developing seeds/silique of Arabidopsis.

CYP72A is highlighted in red. The arrow between GA_{19} and GA_{20} (in Arabidopsis pathway) is thinner than the others to show the low catalytic efficiency of this reaction in Arabidopsis, at least in developing seeds/silique. Uncharacterized oxidase(s) catalyzed the conversion from GA_{34} to GA_8.