Check for updates

the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2025) **124,** e70488 doi: 10.1111/tpj.70488

NGATHA carpel development genes evolved in the common ancestor of seed plants

Ignacio Cota[1,*], Silvia Moschin[2,3], Elisabetta Offer[2,3], Irene Martínez-Fernández[4], Francesco Magnanimi[5], Barbara Ambrose[6], Sebastiano Nigris[2,3], Barbara Baldan[2,3], Cristina Ferrándiz[4] and Soraya Pelaz[1,7,*]

[1]*Centre for Research in Agricultural Genomics, CSIC-IRTA-UAB-UB, Barcelona, Spain,*
[2]*Department of Biology, University of Padova, Padova, Italy,*
[3]*Botanical Garden of Padova, University of Padova, Padova, Italy,*
[4]*Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas, Universitat Politècnica de València, Valencia, Spain,*
[5]*Department of Biology and Biotechnology, Sapienza Università di Roma, Rome, Italy,*
[6]*The New York Botanical Garden, Bronx, New York, USA, and*
[7]*Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain*

Received 24 April 2025; revised 29 August 2025; accepted 11 September 2025.
*For correspondence (e-mail ignacio.cota@cragenomica.es and soraya.pelaz@cragenomica.es).

SUMMARY

The evolution of the carpel, the defining feature of angiosperms, remains a fundamental question in plant biology. Understanding how this organ originated is crucial because it underpins the reproductive success and diversity of flowering plants. Here, we investigated the functional conservation between gymnosperms and angiosperms of key transcription factors involved in carpel development. We found that *Ginkgo biloba* homologs can functionally substitute for their angiosperm counterparts in stigma development. We discovered that *GbRAV5* is related to angiosperm *NGA* genes, challenging previous notions that these are exclusive to angiosperms, and we found a parallel loss of the AP2 domain in gymnosperms providing a rare snapshot of how protein families evolve. Conserved protein interactions and overlapping expression patterns of *GbRAV5* and *GbHEC* in *Ginkgo* ovules suggest that the molecular toolkit for carpel development was largely present in the last common ancestor of seed plants, offering new insights into the evolution of reproductive structures.

Keywords: *NGATHA*, *RAV*, carpel development, gymnosperms, protein evolution, Ginkgo, gene family evolution.

INTRODUCTION

Angiosperms, also called flowering plants, are arguably the most diverse and successful group of land plants (Christenhusz & Byng, 2016). They appeared late in evolutionary history compared with all other land plant lineages and rapidly diversified (Christenhusz & Byng, 2016; Dimitrov et al., 2023; Magallón & Castillo, 2009). Part of their extraordinary success has been attributed to the evolution of a highly specialized reproductive structure, the flower, and most specifically to the carpel, the organ that contains and protects the ovules. The carpel shows adaptations that facilitate pollination and growth of the pollen tube and enables overall reproductive success (Ferrándiz et al., 2010). The assembly of carpels is called the gynoecium, which is usually divided into an apical stigma, a medial style, and a basal ovary, connected to the rest of the flower by the gynophore. The carpel is in fact the defining characteristic of the angiosperms, whose name derives from the Greek words *angeion* (vessel) and *sperma* (seed). Hence, the word angiosperm could be loosely translated as 'enclosed seed' (Wang, 2023).

Because of the rapid evolution and sudden appearance of angiosperms in the fossil record, the origin of their evolutionary innovations is unclear. Gymnosperms (literally 'naked seeds') are the only other extant group of seed plants, sister to angiosperms (Bowe et al., 2000; Chaw et al., 2000). They do not produce flowers and therefore lack carpels; their ovules are exposed and usually develop in the axils of leaf-like organs called megasporophylls. Despite these differences, many genes related to flower organ development have gymnosperm orthologs (Melzer et al., 2010; Theißen & Becker, 2004; Zhang et al., 2004).

© 2025 The Author(s).
The Plant Journal published by Society for Experimental Biology and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

However, gymnosperms have long life cycles (Bond, 1989) and large complex genomes (Wan et al., 2022), and they are not amenable to genetic manipulation, hindering their use in functional studies despite their key phylogenetic position. Even though recent advances have allowed the transformation and genome editing of several gymnosperm species (Cui et al., 2021; Nanasato et al., 2021; Poovaiah et al., 2021), heterologous expression of gymnosperm genes in the model angiosperm *Arabidopsis thaliana* (Arabidopsis) remains, despite its limitations (Kramer, 2015), a useful tool to infer gymnosperm gene function and study angiosperm evolution.

Diverse transcription factor families in Arabidopsis participate in carpel development (Alvarez & Smyth, 1999; Ballester et al., 2021; Crawford & Yanofsky, 2011; Gremski et al., 2007; Mutte et al., 2018; Pfannebecker et al., 2017). The four *NGATHA* (*NGA*) genes in Arabidopsis have a redundant role in the formation of style and stigma: individual *nga* mutants are relatively unaffected, whereas in higher order mutants the formation of style and stigma is completely disrupted (Alvarez et al., 2009; Trigueros et al., 2009). *NGA* genes belong to the *RAV* family, which is characterized by the presence of a B3 DNA-binding domain. There are 13 *RAV* genes in Arabidopsis: six belong to the group of canonical *RAV* genes or class I, defined by the possession of an additional AP2 domain, and seven (including four *NGA* and three related *NGA*-like genes) belong to class II, devoid of AP2 (Romanel et al., 2009). *NGA* genes are considered exclusive to angiosperms and have only been described properly in eudicots, where their function is widely conserved (Fourquin & Ferrándiz, 2014), whereas class I *RAV* genes appear in all streptophytes, including streptophyte algae, bryophytes, lycophytes, ferns, and seed plants (Mutte et al., 2018; Romanel et al., 2009).

The three *HECATE* (*HEC*) genes in Arabidopsis also act redundantly: individual *hec* mutants have unaltered fertility (Gremski et al., 2007), but triple *hec1 hec2 hec3* mutants completely lack stigmatic tissue and transmitting tract, and are fully sterile (Crawford & Yanofsky, 2011). Conversely, overexpression of *HEC* genes causes the formation of ectopic stigmatic tissue (Gremski et al., 2007). *HEC* genes belong to the family of basic helix–loop–helix (bHLH) transcription factors and are exclusive to seed plants (Pfannebecker et al., 2017).

SPATULA (*SPT*) is another member of the bHLH family. An Arabidopsis *spt* mutant shows defects in apical carpel fusion, reduced stigmatic tissue, and loss of transmitting tract (Alvarez & Smyth, 1999). Putative *SPT* orthologs have been identified in gymnosperms (Pfannebecker et al., 2017). The NGA and HEC transcription factors form a protein complex that directs stigma development in Arabidopsis. This complex activates the expression of INDEHISCENT (a close paralog of HEC present in Brassicaceae), whose protein is incorporated into the complex. The complex then activates the expression of *SPT*, and the protein integrates into the complex to regulate other downstream genes involved in stigma development (Ballester et al., 2021).

Another factor involved in the formation of the carpel is *CRABS CLAW* (*CRC*). A *crc* mutant has a short gynoecium with carpels unfused at the apex and a reduced style and lacks nectaries (Alvarez & Smyth, 1999; Bowman & Smyth, 1999). *CRC* belongs to a small family of transcription factors called *YABBY*, with six members in Arabidopsis (Romanova et al., 2021) involved in the establishment of abaxial–adaxial polarity (Bowman, 2000; Bowman et al., 2002; Meister et al., 2005). *YABBY* genes were long thought to be exclusive to seed plants but have recently been reported in lycophytes and hornworts (Romanova et al., 2021).

Members of these families have been subjected to phylogenetic and *in situ* hybridization studies in gymnosperms (D'Apice et al., 2022; Pabón-Mora et al., 2014; Pfannebecker et al., 2017). However, few studies have assessed their functional properties, or how the molecular evolution of regulatory proteins underlies the diversification of their functional roles. Indeed, protein evolution, from small modifications to drastic rearrangements of domains (Cheng, 2007; Cheng et al., 2006; Jaenicke, 1987), is a fundamental driver of adaptive innovations, allowing the acquisition of novel capabilities and roles within an organism (Ferrándiz et al., 2010), mostly through speciation-associated bursts of evolution (Eldredge & Gould, 1972; Pagel et al., 2006; Stolyarova et al., 2019; Webster et al., 2003). In this work, we have tested the ability of gymnosperm homologs to function as their angiosperm counterparts and provided evidence to examine the evolution of gene families critical for the formation of the carpel.

RESULTS

GbRAV5 phenocopies *NGA* and restores stigma formation in a *nga1 nga3* mutant background

NGA genes have been generally considered exclusive to angiosperms (Alvarez et al., 2009; Becker, 2020; Fourquin & Ferrándiz, 2014; Trigueros et al., 2009), except for a phylogenetic study that placed the origin of *NGA* genes prior to the split between gymnosperms and angiosperms (Pfannebecker et al., 2017). Thus, we searched broadly for *RAV* family homologs in the genome of the gymnosperm *Ginkgo biloba* (Ginkgo) (Guan et al., 2019). No class II *RAV* (B3-only) genes were identified, but five class I *RAV* (AP2+B3) genes were found: *Gb_18035* (*GbRAV1*), *Gb_15838* (*GbRAV2*), *Gb_33917* (*GbRAV3*), *Gb_25573* (*GbRAV4*), and *Gb_38819* (*GbRAV5*). These were subsequently used for heterologous expression in Arabidopsis.

© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Among them, only *Gb_38819* (*GbRAV5*) resulted in a phenotype reminiscent of the overexpression of *NGA* genes (Alvarez et al., 2009; Trigueros et al., 2009) and was studied further.

Most of the 35S::*GbRAV5* plants had extreme growth reduction and died early. The three surviving hemizygous T1 plants were dwarf with narrow epinastic leaves (Figures 1a, 2e, and 6a,c). Multiple defects were found in flower development: fasciated stem, reduced number of petals (Figures 1c, 2i, and 6f), and elongated gynophores (Figures 1d, 2i, 5e, and 6l). Siliques were much shorter than the wild-type and displayed a characteristic shape with 'bumpy shoulders' (Figures 1d,g, 2l, 5e, and 6l).

NGA genes function redundantly, and the formation of the stigma is severely impaired in a *nga1-4 nga3-3* mutant (*nga1 nga3*) (Figures 1f and 6i). This double mutant closely resembles the quadruple mutant but is still able to produce a small number of seeds (Alvarez et al., 2009; Trigueros et al., 2009). To test if *GbRAV5* could complement the absence of endogenous *NGA* genes, we introduced 35S::*GbRAV5* into a *nga1 nga3* mutant, both by crossing and transformation of the *nga1 nga3* plants. In the crosses of two independent lines and in the seven independently transformed plants, the expression of *GbRAV5* restored the formation of the stigma (Figure 1g), albeit also producing the characteristic carpel expansion in the distal part of the ovary, indicating that GbRAV5 is functionally equivalent to NGA factors.

GbHEC induces ectopic stigmatic tissue and partially restores stigmatic tissue formation in a *hec1 hec2 hec3* mutant background

A single *HEC* gene, *Gb_27869* (*GbHEC*), was identified in Ginkgo and expressed in Arabidopsis. Plants of two 35S::*GbHEC* lines with more extreme phenotypes formed inflorescences consisting of masses of stigmatic tissue (Figure 1h), with the inflorescence often ending in an individual stigma. Eleven lines with milder phenotypes produced flowers with persistent carpelloid sepals covered in stigmatic papillae (Figure 1i,j), especially in the distal part, often associated with ectopic ovules (Figure 1j). These carpelloid sepals did not undergo abscission after flowering and were as firmly fixed to the pedicel as the carpels themselves (Figure 1k). The stigmatic papillae sometimes also appeared in petals and stamens (Figure 1i,j).

To assess the ability of *GbHEC* to complement *hec* mutants, and because the hec1 hec2 hec3 mutant is fully sterile (Schuster et al., 2015), we transformed the progeny of a *hec1 hec2*/+ *hec3* mutant and identified the triple homozygous *hec1 hec2 hec3* among the progeny. The *hec1 hec2 hec3* flowers completely lack stigmatic tissue (Schuster et al., 2015) (Figure 1m). In contrast, five *hec1 hec2 hec3* 35S::*GbHEC* transformants developed stigmatic tissue, not only in the stigmatic region but also in the sepals (Figure 1n,o), showing that *GbHEC* and Arabidopsis *HEC* genes are functionally comparable.

The interactions needed for the formation of the NGA-HEC protein complex are conserved in Ginkgo

NGA and HEC proteins interact in Arabidopsis to trigger the genetic pathway for the development of the stigma, and this dimer later forms larger complexes with the proteins encoded by their targets, the HEC3 paralog INDEHISCENT (IND) (Girin et al., 2011) and SPT (Ballester et al., 2021). Since GbRAV5 and GbHEC were able to functionally substitute for NGA and HEC in Arabidopsis, it could be predicted that these interactions were conserved. Bimolecular fluorescence complementation (BiFC) assays showed a positive interaction between GbRAV5 and AtHEC3 (Figure 2a). Likewise, GbHEC was shown to interact with AtNGA1 and AtNGA3 (Figure 2c). We also detected a positive interaction between GbRAV5 and both GbHEC and GbSPT (Figure 2b), and between GbRAV5 and both AtIND and AtSPT (Figure 2a). No interactions with either Arabidopsis FRUITFUL (AtFUL) or PISTILLATA (AtPI), which were used as negative controls (Figure S1), were observed.

The phenotypes of constitutive expression of *NGA* genes have been shown to be dependent on the presence of *HEC* genes in Arabidopsis (Ballester et al., 2021). To investigate if this was true for *GbRAV5*, we crossed 35S::*GbRAV5* plants from two independent lines with a *hec1 hec3* mutant.

In the resulting 35S::*GbRAV5 hec1 hec3* plants, the abnormal reproductive phenotypes that were caused by 35S::*GbRAV5* expression were no longer observed, with the exception of silique size, which is coincidentally also reduced in *hec1 hec3* plants. The inflorescences and siliques closely resembled the ones in *hec1 hec3* plants (Figure 2h–l). The *hec1 hec3* mutations also caused major alleviation of the severe vegetative phenotypes caused by 35S::*GbRAV5* alone (Figure 2d–g). Similar to Arabidopsis *HEC* and *NGA*, this data show the dependence that *GbRAV5* has for *HEC* in stigma identity.

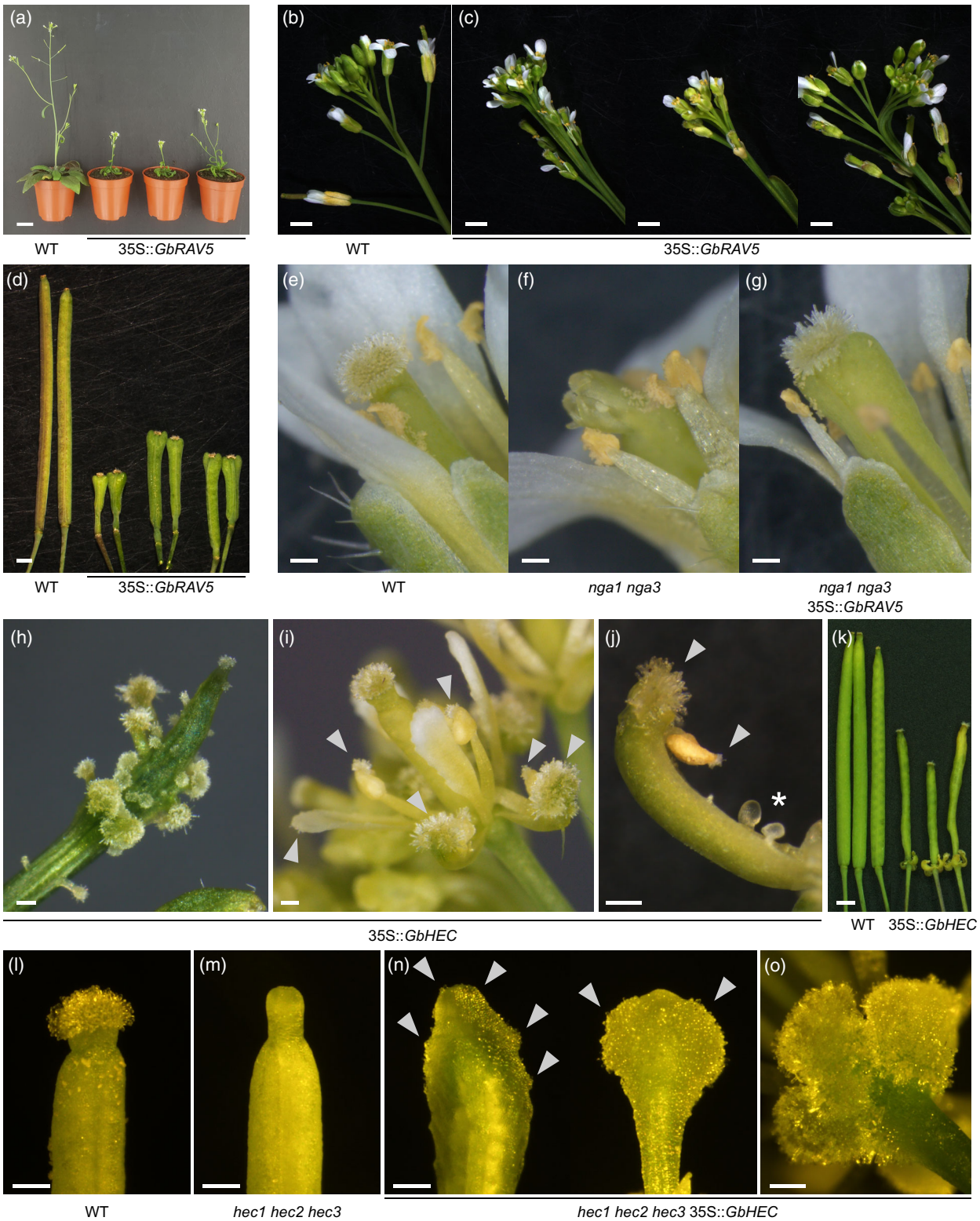
Spatial and temporal expression of *GbRAV5* and *GbHEC* overlap in *Ginkgo* ovules

We wondered if the observed interaction between GbRAV5 and GbHEC could happen in Ginkgo. To visualize their expression pattern in Ginkgo ovules, we performed *in situ* hybridization experiments using gene-specific probes for *GbRAV5* and *GbHEC*. Several time points were analyzed, including bud, pre-pollination, and pollination stages (D'Apice et al., 2021), but the expression of *GbRAV5* was only found shortly after pollination (Figure S2). The *GbRAV5* signal was clearly visible in the basal part of the nucellus around the female gametophyte and in the inner layers of the integument that at maturity will differentiate into the sclerotesta and the endotesta

© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Figure 2m,n). Expression of *GbHEC* was analyzed in pre-pollination, pollination, and post-pollination stages (Figure S2), and it was found that its signal overlaps spatially and temporally with the one obtained for *GbRAV5*, but is lower in intensity (Figure 2o,p). These results suggest the involvement of *GbRAV5* and *GbHEC* in



© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Figure 1. Expression of *GbRAV5* and *GbHEC* causes pleiotropic phenotypes and respectively restores stigma formation in *nga1 nga3* and *hec1 hec2 hec3* mutant backgrounds.
(a) 28-day-old plants of wild-type and three independent 35S::*GbRAV5* lines.
(b, c) Inflorescences of wild-type (b) and three independent 35S::*GbRAV5* lines (c).
(d) Siliques of wild-type and three independent 35S::*GbRAV5* lines.
(e–g) View of the stigma of a wild-type (e), *nga1 nga3* (f), and *nga1 nga3* 35S::*GbRAV5* (g) flower.
(h) Inflorescence of a 35S::*GbHEC* plant with a strong phenotype.
(i) Flower of a 35S::*GbHEC* plant with ectopic stigmatic tissue present in sepals, petals, and stamens (indicated by arrowheads).
(j) Close-up of a 35S::*GbHEC* carpelloid sepal with an associated stamen, both with stigmatic tissue in the apical region (indicated by arrowheads), and associated ectopic ovules (indicated by an asterisk).
(k) Siliques of wild-type and 35S::*GbHEC* plants, showing smaller size and persistent carpelloid sepals.
(l–n) View of the stigma of a wild-type (l), *hec1 hec2 hec3* showing a complete lack of stigmatic tissue (m) and *hec1 hec2 hec3* 35S::*GbHEC* with partially differentiated stigmatic tissue (n).
(o) Sepals of a *hec1 hec2 hec3* 35S::*GbHEC* flower covered in stigmatic tissue. Scale bars = 2 cm in (a), 2 mm in (b, c), 1 mm in (d, k), 200 μm in (e–j, l–o).

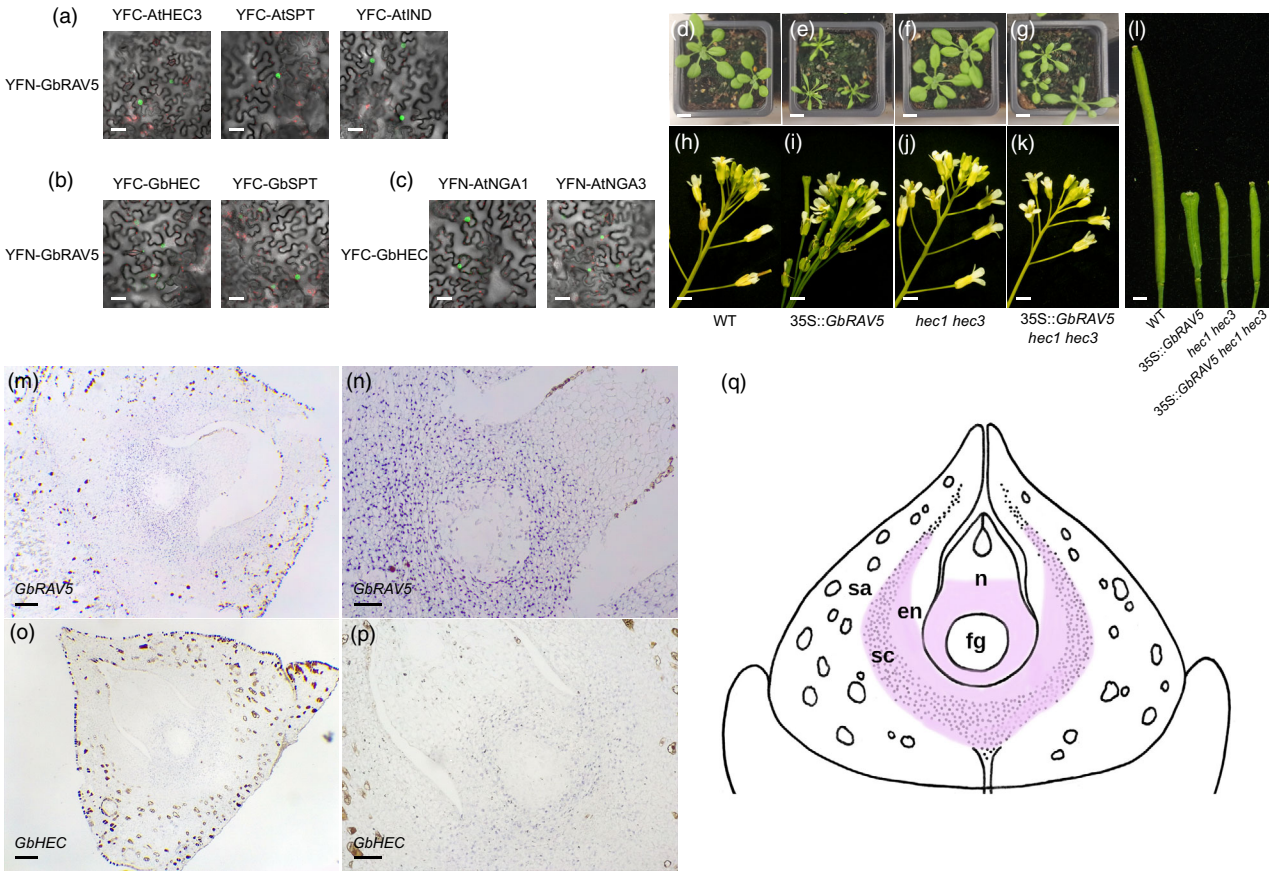


Figure 2. Interactions needed for the formation of a transcriptional complex regulating style and stigma development in Arabidopsis are conserved in Ginkgo, where *GbRAV5* and *GbHEC* are co-expressed.
(a–c) BiFC experiments showing that GbRAV5 is able to interact physically with AtHEC3, AtSPT, and AtIND (a) and also with GbSPT and GbHEC (b), and that GbHEC is able to interact physically with AtNGA1 and AtNGA3 (c).
(d–g) 18-day-old plants of wild-type (d), 35S::*GbRAV5* (e), *hec1 hec3* (f), and 35S::*GbRAV5 hec1 hec3* (g).
(h–k) Inflorescences of wild-type (h), 35S::*GbRAV5* (i), *hec1 hec3* (j) and 35S::*GbRAV5 hec1 hec3* (k).
(l) Siliques of wild-type, 35S::*GbRAV5*, *hec1 hec3*, and 35S::*GbRAV5 hec1 hec3*. Persistent stamens, petals, and sepals were eliminated in 35S::*GbRAV5* to allow better visualization of the elongated gynophore.
(m) *GbRAV5* hybridization signal, visible in the basal part of the nucellus and in the inner part of the integument.
(n) Detail of the *GbRAV5* hybridization signal, present in the lower part of the nucellus, around, but not inside, the female gametophyte.
(o) *GbHEC* hybridization signal showing a similar pattern to the one obtained for *GbRAV5* but lower in intensity.
(p) Detail of the *GbHEC* hybridization signal.
(q) Schematic representation of the gene expression domain of *GbRAV5* and *GbHEC* in the ovule of Ginkgo after pollination time. The region of expression is highlighted in pink and includes the basal part of the nucellus and the inner layers of the integument that at maturity will differentiate the sclerotesta and the endotesta. en: endotesta; fg: female gametophyte; n: nucellus; sa: sarcotesta; sc: sclerotesta. Scale bars = 50 μm in (a–c), 1 cm in (d–g), 2 mm in (h–k), 1 mm in (l), 500 μm in (m, o), 200 μm in (n, p).

© 2025 The Author(s).

13653313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the network of regulatory genes that control female reproductive structures (Figure 2q).

GbSPT and *GbYAB* show diverse degrees of functional similarity to the homologous genes in Arabidopsis

As mentioned above, the NGA-HEC dimer will later form a tetramer with SPT and IND in Arabidopsis after their activation, driving the development of the stigma (Ballester et al., 2021). The single Ginkgo *SPT* homolog, *Gb_09709* (*GbSPT*), was expressed in Arabidopsis. Nine independent T1 plants expressing *GbSPT* did not show dramatic alterations in the structure of the flowers (Figure 3c). The *spt* mutant has unfused carpels and a drastic reduction of stigmatic tissue (Alvarez & Smyth, 1999) (Figure 3b). We transformed *spt-12* (*spt*) plants with 35S::*GbSPT* to assess the ability of *GbSPT* to complement the absence of the endogenous *SPT* gene. In six independent lines, we found that the expression of *GbSPT* could completely restore the formation of the stigma (Figure 3d). The severe fruit phenotype of the *spt* mutant was also markedly alleviated with the expression of *GbSPT* (Figure 3e). This shows that *GbSPT* can largely substitute for the Arabidopsis *SPT*.

Because the *CRC* gene lineage is found only in angiosperms (Pfannebecker et al., 2017; Yamada et al., 2011), we decided to investigate other *YABBY* genes in Ginkgo. Four have been described: *GbYABA*, *GbYAB1B*, *GbYAB2B*, and *GbYABC* (D'Apice et al., 2022; Finet et al., 2016). Complementation of a crc mutation by the expression of *GbYAB* genes was assayed in a Ler *crc-1* mutant (*crc*). The Ler *crc* mutant produces short siliques with a large gap in the carpel apex, dividing the stigma into two sections (Alvarez & Smyth, 1999; Bowman, 2000) (Figure 3r–u). A large number of independent lines were analyzed in each case: 96 Ler *crc* 35::*GbYABA*, 47 Ler *crc* 35S::*GbYAB1B*, 96 Ler *crc* 35S::*GbYAB2B*, and 97 Ler *crc* 35S::*GbYABC*. Ler *crc* plants expressing each of the four

Figure 3. Expression of *GbSPT* restores stigma formation and partial fertility in a *spt* background, but *GbYAB* genes do not complement a *crc* mutant.
(a–d) View of the stigma of a wild-type (a), *spt* (b), 35S::*GbSPT* (c), and *spt* 35S::*GbSPT* (d).
(e) Siliques of wild-type, *spt*, 35S::*GbSPT* and *spt* 35S::*GbSPT*.
(f–q) Plants of wild-type (f, i, l, o), *crc* (g, j, m, p), *crc* 35S::*GbYABA* (h), *crc* 35S::*GbYAB1B* (k), *crc* 35S::*GbYAB2B* (n), and *crc* 35S::*GbYABC* (q).
(r–u) Siliques of wild-type, *crc*, *crc* 35S::*GbYABA* (r), *crc* 35S::*GbYAB1B* (s), *crc* 35S::*GbYAB2B* (t), *and crc* 35S::*GbYABC* (u). Scale bars = 200 μm in (a–d), 2 mm in (e), 1 cm in (f–q), 1 mm in (r–u).

© 2025 The Author(s).

13653313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

GbYAB genes exhibited remarkably similar phenotypic traits associated with the expression of YABBY genes, including stunted growth and narrow twisted leaves (Hou et al., 2020; Siegfried et al., 1999) (Figure 3h,k,n,q), but did not recover the length of the siliques (Figure 3r–u). The gap at the carpel apex was slightly reduced in some cases but not completely closed, showing very limited complementation.

Genes similar to *GbRAV5* are found in other gymnosperms and form a clade with angiosperm *NGA*

The fact that the expression of *GbRAV5* produced the same phenotypes as an *NGA* gene was surprising since these were thought to be exclusive to angiosperms (Becker, 2020; Fourquin & Ferrándiz, 2014; Trigueros et al., 2009) and the existence of *NGA genes* in gymnosperms was at least disputed (Pfannebecker et al., 2017). Furthermore, NGA proteins are very different in their domain structure from GbRAV5, since NGA lack an AP2 domain. Therefore, we investigated if genes related to *GbRAV5* were present in other gymnosperms by performing a Basic Local Alignment Search Tool (BLAST) search using published gymnosperm genomes.

A phylogenetic tree was constructed with predicted protein sequences from Ginkgo and 10 other gymnosperm species (Figure S3): a cycad (*Cycas panzhihuaensis*), five species of Cupressales (*Sequoia sempervirens*, *Sequoiadendron giganteum*, *Cryptomeria japonica*, *Thuja plicata*, and *Torreya grandis*), and four species of Pinales (*Pseudotsuga menziesii*, *Larix kaempferi*, *Picea abies*, and *Pinus taeda*). To analyze the evolutionary relationship between *GbRAV5*, its homologs in other gymnosperm species, and *NGA*, we included sequences from four species of angiosperms (*Amborella trichopoda*, *Aristolochia fimbriata*, *Eschscholzia californica*, and Arabidopsis) and the fern *Ceratopteris richardii* as an outgroup. Interestingly, 37 gymnosperm RAV proteins, including cases from all the species tested, were found to group together with NGA instead of class I RAV proteins (Figure 4a).

We noticed that the sequences from Ginkgo and *C. panzhihuaensis* had a complete AP2 domain, but in the rest of gymnosperms (Pinopsida), the predicted proteins that grouped with *NGA* had either partially degenerated (23) or absent (12) AP2 domains (Figure 4; Figure S3). Remains of partially degenerated AP2 domains were generally not recognized by automatic tools such as Expasy

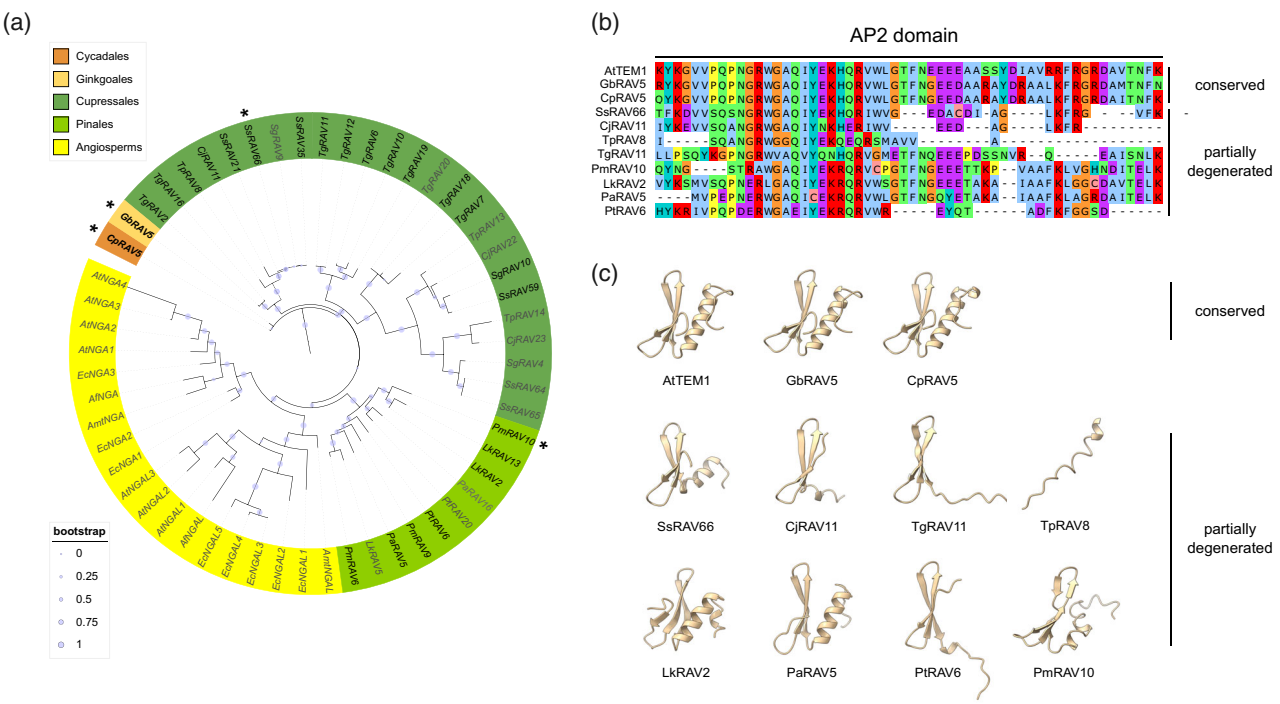


Figure 4. Gymnosperm genes related to *GbRAV5* cluster together with *NGA* genes, and those in Pinopsida have partially degenerated or absent AP2 domains.
(a) Phylogenetic analysis of *GbRAV5*-related genes in gymnosperms and angiosperms. The genes cloned and expressed in Arabidopsis in this study are marked by asterisks. Genes are shown in boldface (conserved AP2 domain), regular font (partially degenerated AP2 domain), and gray font (no remains of an AP2 domain). The tree was generated using NGphylogeny.fr and iTOL.
(b) Protein sequence alignment of representative *GbRAV5*-related proteins containing partially degenerated AP2 domains. AtTEM1 (AT1G25660.1) is included as a reference for a conserved AP2 domain. The alignment was produced using Clustal Omega, manually improved, and visualized using Jalview.
(c) Predicted protein structure of the region corresponding to the AP2 domain in several GbRAV5-related gymnosperm proteins and AtTEM1 (AT1G25660.1) as a reference. Domain structures were generated by Alphafold and visualized using ChimeraX.

© 2025 The Author(s).

13653313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Prosite (Sigrist et al., 2013), but residues with a high level of conservation in AP2 domains could be identified (Figure 4b).

To better visualize the loss of the AP2 domain, we modeled the structure of these proteins using Alphafold (Jumper et al., 2021). While the predicted proteins from Ginkgo and *Cycas* had perfectly conserved AP2 domains with the characteristic alpha helix and three antiparallel beta sheets (Allen et al., 1998; Magnani et al., 2004), highly similar to class I RAV proteins such as AT1G25560 (AtTEM1) (Osnato et al., 2012), other proteins had incomplete and distorted AP2 domains, although still recognizable in the structure of the protein (Figure 4c).

To investigate the functional properties of these *GbRAV5* homologs and include a diversity of extant gymnosperms (Yang et al., 2022), three additional genes were expressed in Arabidopsis. A *Cycas sphaerica* gene corresponding to the transcript GGZS01014648.1 (*CsRAV5*) was cloned instead of the *C. panzhihuaensis CYCAS_021568* gene (*CpRAV5*) used in the phylogenetic analysis. We had identified and cloned *CsRAV5* before the publication of the genome of *C. panzhihuaensis*, but since the two proteins differ only in eight amino acid residues out of 654, we made the assumption that they are encoded by orthologous genes in these closely related species.

Two other genes, both with partially degenerated AP2 domains, were included representing two different lineages of Pinopsida: *SESE_000652* (*SsRAV66*) from *S. sempervirens* (Cupressales) and *PSME_04433* (*PmRAV10*) from *P. menziesii* (Pinales). The plants of 14 independent lines expressing *SsRAV66* shared all the phenotypes with 35S::*GbRAV5*, showing reduced growth, narrow epinastic leaves, fasciation of the inflorescence stem, a reduction in the number of petals, an elongated gynophore, and the characteristic shape of the silique (Figure 5c,e). The result was similar in plants of 11 independent lines expressing *PmRAV10* (Figure 5d,e), although some phenotypes were not as obvious: the inflorescence stem was also fasciated, but the shape of the silique was not as distinct, and the gynophores were only minimally elongated (Figure 5e). In contrast to the plants expressing *GbRAV5* or Pinopsida genes, plants of six 35S::*CsRAV5* lines displayed normal inflorescences and flowers (Figure 5b). The siliques were short but displayed none of the defects found clearly in plants expressing *GbRAV5* or *SsRAV66* (Figure 5e).

To assess the ability of these genes to compensate for the absence of *NGA* genes, we transformed the *nga1 nga3* mutant with *CsRAV5*, *SsRAV66*, and *PmRAV10*. Eleven *nga1 nga3* 35S::*SsRAV66* and seven *nga1 nga3* 35S::*PmRAV10* independent lines in which style and stigma formation were restored were found (Figure 5i,j). Restoration of style and stigma in *nga1 nga3* 35S::*CsRAV5* was found in three plants with extreme growth reduction and was not as complete (Figure 5h).

Deletion of the AP2 domain in *GbRAV5* does not affect its function in carpel formation

Overexpression of *NGA* genes, which lack an AP2 domain, and *GbRAV5*, which has both AP2 and B3 domains, produces similar phenotypes in Arabidopsis (Alvarez et al., 2009; Trigueros et al., 2009) (Figures 1a–g and 5), and homologs of *GbRAV5* in gymnosperms other than

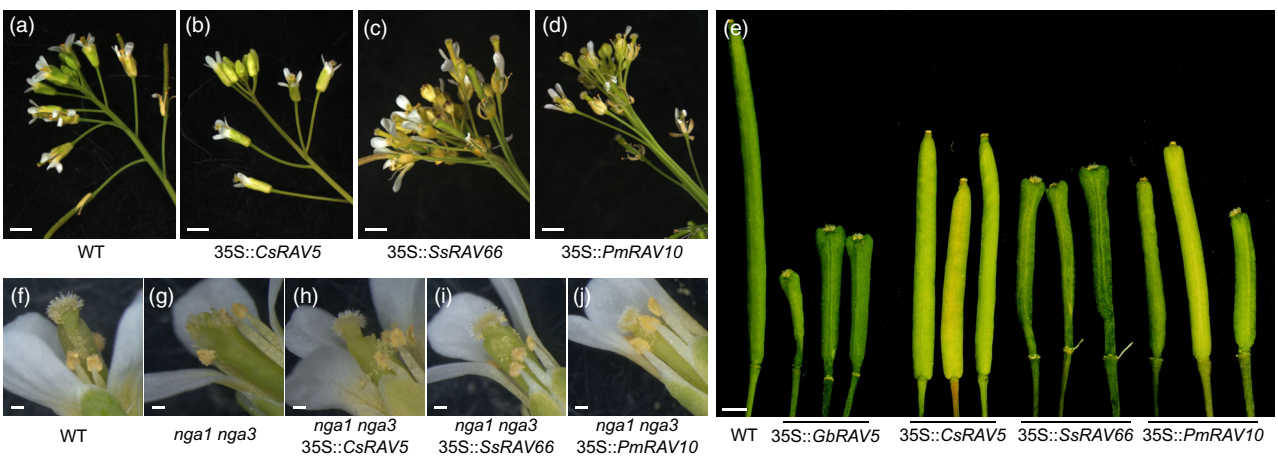


Figure 5. Other gymnosperms *GbRAV5* homologous genes produce similar phenotypes in Arabidopsis and restore the phenotypic defect of the *nga1 nga3* stigma.
(a–d) Inflorescences of wild-type (a), 35S::*CsRAV5* (from *Cycas sphaerica*) (b), 35S::*SsRAV66* (from *Sequoia sempervirens*) (c), and 35S::*PmRAV10* (from *Pseudotsuga menziesii*) (d).
(e) Siliques of wild-type and three independent lines of 35S::*GbRAV5*, 35S::*CsRAV5*, 35S::*SsRAV66*, and 35S::*PmRAV10*. Persistent stamens, petals, and sepals were eliminated in 35S::*GbRAV5*, 35S::*SsRAV66*, and 35S::*PmRAV10* to allow better visualization of the elongated gynophores.
(f–j) View of the stigma of a wild-type (f), *nga1 nga3* (g), *nga1 nga3* 35S::*CsRAV5* (h), *nga1 nga3* 35S::*SsRAV66* (i), and *nga1 nga3* 35S::*PmRAV10* (j) flower. Scale bars = 2 mm in (a–d), 1 mm in (e) and 200 μm in (f–j).

© 2025 The Author(s).

13653313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

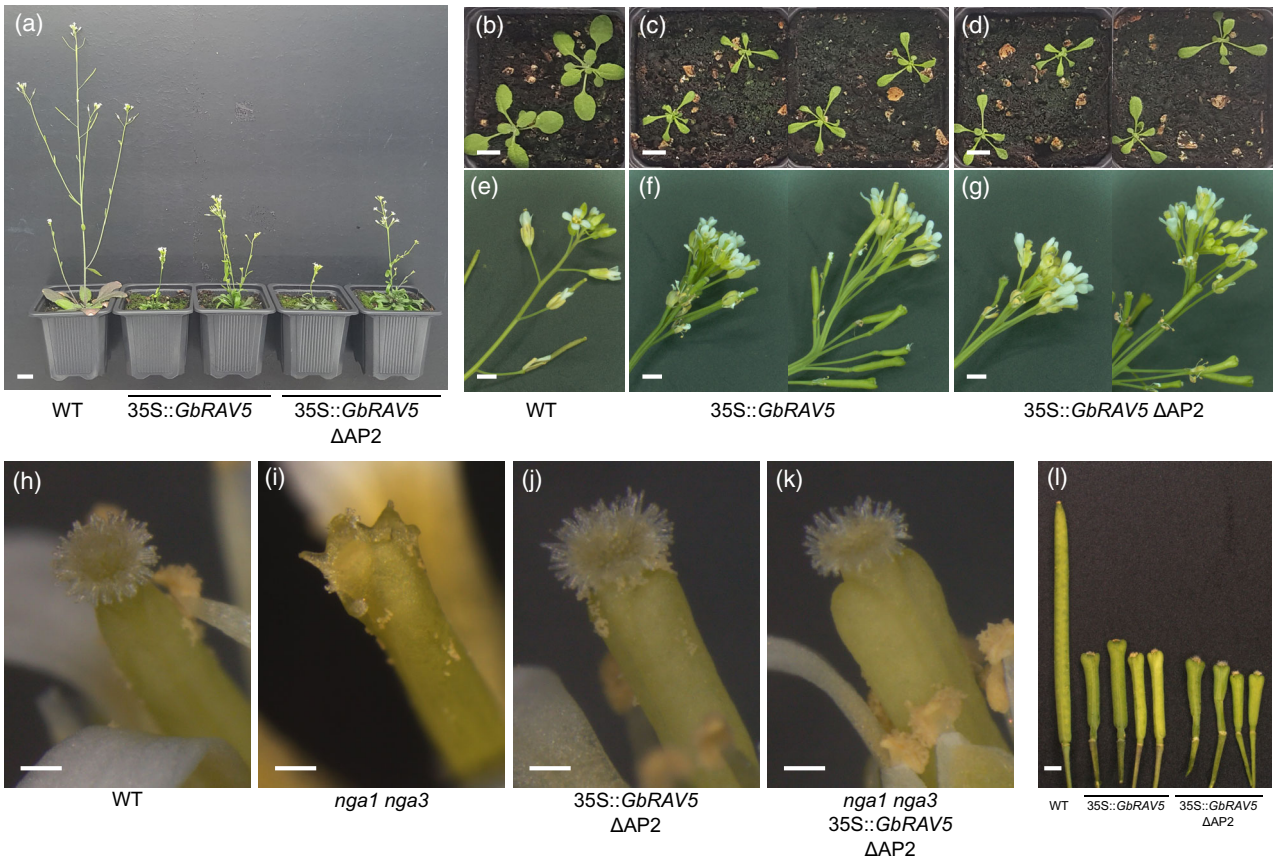


Figure 6. The deletion of the AP2 domain of *GbRAV5* does not alter its effects in Arabidopsis.
(a) 30-day-old plants of wild-type, two independent lines of 35S::*GbRAV5* and two independent lines of 35S::GbRAV5 ΔAP2.
(b–d) 15-day-old plants of wild-type (b), two independent lines of 35S::*GbRAV5* (c), and two independent lines of 35::*GbRAV5* ΔAP2 (d).
(e–g) Inflorescences of wild-type (e), two independent lines of 35S::*GbRAV5* (f), and two independent lines of 35::*GbRAV5* ΔAP2 (g).
(h–k) View of the stigma of a wild-type (h), *nga1 nga3* (i), 35::*GbRAV5* ΔAP2 (j) and *nga1 nga3* 35::*GbRAV5* ΔAP2 (k) flower.
(l) Siliques of wild-type, two independent lines of 35S::*GbRAV5* (two siliques each) and two independent lines of 35S::*GbRAV5* ΔAP2 (two siliques each). Scale bars = 2 cm in (a), 1 cm in (b–d), 2 mm in (e–g), 200 μm in (h–k), 1 mm in (l).

cycads lack a conserved AP2 domain (Figure 4). We wondered if the AP2 domain could be dispensable for mimicking *NGA* overexpression. To investigate this, we constructed a version of *GbRAV5* in which the AP2 domain was deleted and introduced it into Arabidopsis. Plants of 19 independent 35S::*GbRAV5* ΔAP2 transgenic lines were almost indistinguishable from the plants expressing the intact *GbRAV5* in their vegetative and reproductive phenotypes (Figure 6a–g,l).

To test for complementation, two independent 35S::*GbRAV5* ΔAP2 lines were crossed with *nga1 nga3*. In the resulting *nga1 nga3* 35S::*GbRAV5* Δ*AP2*, the stigma was recovered (Figure 6k), showing that the AP2 domain is not essential for the function of *GbRAV5* in stigma identity.

DISCUSSION

The carpel, the specialized organ enclosing and protecting the ovules, is one of the major evolutionary innovations of angiosperms, but its origin remains unclear. The study of gymnosperms is paramount to the evolution of angiosperms, a major transitional point in the evolution of land plants. However, little research has been done using gymnosperm genes related to carpel development. In this study, we have used heterologous expression in Arabidopsis to evaluate the functional conservation of gymnosperm homologs to Arabidopsis *NGA*, *HEC*, *SPT*, and *CRC* genes, involved in the development of carpel marginal tissues.

The use of the 35S promoter has some caveats (Amack & Antunes, 2020), as the expression derived from it does not correspond to the native spatial or temporal expression patterns of endogenous genes. However, we believe that the robust and constitutive expression achieved under its control allows us to draw meaningful conclusions about the effects of different genes in our experimental context.

We have identified *GbRAV5*, a Ginkgo AP2+B3 *RAV* gene able to phenocopy the effects of overexpressing Arabidopsis B3-only *NGA* genes (Alvarez et al., 2009; Shao

© 2025 The Author(s).

13653313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2012; Trigueros et al., 2009) and compensate their absence (Figure 1). *NGA* genes were generally thought to be specific to angiosperms, probably because previous searches were based on their domain structure, assumed to be similar to that of NGA proteins in angiosperms (Trigueros et al., 2009). However, other researchers placed the origin of the *NGA* clade in the ancestor of seed plants (Pfannebecker et al., 2017), and indeed identified three of the same genes found in our study (*GbRAV5* from Ginkgo, which they named *GibimiNGAABS2L*, as well as *PaRAV5* (*PiabNGAABS2L2*) and *PaRAV16* (*PiabNGAABS2L1*) from *P. abies*) as members of the *NGA* clade.

Curiously, *GibimiNGAABS2L* (GenBank EX932950.1) is actually a partial sequence corresponding to *GbRAV5* that is missing the AP2 domain. *PaRAV16* is a B3-only gene with no identifiable remains of the AP2 domain, and only careful visual inspection could have identified the remains of the AP2 domain in *PaRAV5*, so for all intents and purposes, this phylogenetic tree did not seem to contradict the notion that neither *NGA* genes nor their putative homologs in gymnosperms had an AP2 domain.

Our findings support the notion that *NGA* evolved in the common ancestor of gymnosperms and angiosperms, showing that *GbRAV5* and other gymnosperm genes cluster together with the *NGA* genes (Figure 4a; Figure S3). Strikingly, all the gymnosperm proteins that group together with the canonical NGA proteins, except the ones from Ginkgo and *Cycas*, show partially degenerated or absent AP2 domains (Figure 4), which make them increasingly similar in their domain organization to angiosperm NGA proteins.

It could be argued that the presence of B3-only *RAV* genes in gymnosperms might result from mistakes in annotation. Gymnosperm genomes are extraordinarily complex, characterized by long introns and highly repetitive sequences (Wan et al., 2022; Zhu et al., 2023), making genome assembly and annotation especially cumbersome. Indeed, predicted sequence databases are often contaminated with incomplete or chimeric protein sequences (Goudey et al., 2022; Rembeza & Engqvist, 2021; Schnoes et al., 2009; Steinegger & Salzberg, 2020), and the majority of proteins identified with missing domains are due to annotation artifacts (Kress et al., 2023). Although this could be true in specific cases, it is difficult to completely dismiss the evolution of AP2-less *RAV* genes in gymnosperms after we have shown that genes can be identified with increasingly degenerated but still recognizable AP2 domain remains.

In view of this evidence, we believe that the traditional classification of *RAV* genes based solely on the presence or absence of the AP2 domain does not properly capture the more complex evolutionary history of the *NGA* lineage. We propose reclassifying *GbRAV5* and related gymnosperm genes as part of the *NGA* family in the clade *gymnoNGATHA* (*gymNGA*) and have renamed them accordingly (Table S1). This implies that the *NGA* family is no longer defined by the absence of an AP2 domain, but rather that AP2 loss occurred after *NGA* function evolved in the common ancestor of seed plants. No complete AP2 has been found in Pinopsida *gymNGA*, suggesting that it had already degenerated in their last common ancestor. Thus, the AP2 domain has been independently lost at least twice in seed plants: in angiosperms and in Pinopsida. Rapid diversification and higher rates of molecular evolution (De La Torre et al., 2017) might explain why there are no remnants of AP2 in angiosperm *NGA*. An alternative hypothesis is that the loss of the AP2 domain occurred in different ways in each clade: we have shown that in Pinopsida it involves the mutational erosion of the AP2 domain, while the loss of the AP2 in angiosperms might be due to a different mechanism. It is tempting to speculate that the independent loss of the AP2 domain in Pinopsida and angiosperms could be related to some shared reproductive characteristics.

Domain losses are common in eukaryotic evolution (Zmasek & Godzik, 2011). Loss of an active phytochrome binding (APB) domain in a PIF (phytochrome-interacting factor)-like ancestor resulted in the evolution of SPT. Interestingly, loss of the APB domain also happened independently in a different protein of another PIF clade, long hypocotyl in far red (HFR) (Reymond et al., 2012). In snakes, loss of a membrane-anchoring domain paved the way for the evolution of one of the most important families of snake toxins following an apparently arbitrary change in molecular ecology which predisposed the protein for further evolution (Koludarov et al., 2023). Strikingly, the loss of the transmembrane domain also occurred independently in different clades.

Why both *CsRAV5* and *GbRAV5* contain apparently intact AP2 domains could be due to these genes having additional capabilities other than those required for *NGA* function that remained important in this clade. Many proteins are known to be promiscuous, having a dominant function and one or several secondary functions (Khersonsky et al., 2006). Intriguingly, CsRAV5 appears to have lost most of the functional properties that seem to be conserved in the other tested gymNGA, perhaps keeping the hypothetical unrelated function. Contrary to the Ginkgo lineage, which seems to have suffered little morphological and ecological change for 200 million years and is now reduced to a single species (Major, 1967; Royer et al., 2003; Zhou & Zheng, 2003), cycads have recently undergone a major adaptive radiation (Condamine et al., 2015; Nagalingum et al., 2011) and appear to show no sign of morphological stasis (Coiro & Seyfullah, 2024). Rates of evolution are correlated with speciation events, rather than time itself (Fitch & Beintema, 1990; Janzen et al., 2022; Pagel et al., 2006; Webster et al., 2003). Furthermore, species of

© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the genus *Cycas* have very different female reproductive structures compared with related cycad species, since they do not produce a true female cone as all the other cycads and most gymnosperms do, but a loose cluster of macrosporophylls (Liu et al., 2022). Thus, it is possible that CsRAV5 has deviated significantly from its function in the most common recent ancestor of seed plants.

In any case, the fact that *GbRAV5* and other *gymNGAs*, especially *SsRAV66*, behave in such a similar way when expressed in Arabidopsis, each one being at virtual opposite ends of the gymnosperm family tree, suggests that these, and not *CsRAV5*, have kept more of the original biochemical capabilities these genes had in the most recent common ancestor of gymnosperms.

The *gymNGA+NGA* clade does not include any sequences in seedless plants, and all *RAV* genes that we know of outside seed plants include an AP2 domain. Intriguingly, it was previously published that the bryophyte *Marchantia polymorpha* has a *RAV* gene with two B3 domains and no AP2 domain (Swaminathan et al., 2008). However, the GenBank number in the original publication does not correspond to any kind of *RAV* gene, and we could only identify one *Marchantia RAV* gene (*Mapoly0072s0102.1*) (previously described as *MpAP2B3-1-/MpRAV*; Flores-Sandoval et al., 2018) with the archetypical AP2+B3 domain architecture.

In summary, we have shown that *GbRAV5*, *GbHEC*, and *GbSPT* can produce phenotypes and rescue the corresponding Arabidopsis mutants similar to angiosperm counterparts (Alvarez et al., 2009; Gremski et al., 2007; Groszmann et al., 2008; Hidvégi et al., 2020; Trigueros et al., 2009) (Figures 1 and 3). Additionally, the protein–protein interactions between HEC and NGA are conserved between both species (Figure 2a–c), with *GbHEC* and GbRAV5 being co-expressed in Ginkgo ovules only after pollination (Figure 2m–p), suggesting that their function could be related to pollen growth and fertilization. Despite these remarkable similarities, there are no structures reminiscent of carpels or stigmas in gymnosperms. One possible explanation is that CRC plays a crucial role in this process. While *GbYAB* genes produce phenotypes typically associated with ectopic expression of *YABBY* genes in Arabidopsis (Hou et al., 2020; Zhang et al., 2013; Zhao et al., 2006), they fail to restore proper carpel development in a *crc* mutant. This observation, combined with the significant divergence of *CRC* from other *YABBY* genes (Fourquin et al., 2005; Lee et al., 2005) and its strong purifying selection in angiosperms (Bartholmes et al., 2012; Lee et al., 2005), points to *CRC* as a potential key innovation in the evolution of the carpel, as previous studies have suggested (Liu et al., 2023; Trigueros et al., 2009). Strikingly, the mutation of the *CRC* ortholog in the angiosperm *Physalis floridana* (*PFCRC*) results in the formation of an indeterminate gynoecium with spirally arranged flat flake-like carpels, bearing one or two exposed ovules at the base. This is an arrangement surprisingly similar to the cluster of seminiferous scales of a female strobilus in gymnosperms (Gong et al., 2021; Liu et al., 2023). Ovules are sometimes found in the outer surface of carpels in Arabidopsis *crc* mutants (Alvarez & Smyth, 2002; Eshed et al., 1999).

The evolution of new genes is essential to explain the diversity of life forms. Although some originate from non-coding DNA (Levy, 2019; Long et al., 2013; McLysaght & Guerzoni, 2015; Wu et al., 2011), most are modifications of existing genes associated with evolutionary transitions (Janzen et al., 2022). However, the origins of novel proteins are typically obscure (Bornberg-Bauer et al., 2010; Levine et al., 2006) and only very rarely can the process be reconstructed (Long et al., 2003). We have provided a beautiful example of the evolution of a new family of genes thanks to the presence of intermediate steps showing various degrees of domain degeneration in different species of gymnosperms, perhaps recapitulating a similar process occurring in angiosperms for which no traces have been found.

Our study reveals the shared evolutionary history of seed plant genes involved in carpel development, offering insights into broader mechanisms of protein evolution and pointing to potential key genetic factors in the evolution of the flower. While further research is needed to include other gymnosperm species and additional genes, our results provide better understanding of how gene functions evolve across millions of years of plant evolution.

METHODS

Plant material and growth conditions

The wild-type plant used in this study was *A. thaliana* ecotype Columbia 0 (Col-0), as were all the mutants and transformed lines derived from them, with one exception: the *crc-1* mutant, which belongs to the ecotype *Landsberg erecta* (Ler), and its derivatives. Seeds from mutants were kindly sent to us by Lars Østergaard (*spt-12*) (Ichihashi et al., 2010) and Charles P. Scutt (Ler *crc-1*) (Alvarez & Smyth, 1999).

To grow Arabidopsis, seeds were stratified for 3 days at 4°C and plants were grown in controlled conditions in a 6:1:1 (v/v/v) mix of soil:perlite:vermiculite at 22°C, under a long day photoperiod (16 h light/8 h dark).

Young leaves were collected from a Ginkgo tree growing close to the CRAG building (Cerdanyola del Vallès, Barcelona, Spain), immediately frozen in liquid nitrogen and stored at −80°C until further processing.

Generation of transgenic lines

To generate overexpression lines of Ginkgo genes, RNA was extracted from Ginkgo leaves using Maxwell® RSC Plant RNA Kit (Promega, Madison, WI, USA) and subjected to reverse transcription using SuperScript™ III Reverse Transcriptase (Thermo Fischer Scientific, Waltham, MA, USA). The resulting cDNA was used as a

© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

template for polymerase chain reaction (PCR) reactions with specific primers (Table S2). Amplified fragments were cloned into either pENTR D-TOPO or pENTR 3C.

CsRAV5 from *C. sphaerica*, *SsRAV66* from *S. sempervirens*, *PmRAV10* from *P. menziesii* and *GbYAB2B* from Ginkgo were directly synthesized by Twist Bioscience into the entry vector pTwist ENTR. The sequences for *C. sphaerica* (*GGZS01011568.1*) and *S. sempervirens* (*GIDF01065291.1*) were obtained from the assembly of transcriptomic data publicly available in the National Center for Biotechnology.

In the case of *GbSPT* (*Gb_09709*), *GbYABA* (*Gb_36880*), and *GbYAB1B* (*Gb_22423*), some discrepancies were found between the available sequences and the ones we manually cloned. The sequences of all genes used in this study are available in Data S1.

LR reactions were performed using Gateway™ LR Clonase™ II Enzyme mix to introduce the genes into the pAlligator2 vector (Bensmihen et al., 2004), which allows expression of the inserted sequence via a double cauliflower mosaic virus promoter 35S. In addition, a triple HA tag is translationally fused to the N-terminus of the produced protein. The resulting vectors based on pAlligator2 containing the different genes of interest were introduced into *Agrobacterium tumefaciens* GV3101, which were then used to transform Arabidopsis plants by floral dip (Clough & Bent, 1998). Transformed plants were selected by observing GFP fluorescent seeds among the progeny of the plants subjected to the floral dip procedure.

Bimolecular fluorescence complementation (BiFC)

Open reading frames of full-length *AtHEC3*, *AtSPT*, *AtIND*, *AtNGA1*, *AtNGA3*, *GbSPT*, *GbHEC*, and *GbRAV5* CDS were cloned into vectors pYFPN43 and pYFPC43 (https://www.ibmcp.upv.es/FerrandoLabVectors), and BiFC was performed as previously described (Belda-Palazón et al., 2012) with minor modifications: Overnight grown cultures of *A. tumefaciens C58*/pMP90 of about 2.0 OD_{600} units were collected and resuspended in similar volume of infiltration buffer ($MgCl_2$ 10 mM, MES 10 mM pH 5.6, acetosyringone 200 μM) and incubated in a rocking platform at room temperature for 3 h at 50 rpm. To suppress gene silencing, *A. tumefaciens* cells expressing the p19 protein of the tomato bushy stunt virus were used in the co-infiltration procedure. A mixture of *Agrobacterium* strains containing the fluorescent translational fusion constructs pYFPN43, pYFPC43 and the p19 plasmid (1:1:0.5) was prepared for co-infiltration into the abaxial face of *Nicotiana benthamiana* leaves with a needleless syringe. Epidermal cell layers of at least three transformed leaves were assayed for fluorescence under confocal microscope 3 days after infiltration. The experiments were repeated at least three times for every combination.

Sample preparation and *in situ* hybridization of *GbRAV5* and *GbHEC* in Ginkgo ovules

Fresh samples were fixed, dehydrated, and embedded as previously described (D'Apice et al., 2022; Douglas et al., 2007). Sections of 8 μm were cut on the microtome Leica RM 2125 RTS (Leica Biosystems, Milan, Italy).

RNA probes for *in situ* hybridizations were designed on the target Ginkgo coding sequences (*GbRAV5*: *Gb_38819*; *GbHEC*: *Gb_27869*) and *in vitro* synthesized from PCR-derived DNA templates. DNA templates were amplified by Wonder Taq Polymerase (EuroClone, Milan, Italy), using primers listed in Table S. RNA Digoxigenin (DIG)-labeled antisense and sense probes were synthesized from purified PCR amplicons by using T7 RNA polymerases (Roche, Rotkreuz, Switzerland) according to the manufacturer's protocol. The reaction mix contained the DIG RNA labeling mix (Roche, Rotkreuz, Switzerland) and the RNase inhibitor RNaseOUT (Invitrogen, Waltham, MA, USA). The probe hybridizations were performed at 55°C overnight in 50% formamide-humidified boxes (Ambrose et al., 2000). The antibody Anti-Digoxigenin-AP Fab fragments (Roche, Rotkreuz, Switzerland) was diluted 1:700 and incubated for 2 h at room temperature, and the detection with the two NBT/BCIP chromogens (Roche, Rotkreuz, Switzerland) was performed overnight. After staining was stopped, the slides were dehydrated, dried, permanently mounted using Eukitt (Bio-Optica, Milan, Italy), and photographed with a Leica DM500 optical microscope (Leica Biosystems, Milan, Italy).

Phylogenetic analyses of *RAV*

The amino acid sequence of *GbRAV5* was used as a query to construct the data set of possible *GbRAV5* homologs through BLAST (Altschul et al., 1990). Hits with an *E* value $<10^{-30}$ were initially selected, as we found by preliminary observation that this was an appropriate cutoff to include known members of the *RAV* family. The resulting hits were then manually curated using the following criteria: (i) hits were eliminated when the reciprocal first hit in Arabidopsis did not belong to the RAV family, and (ii) hits with only an AP2 and no B3 domain were eliminated. We assumed that they did not correspond to complete genes, since there were few and often perfectly matched partial sequences from other genes in the same species. No members of the RAV family have been described without a B3 domain. Furthermore, the main objective of this was to identify gymnosperm sequences that were similar or related to *GbRAV5*.

For the sake of simplicity, when several transcripts were described for the same gene, only the primary transcript was used for further analyses.

The following databases were used to identify proteins of interest in the selected species: *C. richardii* (https://phytozome-next.jgi.doe.gov/info/Crichardii_v2_1) (Marchant et al., 2022), *G. biloba* (http://gigadb.org/dataset/100613) (Guan et al., 2016), *C. panzhihuaensis* (https://db.cngb.org/codeplot/datasets/PwRftGHfPs5qG3gE) (Liu et al., 2022), *S. sempervirens* (https://treegenesdb.org/org/Sequoia-sempervirens) (Neale et al., 2022), *S. giganteum* (https://treegenesdb.org/org/Sequoiadendron-giganteum), *C. japonica* (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_030272615.1/) (Fujino et al., 2023), *T. plicata* (https://phytozome-next.jgi.doe.gov/info/Tplicata_v3_1) (Shalev et al., 2022), *T. grandis* (https://figshare.com/articles/dataset/Genome_assembly_and_annotation_of_Torreya_grandis/21089869) (Lou et al., 2023), *P. menziesii* (https://treegenesdb.org/org/Pseudotsuga-menziesii) (Neale et al., 2017), *L. kaempferi* (http://www.larixgd.cn/#/download) (Sun et al., 2022), *P. abies* (https://treegenesdb.org/org/Picea-abies) (Nystedt et al., 2013), *P. taeda* (https://treegenesdb.org/org/Pinus_taeda) (Zimin et al., 2017), *A. trichopoda* (https://phytozome-next.jgi.doe.gov/info/Atrichopodavar_SantaCruz_75HAP1_v2_1) (DePamphilis, 2013), *A. fimbriata* (https://ngdc.cncb.ac.cn/gwh/Assembly/21819/show) (Qin et al., 2021), *E. californica* (https://eschscholzia.kazusa.or.jp/) (Hori et al., 2018), *A. thaliana* (https://phytozome-next.jgi.doe.gov/info/Athaliana_TAIR10) (Lamesch et al., 2012).

The phylogenetic tree for *RAV* genes related to *GbRAV5* was constructed using NGPhylogeny.fr (Lemoine et al., 2019) with default one-click options and the PhyML+SMS maximum likelihood method. The resulting tree was visualized using iTOL (Letunic & Bork, 2024).

© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Protein structure prediction

The structure of proteins was modeled using the ColabFold implementation of AlphaFold2 (Jumper et al., 2021; Mirdita et al., 2022; Skolnick et al., 2021) with standard parameters and visualized using ChimeraX (Goddard et al., 2018).

ACKNOWLEDGMENTS

We thank Lars Østergaard and Charles P. Scutt for providing seed material, Michael Purugganan for critical reading of the manuscript, and Lucía Campos-Domínguez for advice in bioinformatic analysis and support. This project was financially supported by grant PID2021-127095NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by 'ERDF/EU'; grant FJCI-2017-32273 funded by MICIU/AEI/10.13039/501100011033 (to IC); 2021 SGR 00792 from the CERCA Programme/Generalitat de Catalunya; RED2022-134917-T MICIN/AEI; and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 101007738. We acknowledge financial support from the Grant CEX2019-000902-S funded by MICIU/AEI/10.13039/501100011033.

AUTHOR CONTRIBUTIONS

IC and SP designed the study. IC conducted most of the experimental work. SM and EO performed the *in situ* hybridization assays in Ginkgo ovules. BA helped with phylogenetic studies. IM-F performed the BiFC assays. FM generated the 35S::*GbRAV5* ΔAP2 lines. IC, BB, SN, BA, CF, and SP analyzed the data. IC and SP wrote the manuscript with contributions from all co-authors.

CONFLICT OF INTEREST

The authors have not declared a conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Data S1. List of genes cloned in this study.

Figure S1. Positive and negative controls used in the BiFC experiments.

Figure S2. Negative controls (sense probes) for *GbRAV5* and *GbHEC* in the post-pollination stage and lack of expression in other stages.

Figure S3. Phylogenetic tree of RAV genes in gymnosperms and angiosperms.

Table S1. List of genes of the phylogenetic tree shown in Figure S1.

Table S2. List of primers used in this study.

REFERENCES

Allen, M.D., **Yamasaki, K.**, **Ohme-Takagi, M.**, **Tateno, M.** & **Suzuki, M.** (1998) A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO Journal*, **17**, 5484–5496.

Altschul, S.F., **Gish, W.**, **Miller, W.**, **Myers, E.W.** & **Lipman, D.J.** (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Alvarez, J. & **Smyth, D.R.** (1999) CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS. *Development (Cambridge, England)*, **126**, 2377–2386.

Alvarez, J. & **Smyth, D.R.** (2002) CRABS CLAW and SPATULA genes regulate growth and pattern formation during gynoecium development in *Arabidopsis thaliana*. *International Journal of Plant Sciences*, **163**, 17–41.

Alvarez, J.P., **Goldshmidt, A.**, **Efroni, I.**, **Bowman, J.L.** & **Esheda, Y.** (2009) The NGATHA distal organ development genes are essential for style specification in Arabidopsis. *Plant Cell*, **21**, 1373.

Amack, S.C. & **Antunes, M.S.** (2020) CaMV35S promoter – a plant biology and biotechnology workhorse in the era of synthetic biology. *Current Plant Biology*, 24. https://doi.org/10.1016/j.cpb.2020.100179

Ambrose, B.A., **Lerner, D.R.**, **Ciceri, P.**, **Padilla, C.M.**, **Yanofsky, M.F.** & **Schmidt, R.J.** (2000) Molecular and genetic analyses of the silky1 gene reveal conservation in floral organ specification between eudicots and monocots. *Molecular Cell*, **5**, 569–579.

Ballester, P., **Martínez-Godoy, M.A.**, **Ezquerro, M.**, **Navarrete-Gómez, M.**, **Trigueros, M.**, **Rodríguez-Concepción, M.** ***et al.*** (2021) A transcriptional complex of NGATHA and bHLH transcription factors directs stigma development in Arabidopsis. *Plant Cell*, **33**, 3645–3657.

Bartholmes, C., **Hidalgo, O.** & **Gleissberg, S.** (2012) Evolution of the YABBY gene family with emphasis on the basal eudicot *Eschscholzia californica* (Papaveraceae). *Plant Biology (Stuttgart, Germany)*, **14**, 11–23.

Becker, A. (2020) A molecular update on the origin of the carpel. *Current Opinion in Plant Biology*, **53**, 15–22.

Belda-Palazón, B., **Ruiz, L.**, **Martí, E.**, **Tárraga, S.**, **Tiburcio, A.F.**, **Culiáñez, F.** ***et al.*** (2012) Aminopropyltransferases involved in polyamine biosynthesis localize preferentially in the nucleus of plant cells. *PLoS One*, **7**, e46907.

Bensmihen, S., **To, A.**, **Lambert, G.**, **Kroj, T.**, **Giraudat, J.** & **Parcy, F.** (2004) Analysis of an activated ABI5 allele using a new selection method for transgenic Arabidopsis seeds. *FEBS Letters*, **561**, 127–131.

Bond, W.J. (1989) The tortoise and the hare: ecology of angiosperm dominance and gymnosperm persistence. *Biological Journal of the Linnean Society*, **36**, 227–249.

Bornberg-Bauer, E., **Huylmans, A.K.** & **Sikosek, T.** (2010) How do new proteins arise? *Current Opinion in Structural Biology*, **20**, 390–396.

Bowe, L.M., **Coat, G.** & **DePamphilis, C.W.** (2000) Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 4092–4097.

Bowman, J.L. (2000) The YABBY gene family and abaxial cell fate. *Current Opinion in Plant Biology*, **3**, 17–22.

Bowman, J.L., **Eshed, Y.** & **Baum, S.F.** (2002) Establishment of polarity in angiosperm lateral organs. *Trends in Genetics*, **18**, 134–141.

Bowman, J.L. & **Smyth, D.R.** (1999) CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development*, **126**, 2387–2396.

Chaw, S.M., **Parkinson, C.L.**, **Cheng, Y.**, **Vincent, T.M.** & **Palmer, J.D.** (2000) Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 4086–4091.

Cheng, J. (2007) DOMAC: an accurate, hybrid protein domain prediction server. *Nucleic Acids Research*, **35**, W354–W356.

Cheng, J., **Sweredoski, M.J.** & **Baldi, P.** (2006) DOMpro: protein domain prediction using profiles, secondary structure, relative solvent accessibility, and recursive neural networks. *Data Mining and Knowledge Discovery*, **13**, 1–10.

Christenhusz, M.J.M. & **Byng, J.W.** (2016) The number of known plants species in the world and its annual increase. *Phytotaxa*, **261**, 201–217.

Clough, S.J. & **Bent, A.F.** (1998) Floral dip: a simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, **16**, 735–743.

© 2025 The Author(s).

13653313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Coiro, M. & **Seyfullah, L.J.** (2024) Disparity of cycad leaves dispels the living fossil metaphor. *Communications Biology*, **7**(1), 1–11.

Condamine, F.L., **Nagalingum, N.S.**, **Marshall, C.R.** & **Morlon, H.** (2015) Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC Evolutionary Biology*, **15**, 1–18.

Crawford, B.C.W. & **Yanofsky, M.F.** (2011) HALF FILLED promotes reproductive tract development and fertilization efficiency in *Arabidopsis thaliana*. *Development*, **138**, 2999–3009.

Cui, Y., **Zhao, J.**, **Gao, Y.**, **Zhao, R.**, **Zhang, J.** & **Kong, L.** (2021) Efficient multi-sites genome editing and plant regeneration via somatic embryogenesis in *Picea glauca*. *Frontiers in Plant Science*, **12**, 751891.

D'Apice, G., **Moschin, S.**, **Araniti, F.**, **Nigris, S.**, **Di Marzo, M.**, **Muto, A.** ***et al.*** (2021) The role of pollination in controlling *Ginkgo biloba* ovule development. *New Phytologist*, **232**, 2353–2368.

D'Apice, G., **Moschin, S.**, **Nigris, S.**, **Ciarle, R.**, **Muto, A.**, **Bruno, L.** ***et al.*** (2022) Identification of key regulatory genes involved in the sporophyte and gametophyte development in *Ginkgo biloba* ovules revealed by *in situ* expression analyses. *American Journal of Botany*, **109**, 887–898.

De La Torre, A.R., **Li, Z.**, **Van De Peer, Y.** & **Ingvarsson, P.K.** (2017) Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. *Molecular Biology and Evolution*, **34**, 1363–1377.

DePamphilis, C.W. (2013) The Amborella genome and the evolution of flowering plants. *Science*, **342**, 1241089.

Dimitrov, D., **Xu, X.**, **Su, X.**, **Shrestha, N.**, **Liu, Y.**, **Kennedy, J.D.** ***et al.*** (2023) Diversification of flowering plants in space and time. *Nature Communications*, **14**(1), 1–16.

Douglas, A.W., **Stevenson, D.W.** & **Little, D.P.** (2007) Ovule development in *Ginkgo biloba* L., with emphasis on the collar and nucellus. *International Journal of Plant Sciences*, **168**, 1207–1236. Available from: https://doi.org/10.1086/521693

Eldredge, N. & **Gould, S.J.** (1972) Punctuated equilibria: an alternative to phyletic gradualism. *Models in Paleobiology*. Available from: https://philpapers.org/rec/ELDOPE

Eshed, Y., **Baum, S.F.** & **Bowman, J.L.** (1999) Distinct mechanisms promote polarity establishment in carpels of Arabidopsis. *Cell*, **99**, 199–209.

Ferrándiz, C., **Fourquin, C.**, **Prunet, N.**, **Scutt, C.P.**, **Sundberg, E.**, **Trehin, C.** ***et al.*** (2010) Carpel Development. *Advances in Botanical Research*, **55**, 1–73.

Finet, C., **Floyd, S.K.**, **Conway, S.J.**, **Zhong, B.**, **Scutt, C.P.** & **Bowman, J.L.** (2016) Evolution of the YABBY gene family in seed plants. *Evolution & Development*, **18**, 116–126.

Fitch, W.M. & **Beintema, J.J.** (1990) Correcting parsimonious trees for unseen nucleotide substitutions: the effect of dense branching as exemplified by ribonuclease. *Molecular Biology and Evolution*, **7**, 438–443.

Flores-Sandoval, E., **Romani, F.** & **Bowman, J.L.** (2018) Co-expression and transcriptome analysis of *Marchantia polymorpha* transcription factors supports class c arfs as independent actors of an ancient auxin regulatory module. *Frontiers in Plant Science*, **9**, 408242.

Fourquin, C. & **Ferrándiz, C.** (2014) The essential role of NGATHA genes in style and stigma specification is widely conserved across eudicots. *New Phytologist*, **202**, 1001–1013.

Fourquin, C., **Vinauger-Douard, M.**, **Fogliani, B.**, **Dumas, C.** & **Scutt, C.P.** (2005) Evidence that CRABS CLAW and TOUSLED have conserved their roles in carpel development since the ancestor of the extant angiosperms. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 4649–4654.

Fujino, T., **Yamaguchi, K.**, **Yokoyama, T.T.**, **Hamanaka, T.**, **Harazono, Y.**, **Kamada, H.** ***et al.*** (2023) A chromosome-level genome assembly of a model conifer plant, the Japanese cedar, *Cryptomeria japonica* D. Don. *bioRxiv*, 2023.02.24.529822. https://doi.org/10.1101/2023.02.24.529822

GigaDB Dataset. *Updated genome assembly of Ginkgo biloba*. Available from: http://gigadb.org/dataset/100613; https://doi.org/10.5524/100613

Girin, T., **Paicu, T.**, **Stephenson, P.**, **Fuentes, S.**, **Körner, E.**, **O'Brien, M.** ***et al.*** (2011) INDEHISCENT and SPATULA interact to specify carpel and valve margin tissue and thus promote seed dispersal in Arabidopsis. *Plant Cell*, **23**, 3641–3653.

Goddard, T.D., **Huang, C.C.**, **Meng, E.C.**, **Pettersen, E.F.**, **Couch, G.S.**, **Morris, J.H.** ***et al.*** (2018) UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Science*, **27**, 14–25.

Gong, P., **Song, C.**, **Liu, H.**, **Li, P.**, **Zhang, M.**, **Zhang, J.** ***et al.*** (2021) *Physalis floridana* CRABS CLAW mediates neofunctionalization of GLOBOSA genes in carpel development. *Journal of Experimental Botany*, **72**, 6882–6903.

Goudey, B., **Geard, N.**, **Verspoor, K.** & **Zobel, J.** (2022) Propagation, detection and correction of errors using the sequence database network. *Briefings in Bioinformatics*, **23**, bbac416.

Gremski, K., **Ditta, G.** & **Yanofsky, M.F.** (2007) The HECATE genes regulate female reproductive tract development in *Arabidopsis thaliana*. *Development (Cambridge, England)*, **134**, 3593–3601.

Groszmann, M., **Paicu, T.** & **Smyth, D.R.** (2008) Functional domains of SPATULA, a bHLH transcription factor involved in carpel and fruit development in Arabidopsis. *The Plant Journal*, **55**, 40–52.

Guan, R., **Zhao, Y.**, **Zhang, H.**, **Fan, G.**, **Liu, X.**, **Zhou, W.** ***et al.*** (2016) Draft genome of the living fossil *Ginkgo biloba*. *GigaScience*, **5**, 49.

Guan, R., **Zhao, Y.**, **Zhang, H.**, **Fan, G.**, **Liu, X.**, **Zhou, W.** ***et al.*** (2019) Updated genome assembly of Ginkgo biloba [Data set]. GigaScience Database. https://doi.org/10.5524/100613

Hidvégi, N., **Gulyás, A.**, **Teixeira da Silva, J.A.**, **Posta, K.** & **Kiss, E.** (2020) Complementation of wild strawberry (*Fragaria vesca* L.) SPATULA (FvSPT) and SPIRAL (FvSPR) genes in *Arabidopsis thaliana*. *COLUMELLA – Journal of Agricultural and Environmental Sciences*, **7**, 23–34.

Hori, K., **Yamada, Y.**, **Purwanto, R.**, **Minakuchi, Y.**, **Toyoda, A.**, **Hirakawa, H.** ***et al.*** (2018) Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in California poppy using a draft genome sequence. *Plant & Cell Physiology*, **59**, 222–233.

Hou, H., **Lin, Y.** & **Hou, X.** (2020) Ectopic expression of a Pak-choi YABBY gene, BcYAB3, causes leaf curvature and flowering stage delay in *Arabidopsis thaliana*. *Genes (Basel)*, **11**, 370.

Ichihashi, Y., **Horiguchi, G.**, **Gleissberg, S.** & **Tsukaya, H.** (2010) The bHLH transcription factor SPATULA controls final leaf size in *Arabidopsis thaliana*. *Plant & Cell Physiology*, **51**, 252–261.

Jaenicke, R. (1987) Folding and association of proteins. *Progress in Biophysics and Molecular Biology*, **49**, 117–237.

Janzen, T., **Bokma, F.** & **Etienne, R.S.** (2022) Nucleotide substitutions during speciation may explain substitution rate variation. *Systematic Biology*, **71**, 1244–1254.

Jumper, J., **Evans, R.**, **Pritzel, A.**, **Green, T.**, **Figurnov, M.**, **Ronneberger, O.** ***et al.*** (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**(7873), 583–589.

Khersonsky, O., **Roodveldt, C.** & **Tawfik, D.S.** (2006) Enzyme promiscuity: evolutionary and mechanistic aspects. *Current Opinion in Chemical Biology*, **10**, 498–508.

Koludarov, I., **Senoner, T.**, **Jackson, T.N.W.**, **Dashevsky, D.**, **Heinzinger, M.**, **Aird, S.D.** ***et al.*** (2023) Domain loss enabled evolution of novel functions in the snake three-finger toxin gene superfamily. *Nature Communications*, **14**, 4861.

Kramer, E.M. (2015) A stranger in a strange land: the utility and interpretation of heterologous expression. *Frontiers in Plant Science*, **6**, 734.

Kress, A., **Poch, O.**, **Lecompte, O.** & **Thompson, J.D.** (2023) Real or fake? Measuring the impact of protein annotation errors on estimates of domain gain and loss events. *Frontiers in Bioinformatics*, **3**, 1178926.

Lamesch, P., **Berardini, T.Z.**, **Li, D.**, **Swarbreck, D.**, **Wilks, C.**, **Sasidharan, R.** ***et al.*** (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, **40**, D1202–D1210.

Lee, J.Y., **Baum, S.F.**, **Alvarez, J.**, **Patel, A.**, **Chitwood, D.H.** & **Bowman, J.L.** (2005) Activation of CRABS CLAW in the nectaries and carpels of Arabidopsis. *Plant Cell*, **17**, 25–36.

Lemoine, F., **Correia, D.**, **Lefort, V.**, **Doppelt-Azeroual, O.**, **Mareuil, F.**, **Cohen-Boulakia, S.** ***et al.*** (2019) NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Research*, **47**, W260–W265.

Letunic, I. & **Bork, P.** (2024) Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, **52**, W78–W82.

Levine, M.T., **Jones, C.D.**, **Kern, A.D.**, **Lindfors, H.A.** & **Begun, D.J.** (2006) Novel genes derived from noncoding DNA in Drosophila melanogaster are frequently X-linked and exhibit testis-biased expression. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 9935–9939.

© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Levy, A. (2019) How evolution builds genes from scratch. *Nature*, **574**, 314–316.

Liu, H., **Li, J.**, **Gong, P.** & **He, C.** (2023) The origin and evolution of carpels and fruits from an evo-devo perspective. *Journal of Integrative Plant Biology*, **65**, 283–298.

Liu, Y., **Wang, S.**, **Li, L.**, **Yang, T.**, **Dong, S.**, **Wei, T.** ***et al.*** (2022) The Cycas genome and the early evolution of seed plants. *Nature Plants*, **8**(4), 389–401.

Long, M., **Betrán, E.**, **Thornton, K.** & **Wang, W.** (2003) The origin of new genes: glimpses from the young and old. *Nature Reviews. Genetics*, **4**, 865–875.

Long, M., **Vankuren, N.W.**, **Chen, S.** & **Vibranovski, M.D.** (2013) New gene evolution: little did we know. *Annual Review of Genetics*, **47**, 307–333.

Lou, H., **Song, L.**, **Li, X.**, **Zi, H.**, **Chen, W.**, **Gao, Y.** ***et al.*** (2023) The *Torreya grandis* genome illuminates the origin and evolution of gymnosperm-specific sciadonic acid biosynthesis. *Nature Communications*, **14**(1), 1–15.

Magallón, S. & **Castillo, A.** (2009) Angiosperm diversification through time. *American Journal of Botany*, **96**, 349–365.

Magnani, E., **Sjölander, K.** & **Hake, S.** (2004) From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. *Plant Cell*, **16**, 2265–2277.

Major, R.T. (1967) The ginkgo, the most ancient living tree. The resistance of *Ginkgo biloba* L. to pests accounts in part for the longevity of this species. *Science*, **157**, 1270–1273.

Marchant, D.B., **Chen, G.**, **Cai, S.**, **Chen, F.**, **Schafran, P.**, **Jenkins, J.** ***et al.*** (2022) Dynamic genome evolution in a model fern. *Nature Plants*, **8**(9), 1038–1051.

McLysaght, A. & **Guerzoni, D.** (2015) New genes from non-coding sequence: the role of de novo protein-coding genes in eukaryotic evolutionary innovation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **370**, 20140332.

Meister, R.J., **Oldenhof, H.**, **Bowman, J.L.** & **Gasser, C.S.** (2005) Multiple protein regions contribute to differential activities of YABBY proteins in reproductive development. *Plant Physiology*, **137**, 651–662.

Melzer, R., **Wang, Y.Q.** & **Theißen, G.** (2010) The naked and the dead: the ABCs of gymnosperm reproduction and the origin of the angiosperm flower. *Seminars in Cell & Developmental Biology*, **21**, 118–128.

Mirdita, M., **Schütze, K.**, **Moriwaki, Y.**, **Heo, L.**, **Ovchinnikov, S.** & **Steinegger, M.** (2022) ColabFold: making protein folding accessible to all. *Nature Methods*, **19**, 679–682.

Mutte, S.K., **Kato, H.**, **Rothfels, C.**, **Melkonian, M.**, **Wong, G.K.S.** & **Weijers, D.** (2018) Origin and evolution of the nuclear auxin response system. *eLife*, **7**, e33399.

Nagalingum, N.S., **Marshall, C.R.**, **Quental, T.B.**, **Rai, H.S.**, **Little, D.P.** & **Mathews, S.** (2011) Recent synchronous radiation of a living fossil. *Science*, **334**, 796–799.

Nanasato, Y., **Mikami, M.**, **Futamura, N.**, **Endo, M.**, **Nishiguchi, M.**, **Ohmiya, Y.** ***et al.*** (2021) CRISPR/Cas9-mediated targeted mutagenesis in Japanese cedar (*Cryptomeria japonica* D. Don). *Scientific Reports*, **11**(1), 1–12.

Neale, D.B., **McGuire, P.E.**, **Wheeler, N.C.**, **Stevens, K.A.**, **Crepeau, M.W.**, **Cardeno, C.** ***et al.*** (2017) The Douglas-fir genome sequence reveals specialization of the photosynthetic apparatus in Pinaceae. *G3: Genes, Genomes, Genetics*, **7**, 3157–3167.

Neale, D.B., **Zimin, A.V.**, **Zaman, S.**, **Scott, A.D.**, **Shrestha, B.**, **Workman, R.E.** ***et al.*** (2022) Assembled and annotated 26.5 Gbp coast redwood genome: a resource for estimating evolutionary adaptive potential and investigating hexaploid origin. *G3 (Bethesda, Md.)*, **12**, jkab380.

Nystedt, B., **Street, N.R.**, **Wetterbom, A.**, **Zuccolo, A.**, **Lin, Y.C.**, **Scofield, D.G.** ***et al.*** (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, **497**, 579–584.

Osnato, M., **Castillejo, C.**, **Matías-Hernández, L.** & **Pelaz, S.** (2012) TEMPRANILLO genes link photoperiod and gibberellin pathways to control flowering in Arabidopsis. *Nature Communications*, **3**(1), 1–8.

Pabón-Mora, N., **Wong, G.K.S.** & **Ambrose, B.A.** (2014) Evolution of fruit development genes in flowering plants. *Frontiers in Plant Science*, **5**, 300.

Pagel, M., **Venditti, C.** & **Meade, A.** (2006) Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science*, **314**, 119–121.

Pfannebecker, K.C., **Lange, M.**, **Rupp, O.**, **Becker, A.** & **Purugganan, M.** (2017) Seed plant-specific gene lineages involved in carpel development. *Molecular Biology and Evolution*, **34**, 925–942.

Poovaiah, C., **Phillips, L.**, **Geddes, B.**, **Reeves, C.**, **Sorieul, M.** & **Thorlby, G.** (2021) Genome editing with CRISPR/Cas9 in *Pinus radiata* (D. Don). *BMC Plant Biology*, **21**, 363.

Qin, L., **Hu, Y.**, **Wang, J.**, **Wang, X.**, **Zhao, R.**, **Shan, H.** ***et al.*** (2021) Insights into angiosperm evolution, floral development and chemical biosynthesis from the *Aristolochia fimbriata* genome. *Nature Plants*, **7**(9), 1239–1253.

Rembeza, E. & **Engqvist, M.K.M.** (2021) Experimental and computational investigation of enzyme functional annotations uncovers misannotation in the EC 1.1.3.15 enzyme class. *PLoS Computational Biology*, **17**, e1009446.

Reymond, M.C., **Brunoud, G.**, **Chauvet, A.**, **Martínez-Garcia, J.F.**, **Martin-Magniette, M.L.**, **Monéger, F.** ***et al.*** (2012) A light-regulated genetic module was recruited to carpel development in Arabidopsis following a structural change to SPATULA. *Plant Cell*, **24**, 2812–2825.

Romanel, E.A.C., **Schrago, C.G.**, **Couñago, R.M.**, **Russo, C.A.M.** & **Alves-Ferreira, M.** (2009) Evolution of the B3 DNA binding superfamily: new insights into REM family gene diversification. *PLoS One*, **4**, e5791.

Romanova, M.A., **Maksimova, A.I.**, **Pawlowski, K.** & **Voitsekhovskaja, O.V.** (2021) YABBY genes in the development and evolution of land plants. *International Journal of Molecular Sciences*, **22**, 4139.

Royer, D.L., **Hickey, L.J.**, **Wing, S.L.** & **Hickey, L.** (2003) Ecological conservatism in the 'living fossil' ginkgo. *Paleobiology*, **29**, 84–104.

Schnoes, A.M., **Brown, S.D.**, **Dodevski, I.** & **Babbitt, P.C.** (2009) Annotation error in public databases: misannotation of molecular function in enzyme superfamilies. *PLoS Computational Biology*, **5**, e1000605.

Schuster, C., **Gaillochet, C.** & **Lohmann, J.U.** (2015) Arabidopsis HECATE genes function in phytohormone control during gynoecium development. *Development (Cambridge)*, **142**, 3343–3350.

Shalev, T.J., **Gamal el-Dien, O.**, **Yuen, M.M.S.**, **Shengqiang, S.**, **Jackman, S.D.**, **Warren, R.L.** ***et al.*** (2022) The western redcedar genome reveals low genetic diversity in a self-compatible conifer. *Genome Research*, **32**, 1952–1964.

Shao, J., **Liu, X.**, **Wang, R.**, **Zhang, G.** & **Yu, F.** (2012) The over-expression of an Arabidopsis B3 transcription factor, ABS2/NGAL1, leads to the loss of flower petals. *PLoS One*, **7**, e49861.

Siegfried, K.R., **Eshed, Y.**, **Baum, S.F.**, **Otsuga, D.**, **Drews, G.N.** & **Bowman, J.L.** (1999) Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. *Development*, **126**, 4117–4128.

Sigrist, C.J.A., **Sigrist, C.J.**, **de Castro, E.**, **Cerutti, L.**, **Cuche, B.A.**, **Hulo, N.** ***et al.*** (2013) New and continuing developments at PROSITE. *Nucleic Acids Research*, **41**, D344–D347.

Skolnick, J., **Gao, M.**, **Zhou, H.** & **Singh, S.** (2021) AlphaFold 2: why it works and its implications for understanding the relationships of protein sequence, structure, and function. *Journal of Chemical Information and Modeling*, **61**, 4827–4831.

Steinegger, M. & **Salzberg, S.L.** (2020) Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. *Genome Biology*, **21**, 115.

Stolyarova, A.V., **Bazykin, G.A.**, **Neretina, T.V.** & **Kondrashov, A.S.** (2019) Bursts of amino acid replacements in protein evolution. *Royal Society Open Science*, **6**, 181095.

Sun, C., **Xie, Y.H.**, **Li, Z.**, **Liu, Y.J.**, **Sun, X.M.**, **Li, J.J.** ***et al.*** (2022) The *Larix kaempferi* genome reveals new insights into wood properties. *Journal of Integrative Plant Biology*, **64**, 1364–1373.

Swaminathan, K., **Peterson, K.** & **Jack, T.** (2008) The plant B3 superfamily. *Trends in Plant Science*, **13**, 647–655.

Theißen, G. & **Becker, A.** (2004) Gymnosperm orthologues of class B floral homeotic genes and their impact on understanding flower origin. *Critical Reviews in Plant Sciences*, **23**, 129–148.

Trigueros, M., **Navarrete-Gómez, M.**, **Sato, S.**, **Christensen, S.K.**, **Pelaz, S.**, **Weigel, D.** ***et al.*** (2009) The NGATHA genes direct style development in the Arabidopsis gynoecium. *Plant Cell*, **21**, 1394–1409.

Wan, T., **Gong, Y.**, **Liu, Z.**, **Zhou, Y.**, **Dai, C.** & **Wang, Q.** (2022) Evolution of complex genome architecture in gymnosperms. *GigaScience*, **11**, 1–10.

Wang, X. (2023) Origin of angiosperms: problems, challenges, and solutions. *Life*, **13**, 2029.

© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Webster, A.J., **Payne, R.J.H.** & **Pagel, M.** (2003) Molecular phylogenies link rates of evolution and speciation. *Science*, **301**, 478.

Wu, D.D., **Irwin, D.M.** & **Zhang, Y.P.** (2011) De novo origin of human protein-coding genes. *PLoS Genetics*, **7**, e1002379.

Yamada, T., **Yokota, S.**, **Hirayama, Y.**, **Imaichi, R.**, **Kato, M.** & **Gasser, C.S.** (2011) Ancestral expression patterns and evolutionary diversification of YABBY genes in angiosperms. *The Plant Journal*, **67**, 26–36.

Yang, Y., **Ferguson, D.K.**, **Liu, B.**, **Mao, K.S.**, **Gao, L.M.**, **Zhang, S.Z.** ***et al.*** (2022) Recent advances on phylogenomics of gymnosperms and a new classification. *Plant Diversity*, **44**, 340–350.

Zhang, P., **Tan, H.T.W.**, **Pwee, K.H.** & **Kumar, P.P.** (2004) Conservation of class C function of floral organ development during 300 million years of evolution from gymnosperms to angiosperms. *The Plant Journal*, **37**, 566–577.

Zhang, X.L., **Yang, Z.P.**, **Zhang, J.** & **Zhang, L.G.** (2013) Ectopic expression of BraYAB1-702, a member of YABBY gene family in Chinese cabbage, causes leaf curling, inhibition of development of shoot apical meristem and flowering stage delaying in *Arabidopsis thaliana*. *International Journal of Molecular Sciences*, **14**, 14872–14891.

Zhao, W., **Su, H.Y.**, **Song, J.**, **Zhao, X.Y.** & **Zhang, X.S.** (2006) Ectopic expression of TaYAB1, a member of YABBY gene family in wheat, causes the partial abaxialization of the adaxial epidermises of leaves and arrests the development of shoot apical meristem in Arabidopsis. *Plant Science*, **170**, 364–371.

Zhou, Z. & **Zheng, S.** (2003) The missing link in ginkgo evolution. *Nature*, **423**(6942), 821–822.

Zhu, P., **He, T.**, **Zheng, Y.** & **Chen, L.** (2023) The need for masked genomes in gymnosperms. *Frontiers in Plant Science*, **14**, 1309744.

Zimin, A.V., **Stevens, K.A.**, **Crepeau, M.W.**, **Puiu, D.**, **Wegrzyn, J.L.**, **Yorke, J.A.** ***et al.*** (2017) An improved assembly of the loblolly pine mega-genome using long-read single-molecule sequencing. *GigaScience*, **6**, 1.

Zmasek, C.M. & **Godzik, A.** (2011) Strong functional patterns in the evolution of eukaryotic genomes revealed by the reconstruction of ancestral protein domain repertoires. *Genome Biology*, **12**, R4.

© 2025 The Author(s).

1365313x, 2025, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tpj.70488 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [28/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License