



HAL
open science

DELLA family duplication events lead to different selective constraints in angiosperms

Jean Keller, Pauline Delcros, Cyril Libourel, Francisco Cabello-Hurtado,
Abdelkader Aïnouche

► **To cite this version:**

Jean Keller, Pauline Delcros, Cyril Libourel, Francisco Cabello-Hurtado, Abdelkader Aïnouche. DELLA family duplication events lead to different selective constraints in angiosperms. *Genetica*, 2020, 148 (5-6), pp.243-251. 10.1007/s10709-020-00102-6 . hal-02929503

HAL Id: hal-02929503

<https://hal.science/hal-02929503>

Submitted on 9 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **DELLA family duplication events lead to different selective constraints in angiosperms.**

2 Authors: Keller J.^{1,2}, Delcros P.¹, Libourel C.², Cabello-Hurtado F.¹, Ainouche A.¹

3 Affiliation : ¹UMR CNRS 6553 Ecobio, OSUR (Observatoire des Sciences de l'Univers de
4 Rennes), Université de Rennes 1, 35042 Rennes, France ; ²LRSV, Université de Toulouse,
5 CNRS, UPS, Castanet-Tolosan, France

6 Corresponding author: abdelkader.ainouche@univ-rennes1.fr

7

8 **Acknowledgement**

9 We acknowledge the GenOuest bioinformatics core facility (<https://www.genouest.org>) for providing the
10 computing infrastructure. Part of this work was conducted at the LRSV laboratory, which belongs to the
11 TULIP Laboratoire d'Excellence (LABEX) (ANR-10-LABX-41). JK and CL were supported by the research
12 project Engineering Nitrogen Symbiosis for Africa (ENSA), which is funded through a grant to the University
13 of Cambridge by the Bill & Melinda Gates Foundation (OPP1172165). JK was supported by a doctoral
14 research grant from the University of Rennes 1 – French Ministry of Higher Education and Research. This
15 work benefited from the International Associated Laboratory “Ecological Genomics of Polyploidy”
16 supported by CNRS (INEE, UMR CNRS 6553 Ecobio), University of Rennes 1, Iowa State University
17 (Ames, USA).

18

19

20

21 **Abstract**

22 Gibberellic acid (GA) is a major plant hormone involved in several biological processes from the flowering
23 to the symbiosis with microorganisms. Thus, the GA regulation is crucial for plant biology. This regulation
24 occurs *via* the DELLA proteins that belong to the GRAS transcription factor family. DELLA proteins are
25 characterised by a DELLA N-terminal and a GRAS C-terminal domains. It is well known that DELLA
26 activity appears after the bryophytes divergence and then evolved in the vascular plant lineages. Here we
27 present the phylogeny of DELLA across 75 species belonging to various lineages from algae, liverworts and
28 angiosperms. Our study confirmed two main duplication events, the first occurring before the angiosperms
29 divergence and the other specific to the eudicots lineage. Comparative analysis of DELLA subclades in
30 angiosperms revealed the loss in Poaceae and strong alteration in other species of the DELLA functional
31 domain in the DELLA2 clade. In addition, molecular evolution analysis suggests that each of the clades
32 (named DELLA1.1, DELLA1.2 and DELLA2) evolved differently but copies of each subclade are under
33 strong purifying selection. This also suggests that, although the DELLA functional domain is altered in
34 DELLA2, DELLA2 orthologs are still functional and operate in a different way compared to DELLA1 copies.
35 In angiosperms, additional duplication events occurred and led to duplicate copies in species, genus or family
36 such as in the Fabaceae subfamily Papilionoideae. This duplication led to the formation of additional paralogs
37 in the DELLA1.2 subclade (DELLA1.2.1 and DELLA1.2.2). Interestingly, both copies appeared to be under
38 relaxing selection revealing different evolutionary fate of the DELLA duplicated copies.

39

40 **Keywords:** DELLA, evolutionary history, gene duplication, Fabaceae, molecular evolution

41

42

43 **Introduction**

44 Plant hormones play essential roles in various plant biology aspects and a huge diversity of these compounds
45 have been identified so far (Davies 2010). Among them, the gibberellic acid (GA), a tetracyclic di-terpenoid
46 molecule, is particularly involved in plant growth, seed germination and development or flowering (Gupta
47 and Chakrabarty 2013). It also plays a role in responses to environmental stimuli (water, light, temperature)
48 as well as during the symbiosis with microorganisms such as nitrogen-fixing bacteria and arbuscular
49 mycorrhizal fungi (Lievens et al. 2005; Davies 2010; Ferguson et al. 2011; Gupta and Chakrabarty 2013;
50 Takeda et al. 2015; Martín-Rodríguez et al. 2016)

51 Regulation of GA activity occurs through the DELLA (named according the conserved DELLA amino acid
52 motif at the protein N-terminal domain) proteins which form a complex with the GA receptor GID1
53 repressing the expression of downstream GA-responsive genes (Ueguchi-Tanaka et al. 2007; Sun 2011).
54 Fixation of GA to the GID1 receptor induces the proteasome-dependant degradation of DELLA proteins (Fu
55 et al. 2002). This GA-GID1-DELLA complex has been observed in all the investigated vascular plants but is
56 absent from bryophytes although a DELLA copy has been identified in *Physcomitrella patens* (Yasumura et
57 al. 2007; Sun 2011). This suggests that the DELLA repressor activity of GA evolved after the bryophyte
58 divergence (~430 MYA) and there is no functional data about the DELLA role and mode of action prior this
59 divergence (Kenrick and Crane 1997; Yasumura et al. 2007; Wang and Deng 2014). In angiosperms, DELLA
60 proteins are known to be involved in the regulation of a wide range of fundamental biological processes such
61 as flowering, root and shoot development, and is a key regulator of both arbuscular mycorrhizal and root
62 nodule nitrogen-fixing symbiosis (Maekawa et al. 2009; Ferguson et al. 2011; Floss et al. 2013; Yu et al.
63 2014; Pimprikar et al. 2016). For example, during nitrogen-fixing symbiosis, DELLA is thought to be a key
64 regulator of the infection process by directly regulating symbiotic genes *NSP1/NSP2*, *ERN1*, *NF-YA1* and by
65 enhancing the binding between CYCLOPS and NSP2 (Fonouni-Farde et al. 2016; Jin et al. 2016). Recently,
66 DELLA proteins were also shown to be likely involved in the cytokinin genetic pathway regulation during
67 nodulation in *Pisum sativum* L. (Dolgikh et al. 2019).

68 *DELLA* constitutes a subfamily of the GRAS transcription factors superfamily and DELLA proteins are
69 characterised by a N- and C- terminal conserved DELLA and GRAS domains respectively (Pysh et al. 1999).
70 The GRAS family encompasses between 8 and 13 subfamilies according to authors and surveyed species
71 (Cenci and Rouard 2017). Although evolutionary history of the GRAS family has been recently investigated
72 in representative angiosperms model species (Cenci and Rouard 2017), there is little data about the DELLAs
73 evolution in angiosperms.

74 In this context, we performed a large screening of DELLA proteins in more than 50 species representing all
75 the main plant orders. This allowed us to confirm different duplication events in the DELLA subfamily, with
76 a subclade exhibiting severe alterations of the DELLA functional domain. In addition, molecular evolution
77 analysis revealed that the angiosperms DELLA subclades have experienced different selective constraints
78 with currently a strong purifying constraint within each of these subclades.

79 **Material & Methods**

80 Data mining

81 Proteomes of 75 plant species covering the main Angiosperms lineages (list available in Supplementary Table
82 S1) were screened against the Pfam database (Finn et al. 2016) and hits with an e-value lower than 0.01 and
83 both GRAS (PF12041.8) and DELLA (PF03514.14) domains were retained. To complement this search and
84 take into account potential missing domains (in cases of losses or miss-annotated proteins), published and
85 well characterized DELLA protein from *Arabidopsis thaliana*, *Oryza sativa*, *Lupinus angustifolius*,
86 *Medicago truncatula* and *Amborella trichopoda* (Sato et al. 2014; Cenci and Rouard 2017; Hane et al. 2017)
87 were used as queries for a tBLASTn v2.7.1+ search against the 75 plant species with default parameters
88 (Camacho et al. 2009) and an e-value threshold of 1^{-60} . Finally, protein sequences shorter than 200 amino
89 acids were removed from the dataset to eliminate partial assembly and putative remnants (retained and
90 removed sequences are available in Supplementary Data S1 and S2 respectively).

91 In order to explore the absence of DELLA proteins in algae, the Algae-PrAS database containing 31 genomes
92 (<http://alga-pras.riken.jp/>) was also screened with the reference DELLA sequences from *A. thaliana* and *O.*
93 *sativa* as queries using the BLASTp v2.7.1+ algorithm and an e-value threshold of 0.01. Until recently, scarce
94 data were available for Zygnematophyceae algae, the likely closest lineage of land plants. Two
95 Zygnematophyceae genomes (*Mesotaenium endlicheranium* and *Spirogloea muscicola*, this latter belonging
96 to a new subclass likely to be the closest from the Zygnematophyceae/embryophytes separation) have been
97 published and were screened for DELLA homologs (Cheng et al. 2019).

98 Phylogenetic analysis

99 Retained protein sequences were aligned using MUSCLE (Edgar 2004) with default parameters. Alignment
100 was subjected to maximum likelihood analysis using IQ-TREE v1.5.5 (Nguyen et al. 2015). The best-fitted
101 evolutionary model was determined using ModelFinder as implemented in IQ-TREE (Kalyaanamoorthy et
102 al. 2017). Maximum likelihood analysis was conducted with 10,000 replicates Ultra Fast Bootstraps (Hoang
103 et al. 2018). Trees were visualized and annotated using the Interactive Tree of Life platform (Letunic and Bork
104 2016).

105 Molecular evolution analyses

106 For each clade of interest identified through the phylogenetic analysis, sequence logos were generated from
107 protein alignments using the WebLogo v3.5 platform with default parameters (Crooks 2004). In addition, to
108 compare signature of selection among DELLA homolog clades identified in phylogeny, we estimated dN/dS
109 ratios from all pairwise comparison of protein-coding sequences using *yn00* module from the PAML package
110 v4.9g (Yang and Nielsen 2000; Yang 2007). We also used the RELAX hypothesis testing framework from
111 HyPhy (Pond et al. 2005; Wertheim et al. 2015) to look for relaxation ($K < 1$) or intensification ($K > 1$) of the
112 signature of selection. To estimate signature of selection, codons were aligned according to the protein
113 alignment using the pal2nal v14 Perl script (Suyama et al. 2006) and all gapped positions were removed. We

114 used the following command line for each clade tested (SupplementaryTable2): *hyphy relax CPU=1 --*
115 *alignment input_file.fna --code Universal --type Nucleotide --keyword value Foreground.*

116

117 **Results and discussion**

118 *DELLA family emerged with land plants and experienced multiple duplication events*

119 To understand the origin of the DELLA family, the Algae-PrAS database that contains more than 30 genomes
120 of non-land plants species (green and red algae, Glaucophyceae) and other organisms such as Oomycetes and
121 diatoms was mined using DELLA protein sequences from *Arabidopsis thaliana* and *Oryza sativa*. Results
122 confirmed the absence of DELLA orthologs in these species. In the two Zygnematophyceae genomes, no
123 DELLA were detected but other GRAS family members were identified and are likely resulting from
124 horizontal gene transfer from bacteria (Cheng et al. 2019). These findings are consistent with results from
125 Yasamura et al. (2007) who suggested an emergence of DELLA proteins along with the land colonization by
126 plants. To improve our understanding of the evolutionary dynamics of DELLA family, 75 land plants
127 genomes were investigated and 293 DELLA orthologs were identified by the phylogenetic analysis
128 (excluding putative remnants and misassembled sequences). Among them, one sequence with a DELLA
129 functional domain was identified in the liverwort *Marchantia polymorpha* (Marpol_Mapoly0058s0044.1)
130 consistently with the hypothesis that the DELLA family originated at least in the common ancestor of the
131 vascular and non-vascular plants (Hernández-García et al. 2019). Previous findings revealed that the GID1-
132 DELLA complex does not exist in the bryophytes lineage (Yasamura et al. 2007; Hirano et al. 2007). In
133 bryophytes, DELLA has been recently proposed as transcriptional co-activator that could have been exapted
134 in vascular plants (Hernández-García et al. 2019). Previously published results based on RNAseq data and
135 putative DELLA's interactors in *Physcomitrella patens* (no DELLA-GID1 interaction), *Solanum*
136 *lycopersicum* and *Arabidopsis thaliana* suggest that DELLA evolution is correlated with an increasing
137 coordination of transcriptional programs between bryophytes and vascular plants (Briones-Moreno et al.
138 2017).

139 Phylogenetic analysis shows that DELLA family underwent multiple duplication events at different points
140 of the land plant evolution as recently shown by Hernández-García and colleagues (2019). Indeed, results
141 revealed a first duplication event that occurred before the divergence of angiosperms (Fig.1, with branch
142 support values to be retrieved in the Supplementary Data S3) leading to the formation of two main clades,
143 called DELLA2 (blue clade in Fig.1; corresponding to DELLA3 in Hernández-García et al. 2019) and
144 DELLA1 (orange, green and red clades) accordingly to Cenci and Rouard (2017). DELLA sequences from
145 *Marchantia polymorpha* and the lycophyte *Selaginella moellendorffii* seem to belong to the DELLA2 clade;
146 however, these branches are not supported (UFB value < 50%) and these sequences are likely at the base of
147 the whole DELLA family. As demonstrated in recent phylogeny and in this study the DELLA1/DELLA2
148 clades emergence is probably due to the Whole Genome Duplication (WGD) event that occurred prior the
149 seed plants (c.a 319 MYA) (Jiao et al. 2011; Hernández-García et al. 2019). Another level of duplication has

150 been identified in the DELLA1 clade. Indeed, this latter, with the basal Angiosperms *Amborella trichopoda*
151 at its base, can be divided into three main subclades (Fig.1). The first, in green, is exclusively composed by
152 Monocot species whereas the two others only contain Eudicot species. These two Eudicot subclades emerged
153 from the core Eudicot WGD event that occurred approximately 125 MYA (Van de Peer et al. 2017) and were
154 arbitrarily named DELLA1.1 (orange clade in Fig.1) and DELLA1.2 (red clade). Finally, additional
155 duplication events limited to species or subclades were identified such as in Brassicaceae DELLA1.1 or the
156 DELLA1.2 of Papilionoideae (WGD of the Papilionoideae common ancestor, around 60 MYA, see below).

157 For each sequence, gene structure was extrapolated from the annotation file, when available, and results were
158 mapped on the phylogeny (Fig. 1; outer circle). Excepting few sequences, most of the genes present a non-
159 intronic structure and a length ranked between 1000 and 2000 bp.

160 Analysis of the functional domains revealed alterations in DELLA2

161 DELLA proteins are characterised by the presence of two functional domains. The first (DELLA) is in the
162 N-terminal region of the protein and is specific to the DELLA family whereas the second (GRAS) is shared
163 with other members of the GRAS superfamily. Functional analysis of the DELLA domain in the model plant
164 *Arabidopsis thaliana* characterised two motifs required for the interaction with GID1 (Murase et al. 2008):
165 DELLA§LXYXVXXXXMAXVAXXLEXLEX§ and TVHYNPXXLXXWXXXM (where X and § stand for
166 any and non-polar amino acids respectively).

167 DELLA and GRAS functional domains identified using the Pfam search were mapped on the phylogenetic
168 tree (Fig.1, second outer circle) and revealed that most of the sequences possess both domains with few
169 exceptions that can be explained by assembly and annotation issues or putative pseudogenes. However, it
170 appears that all the six Poaceae from the DELLA2 clade have lost the DELLA functional domain (absence
171 of red right pointing pentagram in Fig.1, and Supplementary Figure 1). The loss of the DELLA domain has
172 been previously reported in rice (Itoh et al. 2005; Cenci and Rouard 2017) and our results indicate that this
173 loss is likely to have occurred before diversification of the Poaceae from their common ancestor. Indeed,
174 Monocots from other orders such as *Spirodela polyrhiza*, *Zostera marina* (two Alismatales) and *Musa*
175 *accuminata* (Zingiberales) retained both DELLA and GRAS domains (Supplementary Figure 1). At the
176 opposite, all Monocots DELLA1 proteins have a full DELLA domain. Previous functional analysis of the
177 two DELLA2 copies of rice (also called SLRL1 and SLRL2 for SLR-like1 and 2) revealed that SLRL1 is
178 positively affected by GA treatment whereas SLRL2 is not. In addition, trans-complementation of *slr1*
179 (DELLA1) mutant by SLRL1 rescued the phenotype (Itoh et al. 2005). Since the N-terminal DELLA domain
180 is missing in SLRL1 and 2, proteins were not degraded following the GA treatment but also exhibit a lower
181 activity against GA than a SLR1 truncated copy missing the DELLA domain (Itoh et al. 2005). Itoh and
182 colleagues suggest that SLR1 and SLRL1 can function together and are required for a fine-tuning of GA
183 signalling.

184 In a second approach, the focus has been made on the DELLA functional domain extracted from the protein
185 alignment and amino acids conservation represented using the WebLogo tool (Crooks 2004) for each main

186 clade of DELLA (Fig.2). Results revealed a high conservation of amino acids required for interaction with
187 GID1 across the different DELLA1 subclades (DELLA1.1, DELLA1.2 and DELLA1 from Monocots).
188 However, severe disruption of the DELLA motifs can be observed in the DELLA2 clade (Fig.2; red arrows).
189 Indeed, the methionine from the MAXVA motif and at the end of the second DELLA subdomains is not
190 conserved anymore among the DELLA2 sequences. Other disruptions of the DELLA domain are also
191 characteristic of the DELLA2 clade such as the low conservation of the first leucine in the DELLA first motif
192 or the alteration of the TV motif in the second DELLA motif (Fig.2). This low-conservation pattern observed
193 in the DELLA2 clade could be explained by different hypothesis such as the pseudogenisation of the
194 DELLA2 copies, a new function of these proteins, or another way for these proteins to interact with GID1.

195 To test whether the DELLA2 clade is under relaxed selective constraints, different evolutionary tests were
196 applied. To this aim, only sequences with entire DELLA domains were retained and evolutionary tests
197 performed on the domain alignment. The analysis of non-synonymous *versus* synonymous (dN/dS) revealed
198 a higher dN/dS ratio when comparing DELLA2 to the other DELLA1 subclades, whereas these latter had
199 similar dN/dS ratios (Fig.3). To determine if relaxed selection is acting on the DELLA2 clade, a branch model
200 (RELAX) was applied with the DELLA2 clade marked as foreground branch (Supplementary Table 2).
201 According to pairwise comparisons, estimating the selective intensity parameter (K) suggests that the
202 DELLA2 clade is subjected to an ongoing process of relaxing selection ($K=0.81$, $p\text{-val}\approx 0.00074$, $LR=11.39$).
203 Conversely, the higher value found for DELLA1 ($K=1.28$, $p\text{-val}\approx 0.00038$, $LR=12.61$) suggests
204 intensification of the purifying selection ($K>1$) to maintain the proper role of DELLA1. In addition, logo and
205 amino acid frequency for each position of the DELLA motifs (Fig.2 and Supplementary Table S3) show that
206 mutations of essential amino acid in DELLA2 clade are shared by most of the sequences in this clade. For
207 example, methionine from MAXVA stretch is replaced by leucine in 100% of sequences. Similarly, almost
208 90% of DELLA2 sequences share a glycine instead of a glutamate normally expected at the second position
209 of the DELLA first motif. Altogether, these analyses suggest that modification of DELLA2 sequences
210 occurred in the DELLA2 ancestor and then, have been maintained. These results are in line with a previous
211 report suggesting that positive selection acted on the DELLA family during its evolution (Chen et al. 2013).
212 To date, there is no detailed studies of DELLA2 proteins from Eudicot lineage. However, it is possible that,
213 similarly to Poaceae DELLA2, these proteins function along with their DELLA1 counterparts and act as
214 weaker GA signalling repressors without being degraded. Interestingly, it has been shown that the RGL1
215 protein of *Arabidopsis thaliana* (DELLA1.2 subclade) works without being degraded upon GA application
216 (Wen and Chang 2002). All together, these results suggest that the multiple rounds of duplication in the
217 DELLA family led to a slight subfunctionalization of DELLA copies, all involved in the GA regulation
218 pathway. In their work, Yasumura and colleagues (2007), proposed a stepwise evolution of the DELLA-GA
219 complex in plants with the GID1-DELLA interaction evolving in vascular plants. Here, we show that the
220 DELLA family pursued its evolution in Angiosperms by experiencing several rounds of duplication, leading
221 to sub-clades under different selective constraint. In addition, our results revealed that the loss of DELLA
222 domain is shared among the DELLA2 proteins of Poaceae, supporting the hypothesis of a different
223 mechanism (degradation-independent for example) for DELLA in this family (Itoh et al. 2005).

224 *DELLAs from the Fabaceae show similar evolutionary pattern compared to that of the DELLA family*

225 DELLA proteins play major role in various biological aspects of plant life. In legumes, they also have a
226 critical role for the establishment of the root nodule nitrogen-fixing symbiosis. Recently, DELLA proteins
227 have been proposed to be also involved in the regulation of cytokinin genetic pathway through interaction
228 with KNOX and BELL transcription factors (Dolgikh et al. 2019). Thus, a focused analysis on the DELLA
229 sequences from the Fabaceae family has been performed and showed similar evolutionary pattern compared
230 to results discussed above.

231 First, the evolutionary pattern of DELLAs confirmed that the three main DELLA clades (DELLA2,
232 DELLA1.1 and DELLA1.2) are conserved among the Fabaceae (Fig.4A). DELLA1 and DELLA2 copies
233 both present mutations in amino acids involved in the interaction with GID1 (Fig.4B). The branch model
234 analysis on both clades showed that DELLA1 and DELLA2 copies have a similar pattern of selection with
235 no sign of relaxation/intensification (Supplementary Table S2). However, no change in the selective
236 constraint was detected among the different clades suggesting that, also in Papilionoideae, there is no positive
237 selection acting on them.

238 Second, in the DELLA1.2 subclade, an additional duplication event occurred leading to the formation of two
239 subclades arbitrarily called DELLA1.2.1 and DELLA1.2.2. These two subclades likely originated from the
240 Papilionoideae WGD that occurred 60MYA. Analysis of the selective constraints strongly suggests that both
241 paralogs ($K=0.25$, $p\text{-val}=7.3E-08$, $LR=28.98$ and $K=0.26$, $p\text{-val}=7.2E-08$, $LR=29$ for DELLA1.2.1 and
242 DELLA1.2.2 respectively) are under pseudogenization process. This is also supported by the apparent loss
243 of DELLA1.2.1 copy in seven out of the 15 legume samples investigated.

244

245 **Conclusion**

246 DELLA proteins have a critical role in numerous aspects of plant life, from development to flowering but
247 also in specific aspects such as in the nitrogen-fixing symbiosis. Phylogenetic analysis of DELLA family
248 including more than 70 species confirmed its emergence in land plants along with its absence in algae. Results
249 of phylogenetic analysis of the DELLA family is consistent with previous reports and show the existence of
250 two main duplication events leading to the formation of three clades: DELLA1 and DELLA2 clades emerged
251 from a duplication event that occurred before the Angiosperms divergence and two additional DELLA1
252 subclades (DELLA1.1 and DELLA1.2) emerged following the Eudicots whole genome duplication. A
253 detailed analysis of the evolution of the three DELLA clades revealed that they are all under purifying
254 selection although their ancestors seem to have been subjected to a relaxed selective constraint. In the
255 DELLA2 clade, our results strongly support the loss of the DELLA domain before diversification of Poaceae
256 and severe alterations of this motif in other DELLA2 proteins suggesting that DELLA2 protein could function
257 in a degradation-independent manner. Finally, DELLA proteins from the Fabaceae family revealed a similar
258 evolutionary history with additional duplication due to the Whole Genome Duplication of the Papilionoideae

259 ancestor leading to additional DELLA1.2.1 and DELLA1.2.2 clades. Both clades are under relaxing selection
260 suggesting they are likely to be pseudogenized or neofunctionalized.

261 All together this study supports the hypothesis of a complex evolutionary history of DELLA proteins in
262 Angiosperms with multiple changes in the selective constraint. To extend these results, functional data on
263 the different DELLA copies from both Monocots and Eudicots would be of a great interest to better
264 understand the biological consequences of these differential selective constraints.

265 **Conflict of interest**

266 The authors declare that they have no conflict of interest.

267

268 **References**

269 Briones-Moreno A, Hernández-García J, Vargas-Chávez C, et al (2017) Evolutionary Analysis
270 of DELLA-Associated Transcriptional Networks. *Front Plant Sci* 8:626.
271 <https://doi.org/10.3389/fpls.2017.00626>

272 Camacho C, Coulouris G, Avagyan V, et al (2009) BLAST+: architecture and applications.
273 *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>

274 Cenci A, Rouard M (2017) Evolutionary Analyses of GRAS Transcription Factors in
275 Angiosperms. *Front Plant Sci* 8:. <https://doi.org/10.3389/fpls.2017.00273>

276 Chen J, Cheng T, Wang P, et al (2013) Genome-wide bioinformatics analysis of DELLA-family
277 proteins from plants. *Plant Omics Journal*, POJ 6(3):201–207

278 Cheng S, Xian W, Fu Y, et al (2019) Genomes of Subaerial Zygnematophyceae Provide
279 Insights into Land Plant Evolution. *Cell* 179:1057-1067.e14.
280 <https://doi.org/10.1016/j.cell.2019.10.019>

281 Crooks GE (2004) WebLogo: A Sequence Logo Generator. *Genome Res* 14:1188–1190.
282 <https://doi.org/10.1101/gr.849004>

283 Davies PJ (ed) (2010) *Plant Hormones*. Springer Netherlands, Dordrecht

284 Dolgikh AV, Kirienko AN, Tikhonovich IA, et al (2019) The DELLA Proteins Influence the
285 Expression of Cytokinin Biosynthesis and Response Genes During Nodulation. *Front*
286 *Plant Sci* 10: 432. <https://doi.org/10.3389/fpls.2019.00432>

287 Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high
288 throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>

289 Ferguson BJ, Foo E, Ross JJ, Reid JB (2011) Relationship between gibberellin, ethylene and
290 nodulation in *Pisum sativum*. *New Phytol* 189:829–842. <https://doi.org/10.1111/j.1469-8137.2010.03542.x>
291

- 292 Finn RD, Cogill P, Eberhardt RY, et al (2016) The Pfam protein families database: towards a
293 more sustainable future. *Nucleic Acids Res* 44:D279–D285.
294 <https://doi.org/10.1093/nar/gkv1344>
- 295 Floss DS, Levy JG, Levesque-Tremblay V, et al (2013) DELLA proteins regulate arbuscule
296 formation in arbuscular mycorrhizal symbiosis. *Proc Natl Acad Sci* 110:E5025–E5034.
297 <https://doi.org/10.1073/pnas.1308973110>
- 298 Fonouni-Farde C, Tan S, Baudin M, et al (2016) DELLA-mediated gibberellin signalling
299 regulates Nod factor signalling and rhizobial infection. *Nat Commun* 7:.
300 <https://doi.org/10.1038/ncomms12636>
- 301 Fu X, Richards DE, Ait-ali T, et al (2002) Gibberellin-Mediated Proteasome-Dependent
302 Degradation of the Barley DELLA Protein SLN1 Repressor. *Plant Cell* 14:3191–3200.
303 <https://doi.org/10.1105/tpc.006197>
- 304 Gupta R, Chakrabarty SK (2013) Gibberellic acid in plant: Still a mystery unresolved. *Plant*
305 *Signal Behav* 8:e25504. <https://doi.org/10.4161/psb.25504>
- 306 Hane JK, Ming Y, Kamphuis LG, et al (2017) A comprehensive draft genome sequence for
307 lupin (*Lupinus angustifolius*), an emerging health food: insights into plant-microbe
308 interactions and legume evolution. *Plant Biotechnol J* 15:318–330.
309 <https://doi.org/10.1111/pbi.12615>
- 310 Hernández-García J, Briones-Moreno A, Dumas R, Blázquez MA (2019) Origin of Gibberellin-
311 Dependent Transcriptional Regulation by Molecular Exploitation of a Transactivation
312 Domain in DELLA Proteins. *Mol Biol Evol* 36:908–918.
313 <https://doi.org/10.1093/molbev/msz009>
- 314 Hirano K, Nakajima M, Asano K, et al (2007) The GID1-Mediated Gibberellin Perception
315 Mechanism Is Conserved in the Lycophyte *Selaginella moellendorffii* but Not in the
316 Bryophyte *Physcomitrella patens*. *Plant Cell* 19:3058–3079.
317 <https://doi.org/10.1105/tpc.107.051524>
- 318 Hoang DT, Chernomor O, von Haeseler A, et al (2018) UFBoot2: Improving the Ultrafast
319 Bootstrap Approximation. *Mol Biol Evol* 35:518–522.
320 <https://doi.org/10.1093/molbev/msx281>
- 321 Itoh H, Shimada A, Ueguchi-Tanaka M, et al (2005) Overexpression of a GRAS protein lacking
322 the DELLA domain confers altered gibberellin responses in rice: Rice SLR1-like genes.
323 *Plant J* 44:669–679. <https://doi.org/10.1111/j.1365-313X.2005.02562.x>
- 324 Jiao Y, Wickett NJ, Saravanaraj A, et al (2011) Data from: Ancestral polyploidy in seed plants
325 and angiosperms. *Nature* 473 (7345): 97–100. <https://doi.org/10.1038/nature09916>.
- 326 Jin Y, Liu H, Luo D, et al (2016) DELLA proteins are common components of symbiotic
327 rhizobial and mycorrhizal signalling pathways. *Nat Commun* 7:(1).
328 <https://doi.org/10.1038/ncomms12433>
- 329

- 330 Kalyaanamoorthy S, Minh BQ, Wong TKF, et al (2017) ModelFinder: fast model selection for
331 accurate phylogenetic estimates. *Nat Methods* 14:587–589.
332 <https://doi.org/10.1038/nmeth.4285>
- 333 Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. *Nature* 389
334 (6646): 33–39. <https://doi.org/10.1038/37918>
- 335 Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and
336 annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:W242–W245.
337 <https://doi.org/10.1093/nar/gkw290>
- 338 Lievens S, Goormachtig S, Den Herder J, et al (2005) Gibberellins Are Involved in Nodulation
339 of *Sesbania rostrata*. *Plant Physiol* 139:1366–1379.
340 <https://doi.org/10.1104/pp.105.066944>
- 341 Maekawa T, Maekawa-Yoshikawa M, Takeda N, et al (2009) Gibberellin controls the
342 nodulation signaling pathway in *Lotus japonicus*. *Plant J* 58:183–194.
343 <https://doi.org/10.1111/j.1365-313X.2008.03774.x>
- 344 Martín-Rodríguez JA, Huertas R, Ho-Plágaro T, et al (2016) Gibberellin–Abscisic Acid
345 Balances during Arbuscular Mycorrhiza Formation in Tomato. *Front Plant Sci* 7:.
346 <https://doi.org/10.3389/fpls.2016.01273>
- 347 Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A Fast and Effective
348 Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol*
349 32:268–274. <https://doi.org/10.1093/molbev/msu300>
- 350 Pimprikar P, Carbonnel S, Paries M, et al (2016) A CCaMK-CYCLOPS-DELLA Complex
351 Activates Transcription of RAM1 to Regulate Arbuscule Branching. *Curr Biol* 26:987–
352 998. <https://doi.org/10.1016/j.cub.2016.01.069>
- 353 Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies.
354 *Bioinformatics* 21:676–679. <https://doi.org/10.1093/bioinformatics/bti079>
- 355 Pysh LD, Wysocka-Diller JW, Camilleri C, et al (1999) The GRAS gene family in Arabidopsis:
356 sequence characterization and basic expression analysis of the SCARECROW-LIKE
357 genes. *Plant J* 18:111–119. <https://doi.org/10.1046/j.1365-313X.1999.00431.x>
- 358 Sato T, Miyanoiri Y, Takeda M, et al (2014) Expression and purification of a GRAS domain of
359 SLR1, the rice DELLA protein. *Protein Expr Purif* 95:248–258.
360 <https://doi.org/10.1016/j.pep.2014.01.006>
- 361 Sun T (2011) The Molecular Mechanism and Evolution of the GA–GID1–DELLA Signaling
362 Module in Plants. *Curr Biol* 21:R338–R345. <https://doi.org/10.1016/j.cub.2011.02.036>
- 363 Suyama M, Torrents D, Bork P (2006) PAL2NAL: robust conversion of protein sequence
364 alignments into the corresponding codon alignments. *Nucleic Acids Res* 34:W609–
365 W612. <https://doi.org/10.1093/nar/gkl315>
- 366 Takeda N, Handa Y, Tsuzuki S, et al (2015) Gibberellin regulates infection and colonization of
367 host roots by arbuscular mycorrhizal fungi. *Plant Signal Behav* 10:e1028706.
368 <https://doi.org/10.1080/15592324.2015.1028706>

- 369 Ueguchi-Tanaka M, Nakajima M, Motoyuki A, Matsuoka M (2007) Gibberellin Receptor and
 370 Its Role in Gibberellin Signaling in Plants. *Annu Rev Plant Biol* 58:183–198.
 371 <https://doi.org/10.1146/annurev.arplant.58.032806.103830>
- 372 Van de Peer Y, Mizrachi E, Marchal K (2017) The evolutionary significance of polyploidy. *Nat*
 373 *Rev Genet* 18:411–424. <https://doi.org/10.1038/nrg.2017.26>
- 374 Wang Y, Deng D (2014) Molecular basis and evolutionary pattern of GA–GID1–DELLA
 375 regulatory module. *Mol Genet Genomics* 289:1–9. [https://doi.org/10.1007/s00438-013-](https://doi.org/10.1007/s00438-013-0797-x)
 376 [0797-x](https://doi.org/10.1007/s00438-013-0797-x)
- 377 Wen C-K, Chang C (2002) Arabidopsis *RGL1* Encodes a Negative Regulator of Gibberellin
 378 Responses. *Plant Cell* 14:87–100. <https://doi.org/10.1105/tpc.010325>
- 379 Wertheim JO, Murrell B, Smith MD, et al (2015) RELAX: Detecting Relaxed Selection in a
 380 Phylogenetic Framework. *Mol Biol Evol* 32:820–832.
 381 <https://doi.org/10.1093/molbev/msu400>
- 382 Yang Z (2007) PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol*
 383 24:1586–1591. <https://doi.org/10.1093/molbev/msm088>
- 384 Yang Z, Nielsen R (2000) Estimating Synonymous and Nonsynonymous Substitution Rates
 385 Under Realistic Evolutionary Models. *Mol Biol Evol* 17:32–43.
 386 <https://doi.org/10.1093/oxfordjournals.molbev.a026236>
- 387 Yasumura Y, Crumpton-Taylor M, Fuentes S, Harberd NP (2007) Step-by-Step Acquisition of
 388 the Gibberellin-DELLA Growth-Regulatory Mechanism during Land-Plant Evolution.
 389 *Curr Biol* 17:1225–1230. <https://doi.org/10.1016/j.cub.2007.06.037>
- 390 Yu N, Luo D, Zhang X, et al (2014) A DELLA protein complex controls the arbuscular
 391 mycorrhizal symbiosis in plants. *Cell Res* 24:130–133.
 392 <https://doi.org/10.1038/cr.2013.167>

393

394 **Figure legends**

395 **Fig. 1:** Maximum likelihood tree of DELLA proteins (model JTT+F+R10). Tree was rooted using three
 396 SCARECROW (SCR, belonging to another subfamily of GRAS transcription factor) sequences from
 397 liverworts *Marchantia polymorpha* and are collapsed in the “outgroup” clade. DELLA2, Monocots
 398 DELLA1, Eudicots DELLA1.1 and DELLA1.2 are indicated by coloured branches as follow: blue, green,
 399 orange, red, respectively. The first outer circle indicates the main Angiosperm lineages as described below
 400 the tree. The second outer circle represents the presence/absence and the position of DELLA (red right
 401 pointing pentagram) and GRAS (dark blue rectangle) domains. Size and position of symbols is proportional
 402 to their length and position in the protein. The third outer circle illustrates the gene structure, extrapolated
 403 from the annotation file when available. Blue squares represent gene exons, their size and position are
 404 proportional to their length and positions in the gene. Branch support was tested using the UFB2 method and

405 were masked from the tree for readability reason. Support values can be retrieved in the Supplementary Data
406 S3 corresponding to the rooted tree shown here in Newick format.

407 **Fig. 2:** Logos of functional DELLA domains generated for each phylogenetic clade. For DELLA2, strongly
408 affected sites are indicated by red arrows. Size of letter is proportional to the amino acid conservation at a
409 given site. Blank in the logo represents highly gapped sites in the alignment. Theoretical functional models
410 are written in black above logos and essential amino acids for the interaction with GID1, as defined by Murase
411 et al. (2008), are indicated in bold.

412 **Fig. 3:** Boxplot representing pairwise dN/dS comparisons among the DELLA clades. All pairwise
413 comparison between all DELLA sequences were calculated. For each clade (X-axis), comparison with other
414 clades and itself (white waves) are shown separately. The dN/dS ratios were estimated using the Yang &
415 Nielsen method (Yang and Nielsen 2000).

416 **Fig. 4:** Maximum likelihood tree of DELLA proteins (model JTT+ R5) in the Papilionoideae clade (A).
417 Alignment of DELLA functional domains mapped on the phylogenetic tree. Blue background is proportional
418 to identity level for each site. Functional DELLA subdomains are indicated by red rectangles (B).

419

420 **Supplementary Data**

421 Supplementary Data S1: Protein sequences used for phylogenetic analysis

422 Supplementary Data S2: Protein sequences discarded prior the phylogenetic analysis (less than 200 amino
423 acids)

424 Supplementary Data S3: Rooted tree presented in Figure 1 in Newick format.

425 Supplementary Table S1: Summary table of DELLA sequences repartition. For each species, sequence ID is
426 indicated for each DELLA clade. Absence of sequence for a given species/clade is indicated by empty cells.
427 In addition, order, family, reference and abbreviation used for each species are indicated in the last four
428 columns. For *Vitis vinifera*, the DELLA2 sequence is indicated but not retained in the analysis since it seems
429 that this sequence is missassembly.

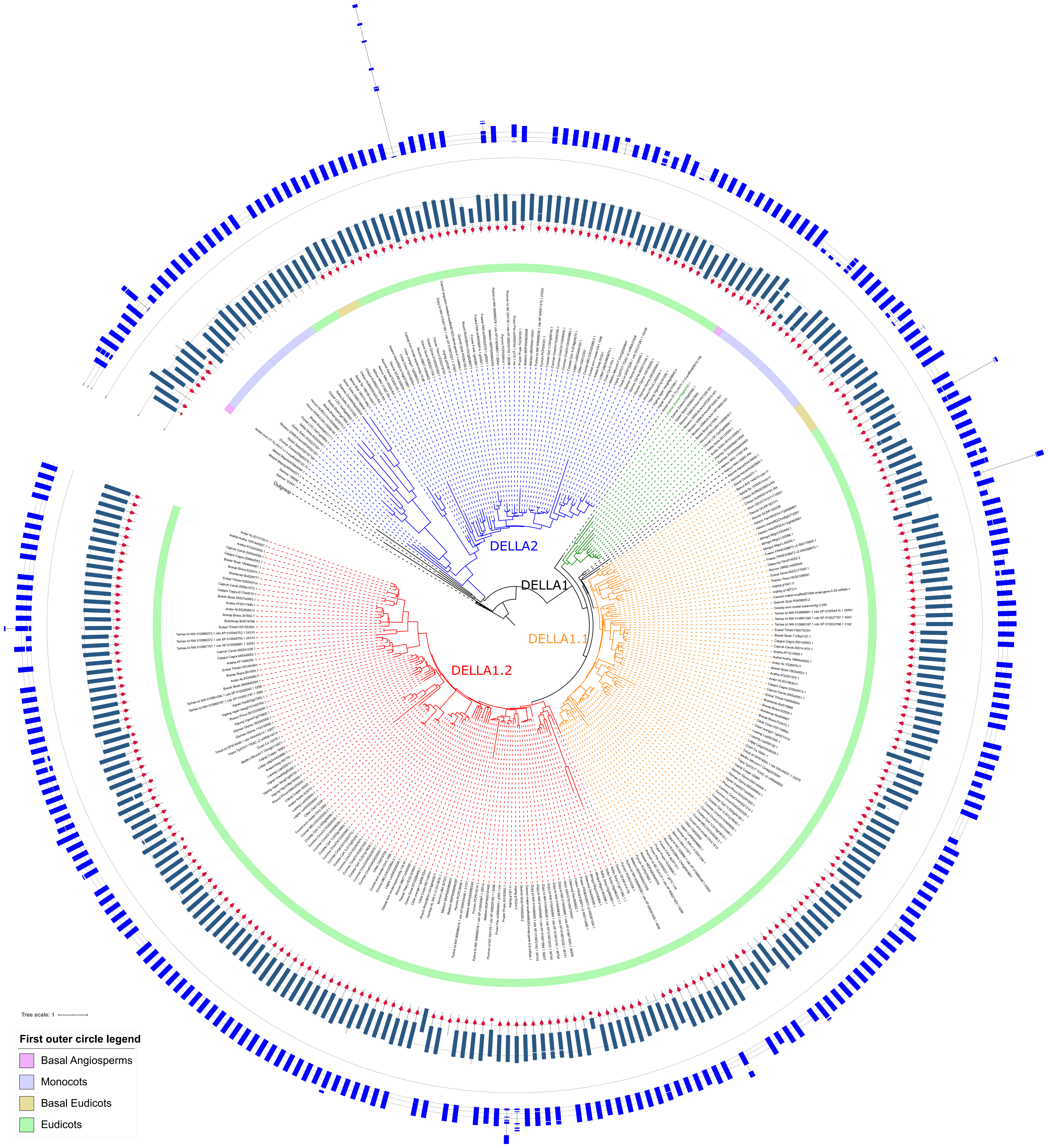
430 Supplementary Table S2: Results of the branch analysis for signature of selection acting on the different
431 DELLA clades.

432 Supplementary Table S3: Amino acid frequency matrices for each DELLA copy. Numbers represent
433 percentage of each amino acid at each position in the DELLA functional domain. First column indicates the
434 consensus of DELLA domain.

435

436 Supplementary Figure 1: Alignment of DELLA domains in DELLA2 Monocots and *Medicago truncatula* as
437 reference. Blue-coloured names represent sequences with a predicted DELLA domain by Pfam. The DELLA
438 functional domains are indicated by black squares with the theoretical sequence indicated below. *Musa*
439 *acuminata*, *Siprodela polyrhiza* and *Zostera marina* belong to diverse lineages of Monocots outside the
440 Poaceae clade.

441



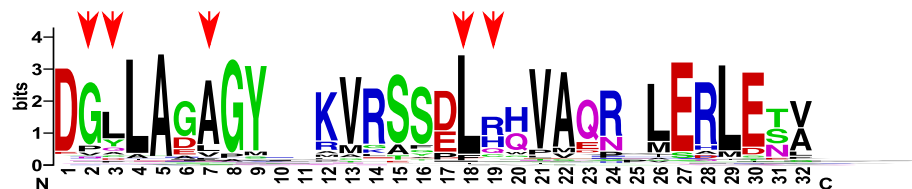
Tree scale: 1

First outer circle legend

- Basal Angiosperms
- Monocots
- Basal Eudicots
- Eudicots

DELLA \S LxYxVxxxxMAxVaxxLExLEx \S

DELLA2



DELLA1
Monocots



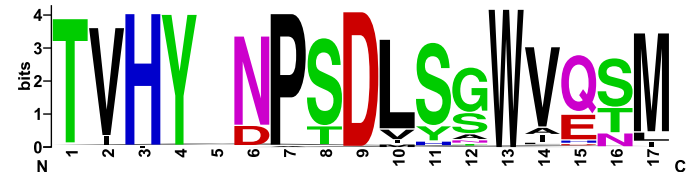
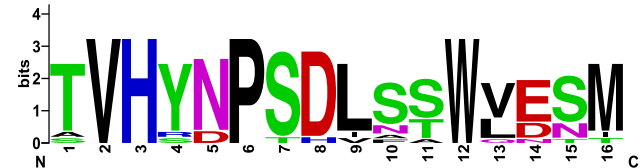
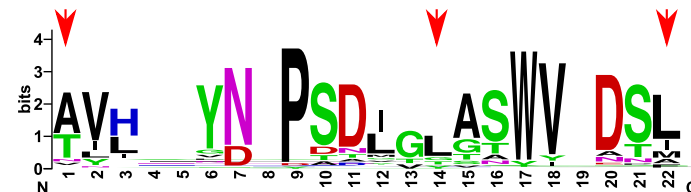
DELLA1.1
Eudicots

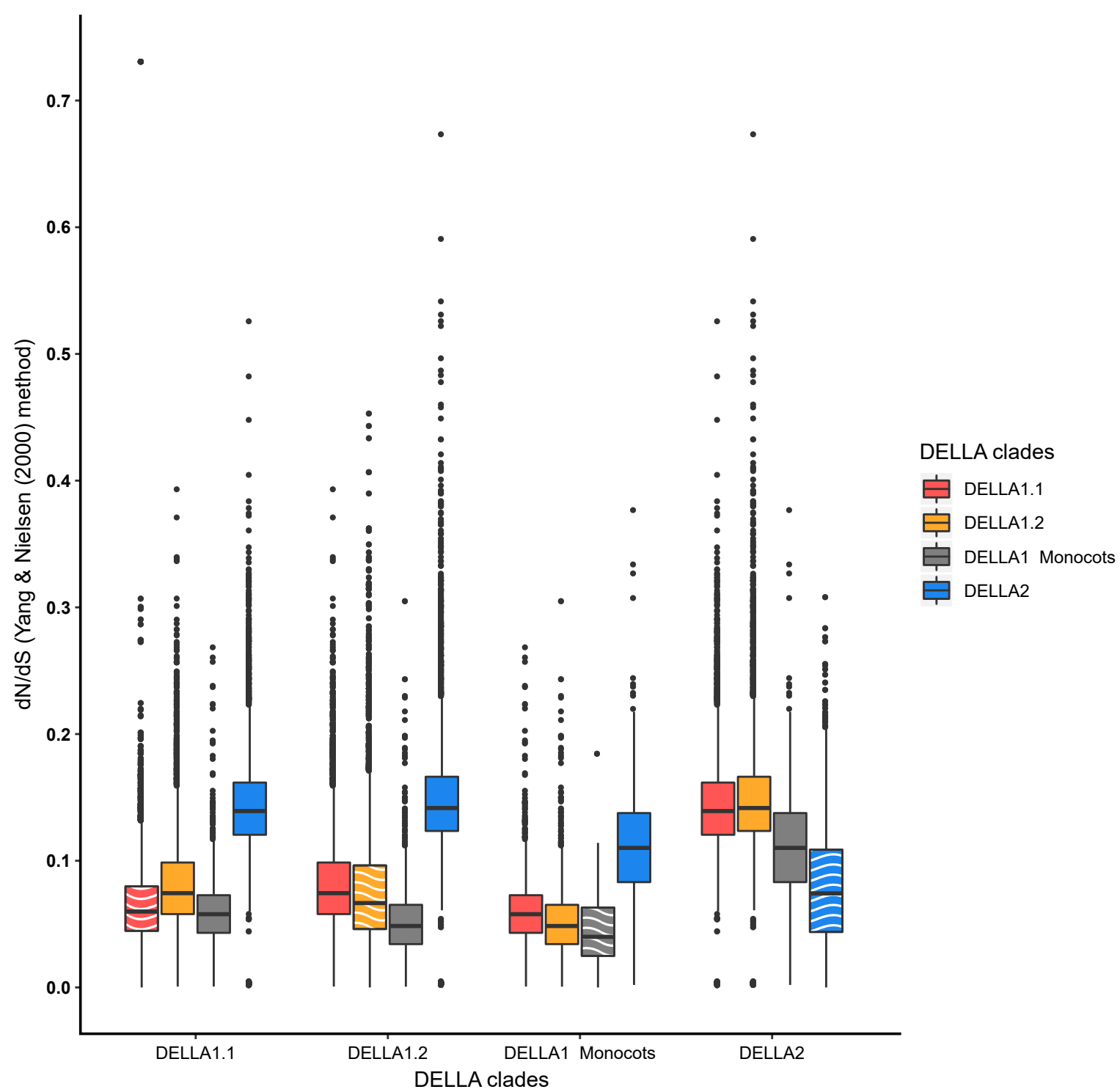


DELLA1.2
Eudicots

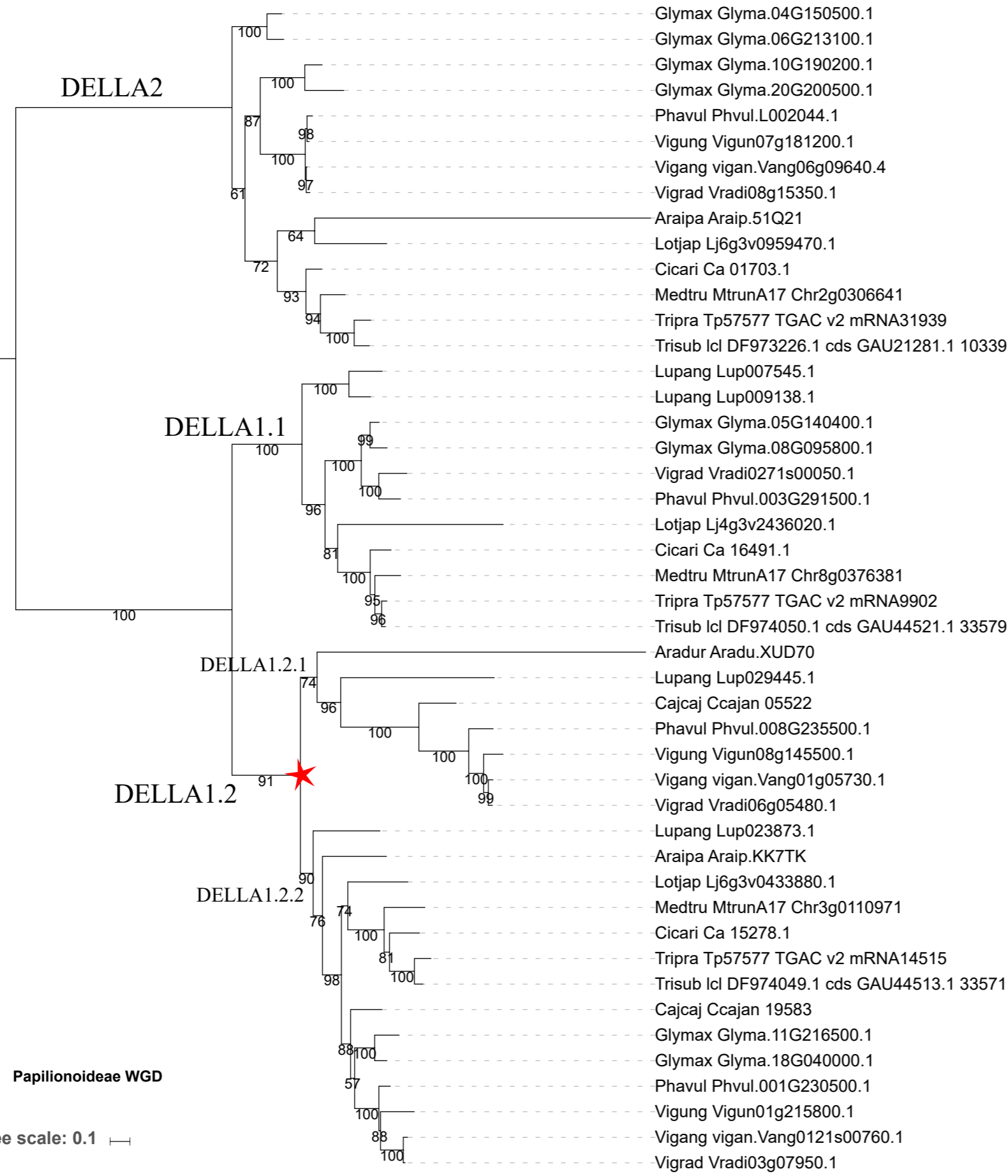


TVHYNPxxLxxWxxxM





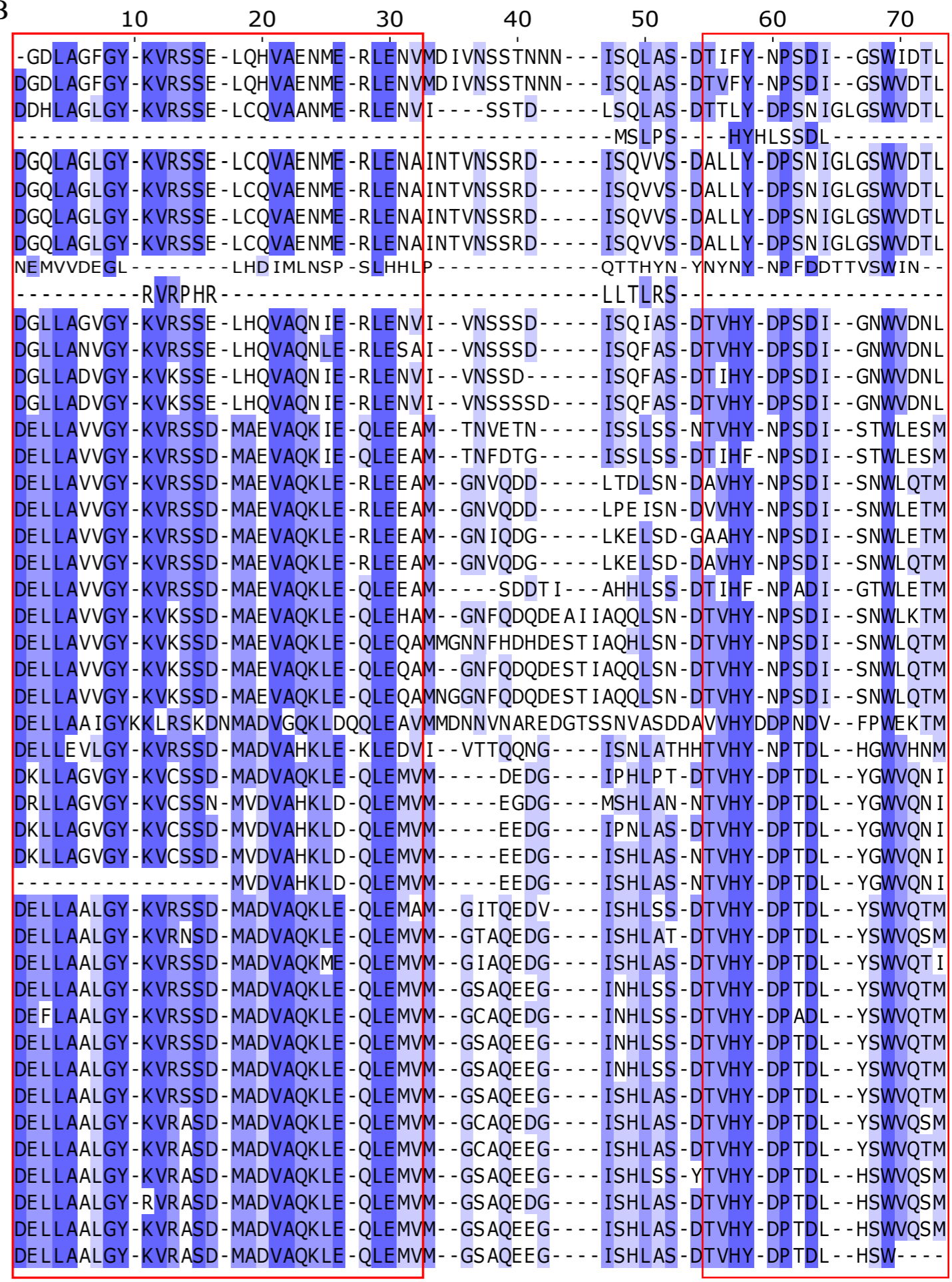
A



★ Papilionoideae WGD

Tree scale: 0.1

B



Consensus

DE L L A A L G Y K K V R S S D N M A D V A Q K L E Q Q L E M V M N T G N A Q E D G E S T I I S H L A S + D T V H Y + D P S D I G L Y S W V Q T M