



HAL
open science

Adaptive evolution of the enigmatic *Takakia* now facing climate change in Tibet

Ruoyang Hu, Xuedong Li, Yong Hu, Runjie Zhang, Qiang Lv, Min Zhang, Xianyong Sheng, Feng Zhao, Zhijia Chen, Yuhan Ding, et al.

► **To cite this version:**

Ruoyang Hu, Xuedong Li, Yong Hu, Runjie Zhang, Qiang Lv, et al.. Adaptive evolution of the enigmatic *Takakia* now facing climate change in Tibet. *Cell*, 2023, 186 (17), pp.3558-3576.e17. 10.1016/j.cell.2023.07.003 . hal-04188574

HAL Id: hal-04188574

<https://hal.science/hal-04188574v1>

Submitted on 14 Nov 2023

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 Research Article

2

3 **Adaptive evolution of the enigmatic *Takakia* now facing climate change**
4 **in Tibet**

5

6 **AUTHORS**

7 Ruoyang Hu^{#1,14†}, Xuedong Li^{#1}, Yong Hu¹, Runjie Zhang¹, Qiang Lv¹, Min Zhang¹, Xianyong Sheng¹,
8 Feng Zhao¹, Zhijia Chen¹, Yuhan Ding¹, Huan Yuan¹, Xiaofeng Wu¹, Shuang Xing¹, Xiaoyu Yan¹,
9 Fang Bao¹, Ping Wan¹, Lihong Xiao^{1,19†}, Xiaoqin Wang¹, Wei Xiao¹, Eva L. Decker², Nico van Gessel²,
10 Hugues Renault^{2,20†}, Gertrud Wiedemann^{2,21†}, Nelly A. Horst^{2,22†}, Fabian B. Haas³, Per K.I.
11 Wilhelmsson³, Kristian K. Ullrich³, Eva Neumann³, Bin Lv^{4,23†}, Chengzhi Liang⁵, Huilong Du^{5,24†},
12 Hongwei Lu^{5,25†}, Qiang Gao^{5,8†}, Zhukuan Cheng^{5,26†}, Hanli You^{5,26†}, Peiyong Xin⁵, Jinfang Chu⁵,
13 Chien-Hsun Huang⁶, Yang Liu^{7,27,28†}, Shanshan Dong²⁷, Liangsheng Zhang⁸, Fei Chen⁹, Lei Deng¹⁰,
14 Fuzhou Duan¹⁰, Wenji Zhao¹⁰, Kai Li¹¹, Zhongfeng Li¹¹, Xingru Li¹¹, Hengjian Cui¹², Yong Zhang¹³,
15 Chuan Ma¹⁴, Ruiliang Zhu¹⁵, Yu Jia¹⁶, Meizhi Wang¹⁶, Mitsuyasu Hasebe¹⁷, Jinzhong Fu⁴, Bernard
16 Goffinet⁷, Hong Ma¹⁸, Stefan A. Rensing³, Ralf Reski^{2,29*}, Yikun He^{1*}

17 # Contributed equally.

18 † Present addresses.

19 * Correspondence: ralf.reski@biologie.uni-freiburg.de (Ralf Reski), yhe@cnu.edu.cn (Yinkun He)

20

21 **AFFILIATIONS**

- 22 1. Beijing Key Laboratory of Plant Gene Resources and Biotechnology for Carbon Reduction and
23 Environmental Improvement, College of Life Sciences, Capital Normal University (CNU), Beijing,
24 Beijing 100048, China
- 25 2. Plant Biotechnology, Faculty of Biology, University of Freiburg, 79104 Freiburg, Germany
- 26 3. Department of Biology, University of Marburg, 35043 Marburg, Germany
- 27 4. Department of Integrative Biology, University of Guelph, Guelph, Ontario N1G 2W1, Canada
- 28 5. Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, Beijing
29 100101, China

- 30 6. State Key Laboratory of Genetic Engineering and Collaborative Innovation Center of Genetics
31 and Development, Ministry of Education Key Laboratory of Biodiversity and Ecological
32 Engineering, Institute of Plant Biology, Center of Evolutionary Biology, School of Life Sciences,
33 Fudan University, Shanghai, Shanghai 200433, China
- 34 7. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut
35 06269, USA
- 36 8. Genomics and Genetic Engineering Laboratory of Ornamental Plants, College of Agriculture and
37 Biotechnology, Zhejiang University, Hangzhou, Zhejiang 310058, China
- 38 9. Hainan Yazhou Bay Seed Laboratory, College of Tropical Crops, Sanya Nanfan Research
39 Institute, Hainan University, Sanya, Hainan 572025, China
- 40 10. College of Resource Environment and Tourism, CNU, Beijing, Beijing 100048, China
- 41 11. Department of Chemistry, CNU, Beijing, Beijing 100048 China
- 42 12. School of Mathematical Sciences, CNU, Beijing, Beijing 100048, China
- 43 13. Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy
44 of Sciences, Beijing, Beijing 100101, China
- 45 14. Institute of Apicultural Research, Chinese Academy of Agricultural Sciences, Beijing, Beijing
46 100193, China
- 47 15. Department of Biology, School of Life Sciences, East China Normal University, Shanghai,
48 Shanghai 200241, China
- 49 16. State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese
50 Academy of Sciences, Beijing, Beijing 100093, China
- 51 17. Division of Evolutionary Biology, National Institute for Basic Biology, Okazaki, Aichi 444-8585,
52 Japan
- 53 18. Department of Biology, Huck Institutes of the Life Sciences, Pennsylvania State University,
54 University Park, Pennsylvania 16802, USA
- 55 19. State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, Zhejiang
56 311300, China
- 57 20. Institut de Biologie Moléculaire des Plantes, CNRS, University of Strasbourg, 67084 Strasbourg,
58 France

- 59 21. Inselspital, University of Bern, 3010 Bern, Switzerland
- 60 22. MetaSystems Hard & Software GmbH, 68804 Altlußheim, Germany
- 61 23. Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, Sichuan 610041, China
- 62 24. School of Life Sciences, Institute of Life Sciences and Green Development, Hebei University,
63 Baoding, Hebei 071002, China
- 64 25. Department of Biology, School of Life Sciences, Southern University of Science and Technology,
65 Shenzhen, Guangdong 518055, China
- 66 26. Key Laboratory of Plant Functional Genomics of the Ministry of Education, Jiangsu Co-Innovation
67 Center for Modern Production Technology of Grain Crops, Yangzhou University, Yangzhou,
68 Jiangsu 225009, China
- 69 27. Key Laboratory of Southern Subtropical Plant Diversity, Fairy Lake Botanical Garden, Shenzhen
70 & Chinese Academy of Sciences, Shenzhen, Guangdong 518004, China
- 71 28. State Key Laboratory of Agricultural Genomics, BGI-Shenzhen, Shenzhen, Guangdong 518085,
72 China
- 73 29. Signalling Research Centres BIOS and CIBSS, University of Freiburg, 79104 Freiburg,
74 Germany

75

76 **KEY WORDS**

77 abiotic stress; adaptive evolution; development; DNA repair; fast-evolving genome; gene expansion;
78 global warming; molecular clock; UV radiation; palaeontology

79

80 **SUMMARY**

81 The most extreme environments are the most vulnerable to transformation under a rapidly changing
82 climate. These ecosystems harbor some of the most specialized species, which will likely suffer the
83 highest extinction rates. We document the steepest temperature increase (2010–2021) on record at
84 altitudes of above 4,000 meters, triggering a decline of the relictual and highly adapted moss *Takakia*
85 *lepidozoides*. Its *de novo* sequenced genome with 27,467 protein-coding genes includes unique
86 adaptations to abiotic stresses and comprises the largest number of fast-evolving genes under
87 positive selection. The uplift of the study site in the last 65 million years has resulted in life-
88 threatening UV-B radiation and drastically reduced temperatures, and we detected several of the

89 molecular adaptations of *Takakia* to these environmental changes. Surprisingly, peculiar
90 morphological features likely occurred earlier than 165 Mya. Following nearly 400 million years of
91 evolution and resilience this species is now facing extinction.

92

93 **INTRODUCTION**

94 About 500 million years ago (Mya), when the ancestors of humans were tiny water-borne animals¹,
95 the Earth's rocky land masses were colonized by plants that evolved from freshwater algae. A series
96 of biochemical and morphological innovations helped these first land plants to cope with and diversify
97 in the new and harsher terrestrial environments, triggering a massive change in global carbon cycling
98 through rock weathering and photosynthetic activity^{2,3}. This drastically decreased atmospheric CO₂,
99 resulting in an ice age⁴, and elevated atmospheric O₂ concentration similar to modern levels, thereby
100 facilitating the evolution of complex animals on land^{5,6}. Stomata were one such critical innovations.
101 These microscopic pores on plant surfaces, originated over 400 Mya, facilitating the greening of the
102 land masses by permitting plant-atmosphere gas exchange and plant hydration^{7,8}.

103 Starting 65 Mya, the India-Asian collision created the Himalayas with an average altitude
104 exceeding 4,500 m⁹, resulting in one of the harshest environments on Earth¹⁰. A part of it, the Tibetan
105 Plateau, is often referred to as "Roof of the World" where global warming drives upslope tree
106 expansion and habitat loss of the endemic flora¹¹. Due to their ecophysiological features and
107 environmental sensitivity, mosses are considered as indicators of climate change¹².

108 Mosses, together with liverworts and hornworts, compose the bryophytes, the sister lineage to
109 vascular plants¹³. They likely share with the first land plants a gametophytic life style with freely
110 moving sperm cells and a lack of roots, true vasculature, flowers and seeds¹⁴. Many bryophytes are
111 pioneer plants¹⁵ well adapted to harsh environments^{16,17}. Several genome sequences of model
112 bryophytes are available and have shed light on the early evolution of land plants¹⁸⁻²².

113 The genus *Takakia* comprises only two species: *T. lepidozoides* and *T. ceratophylla*, and its
114 phylogenetic affinities to either algae, mosses or liverworts²³⁻²⁷ was long ambiguous, until the
115 discovery of their spore-bearing capsules that are similar to those of mosses^{25,28}. *Takakia* is now
116 resolved phylogenetically as the sister group to all other extant mosses²⁹, from which it differs in leaf
117 morphology and the mode of sporangial dehiscence, while their gametophyte resembles those of
118 some liverworts^{23,25}, as some biochemical features do. The possession of some algal features further
119 suggests that *Takakia* may share some characteristics with the last common ancestor (LCA) of land

120 plants and provide critical insights into early land plant evolution.

121 Both species of *Takakia* co-occur only in the Tibetan Plateau, which may be viewed as the
122 modern distribution center of *Takakia*²⁸. In 2005, we found *T. lepidozoides* in the Gawalong glacier
123 area above 4,000 meters elevation, well above the tree line (Figure 1A) where the endemic
124 vegetation is undergoing drastic and profound changes due to local warming under climate change¹¹,
125 but the fate of specific species is unclear. To comprehensively explore the genetic and phenotypic
126 characteristics of *T. lepidozoides* and correlate its adaptive evolution to environmental changes, we
127 set up sample sites in a ~18 km² region, with real-time image and meteorological data recorders to
128 characterize individual plant growth and reproductive development, and population dynamics in a
129 climatic context. Further, we complemented the data of *Takakia* fossil and paleometeorology, and
130 field observations with laboratory experiments on its physiology, its biochemistry and cell biology,
131 and its genome bioinformatics, revealing adaptations to harsh environments, in particular, extreme
132 tolerance to high UV radiation and freezing, resulted from the uplift of our experimental site started
133 65 Mya.

134

135 **RESULTS**

136 **Global warming impacted population dynamics in the past decade**

137 We studied populations of *T. lepidozoides* in the southern part of the Tibetan Plateau, a top hotspot
138 of global biodiversity conservation³⁰. We established 68 sample sites (10×10 cm² each) in three
139 areas at altitudinal intervals of about 100 m (Figures 1A–1C; Movie S1). The matrix properties of
140 sample substrates comprised three types: soil, rock covered with a thin layer of soil and bare rock.
141 The composition of mineral elements was analyzed from 20 soil samples (Data S1A).

142 The average annual temperature in the sampling area increased by 0.32 °C per decade from
143 1961 to 2021 (Figure 1D). The snow-cover-free period for *Takakia* plants spans from June to October
144 for 130 ± 11 days (Figure 1E; Table S1A), during which 60% of the annual average precipitation
145 occurred, and air humidity and soil moisture were the highest (Data S1B). Between 2010 and 2021,
146 the annual average temperature increased by 0.43 °C, which is the highest in documented history,
147 and the Gawalong East Glacier retreated by 49 ± 3 m / year (Data S1C and S1D; Table S1B), above
148 the average recorded for 82 glaciers in this area in 30 years (1970–2000)³¹.

149 Between 2010 and 2021, the average coverage of *T. lepidozoides* populations decreased by
150 1.6% annually (Figure 1F), faster than that of four local common mosses (Figures S1A–S1D). By

151 contrast, the coverage of the liverwort *Herbertus sendtneri* increased by 2.27% / year (Figures S1E–
152 S1G), that of the angiosperm *Diapensia wardii* by 0.88% / year (Figure S1H), and that of the
153 lichenized fungus *Stereocaulon alpinum* by 0.47% / year (Figure S1M). We employed principal
154 component analysis to further explore the major variation feature in populations and found that
155 populations of the liverwort *Plagiochila recurvata* also increased in size (Data S1E; Table S1C).
156 Based on Spearman rank correlation coefficient between changes in population size in these species
157 and environmental factors, shifts in *T. lepidozoioides* were significantly correlated to changes in
158 maximal temperature during the growing season (Data S1E; Table S1D). The mineral content of the
159 substrate supporting versus lacking *T. lepidozoioides* (Data S1A) was significantly richer in N, Mg and
160 Fe and poorer in P, S and Ca (Data S1A), suggesting that these macronutrients shape the local
161 distribution of *T. lepidozoioides*. We detected no significant differences between the dynamics of *T.*
162 *lepidozoioides* populations on different substrates, suggesting that only warming shapes the loss in
163 coverage of this species (Data S1F; Table S1E).

164

165 **Vegetative growth for stable network structures**

166 Spores of *T. lepidozoioides* germinate to form thalloid protonemata, reminiscent of liverwort
167 protonemata. These develop cylindrical rhizomes, which continuously branch or develop erect stems
168 with deeply lobed leaves (phylloids) (Data S1G and S1H). Rhizoids are always lacking. In culture,
169 erect stems grow 1.52 ± 0.03 mm in 130 days (corresponding to the four months snow-free period
170 in the field), compared to 1.35 ± 0.44 mm in the field, while rhizomatous, leafless and stoloniform
171 stems grow by 2.87 ± 0.02 mm over the same period (Data S1H). The distance between two adjacent
172 branches (rhizomes and erect stems) is 1.73 ± 1.56 mm in the field and 2.26 ± 2.74 mm in culture
173 (Data S1H).

174 The branches do not originate from the leaf axil, unlike in angiosperms such as *Arabidopsis*
175 *thaliana*. In contrast, it is similar to the randomness of lateral root-formation position in *A. thaliana*³².
176 Branches originate from epidermal and subepidermal cell layers of the stem. While epidermal cells
177 divide vertically, sub-epidermal cells divide periclinally and anticlinally, forming a dome, before
178 gradually elongating to form a new branch (Figures 2A, 2B and S2; Movie S2). The cells in the middle
179 part of stems are longer, with inclined walls between adjacent cells, and connect vertically at the
180 intersection between the parental stem and the new branch (Figures 2A, 2B and S2). These potential
181 vessel elements facilitate transport of macromolecules (Figure 2C). Thus, the connecting pattern is

182 similar to the connect model of vessel elements in roots (Figures 2D; Data S1I), but different from
183 that in stems of vascular plants such as *Arabidopsis* (Figures 2E; Data S1I). New erect branches
184 grow 5.22 ± 1.93 mm (Data S1J) to achieve sexual maturity in ~3 years (Data S1J). Thus, individual
185 rhizomes branch continuously, forming a population of clonal erect stems that ultimately decay at
186 their base and maintain growth at the tip (Figures 2F and 2G; Data S1H). Erect stems and rhizomes
187 continuously branch and grow in different directions, adjoined with surface mucilage hairs at cross-
188 points (Data S1K), resulting in a network structure (Data S1L) that can withstand heavy snowstorms
189 (Data S1M). Stems of *T. lepidozoides* show weaker negative geotropism (Data S1L, S1N) compared
190 to other surveyed mosses (Data S1O), while positive phototropism is obvious (Data S1N).

191

192 **Reproductive strategy**

193 The sexual organs of *T. lepidozoides*, i.e., the male antheridia and female archegonia, could not be
194 induced in culture. In wild populations, erect stems develop either antheridia or archegonia.
195 Beginning in July, pale red antheridia are initiated and their number peaks in September (Data S1P).
196 Stems carry 9.13 ± 2.07 antheridia (Data S1G and S1P) or 4.63 ± 0.82 archegonia at the stem tip
197 (Data S1Q). One stem can produce antheridia continuously for ~3 years (Data S1P). Fertilization
198 occurs in autumn, and zygotes or early embryos remain dormant under the snow at temperatures of
199 -1 to -4 °C for about eight months. In mosses, embryos grow into unbranched sporophytes with
200 terminal sporogenous capsules. In the field, light green *T. lepidozoides* sporophytes occur
201 synchronously about ten days after snow melt in June. Following the rapid elongation of the capsule
202 stalk (seta) to 3.13 ± 0.29 mm within ~30 days, capsules expand and mature to become deep-brown
203 in ~35 days (Data S1R and S1S; Movie S3). Generally, one stem supports one sporophyte, only
204 occasionally two (Data S1R). At maturity capsules dehisce along a vertical spiral line to release
205 19.5 ± 2.9 thousand spores (Data S1S) in a radius of 1.36 ± 0.69 cm (Data S1S). *T. lepidozoides* capsules
206 lack stomata (Data S1T). In culture, spore germination rates reached 96%. In the field, we have not
207 found new spore-derived populations since 2005. Of 52 samples, only 22 populations had
208 reproductive organs: three with archegonia, nine with antheridia, and ten with mixed male and female
209 stems. These populations did not differ significantly in their dynamics, revealing that *T. lepidozoides*
210 mainly depends on clonal growth in the Tibetan Plateau.

211

212 **Tolerance to extreme UV-B radiation**

213 Exposure to solar UV-B radiation increases with altitudes and hence, at high elevations, mutagenic
214 UV-B³³ is a potent stressor. We exposed *T. lepidozoides*, the model moss *Physcomitrium*
215 (*Physcomitrella*) *patens*³⁴ and the model angiosperm *A. thaliana* to 144 kJ m⁻² UV-B, about ten times
216 the intensity in our sample site, for two hours. Leaves of *A. thaliana* began wilting within 24 hours
217 after treatment, and within 72 hours plants withered and died (Figure S3). Apical leaves of *P. patens*
218 turned pale within 72 hours, and plants gradually died, whereas *T. lepidozoides* plants appeared to
219 be unaffected (Figures 3A–3F). Cellular damage in *P. patens* included disordered chloroplast
220 lamellae, expanded mitochondria, disintegrated nuclear membranes, agglomerated chromatin, and
221 final cell collapse (Data S1U). By contrast, cells of *T. lepidozoides* exhibited only a slight
222 condensation of heterochromatin but harbored significantly more oil bodies in response to UV-B
223 (Data S1U).

224 These oil bodies were rich in neutral lipids and membrane lipids (Data S1V). The total lipid
225 content was 1.91 µg / mg fresh weight (fw), including 61.61% neutral lipids, of which 53.41% are
226 polyunsaturated fatty acids (PUFAs). In contrast, the total lipid content of *P. patens* was 1.01 µg /
227 mg fw, including 19.36% neutral lipids, of which 36.29% are PUFAs (Data S1W). Thus, *T.*
228 *lepidozoides* does not only contain much more overall lipids than *P. patens*, but especially more
229 PUFAs, which are involved in tolerance³⁵ to various biotic and abiotic stresses. Upon UV-B treatment,
230 the contents of diacylglycerol-*N,N,N*-trimethylhomoserine (DGTS), neutral lipids and PUFAs in
231 neutral lipids increased significantly in *T. lepidozoides* (Figure 3G; Data S1X), compared to no DGTS
232 or stagnation of neutral lipids in *P. patens* (Figure 3H; Data S1X). Monogalactoglycerolipid (MGDG)
233 is the most abundant glycolipid in both mosses. Exposure to UV-B decreased its content slightly in
234 *T. lepidozoides*, but significantly in *P. patens* (Data S1X). Phosphatidylcholine (PC), the main
235 component of nonphotosynthetic membrane lipids, is the major phospholipid in both mosses. Upon
236 radiation, the amount of PC remained stable in *T. lepidozoides* but decreased in *P. patens* (Data
237 S1X).

238 After UV-B treatment, amounts of polyphenols and flavonoids doubled to tripled in *T.*
239 *lepidozoides* to a fivefold equivalent of that in *P. patens* (Figures 3I and 3J; Table S1F), which was
240 also supported by the result of non-targeted metabolic profiles (Data S1Y). Likewise, catalase activity
241 increased while levels of the toxic malondialdehyde remained stable in *T. lepidozoides*, in contrast
242 to *P. patens* (Figures 3K and 3L). While the maximum quantum efficiency of PSII photochemistry
243 (Fv/Fm) was about 0.75 in both mosses, it decreased to 20% in *P. patens* and to 50–70% in *T.*

244 *lepidozooides* following UV-B treatment (Figure 3M). Likewise, the actual photochemical quantum
245 efficiency (Fv/Fo) decreased sharply in *P. patens* but only slightly in *T. lepidozooides* (Figure 3N).
246 The increased tolerance of *T. lepidozooides* to UV-B radiation was also visible on protein levels for
247 both photosystems (Figures 3O and 3P). Taken together, these results revealed a strong defense
248 response to UV-B in *T. lepidozooides* aimed at blocking radiation and at scavenging toxic ROS, and
249 maintaining photosynthetic capacity by preventing lipid peroxidation and cell structure damage.

250

251 **Takakia Genome Features**

252 **Assembly, Annotation and Comparison**

253 The haploid genome of *T. lepidozooides* gametophytes that were initiated from a single spore was
254 sequenced (Table S1G) and assembled. It comprises 325 Mbp, with 98.5% anchored to four
255 chromosomes (Figures 4A and S4A), which coincides with our cytological observation (Figures S4B–
256 S4E), and evidence from a Japanese isolate³⁶. These chromosomes (1–4) are 96, 83, 77 and 64
257 Mbp in size, respectively. The genome sequence represents 87% of the 375 Mbp estimated by flow
258 cytometry (Data S2A; Table S1H) and 98% of the K-mer estimation of 331 Mbp (Data S2A), with a
259 contig N50 length of 1.4 Mbp and a scaffold N50 length of 83 Mbp (Table S1I). The completeness of
260 the genome estimated by BUSCO indicated that 80–97% of the conserved core gene sets are
261 present (Data S2B). As in *P. patens*, we were unable to define specific centromeric regions. Gene
262 annotation combined various strategies of *de novo*, homology, and transcriptome predictions. The
263 final set contains 27,467 high confidence protein-coding gene models with mean gene length of
264 3,301 bp, of which 85% have an assigned function (Table S1J). Further BUSCO analysis show that
265 71–89% conserved core gene sets are present (Data S2B).

266 We found that over 52% of the *T. lepidozooides* genome (~172 Mbp) is composed of repetitive
267 sequences. Long terminal repeat retrotransposons (LTRs) are the most abundant transposable
268 elements (TEs), occupying more than 36% of the genome (Table S1K), similar to the 35.6% in the
269 lycophyte *Selaginella moellendorffii*³⁷, but different from the nearly 50% in *P. patens*³⁸ and the 22%
270 in the liverwort *Marchantia polymorpha*¹⁹. Of the 119 Mbp LTRs in *T. lepidozooides*, 84% are from the
271 gypsy superfamily and 5% are from the copia superfamily (Table S1K), different from the 48% and
272 3.5%, respectively, in *P. patens*³⁸. In *T. lepidozooides*, the median and mean age of copia elements
273 is 0.85 Mya and 0.65 Mya, respectively, and of gypsy elements is 0.66 Mya and 0.47 Mya,
274 respectively (Data S2B).

275 Moreover, we established the organellar genomes of *T. lepidozoides* with the plastome
276 (NC_005087.1) and chondriome (NC_007945.1) of *P. patens* as seeds. The *T. lepidozoides*
277 plastome is a circular molecule of 149,011 bp (Data S2C) and similar to the plastome of a Japanese
278 isolate³⁹. The *T. lepidozoides* chondriome is a circular molecule of 139,841 bp (Data S2C).

279

280 **Whole Genome Evolutionary History**

281 Intragenomic analysis detected 437 syntenic blocks containing 2,382 gene pairs with an average of
282 5.45 gene pairs / block (Data S2D). Synonymous substitutions per site (Ks) age distributions showing
283 a peak range from 0.7 to 1.1 (median 0.9) unveiled evidence in *T. lepidozoides* for an ancient
284 diploidized state (Figures 4B; Data S2D). Based on an assumed rate of silent-site substitution of 9.4
285 per silent site per billion years³⁸, the diploidization occurred 58.51–37.23 Mya (median 47.87 Mya),
286 during the period of the Tibetan Plateau uplifting⁹.

287 Regarding diploidization, including whole-genome duplication (WGD) and single-gene
288 duplications, we identified 26,829 duplicated genes that were classified into five categories: 772
289 WGD genes (2.88%), 452 tandem duplicated genes (TD, 1.68%), 1,172 proximal duplicated genes
290 (PD, 4.37%), 6,005 transposed duplicated genes (TRD, 22.38%) and 18,428 dispersed duplicated
291 genes (DSD, 68.69%). To identify the most recent and more ancient duplicated genes in *Takakia*,
292 and how evolutionary forces inferred to affect them, we estimated the Ks values, nonsynonymous
293 substitutions per site (Ka) values and Ka/Ks values for the gene pairs contained in each mode of
294 gene duplication. We compared the Ks and Ka/Ks distribution among different modes of gene
295 duplication. WGD genes showed a Ks peak ranging from 0.75–1.10 (median 0.90) and Ka/Ks peak
296 ranging from 0.15–0.32 (median 0.23); TD genes showed a Ks peak ranging from 0.12–0.65 (median
297 0.23) and Ka/Ks peak ranging from 0.18–0.52 (median 0.32); PD genes showed a Ks peak ranging
298 from 0.06–0.46 (median 0.15) and Ka/Ks peak ranging from 0.20–0.63 (median 0.40); TRD genes
299 showed a Ks peak ranging from 0.05–0.26 (median 0.12) and Ka/Ks peak ranging from 0.18–0.59
300 (median 0.36); DSD genes showed a Ks peak ranging from 0.07–0.54 (median 0.18) and Ka/Ks
301 peak ranging from 0.19–0.57 (median 0.33). Higher Ka/Ks ratios and smaller Ks values were found
302 for PD gene pairs and TRD gene pairs than for other duplication modes (Figures 4C; Data S2E),
303 suggesting a more rapid sequence divergence and a stronger positive selection on genes in these
304 two modes than other duplication modes. These genes may have implications in stress responses
305 involving hormones, such as jasmonic acid signalling (Data S2E). Most retained WGD genes

306 participated in post translational modifications of proteins, such as protein esterification and
307 deacylation, which may be involved in stress responses^{40,41} (Data S2E). According to the rate of
308 silent-site substitution of 9.4 per silent site per billion years³⁸, the WGD, TD, PD, TRD and DSD
309 occurred 58.51–37.23 Mya (median 47.87 Mya), 34.57–6.38 Mya (median 12.23 Mya), 24.47–3.19
310 Mya (median 7.98 Mya), 13.83–2.66 Mya (median 6.38 Mya) and 28.72–3.72 Mya (median 9.57 Mya)
311 respectively, and most of duplications occurred during the Tibetan Plateau uplifting. Furthermore, 27
312 singleton genes occur in *T. lepidozoides*, some of which may participate in DNA repair (Data S2E).

313

314 **Phylogenetic Position**

315 Inferences from 105 concatenated nuclear genes resolve *Takakia* as sister to all other mosses,
316 including *Sphagnum*²⁹, which was previously considered close to *Takakia*⁴², despite their
317 conspicuous morphological dissimilarities. With the sequenced genome of *Takakia*, and genomes of
318 two Charophyte species most closely related to land plants, we re-assessed the phylogenetic
319 relationships of *Takakia* to other bryophytes based on an analysis of 818 nuclear genes from 56 plant
320 species (Table S1L) and resolved *Takakia* as sister to other extant mosses, including *Sphagnum*
321 (Data S2F), congruent with the outcome of our concatenated analysis and previous coalescent
322 analyses²⁹. Molecular clock analyses estimated that *Takakia* diverged from the lineage leading to
323 other mosses ~389.92 Mya, ~38 million years before the separation of *Sphagnum* from other mosses,
324 following the divergence of liverworts and mosses at 484.32 Mya (Figure 4D; Data S2F).

325

326 **Special Genes**

327 To gain further insights about how *Takakia* is related to mosses, and other bryophytes, we compared
328 genes among those conserved orthogroups among bryophytes and also Charophytes. *Takakia*
329 shares with one or more Charophytes 1,118 genes in 438 orthogroups (Data S2G) that are not found
330 in *P. patens*, suggesting that these genes have been retained from the ancestor of land plants, but
331 were lost in *P. patens*. In addition, *Takakia* shares with hornworts and liverworts 1,665 genes in 461
332 orthogroups (Data S2G) that are not detected in *P. patens*, suggesting that they were present in the
333 most recent ancestor to all bryophytes, and lost in a subsequent lineage giving rise to *P. patens*.
334 Although the functions of many of these genes are not yet understood, it is possible that some of
335 them are the basis for *Takakia* having some characteristics similar to liverworts and algae, but not to
336 mosses.

337

338 **Genetic Foundation of Adaptive Features**

339 ***Positive Selection and Fast Evolution***

340 To gain insight into the evolutionary history of *Takakia*, we compiled a data set including orthologs
341 from 12 plant taxa including one hornwort, four liverworts, five mosses and one eudicot. We
342 employed maximum likelihood analysis using the branch-site model in PAML and detected positive
343 selection signals in 83 out of 1,970 genes in *T. lepidozoides*. According to GO clustering, positively
344 selected genes are enriched in the categories DNA repair, response to oxidative damage,
345 photosynthesis and development (Table S1M). To identify fast-evolving genes, we estimated the
346 global and local dN/dS (non-synonymous substitution rates / synonymous substitution rates) ratios,
347 respectively. In total, 121 significantly fast-evolving genes were identified. Three functional
348 annotation clusters ranked among the top 10 results: development, DNA repair and response to
349 abiotic stress (Table S1N). In comparison with the liverwort *H. sendtneri* from the same environment,
350 we identified 59 genes with convergent accelerated evolutionary rate (Table S1O) and 16 shared
351 genes in all data sets (Table S1P). Their functions include metabolic process, energy generation,
352 photosynthesis, oxidation-reduction, and defence response.

353 We calculated the dN/dS ratios for each lineage and compared *T. lepidozoides* to *A. thaliana*,
354 *P. patens*, other mosses (excluding *T. lepidozoides*) and liverworts (Data S2H) and found that the
355 upper boundary in *T. lepidozoides* was higher than of any other taxon group (Data S2H). This
356 suggests that many fast-evolving genes occurred along the *Takakia* lineage. We further calculated
357 the mean dN/dS ratios of GO categories for *T. lepidozoides*, *A. thaliana*, *P. patens*, and the LCA of
358 all mosses (including *T. lepidozoides*) (Data S2H). Compared to *A. thaliana* and *P. patens*, we found
359 a higher number of rapidly evolving GO categories in *T. lepidozoides* (342 in *T. lepidozoides* vs 78
360 in *A. thaliana*, 208 in *T. lepidozoides* vs 124 in *P. patens*). Further, 56 GO categories evolved faster
361 in the LCA of mosses than in *T. lepidozoides*. In contrast, 391 GO categories evolved faster in *T.*
362 *lepidozoides* than in the LCA of all mosses. Among these are 31 GO categories that evolve faster
363 in *T. lepidozoides* than in all other mosses (Table S1Q). These include response to desiccation,
364 osmotic stress, heme biosynthesis, auxin metabolism and photosynthesis. Taken together, *T.*
365 *lepidozoides* has the largest number of fast-evolving genes under positive selection on documented
366 record.

367

368 **Transcription Associated Proteins**

369 Transcription associated proteins (TAPs) comprise transcription factors (TFs) and transcriptional
370 regulators (TRs). Genome-wide, domain-based annotation of TAPs⁴³ identified 1,091 TAPs, of which
371 427 are TFs and 664 TRs, which is more than that of the hornwort *A. angustus* (570) and liverwort
372 *M. polymorpha* (682), but less than that of *P. patens* (1,638). However, the proportion of TAPs to
373 protein-coding genes in *T. lepidozoides* (3.97%) is similar to *A. angustus* (3.89%) and *M. polymorpha*
374 (3.53%) but less than in *P. patens* (4.97%) (Table S1R). In general, most TAP families in *T.*
375 *lepidozoides* have a similarly low number in *A. angustus* and *M. polymorpha* than in *P. patens*. For
376 example, the zf-HD family, which plays a key role in the control of growth and development⁴⁴,
377 comprises a single member in *T. lepidozoides*, *A. angustus* and *M. polymorpha*, but 12 in *P. patens*.
378 Similarly, C2C2-CO-like TFs, which act between the circadian clock and genes controlling meristem
379 identity⁴⁵, are represented by two genes in *T. lepidozoides* and *A. angustus*, three in *M. polymorpha*,
380 but 16 in *P. patens*. Interestingly, *T. lepidozoides* encodes no nucleic acid-binding cold shock domain
381 protein (CSD) (Table S1R), which is an ancient and conserved family of proteins in various
382 prokaryotes and eukaryotes⁴⁶. In *Arabidopsis*, CSD proteins negatively regulate freezing tolerance⁴⁷.

383 By contrast, with 438 genes, plant homeodomain zinc fingers (PHDs) are the most abundant
384 TRs in *T. lepidozoides*, a diversity vastly outnumbering that in *A. angustus* (32), *M. polymorpha* (55)
385 and *P. patens* (45) (Table S1R). PHDs are structurally conserved modules found in plants as
386 epigenome readers that control gene expression through recruitment of multi-protein complexes of
387 chromatin regulators, and play crucial roles in abiotic stress responses⁴⁸. Among the 438 genes, 118
388 are in small clusters of 2–5 genes in close proximity (fewer than 10 genes), with 164 of them lacking
389 introns. Ks analysis showed these intronless genes have a peak ranging from 0.19–0.54 (median
390 0.33), corresponding to their expansion 28.71–10.10 Mya (median 17.50 Mya), suggesting that at
391 least some of them have experienced tandem duplications and retrotranspositions, respectively.

392 Mitochondrial transcription termination factors (mTERFs) are diversified in *T. lepidozoides* (18)
393 compared to *A. angustus* (16), *M. polymorpha* (15) and *P. patens* (16) (Table S1R). mTERFs are
394 localized in mitochondria or chloroplasts and involved in stimulating transcription and stabilizing post-
395 transcriptionally processed mRNA, terminating transcription, promoting tRNA maturation, intron
396 splicing and chloroplast ribosome assembly and translation⁴⁹. They are important for stress
397 responses and development⁵⁰. The mTERFs in *T. lepidozoides* have a Ks peak range from 1.69–
398 2.57 (median 2.07), suggesting that their diversity occurred 136.70–89.89 Mya (median 110.11 Mya).

399 Likewise, auxin/indole-3-acetic acid repressors (Aux/IAAs) are diversified in *T. lepidozoides* (5)
400 compared to *A. angustus* (2), *M. polymorpha* (3) and *P. patens* (4) (Table S1R). The diversity of
401 Aux/IAAs in *T. lepidozoides* occurred 218.62–142.02 Mya (median 218.09 Mya).

402

403 **Protein phosphorylation**

404 Protein phosphorylation is important for regulating protein activity and plays crucial roles in
405 responses to environmental stresses⁵¹. In *Takakia*, 273 genes encode protein kinases, including 92
406 that are in small dense clusters of 2–4 genes (Data S2I), suggesting that they were generated by
407 tandem duplications. Moreover, 32 genes lack introns and are scattered throughout the genome,
408 implying that some of them were acquired by retrotranspositions. According to Ks distribution, they
409 diverged 145.07–84.76 Mya (median 107.26 Mya). However, 40 genes are younger than 30 Mya
410 (Data S2I).

411

412 **PPR proteins and RNA-editing**

413 Nuclear-encoded pentatricopeptide (PPR) proteins localize to mitochondria and chloroplasts, control
414 organellar gene expression, including RNA-editing⁵², and thus impact growth and development in
415 response to stresses^{53,54}. The *T. lepidozoides* genome encodes 1,926 PPRs, far more than that of
416 *P. patens* (105) or *M. polymorpha* (69) (Figure 5A; Table S2A–S2H). An examination of their
417 chromosomal positions revealed at least ~270 small clusters of two or more tandemly repeated
418 genes (total 789 genes), and 271 of them have only 1–3 introns with a Ks value lower than 0.3,
419 suggesting that they result from tandem duplications at about 15 Mya. About 429 PPR genes lack
420 introns and are dispersed throughout the genome; 95% of them have a Ks value below 0.5 and 213
421 of them have a Ks value below 0.3, suggesting that they expanded 30–15 Mya. Some of them might
422 have experienced retro-transposon-mediated duplication (Data S2I). While in most bryophytes PPR
423 proteins are equally distributed between plastids and mitochondria, the majority target mitochondria
424 in *T. lepidozoides* (Figure 5B). RNA-editing occurs rarely in mosses (10 sites in *P. patens*). By
425 contrast, we identified 912 editing sites in the *T. lepidozoides* organellar genomes, which is more
426 than in *A. thaliana* (456) (Figure 5C; Table S2I and S2J). PPR-expression is induced by UV-B
427 treatment (Figure 5D). Taken together, the expansion of PPR proteins and RNA-editing sites in *T.*
428 *lepidozoides* may be linked to adaptation to extreme environments, such as exposure to elevated
429 UV radiation (Figures 5E–5G).

430

431 **DNA repair**

432 UV-B radiation leads to pyrimidine dimers and (6-4) photoproducts, which can be excised and
433 repaired in nuclear DNA by a group of proteins including ERCC1 and RAD2⁵⁵. While *P. patens* and
434 *A. thaliana* each encode one ERCC1 and one RAD2, *T. lepidozoioides* underwent a duplication of
435 each these (Table S2K) about 5.57 Mya and 203.17 Mya, respectively. RAD51 is the eukaryotic
436 homolog of bacterial RecA and a key factor in homologous recombination and DNA damage
437 repair^{56,57}. We identified five RAD51 genes in *T. lepidozoioides* compared to six in *P. patens* (Data
438 S2J). In *T. lepidozoioides* a pair of RAD51 (Tle2c04609.1, Tle4c02755.1) arose at about 20.78 Mya,
439 while paralogs in *P. patens* (Pp3c11_22000V3.1, Pp3c7_7920V3.1) separated at about 44.27 Mya.
440 BRCA1 and BARD1 evolved from a common eukaryotic ancestor and are required for efficient DNA
441 repair in humans and plants⁵⁸. Like all other plants, except *P. patens* which lacks the gene, *T.*
442 *lepidozoioides* encodes one BRCA1 (Table S2L). To trace its evolutionary history, we analysed the
443 BRCA1 and BARD1 homologs in 11 bryophyte species with published genomes (two hornworts, one
444 liverwort, eight mosses), and found that BRCA1 and BARD1 were lost in the ancestor to other
445 mosses after *Takakia* separated from their LCA (Data S2K). Taken together, the genome of *T.*
446 *lepidozoioides* is uniquely equipped with genes enhancing its ability to recover from DNA damage
447 incurred by exposure to high UV radiations associated with its current habitat at high elevations in
448 the Tibetan Plateau.

449

450 **Phenylalanine ammonia-lyases**

451 *T. lepidozoioides* is rich in phenolic compounds, including flavonoids (Figure 3J; Data S1Y), which are
452 products involved in the phenylpropanoid pathway⁵⁹. Phenylalanine ammonia-lyase (PAL) is the first
453 and key enzyme in phenylpropanoid biosynthesis and is expressed in response to various stresses⁶⁰.
454 *T. lepidozoioides* encodes 20 PALs, more than *A. thaliana* (4), *M. polymorpha* (12) or *P. patens* (16).
455 Sixteen of them cluster in a subclade, and nine of them are intronless. From these 20 PALs, 15
456 expanded 50 Mya and the intronless PALs expanded 30 Mya, indicating a recent expansion via
457 segmental or tandem duplications (Data S2L). In *Takakia*, PAL expression increases under UV-B
458 stress and reaches the highest level during the recovery from UV-B exposure (Data S2L).

459

460 **Cytochromes P450**

461 CYP450s are widespread in eukaryotes. They catalyse various oxygenation / hydroxylation reactions
462 that require molecular oxygen and depend on the electron transfer from NADPH via NADPH-P450
463 reductase. CYP450s are involved in a variety of pathways producing primary and specialized
464 metabolites⁶¹. In mosses, they are also important for the formation of the cuticle, a complex of
465 phenol-enriched polymers on the epidermis that enables organogenesis and protects from
466 dehydration⁶². We identified 71 CYP450s, close to the number in *P. patens* (80). Compared to *P.*
467 *patens*, several clades of CYP450s such as CYP703A2 , CYP704 and CYP77, which are involved
468 in fatty acid and hormone metabolism, are expanded in *T. lepidozoides* (Data S2M). This expansion,
469 dated to 145.21–87.77 Mya (median 136.17 Mya), may explain why its fatty acid content is higher
470 than that of *P. patens* under normal conditions and under UV stress.

471

472 **Genetic basis of phenotypes and development**

473 ***Leaf adaxial-abaxial polarity***

474 Most extant land plants have bifacial leaves with adaxial-abaxial (dorsiventral) polarity, which may
475 have been different in the earliest land plants⁶³. *T. lepidozoides* leaves have no such polarity (Data
476 S1G). Polarity establishment in angiosperms is controlled by HD-ZIP III TFs and KAN TFs^{64,65}. While
477 the former are present in *T. lepidozoides*, *P. patens* and *A. thaliana*, the latter are missing in *T.*
478 *lepidozoides* (Table S2M), suggesting that the non-polar leaves resulted from the loss of KAN1
479 (Figure 6A). The *Takakia* sp. fossil records dated to the Middle-Late Jurassic (165 Mya⁶⁶) help us to
480 trace the evolutionary history of morphological features in *T. lepidozoides*. The fossil plant shape
481 (height 12.48 ± 2.46 mm, stem diameter 0.31 ± 0.14 mm; leaf length 0.88 ± 0.17 mm, leaf diameter
482 0.04 ± 0.01 mm) is similar to that of extant *T. lepidozoides* (height 11.14 ± 2.90 mm, stem diameter
483 0.26 ± 0.08 mm; leaf length 0.75 ± 0.16 mm, leaf diameter 0.04 ± 0.01 mm) in the experiment site of
484 Tibet Plateau, and both of them hold similarly lobed terete leaves (Data S2N). Based on data of
485 morphological characteristics of the fossil and extant *Takakia* plants, we suppose that *Takakia* might
486 have lost adaxial-abaxial leaf polarity no later than 165 Mya and after the LCA of all mosses. To
487 better understand the evolutionary history of KAN in mosses, we constructed a timetree using KAN
488 homologs from 11 bryophyte species with two fossil calibrations and found that the KAN diversity
489 within mosses occurred 396.15–240.84 Mya (median 314.86 Mya) (Data S2O), displaying these
490 divergences of KANs did not disturb their function for maintaining the adaxial-abaxial leaf polarity
491 among mosses. The ancestor of *Takakia* lost KAN likely between 389.92 Mya, when it separated

492 from the LCA of all mosses, and 165 Mya, resulting in the unique leaf morphology of *Takakia* among
493 all mosses surveyed so far.

494 In addition, jasmonate signaling is important for leaf expansion⁶³. From these factors, *T.*
495 *lepidozoides* lacks MYC2 and JAZ1 (Data S2P; Table S2P). Another possibility is post-
496 transcriptional regulation of gene expression via microRNAs (miRNAs), as *T. lepidozoides* has only
497 one DICER1 among three DICER proteins (Data S2Q). In *P. patens*, knockout of *DCL1a* or *DCL1b*
498 affects stem and leaf differentiation⁶⁷. Together, these deficiencies may explain the deeply lobed leaf
499 shape in *T. lepidozoides*, which is unique among mosses but found in some liverworts⁶⁸.

500

501 **Rhizoids**

502 Rhizoids are haploid filamentous cells that facilitate interaction with the substrate, including water
503 absorption and nutrient uptake⁶⁹. They share similar regulatory developmental networks to those of
504 root hairs of angiosperms⁷⁰. Unlike other mosses, *T. lepidozoides* lacks rhizoids (Data S1G and
505 S1H). The bHLH TFs RHD6 and RSL1 control root hair development in *A. thaliana* and rhizoid
506 formation in *P. patens*⁷⁰, while null mutants in *P. patens* and *M. polymorpha* lack rhizoids^{71,72}. The
507 Zygnematophyceae *Spirogloea muscicola* and *Mesotaenium endlicherianum*, the likely sister group
508 to land plants that share the same subaerial/terrestrial habitat with the earliest-diverging bryophytes⁷³,
509 lack rhizoids and RHD6 homologs. At least one homolog of these TFs (RHD6 clade) occurs in each
510 of the selected land plants, but none in *T. lepidozoides* (Figures 6B and S5A; Table S2N), suggesting
511 that the loss of rhizoids resulted from the loss of these genes. The *Takakia* sp. fossil lack rhizoids
512 (Data S2N), thus we suggest that *Takakia* lost rhizoids no later than the Middle-Late Jurassic after
513 their ancestor separated from the LCA of all mosses. We constructed a timetree using RHD6
514 homologs from 11 bryophyte species with two fossil calibrations and found that the RHD6 diversity
515 within mosses occurred 326.81—163.40 Mya (median 246.95 Mya) (Data S2R), the lower boundary
516 close to the date of fossil 165 Mya.

517

518 **Stomata**

519 Stomata are microscopic pores on plant surfaces that facilitate gas exchange, hydration and in *P.*
520 *patens* the release of spores⁷. They were one of the first developmental innovations of land plants,
521 at about 400 Mya^{2,74,75}. However, *T. lepidozoides* lacks stomata (Data S1T). PpSCRM1 and PpSMF1,
522 orthologs of bHLH TFs, regulate stomata development in *A. thaliana*, and control stomata formation

523 in *P. patens*, given that their individual knock-out resulted in the lack of stomata⁷. We identified two
524 SCRM-homologs but no SMF-homologs in *T. lepidozoides* (Figures 6C, S5B–S5D; Table S2O),
525 suggesting that the lack of stomata is due to a loss of SMF. The *Takakia* sp. fossil is uninformative
526 for stomata. To better understand the evolutionary history of SMF in mosses, we constructed a
527 timetree using SMF homologs from 11 bryophyte species with two fossil calibrations and found that
528 the SMF diversity within mosses occurred 455.07–278.06 Mya (median 367.43 Mya) (Data S2S),
529 the lower boundary earlier than the date of fossil 165 Mya.

530

531 **Tropisms**

532 Plant tropisms rely on auxin gradients^{76,77}. In contrast to vascular plants, auxin concentrations in *T.*
533 *lepidozoides* were highest at the top and lowest at the base of erect shoots, whereas the
534 concentration remained constant along the rhizomes (Data S2T). The *T. lepidozoides* genome
535 encodes three PIN proteins possibly related to auxin efflux (Data S2T), but only *TIPIN1*
536 (Tle1c08012.1) has the classical domain structure. Ectopic expression of these genes in *A. thaliana*
537 revealed that *TIPIN1* represses root hair growth greatly, *TIPIN2* (Tle4c04210.1) mildly, and *TIPIN3*
538 (Tle1c06076.1) not at all (Data S2T). AtPIN1 antibodies detected TIPIN proteins on plasma
539 membranes in cells from leaves and stems with no polar localization (Data S2U), which is different
540 from leaf cells of *P. patens*⁷⁸ and root cells of *A. thaliana*^{76,77}. The functional divergence of TIPIN1-3
541 occurred 134.59–101.08 Mya. The unusual nonpolar localization of TIPIN1 in *Takakia* may account
542 for its unusual weak negative geotropism, which facilitates the adaptive growth of populations on
543 rugged and steep surfaces in the absence of rhizoids.

544

545 **DISCUSSION**

546 **Evolution of specific morphological traits predated the uplift of the Himalayas**

547 *Takakia* has been an enigmatic genus, as it combines a mixture of morphological traits of algae,
548 liverworts and mosses. Our analysis of the *T. lepidozoides* genome combined with the fossil
549 evidence suggests that its extant shape evolved earlier than 165 Mya, after their ancestor separated
550 from the LCA of all other mosses, in which there were at least two periods of severe high
551 temperatures according to paleometeorological data⁷⁹. This predates the uplift of the Himalayas,
552 which started 65 Mya. Based on fossils and phylogenomics data, it is believed that extant vascular
553 plants and bryophytes originated from a more complex ancestral land plant during the Cambrian,

554 515–494 Mya⁸⁰. Thus, we suggest that adaptive evolution of *Takakia* included the loss of stomata,
555 rhizoids, and leaf polarity, rather than it representing an ancestral state.

556 The determination of leaf shape is controlled by complex molecular regulatory networks, which
557 are involved in leaf origination, polarity fixation and blade expansion^{81,82}. Our analysis of 11 bryophyte
558 genomes revealed that *T. lepidozoides* encodes gene homologs related to leaf origination and blade
559 expansion but lost the leaf polarity-determining TF KAN1 earlier than 165 Mya (Data S2O),
560 suggesting that its ancestor had dorsiventral leaves, and that loss of polarity may be a special
561 adaptive morphological trait. Likewise, all 11 bryophyte genomes, except *T. lepidozoides*, encode
562 RHD6, which controls rhizoids development in bryophytes⁷⁰, and diverged 326.81–163.40 Mya
563 (median 246.95 Mya) (Data S2R). Therefore, we suggest that rhizoid loss is a consequence of the
564 loss of RHD6 in the unique ancestor of *Takakia*. Fossils and phylogenomic data indicate that stomata
565 were present in the LCA of land plants, prior to the divergence of bryophytes and tracheophytes, and
566 subsequently lost multiple times during the diversification of bryophytes⁸. Our analysis of 11
567 bryophyte genomes reveals that all of them, except *T. lepidozoides*, had genetic traces related to
568 the TF SMF, which is associated with stomata development in embryophytes⁷, during
569 455.07–278.06 Mya (median 367.43 Mya) (Data S2S). Therefore, we suggest that the loss of
570 stomata resulted from the loss of SMF in *T. lepidozoides*, rather than this species never encoded
571 SMF.

572 Instead of rhizoids that attach plants to the soil matrix, *T. lepidozoides* evolved a morphology
573 unique among mosses. Plant tropism and branching are closely related to auxin signalling^{76,77}.
574 Aux/IAAs interact with ARFs leading to the repression of ARF-controlled genes, which are involved
575 in tropisms, braching and root formation⁸³⁻⁸⁵. The timing history of Aux/IAA expansion (218.62–
576 142.02 Mya), RHD6 loss (no later than 165 Mya) and PIN divergence (134.59–101.08 Mya) might
577 reveal the reasons why *Takakia* plants display a weak negative geotropism, continued branching
578 and randomness of branching positions. The expansion of Aux/IAA genes might have enhanced the
579 function of Aux/IAA, strengthening the inhibitory effect of ARF-regulated gene expression, which could
580 lead to weaker effects of auxin sigalling. The timing of RHD6 loss and Aux/IAA divergence may
581 overlap and follow with the loss of polar PIN localization. Thus, there is no IAA distribution gradient
582 (Data S2T) in *Takakia* in contrast to the situation in seed plants⁸⁶ and *P. patens*⁸⁷. If such a causal
583 relationship between the three gene families exists could be tested experimentally.

584

585 **The evolution of specific protective traits occurred during the uplift of the Himalayas**

586 While some morphological features characteristic of *Takakia* evolved before 65 Mya, the duplication
587 and divergence of genes with implications in stress protection occurred mainly during the uplift of
588 the Himalayas, namely between 50 Mya to present. This period is characterized by a sharp increase
589 in harmful UV-B radiation and a sharp drop in temperature, both caused by the uplift of the region⁸⁸.

590 The evolutionary deeply conserved RAD51 is a key factor in DNA repair and recombination in
591 prokaryotes and eukaryotes^{56,57}. Its enhanced expression under elevated UV-B levels results in
592 increasing frequencies of homologous recombination to maintain genome stability in *A. thaliana* and
593 potato plants³³. In *Takakia*, two of the five RAD51 genes arose from duplications about 20.78 Mya
594 when the experimental site was being uplifted to about 4,000 m elevation⁸⁸, suggesting that this
595 expansion enhances their protective effect on genome stability in responses to elevated UV-B
596 radiation.

597 UV-B radiation leads to pyrimidine dimers and (6-4) photoproducts, which can be excised and
598 repaired in nuclear DNA by a group of proteins including RAD2 and ERCC1⁵⁵. Compared with one
599 RAD2 and ERCC1 gene in all surveyed plants, two paralogs of each occurred in *Takakia* (Table S2K)
600 from duplication dated to about 203.17 Mya and 5.57 Mya, respectively. While the driving force for
601 RAD2 expansion remains enigmatic, the expansion of ERCC1 may have been driven by severe UV-
602 B radiation at high altitude. BRCA1 and BARD1 evolved from a common eukaryotic ancestor and
603 are required for efficient DNA repair in humans and plants⁵⁸. While all surveyed moss genomes lack
604 BRCA1 and BARD1, the *T. lepidozoides* genome encodes one BRCA1, suggesting that the common
605 ancestor of the other mosses lost BRCA1 and BARD1 after their separation from the lineage that
606 leads to *Takakia* (Data S2K).

607 One strategy of plants to resist UV-B damage is to produce abundant metabolites that effectively
608 absorb radiation or eliminate ROS, such as accumulate flavones to increase UV-tolerance⁸⁹, phenols
609 and unsaturated fatty acids to enhance resistance to UV-B⁹⁰. *T. lepidozoides* is rich in phenolic
610 compounds, including flavonoids (Figure 3J; Data S1Y), which are products involved in the
611 phenylpropanoid pathway⁵⁹. PAL is the first and key enzyme in phenylpropanoid biosynthesis and is
612 expressed in response to various stresses⁶⁰. *T. lepidozoides* encodes 20 PALs, more than all
613 surveyed plants, 15 of them expanded within 50 Mya and intronless PALs expanded 30 Mya,
614 indicating a recent expansion via segmental or tandem duplications (Data S2L) likely in response to
615 the changing environment during the uplifting of the Himalayas. Expression of these PALs increases

616 under UV-B stress and peaks during the recovery from this stress (Data S2L), resulting in the
617 production of several flavonoids (Table S1F).

618 The uplift of the Himalayas has caused the annual average temperature in the area of our
619 experimental site to drop from 27 °C to about 5 °C⁸⁸, forcing *Takakia* plants to withstand freezing.
620 CSD proteins are evolutionary deeply conserved in prokaryotes and eukaryotes⁴⁶. In *Arabidopsis*,
621 CSD proteins negatively regulate freezing tolerance as the null mutants showed higher freezing
622 tolerance than WT when cold-acclimated. The increase in freezing tolerance was associated with
623 up-regulation of CBF transcription factors and their downstream genes⁴⁷. We constructed a timetree
624 using CSD homologs from 11 bryophyte species with two fossil calibrations and found that the CSD
625 diversity within mosses occurred 463.43—338.08 Mya (median 400.37 Mya) (Data S2V), the lower
626 boundary earlier than the date of fossil 165 Mya. So we supposed the ancestor of *Takakia* may lost
627 CSD after their ancestor separated from other mosses LCA, suggesting that CSD loss may help
628 *Takakia* plants to withstand harsh conditions, such as being covered by snow for eight months at -
629 4 °C.

630

631 **Limitations of the study**

632 Our decade long survey of the experimental site at 4 km altitude, together with physiological, cell
633 biological, transcriptomic and genomic data comprehensively reveal the peculiar features of the so
634 far enigmatic genus *Takakia*. A combination of molecular clock analyses and paleometeorological
635 data provide evidence for molecular adaptations to the harsh environment during the uplift of the
636 region that started 65 Mya. Nevertheless, we can not explain the current population decline in
637 response to global warming. Remarkably, many of the morphological adaptations likely occurred
638 earlier than 165 Mya Therefore, similar studies should be performed with relictual *Takakia*
639 populations from different regions, such as Japan and North America. Such studies, together with
640 our planned sequencing of more *Takakia* genomes from our 52 experimental quadrants will provide
641 insights into genetic diversity and population dynamics. Subsequent selective-sweep and genome-
642 wide association studies may lead to a better understanding of the causes underlying the rapid
643 decline of this species in response to current global warming.

644

645 **Figure titles and legends**

646 **Figure 1. Geomorphologic and meteorological characteristics of the area in the Tibetan**
647 **Plateau harboring *T. lepidozoides* populations.**

648 (A) The geomorphologic view of the survey area: exact location of experimental plots (red dots) in
649 the area of ~18 km² near Gawalong East Glacier, Lat 29°45' N, Long 95°42' E, altitude 3,800–4,400
650 meters, at Bomi County, Tibet, China, yellow dotted line marks the treeline position in survey area.
651 (B) A population of *T. lepidozoides* growing on the surface of a cliff in area of Plot C (bar = 1 cm).
652 (C) *Takakia* plants with capsules (red arrow) in one population of Plot B (bar = 5 mm). (D)
653 Temperature change in the survey area from 1960 to 2021, with T° rising at a rate of 0.32°C per
654 decade. (E) Temperature change per year in the sample plot from 2010 to 2021 (mean ± s.d.; n = 6
655 investigated plots), the gray shadow area shows the period during which *Takakia* plants were
656 covered by snow. (F) The coverage of *T. lepidozoides* population decreased continually at a rate of
657 1.6% per year from 2010 to 2021. The vertical axis represents the relative coverage of *T.*
658 *lepidozoides* in observation quadrants (n = 52), and the dotted line marks the relationship between
659 the *T. lepidozoides* population size and year. The central mark of each box indicates the median,
660 and the bottom and top edges of the box indicate the 25th and 75th percentiles respectively. The
661 whiskers extend to the most extreme data points that are not considered as outliers.

662

663 **Figure 2. Branch origin and potential vessel elements connect in stems of *T. lepidozoides***
664 **and in roots of *A. thaliana*.**

665 (A) The dome originates from cells of the epidermis (yellow 8 and white 6) and sub-epidermis (yellow
666 7", 7', 7 and white 5) of parent stem and gradually elongates to form a branch, with potential vessel
667 elements connected vertically (blue arrow), bar = 50 μm. (B) The parent stem and the branch
668 comprise eight row cells, the vertical walls between epidermal cells (yellow arrows), the inclined walls
669 between potential vessel elements in stem (white arrows) and at branch point (blue arrows), bar =
670 50 μm. (C) Evans blue staining in the median longer cells in stem (arrow marks the branch point in
671 the box of G), bar = 50 μm. (D) The connection of vessel elements between main root (white arrow)
672 and lateral root (yellow arrows) in *Arabidopsis*, bar = 30 μm. (E) There is no direct connection of
673 vessel elements between the stem (white arrows) and branch (yellow arrows) in *Arabidopsis*,
674 between main stem (white arrows) and branch (yellow arrows) in *Arabidopsis*, bar = 500 μm. (F) The
675 net structure population, green parts can maintain photosynthesis for 3 ~ 5 years, gradually aging

676 and death (blue arrow), and eventually become humus (white arrow), bar = 0.5 cm. (G) One part of
677 an individual from F (bar = 2 mm), erect stems (red arrows), rhizomes (blue arrows).

678

679 **Figure 3. Physiological and biochemical properties in response to intense UV-B radiation in**
680 **the two mosses *T. lepidozoides* and *P. patens*.**

681 Shoot phenotypes of *T. lepidozoides* (A–C) and *P. patens* (D–F) under normal conditions (A, D) 144
682 kJ m⁻² UV-B treatment for 2 hours (B, E) and recovery (R) on the 3rd day (C, F) (bars = 1 mm). (G,
683 H) Unsaturated fatty acid compositions under UV-B treatment and at different points of recovery in
684 *T. lepidozoides* (G) and *P. patens* (H), red color represents MGDG, blue color represents PC, green
685 color represents neutral lipids, purple color represents DGTS. (I) Polyphenol contents under UV-B
686 treatment and recovery in *T. lepidozoides* (blue) and *P. patens* (red). (J) Flavonoid contents under
687 UV-B treatment and recovery in *T. lepidozoides* (blue) and *P. patens* (red). (K) Active unit of catalase
688 (CAT) under UV-B treatment and recovery in *T. lepidozoides* (blue) and *P. patens* (red). (L)
689 Concentrations of malondialdehyde (MDA) under UV-B treatment and recovery in *T. lepidozoides*
690 (blue) and *P. patens* (red). (M) The maximum quantum efficiency of PSII photochemistry (Fv/Fm)
691 under UV-B treatment and recovery in *T. lepidozoides* (blue) and *P. patens* (red). (N) The actual
692 photochemical quantum efficiency (Fv/Fo) under UV-B treatment and recovery in *T. lepidozoides*
693 (blue) and *P. patens* (red). Western blot of photosystem components in *T. lepidozoides* (O) and *P.*
694 *patens* (P) under normal condition, UV-B treatment, and recovery.

695

696 **Figure 4. Organization and characteristics of the *T. lepidozoides* genome.**

697 (A) Circos plot of the *T. lepidozoides* genome, extension from outer to inner showing the
698 chromosomes with pentatricopeptide repeat regions (black band), GC contents in every 100 kb
699 sliding window (purple histogram), transposable elements densities in every 100 kb sliding window
700 (coral histogram), gene densities in every 100 kb sliding window (blue histogram), differential gene
701 expression in UV radiations (solid line in red represents genes up-regulated, solid line in green
702 represents genes down-regulated) and chromosomal syntenic relationships (solid lines with colors
703 represent inter-chromosome collinearity, solid line in grey represents intra-chromosome collinearity).
704 (B) WGD and divergence of *T. lepidozoides* indicated by Ks distribution. (C) The Ka/Ks ratio
705 distributions of gene pairs derived from different modes of duplication. (D) *T. lepidozoides* time tree,
706 numbers on the nodes represent divergence time, light purple bars show 95% confidence interval

707 length, nodes with black dots show fossil calibration points, the ordinal phylogeny structure with
708 branch length shown on the top left, full phylogeny tree with branch length shown in Data S2F.

709

710 **Figure 5. PPR features in *T. lepidozooides*.**

711 (A) Classification of all PPRs in eight plant genomes based on structural motifs. (B) Word cloud of
712 PPRs subcellular localization prediction shows that most of PPR proteins are targeted to
713 mitochondria or chloroplasts. (C) RNA editing as confirmed by 30x sequence analysis, PPR proteins
714 influence expression of mitochondria or chloroplasts transcripts by altering RNA sequence in *T.*
715 *lepidozooides*. Most abundant RNA edit change is C to U nucleotides. (D) Expression levels of PPRs
716 show significant differences in response to UV-B radiation and during recovery. (E) Gene ontology
717 (GO) enrichment word cloud (biological process) of PPRs in response to UV-B radiation and during
718 recovery. (F) GO enrichment word cloud (cellular component) of PPRs in response to UV-B radiation
719 and during recovery. (G) GO enrichment word cloud (molecular function) of PPRs in response to UV-
720 B radiation and during recovery.

721

722 **Figure 6. Genetic foundation of morphological features in *T. lepidozooides* versus in *P. patens*
723 and *A. thaliana*.**

724 (A) Proteins regulating leaf (sporophytic) and phylloid (gametophytic) development. *T. lepidozooides*
725 has terete leaves (phylloids) with no adaxial-abaxial polarity, probably due to the absence of the
726 transcription factor KAN. Dashed lines with gene names mark the genes absent in the pathway,
727 arrows indicate positive/promoting regulation, lines with perpendicular end bars indicate
728 negative/repressing regulation, and crosses indicate that the pathway is blocked or absent. (B)
729 Proteins regulating root hair (sporophytic) and rhizoid (gametophytic) development. *T. lepidozooides*
730 has no rhizoids, probably due to the lack of the transcription factor RHD6. (C) Proteins regulating
731 stomatal development. *T. lepidozooides* lacks stomata, probably due to the missing transcription
732 factor SMF. In comparison to *A. thaliana*, *P. patens* lacks the transcription factor SCAP, probably
733 resulting in less complex stomata. In both plants, stomata are restricted to the sporophyte; leaves in
734 *A. thaliana* and spore capsules in *P. patens*.

735

736 **STAR Methods**

737 **RESOURCE AVAILABILITY**

738 **Lead contact**

739 Further information and requests for resources and reagents should be directed to and will be fulfilled
740 by the lead contact, Yikun He (yhe@cnu.edu.cn).

741

742 **Materials availability**

743 This study did not generate new unique reagents. Clones of the *Takakia lepidozooides* plant that was
744 sequenced are deposited at the International Moss Stock Center IMSC (<https://www.moss-stock-center.org/>) under accession number 41271.

746

747 **Data and code availability**

748 Whole-genome assembly and raw reads of *T. lepidozooides* have been deposited in the National
749 Center for Biotechnology Information (NCBI, <https://www.ncbi.nlm.nih.gov/>) under BioProject
750 accession number PRJNA796444, the pseudomolecules genome with GenBank accession number
751 JAKLZA000000000, the raw reads of DNA and RNA sequencing with Sequence Read Archive
752 accession number SRP354614 (Illumina PCR-free reads SRR17573914—SRR17573930 and
753 PacBio reads SRR17573931—SRR17573932). The genome assembly and raw data was also
754 deposited in National Genomics Data Center (NGDC, <https://ngdc.cncb.ac.cn/?lang=en>) under
755 BioProject accession number PRJCA014819, the genome with Genome Warehouse accession
756 number GWHBRAK000000000 and the raw sequences data with Genome Sequence Archive
757 accession number CRA009836. Metabolomics raw data of *T. lepidozooides* have been deposited in
758 MetaboLights (<https://www.ebi.ac.uk/metabolights/>) under accession number MTBLS6999. The
759 *Takakia* genome assembly sequences and gene annotations are available at
760 <https://www.takakia.com>. This study does not report original code. All other data and materials are
761 available from the corresponding authors upon reasonable request.

762

763 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

764 Natural samples of *T. lepidozooides* with mature capsules was collected from a shady, moist rock on
765 Galongla Mountain, Bomi County, Tibet, China (29°45' N, 95°42' E) in October 2005 by Xuedong Li
766 and Meizhi Wang. The *T. lepidozooides* tissue culture plantlets were grown in the greenhouses at

767 Beijing Key Laboratory of Plant Gene Resources and Biotechnology for Carbon Reduction and
768 Environmental Improvement, College of Life Sciences, Capital Normal University, Beijing, China
769 (39°56' N, 116°18' E) since 2005. The field study was carried out in Bomi County, Tibet, China (29°45'
770 N, 95°42' E), altitude from 3,800 to 4,400 m since 2005.

771

772 **METHOD DETAILS**

773 **Field study**

774 The study site is located in Bomi County, Tibet, China (29°45' N, 95°42' E), altitude from 3,800 to
775 4,400 m, each plot is 50 m × 50 m with 100 m interval of altitude. At each plot, three east-west
776 transects of 10 m × 50 m with 10 m interval and three altitudinal gradient transects of 10 m × 50 m
777 with 10 m interval were established. All 10 cm × 10 cm observation quadrants of *Takakia*
778 *lepidozoides* were established randomly in the transect. The growing season in the site runs from
779 late June to middle of October, quadrant records were collected twice at the beginning and end of
780 the growing season each year. Real-time image recorders for quadrant and glacier were collected
781 by outdoor wildlife cameras (TimelapseCams, Wingscapes). The meteorological data and sun
782 radiation data were collected real-time by a microclimate station (MAWS830-FC, Hua Yun Instrument;
783 L99, Hangzhou Luge Ltd.) and a radiant heat balance station (HYZY-SP, Hua Yun Instrument).
784 Remote sensing observation data of Resources satellite three (Resources-3) and Gaofen-1 satellite
785 (Gaofen-1) were collected from the National Geographical Information Monitoring Cloud Platform
786 (<http://www.dsac.cn/>). The glacier ablation was calculated as

$$787 |distance| = \frac{\sqrt{(longitude_m - longitude_n)^2 + (latitude_m - latitude_n)^2}}{day_m - day_n} \times 365.$$

788 The Shannon-Wiener index was calculated as $H' = \sum(P_i) (\ln P_i)$, where P_i represented the
789 coverage of species i in an observation quadrant. Spearman's rank correlation was used to analyze
790 possible correlations among species, redundancy and canonical correspondence analysis to
791 determine the effects of climatic variables on the population changes, where only commonly found
792 species were considered. All multiple regressions and backward eliminations were performed by
793 vegan package (version 2.5-7) (<https://CRAN.R-project.org/package=vegan>) in R. The potential
794 distribution of *T. lepidozoides* was predicted using the ecological niche modeling software MaxEnt
795 (version 3.4.1) (https://biodiversityinformatics.amnh.org/open_source/maxent/) combined with data
796 of the global climate from WorldClim (<https://www.worldclim.org/>). The 'Random test percentage' was
797 set as 25, and 'Make pictures of predictions' and 'Do jackknife to measure variable importance' were

798 chosen, the remaining model values were set to default values in the initial model; 'Random seed'
799 was chosen, and 100 replicate models were run in the final model.

800 Soil samples were collected from four experimental plots; a 5-point cross sampling protocol was
801 carried out in each plot. In brief, each plot was divided by a rectangle grid with a spacing distance of
802 1 m, samples were retrieved from the top 5-cm surface soil in the four corner grids and the middle
803 grid, then mixed by plot to measure the soil elemental compositions. Soil elemental compositions
804 were analyzed using Inductively Coupled Plasma–Mass Spectrometry (iCAP TQ ICP–MS, Thermo
805 Scientific), Inductively Coupled Plasma–Optical Emission Spectrometry (iCAP 7400 ICP–OES,
806 Thermo Scientific) and X-ray fluorescence (ARL PERFORM'X, Thermo Scientific) to determine the
807 content of macro- and micronutrients as well as toxic elements^{91,92}.

808

809 **Plant materials**

810 Samples of *T. lepidozoides* with mature capsules were collected from a shady, moist rock on
811 Galongla Mountain, Tibet, China. Healthy capsules were selected and surface-sterilized as follows:
812 (1) five 3-min rinses with sterilized water, (2) five rounds of sterilization for 5 s with 75 % ethanol and
813 washing for 1 min with sterilized water, (3) sterilization with 0.05 % HgCl₂ solution for 2 min and (4)
814 five 1-min washes with sterilized water. The capsules were transferred into sterilized water to prepare
815 a spore suspension, which was inoculated onto the surface of Beneke's medium with some
816 modification (200 mg L⁻¹ NH₄NO₃, 100 mg L⁻¹ KH₂PO₄, 100 mg L⁻¹ MgSO₄·7H₂O, 100 mg L⁻¹
817 CaCl₂·2H₂O, 15 mg L⁻¹ FeSO₄·7H₂O) containing 0.5 % (w/v) glucose (pH 5.8), and were cultivated
818 under the following conditions: 14/10 °C day/night temperature, a 12-h/12-h light - dark photoperiod,
819 illumination at 100 μmol photons m⁻² s⁻¹, and approximately 85 % relative humidity (RH). Plants were
820 transferred to the same medium and incubated under the same light and temperature conditions
821 mentioned above every 2 weeks. The *Physcomitrella patens* wild-type "Gransden 2004" was grown
822 on BCD medium (1.8 mM KH₂PO₄, 1 mM MgSO₄·7H₂O, 45 μM FeSO₄·6H₂O, 10 mM KNO₃, 0.22 μM
823 CuSO₄·5H₂O, 0.19 μM ZnSO₄·7H₂O, 10 μM H₃BO₃, 0.1 μM Na₂MoO₄·2H₂O, 2 μM MnCl₂·4H₂O, 0.23
824 μM CoCl₂·6H₂O, 0.17 μM KI) containing 0.5 % (w/v) glucose (pH 6.5) with a light intensity of 100
825 μmol m² s⁻¹ delivered in a 16-h/8-h light/dark cycle at 25 °C. Plants were transferred to the same
826 medium and incubated under the same light and temperature conditions every 2 weeks, and the 2-
827 weeks-old gametophytes were used as the material for future experiments, unless specifically
828 mentioned. Seeds of *Arabidopsis thaliana* wild-type accession Col-0 lines were germinated on MS

829 medium (Sigma-Aldrich, M2909) containing 2 % (w/v) glucose (pH 6.0), and the plants were
830 subsequently transferred to soil and grown in a growth chamber at 22 °C with 100 $\mu\text{mol photons m}^{-2}$
831 s^{-1} , a photoperiod of 16-h/8-h light/dark, and RH of 75%.

832

833 **UV-B treatment**

834 For UV-B irradiation the following plants were used: *A. thaliana* Col-0 25 days after germination in
835 normal condition, plants of *T. lepidozoides* and *P. patens* grown for 24 h under 50 $\mu\text{mol photons m}^{-2}$
836 s^{-1} with the equivalent temperature and RH described above before starting the treatment. To all
837 plants 20 $\text{W m}^{-2} \text{s}^{-1}$ UV-B was provided (2 x 312 nm XX-15B UV lamp, Spectronics) for 2 h. After UV-
838 B irradiation, controls and UV-B-exposed samples were collected immediately, then plants were
839 transferred to the normal condition, recover samples were collected at 6, 12, 24 hours and 3, 5, 7
840 days after UV-B treatment.

841

842 **Cytological observations**

843 Samples for ultrastructural observation of cells were carried out according to Xiao et al.⁹³ with some
844 modifications. Control and UV-B treated plants were fixed with 2 % glutaraldehyde in 0.1 M
845 Sorensen's phosphate buffer for 24 h, dried under vacuum for 2 h, post-fixed in 1 % (w/v) osmium
846 tetroxide for 2 h, and then dehydrated for 30 min in an ethanol to acetone gradient. The samples
847 were then divided into two portions. One portion was embedded overnight in Spurr's low-viscosity
848 resin. The embedded samples were dried at 40 °C for 2 days and then at 70 °C for 24 h. Sections
849 were prepared with an ultramicrotome (EM UC6, Leica), stained with 1 % aqueous uranyl acetate
850 and 0.1 % aqueous lead citrate for 30 min each, and examined with a transmission electron
851 microscope (H7500, Hitachi). The other sample portion was air-dried overnight in a laminar flow
852 hood, sputter-coated with gold-palladium for 90 s at 2.2 kV and examined with a scanning electron
853 microscope (S4800, Hitachi).

854 The chromosome preparation was conducted according to a previous study⁹⁴. Shoot tips of
855 vigorously growing plants were harvested and fixed with Cannoy Solution (ethanol:acetic acid = 3:1)
856 overnight. After washing with distilled water, the shoots were digested in a mixture of 2% cellulose
857 (Sigma-Aldrich) and 1% pectinase (Sigma-Aldrich) at 37 °C for 90 min. After that, they were quickly
858 rinsed in distilled water and fixed with the same fixation solution. The samples were spread on
859 precooled slides and flame dried. Chromosomes were counterstained with 4',6-diamidino-2-

860 phenylindole (DAPI, Sigma-Aldrich) in an anti-fade solution (Vector Laboratories, H-1200-10).
861 Images were captured under a Zeiss fluorescence microscope with a micro-charge-coupled camera
862 (Imager A2, Zeiss).

863

864 **Immunoblot of photosynthetic proteins**

865 Thylakoid membranes proteins were extracted as described^{95,96}. Plants leaves were collected and
866 ground under liquid nitrogen, then homogenized in an ice-cold isolation buffer containing 125 mM
867 Tris-HCl pH 8.8, 1% (w/v) SDS, 10% (v/v) glycerol, 50 mM Na₂S₂O₅ with a chilled mortar and pestle
868 and filtrated through two layers of cheesecloth. The filtrate was centrifuged at 5,000g for 10 min. The
869 thylakoid pellets were washed with isolation buffer and recentrifuged. The thylakoid membranes
870 were washed in 330 mM sorbitol and 50 mM BisTris-HCl, pH 7.0, and suspended in resuspension
871 buffer (20% glycerol and 25 mM BisTris-HCl, pH 7.0) at 1.0 mg chlorophyll/mL. An equal volume of
872 resuspension buffer containing 2% (w/v) DM was added to the thylakoid suspension in a dropwise
873 manner. After incubation at 4°C for 5 min, insoluble material was removed by centrifugation at
874 12,000g for 10 min. The supernatant was combined with one-tenth volume of 5% Serva blue G in
875 100 mM BisTris-HCl, pH 7.0, 0.5 M 6-amino-n-caproic acid, and 30% (w/v) glycerol and applied to
876 0.75-mm-thick 6 to 12% acrylamide gradient gels in a Hoefer Mighty Small vertical electrophoresis
877 unit connected to a cooling circulator. For two-dimensional analysis, excised BN-PAGE lanes were
878 soaked in SDS sample buffer and 5% β-mercaptoethanol for 30 min and layered onto 1-mm-thick
879 15% SDS polyacrylamide gels containing 6 M urea. After SDS-PAGE, proteins were transferred to
880 nitrocellulose membranes (Amersham Biosciences). The membranes were incubated with specific
881 antibodies and visualized using the enhanced chemiluminescence method. The antibodies of
882 AtCP47 and AtPsaD were from Liu et al.⁹⁷.

883

884 **Immunofluorescence labeling**

885 *T. lepidozoides* plants were fixed in FAA (3% w/v formaldehyde, 5% v/v acetic acid, 50% v/v ethanol)
886 at room temperature overnight, embedded in STEEDMAN's wax after dehydration, and sectioned
887 as described⁹⁸. Rehydrated sections were then subjected to an antigen retrieval procedure⁹⁹. Slides
888 were first incubated in blocking buffer (2% nonfat milk in PBST) for 1 h, then with anti-AtPIN1
889 antibody (PhytoAB, PHY1046S) at a dilution of 1:100 in blocking buffer with 2% normal goat serum
890 for 5 h. Slides were washed with PBST buffer three times, each for 10 min, and then incubated for 2

891 h with an Alexa-fluor-488-conjugated anti-rabbit secondary antibody at a dilution of 1:200 (Invitrogen).
892 After washing, the slides were mounted using ProLong Diamond Antifade solution (Life
893 Technologies). A Leica DM5500 multifunctional microscope equipped with an Andor Zyla
894 monochrome CCD and filter cubes for GFP and chlorophyll were used for fluorescence micrograph
895 capture.

896

897 **Transgene constructs**

898 The *Arabidopsis EXPA7* (AT1G12560) promoter¹⁰⁰ and the 35S promoter were used to drive all PIN
899 related transgenes. The binary vector *pCAMBIA1300-GFP-NOS* was used as the cloning vector.
900 The *ProEXPA7:GFP-NOS* vector was generated by replacing the 35S promoter region with the
901 genomic *EXPA7* promoter fragment by PCR from genomic DNA of *Arabidopsis* in the binary vector
902 *pCAMBIA1300-GFP-NOS* using the primer sets in Table S2Q. The *Pro35S:AtPIN2* (AT5G57090),
903 *ProEXPA7:AtPIN2*, *PpPINB* (Pp3c24_2970V3.1), *TIPIN1* (Tle1c08012.1), *TIPIN2* (Tle4c04210.1),
904 and *TIPIN3* (Tle1c06076.1)-*GFP* constructs were generated by connecting the *Pro35S* and
905 *ProEXPA7* promoter to *PIN* coding regions by PCR from cDNA of corresponding species in the binary
906 vector *pCAMBIA1300-GFP-NOS*. Plants were transformed with the floral dip method¹⁰¹. Root hairs
907 of transgenic plants were observed as described¹⁰². For estimation of root hair length, digital
908 photographs of roots were taken under a stereomicroscope (Zeiss, AXIO Zoom. V16) at 40x to 50x
909 magnifications. The hairs in the hair maturation region were observed.

910

911 **Observation of branching origin**

912 To analyze the characteristics of the origin of the branches (stem and root) and the connection of the
913 conducting elements at the cross-point between stems, the mPS-PI staining technique¹⁰³, in which
914 polysaccharides (especially cell walls) were specifically labeled with propidium iodide, with some
915 modifications was used to reconstruct 3D images of plant organs. In brief, *T. lepidozoides* and *A.*
916 *thaliana* organs were obtained and fixed in fixative (50% methanol and 10% acetic acid) at 4 °C
917 overnight. After washing in ddH₂O for three times, samples were incubated in 1% periodic acid at
918 room temperature for 40 min. Next, samples were rinsed with ddH₂O and incubated in Schiff's
919 reagent (100 mM sodium metabisulphite and 0.15 N HCl) containing 100 mg/mL propidium iodide
920 for 1 to 2 h. Finally, samples were transferred to glass bottom dishes and cleaned with a chloral
921 hydrate solution (4 g chloral hydrate, 1 mL glycerol, and 2 mL water) for additional 1–2 d.

922 Fluorescence images were obtained using a Leica SP8 laser scanning confocal microscope, with
923 the excitation wavelength at 561 nm and emission wavelength from 570 nm to 670 nm.

924

925 **Assay of macromolecule transport**

926 The Evans blue staining technique¹⁰⁴ with some modifications was used to monitor how
927 macromolecules are transported in stems of *T. lepidozooides*. Plants from Tibet were collected and
928 incubated in Evans blue staining buffer (3% w/v Evans Blue, 50 mM NaH₂PO₄) at room temperature
929 for 3 min, followed by washing four times with deionized water. Digital photographs of stems were
930 taken under a stereomicroscope (BA600-4, Motic).

931

932 **Measurement of free IAA**

933 IAA measurements were carried out according to Wang et al.¹⁰⁵ with some modifications. In brief,
934 the tip, middle and base fragments were isolated separately from erects and rhizomes of *T.*
935 *lepidozooides* plants under normal condition for 3 months subculture and collected, 200 mg (fresh
936 weight) per sample were weighted, homogenized under liquid nitrogen, and extracted at -20 °C for
937 24 h with 2 mL of cold methanol containing antioxidant and ²H₂-IAA (internal standard, CDN Isotopes).
938 Purification was performed with Oasis MAX solid-phase extract cartridge (150 mg/6 cc; Waters).
939 Measurement of free IAA was carried out using a UPLC-MS/MS system consisting of a UPLC system
940 (ACQUITY UPLC; Waters) and a triple quadruple tandem mass spectrometer (Quattro Premier XE;
941 Waters).

942

943 **Physiological experiments**

944 To estimate physiological changes under UV-B radiation, control and UV-B treated plants were
945 collected and the contents of polyphenol (A143-1-1, plant total phenol test kit), catalase (A007-1-1,
946 catalase assay kit), flavonoid (A142-1-1, plant flavonoids test kit) and malondialdehyde (A003-3-1,
947 plant malondialdehyde assay kit) were measured using detection kits from the Nanjing Jiancheng
948 Bioengineering Institute. Chlorophyll fluorescence was measured using a chlorophyll fluorometer
949 (IMAGING-PAM M-series, Heinz Walz).

950 Two-dimension thin layer chromatography (2D-TLC) and gas chromatography/mass
951 spectrometry (GC-MS) were employed to detect and identify total fatty acids¹⁰⁶. In brief, plants were
952 grinded in chlorohydrocarbon and a methyl alcohol mixture (1:2) at 4 °C, then mixed with 1M KCl to

953 a final volume ratio at 1:1:0.9, centrifuged for 10 min at 2,500 g. Subsequently, the organic phase
954 was removed with N₂. 20 × 20 cm × 250 μm glass Si gel plates (Silica Gel GF Uniplates, Sigma)
955 were activated for 1 h at 100 °C and stored in a desiccator to increase reproducibility. Crude extract
956 material of lipids applied to form a single small spot (2–4 mm diameter) in a corner 1 cm from either
957 of two sides of a Si gel plate, then TLC plates were placed in a solvent chamber (CHCl₃:CH₃OH:H₂O,
958 65:25:4, v/v/v), and the extract spot was allowed to develop along the left-hand side of the plate.
959 TLC plates were then removed from the solvent chamber immediately before the solvent front
960 reached the top of the plates. These plates were allowed to air-dry and subsequently placed in
961 another solvent chamber (CHCl₃:CH₃OH:NH₃·H₂O, 65:35:5, v/v/v) so that the developed lane of
962 extract was placed horizontally just above the solvent. Each plate was allowed to develop until the
963 solvent front reached 1 cm from the top of the plate and then removed from the chamber and air-
964 dried. Sample spots were collected and mixed with 10 μg N-heptadecanoic acid, 5 μg butylated
965 hydroxytoluene, 1.5 mL sulfuric acid: methyl ester (5:95), 85 °C for 1 h and then with 1 mL H₂O, 2
966 mL N-pentane, centrifuged for 10 min at 3,000 g. Subsequently, the water phase was removed. GC-
967 MS (Trace GC 2000-Trace DSQ, Thermo Finnigan) was performed using a quartz capillary column.
968 The heating program used an initial temperature of 50 °C, increasing at a rate of 10 °C/min to 150 °C
969 and maintain 2 min, then increasing to 280 °C at a rate of 3 °C/min and maintain 10 min, then
970 maintained at 250 °C until the running was completed. Individual compounds were identified by
971 comparing the MS fragmentation patterns with those provided by the National Institute of Standards
972 and Technology (<https://chemdata.nist.gov/>).

973 To classify the oil body components, the oil bodies were separated according to Zhang et al.
974 (2010). In brief, plants were collected and washed three times with 50 ml 0.001% PBS Triton-X100;
975 then washed with 50 ml buffer A (25 mm Tricine, pH 7.6, 250 mm sucrose, and 0.2 mm
976 phenylmethylsulfonyl fluoride); homogenized using a polytron (Cole-Parmer® Labgen™ 125 and
977 700 Tissue Homogenizers) in 10 ml buffer A, and centrifuged at 1,000 × g for 30 s; supernatant was
978 homogenized again by nitrogen cavitation (Ashcroft Duralife Pressure Gauge) at 500 psi for 15 min
979 on ice, and then centrifuged at 1,000 × g for 10 min; nine milliliters of this postnuclear supernatant
980 was loaded into an SW40 tube, and 3 ml buffer B (20 mm HEPES, pH 7.4, 100 mm KCl, and 2 mm
981 MgCl₂) was overlaid on top. The tube was then centrifuged at 12,628 × g for 1 h at 4 °C. The oil body
982 fraction was carefully collected from the top layer and washed three times with 200 μl buffer B. The
983 oil body components were identified using 2D-TLC.

984

985 **Non-targeted metabolome profile**

986 Plants were collected and milled with 80% methyl alcohol and dried with N₂. 25 mg powder per
987 sample were solved with 50% acetonitrile. Chromatographic separation was performed on a Thermo
988 Ultimate 3000 (Thermo Fisher Scientific) with 100 mm x 2.1 mm x 1.8 μm C18 column (ZORBAX
989 SB-Aq, Agilent Technology). The chromatographic conditions were set as follows: flow rate of 0.3
990 mL/min, column temperature of 40 °C, and injection volume of 4 μL. The mobile phase comprised
991 water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B), by using the following
992 solvent gradient system: 0.0–6.0 min, 5–20% B; 6.0–10.0 min, 20–95% B; 10.0–11.0 min, 95–95%
993 B; 11.0–11.5 min, 5% B; 11.5–14.5 min, 5% B.

994 All samples were run on a Q Exactive Orbitrap-mass spectrometer (Thermo Fisher Scientific)
995 equipped with a heated electrospray ionization (ESI) interface (HESI-II). The QC sample was
996 injected every six samples. Data with a mass range of m/z 70–1000 were acquired at positive and
997 negative switching modes with data dependent MS/MS. The full scan and fragment spectra
998 resolution were 70,000 FWHM and 17,500 FWHM, respectively. The ion source parameters were as
999 follows: capillary temperature of 320 °C, vaporizer temperature of 350 °C, spray voltage of
1000 3500/3000 V(+/-), sheath gas flow rate of 40 Arb, and the auxiliary gas flow rate of 5 Arb. The
1001 acquisition mode of stepped NCE (normalized collision energy) was set to 15, 40 and 60, respectively.
1002 Raw LC-MS data were processed by Compound Discoverer 3.2 (Thermo Fisher Scientific).
1003 Compounds were identified by comparing the MS/MS spectra with those in mzCloud
1004 (<https://www.mzcloud.org>) and ChemSpider (<http://www.chemspider.com/>).

1005

1006 **Genome and transcriptome sequencing**

1007 Total DNA for *de novo* genome sequencing was extracted from young gametophytes produced from
1008 a single spore. For transcriptome sequencing we sampled tissues from *T. lepidozoides* including
1009 gametophytes in different growth stages and under different growth conditions, including
1010 protonemata, phylloids (leaves) at 7, 14, 21, 28, 35 and 42 days as well as phylloids under 144 kJ
1011 m⁻², 288 kJ m⁻², 432 kJ m⁻², 576 kJ m⁻², 720 kJ m⁻² UV-B radiations and recovery in normal conditions
1012 for 6 h.

1013 For PacBio sequencing, we prepared approximately 20-kb SMRT cell libraries. A total of 25.9
1014 Gb data was sequenced on the PacBio system. All transcriptome libraries were sequenced using

1015 the Illumina platform. For the Hi-C sequencing and scaffolding, a Hi-C library was created from young
1016 gametophytes of *T. lepidozoides* under normal conditions. In brief, the gametophytes were fixed with
1017 formaldehyde and lysed, and the cross-linked DNA was then digested with Mbol overnight. Sticky
1018 ends were biotinylated and proximity-ligated to form chimeric junctions. Chimeric fragments
1019 representing the original cross-linked long-distance physical interactions were then processed into
1020 paired-end sequencing libraries and 150-bp paired-end reads, which were sequenced on the Illumina
1021 platform.

1022

1023 **Genome assembly**

1024 The contig-level assembly was performed on full PacBio long reads using the Canu package¹⁰⁷
1025 (version 2.0) (<https://github.com/marbl/canu>). We then polished the draft assembly using Arrow
1026 (version 2.3.1) (<https://github.com/PacificBiosciences/GenomicConsensus>). To increase the
1027 accuracy of the assembly, Illumina short reads were recruited for further polishing with the Pilon
1028 program¹⁰⁸ (version 1.22) (<https://github.com/broadinstitute/pilon>). The paired-end reads from Hi-C
1029 were uniquely mapped onto the draft assembly contigs. We used the 3D-DNA pipeline¹⁰⁹ (version
1030 190716) (<https://github.com/aidenlab/3d-dna>) and Juicer¹¹⁰ (version 1.11.08)
1031 (<https://github.com/aidenlab/juicer>) to group them into scaffolds and chromosomes. Additionally, the
1032 10-Kb and 20-Kb mate pair reads were used to connect sequence contigs into scaffolds using
1033 SSPACE¹¹¹ (version 2.1.1) (https://github.com/nsoranzo/sspace_basic) with default parameters. To
1034 minimize errors introduced in the scaffolding step, we only connected the sequence contigs
1035 supported by linkage evidence from mate pair reads and HiC data. The SMRT reads were also
1036 applied to fill the gap in scaffolds using PBJelly¹¹² (version 15.8.24)
1037 (<https://sourceforge.net/projects/pb-jelly/>) with default parameters to generate longer sequence
1038 contigs. The scaffolds clustering into 5 groups initially. As cytological evidence strongly supports 4
1039 chromosomes in *T. lepidozoides*, 2 shortest pseudochromosomes were linked together after error
1040 correction and manual inspection finally. The genome assembly quality was measured using
1041 BUSCO¹¹³ (version 4.0.6) (<https://busco.ezlab.org/>).

1042

1043 **Analysis of repetitive elements**

1044 Repetitive sequences and transposable elements were identified using a combination of *ab initio* and
1045 homology-based methods. In brief, an *ab initio* repeat library for *T. lepidozoides* was predicted using

1046 LTR_FINDER¹¹⁴ (version 1.07) (https://github.com/xzhub/LTR_Finder), RepeatModeler (version
1047 1.0.11) (<http://www.repeatmasker.org/RepeatModeler/>) with default parameters. The library was
1048 aligned to PGSB Repeat Element Database¹¹⁵ (<http://pgsb.helmholtz-muenchen.de/plant/recat/>) to
1049 classify the type of each repeat family. For identification of the repeats throughout the genome,
1050 RepeatMasker (version 4.0.8) (<http://www.repeatmasker.org/>) was applied with both the *ab initio*
1051 repeat databases and Repbase (<http://www.girinst.org/repbase>) using the WU-BLAST¹¹⁶
1052 (<http://blast.wustl.edu/>) search engine. Overlapping transposable elements belonging to the same
1053 repeat class were collated and combined. In addition, we annotated the tandem repeats using
1054 Tandem Repeats Finder¹¹⁷ (version 4.04) (<https://tandem.bu.edu/trf/trf.html>). Solo-LTR were
1055 identified using LTRharvest¹¹⁸ (version 1.6.1) (http://genometools.org/tools/gt_ltrharvest.html). A
1056 customized script was implemented to identify intact LTR/Copia and LTR/Gypsy retrotransposons.
1057 ClustalW¹¹⁹ (version 1.81) (<http://www.clustal.org/>) was applied to align 5' and 3' solo-LTRs to intact
1058 LTR elements. The evolutionary distance of the two LTR sequences was estimated using the Kimura
1059 two-parameter method embedded in BASEML in PAML¹²⁰ (version 4.9j)
1060 (<http://abacus.gene.ucl.ac.uk/software/paml.html>). A substitution rate of 1.3×10^{-8} mutations per site
1061 per year was used to convert evolutionary distance between 5' and 3' solo-LTRs to insertion age of
1062 retrotransposons.

1063

1064 **Gene prediction and functional annotation**

1065 Genscan¹²¹ (version 1.0) (<http://genes.mit.edu/GENSCAN.html>) and Augustus¹²² (version 3.3.3)
1066 (<http://augustus.gobics.de/>) were used to carry out *de novo* predictions with gene model parameters
1067 trained from *Chlamydomonas reinhardtii*, *Marchantia polymorpha*, *P. patens*, *Selaginella*
1068 *moellendorffii* and *A. thaliana*. Furthermore, gene models were *de novo* predicted using MAKER¹²³
1069 (version 3.01.02) (<http://www.yandell-lab.org/software/maker.html>). We then evaluated the genes by
1070 comparing MAKER results with the corresponding transcript evidence to select gene models that
1071 were the most consistent. The gene prediction quality was measured using BUSCO¹¹³ (version 4.0.6)
1072 (<https://busco.ezlab.org/>). Protein-coding genes were retrieved and functionally annotated by blast
1073 searches against databases including UniProtKB/Swiss-Prot and UniProtKB/TrEMBL¹²⁴
1074 (<https://www.uniprot.org/>). They were also subjected to Gene Ontology (GO) annotation and protein
1075 family annotation by the InterProScan pipeline¹²⁵ (version 5.44-79.0) ([https://github.com/ebi-pf-](https://github.com/ebi-pf-team/interproscan)
1076 [team/interproscan](https://github.com/ebi-pf-team/interproscan)). GO terms for each gene were assigned by blast searches against the GO

1077 database¹²⁶ (<http://geneontology.org/>).

1078

1079 **Organellar genomes**

1080 The organellar raw reads of *T. lepidozooides* were identified using BLASTn against the PacBio
1081 subreads with plastome (NC_005087.1) and chondriome (NC_007945.1) of *P. patens* as references,
1082 respectively. The organellar assembly were performed on raw reads using the Canu package¹⁰⁷. The
1083 organellar genome sequences were annotated as previously described¹²⁷. Protein-coding genes,
1084 rRNA genes, exact gene and exon/intron boundaries were annotated using GeSeq¹²⁸
1085 (<https://chlorobox.mpimp-golm.mpg.de/geseq.html>).

1086

1087 **Identification of genome synteny**

1088 We identified homologs using BLAST with default parameters, and scanned syntenic blocks
1089 consisting of homologous genes within the *T. lepidozooides* genome using MCSan¹²⁹ (version 1.0.6)
1090 (<https://github.com/tanghaibao/jcvi>).

1091 We identified WGD events in *T. lepidozooides* genome and two bryophytes (*M. polymorpha* and
1092 *P. patens*) using the WGDI toolkit¹³⁰ (version 0.5.3) (<https://github.com/SunPengChuan/wgdi>). In
1093 brief, delineation of paralogous gene families and one-to-one orthologs starts from all-versus-all
1094 BLASTP similarity searches with “-evaluate 1e-5 -num_alignments 10”. Collinear genes were first
1095 identified with the parameter “-icl” of WGDI within each genome and between genomes, and the
1096 collinear genes dot plots were used to count the syntenic ratios between different species to confirm
1097 the polyploidy level of each species, then gene pairs located on the same chromosome or on the
1098 same scaffold were removed, tandem gene pairs were also removed. Frequencies of synonymous
1099 substitutions per synonymous site (Ks) between colinear genes were estimated using the Nei-
1100 Gojobori approach as implemented in PAML¹²⁰. The median Ks values of each block were selected
1101 to perform the Ks peak fitting by WGDI with the parameter “-pf”. Ks peaks of *T. lepidozooides* vs. two
1102 species (*M. polymorpha* and *P. patens*) syntenic orthologs allows calculating Ks per year (r) following
1103 $r = Ks / (2 \times (\text{divergence time}))^{38}$.

1104 We used DupGen_finder¹³¹ (https://github.com/qiao-xin/DupGen_finder) with default
1105 parameters to search for genome-wide duplications. First, all-versus-all local BLASTP was
1106 performed using protein sequences with “-evaluate 1e-5 -num_alignments 10”. Second, we identified
1107 different modes of gene duplication as whole-genome duplicates (WGD), tandem duplicates (TD),

1108 proximal duplicates (less than 10 gene distance on the same chromosome: PD), transposed
1109 duplicates (transposed gene duplications: TRD), or dispersed duplicates (other duplicates than WGD,
1110 TD, PD and TRD: DSD). Third, the Ka (number of substitutions per nonsynonymous site), Ks, and
1111 Ka/Ks values were estimated for gene pairs generated by different modes of duplication based on
1112 the YN model in KaKs_Calculator¹³² (version 2.0, June 2009)
1113 (<https://sourceforge.net/projects/kakscalculator2/>), after conversion of amino acid alignments into
1114 the corresponding codon alignments with PAL2NAL¹³³ (version 14)
1115 (<http://www.bork.embl.de/pal2nal/>).

1116

1117 **Phylogenomic reconstruction**

1118 To explore the phylogenetic position of *T. lepidozoides*, we constructed data sets from a total of 56
1119 plant taxa (Table S1L), including nine genomes from Ensembl plants
1120 (<http://plants.ensembl.org/index.html>) and JGI phytozome (<http://phytozome.jgi.doe.gov>), 20
1121 transcriptomes from the onekp project
1122 (http://mirrors.iplantcollaborative.org/browse/iplant/home/shared/onekp_pilot), 13 transcriptomes
1123 from the NCBI sequence read archive (<https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?>), genes of
1124 *Klebsormidium nitens* were obtained from the *K. nitens* genome project
1125 (http://www.plantmorphogenesis.bio.titech.ac.jp/~algae_genome_project/klebsormidium/index.html
1126), those of *Chara braunii* from ORCAE (<https://bioinformatics.psb.ugent.be/orcae/overview/Chbra>),
1127 those of *Spirogloea muscicola* and *Mesotaenium endlicherianum* from Figshare
1128 ([https://figshare.com/articles/dataset/Genomes_of_subaerial_Zygnematophyceae_provide_insights
1129 _into_land_plant_evolution/9911876/1](https://figshare.com/articles/dataset/Genomes_of_subaerial_Zygnematophyceae_provide_insights_into_land_plant_evolution/9911876/1)), those from *Anthoceros agrestis* and *A. punctatus* from
1130 University of Zurich (<https://www.hornworts.uzh.ch/en/download.html>), while the remaining data are
1131 from our newly sequenced transcriptomes. Putative orthologous genes were identified using the best
1132 reciprocal blast hits (BRBH) method. Eight genomes (*Anthoceros agrestis*, *Klebsormidium nitens*,
1133 *Mesotaenium endlicherianum*, *Physcomitrella patens*, *Chara braunii*, *Marchantia polymorpha*,
1134 *Penium margaritaceum*, *Takakia lepidozoides*) were selected for inferring low-copy orthologous
1135 genes by OrthoMCL pipeline¹³⁴ (version 2.0.9) (<https://legacy.orthomcl.org/orthomcl.b5/>), then 818
1136 single-copy orthologous genes (OGs) were selected and set as queries to search homologous
1137 sequences from the 56 species (Table S1L) by HaMStR¹³⁵ (version 1.6.0)
1138 (<https://github.com/BIONF/HaMStR>) with an e-value of 1×10^{-5} . Protein sequences of each OGs were

1139 aligned by MUSCLE¹³⁶ (version 3.8.31) (<http://www.drive5.com/muscle>) and poorly aligned regions
1140 were trimmed (-automated1) using trimAl¹³⁷ (version 1.4) (<http://trimal.cgenomics.org>). Each OGs
1141 were reconstructed by IQ-tree¹³⁸ (version 1.6.12) (<http://www.iqtree.org/>) with ultrafast bootstrap
1142 parameter and suggested models, and then the 818 trees analyzed using ASTRAL¹³⁹ (version 5.7.1)
1143 (<https://github.com/smirarab/ASTRAL>) to infer species trees based on coalescent methods.

1144

1145 **Divergence time estimation**

1146 We used a total of nine fossil¹⁴⁰⁻¹⁴² calibrations for TreePL analysis (Table S2R). Considering the
1147 arguments of the maximum age of five fossils, we set two calibrations: in set1, for the #3-7 fossils
1148 only minimum ages were used for calibration, whereas in set2, minimum and maximum age
1149 calibrations were applied to all fossils. The concatenation of the alignments of the 818 OGs was
1150 used to calculate the branch length by IQ-tree, with the 818-gene coalescent tree as the constraint
1151 tree. Divergence times were estimated using the penalized likelihood method¹⁴³ by treePL¹⁴⁴ (version
1152 1.0) (<https://github.com/blackrim/treePL>). The cross-validation tested a set of smoothing parameters
1153 and then an optimal smoothing value of 0.1 was selected.

1154

1155 **Testing positive selection and fast evolution**

1156 To gain insight into the evolutionary history of *T. lepidozooides*, we constructed a data set of orthologs
1157 from 12 plant taxa (*Anthoceros angustus*, *Arabidopsis thaliana*, *Atrichum undulatum*, *Ctenidium*
1158 *capillifolium*, *P. patens*, *Sphagnum palustre*, *T. lepidozooides*, *Conocephalum japonicum*,
1159 *Haplomitrium mnioides*, *Herbertus sendtneri*, *Marchantia polymorpha* and *Selaginella moellendorffii*).
1160 The CDS sequences corresponding to genomes of *P. patens*, *A. thaliana*, and *S. moellendorffii* were
1161 obtained from the Ensembl plants web site (<http://plants.ensembl.org/index.html>), the *A. angustus*
1162 transcriptome data from the NCBI sequence read archive ([https://sra-downloadb.be-](https://sra-downloadb.be-md.ncbi.nlm.nih.gov/sos2/sra-pub-run-15/SRR9662965/SRR9662965.1)
1163 [md.ncbi.nlm.nih.gov/sos2/sra-pub-run-15/SRR9662965/SRR9662965.1](https://sra-downloadb.be-md.ncbi.nlm.nih.gov/sos2/sra-pub-run-15/SRR9662965/SRR9662965.1)), and the remaining data
1164 are from our newly sequenced transcriptomes (Table S2S). Putative orthologous genes were
1165 identified using the best reciprocal blast hits (BRBH) method. The homologous genes were first
1166 detected by the TBLASTX program with an E-value cut-off of 1×10^{-5} . Well-annotated proteins of *A.*
1167 *thaliana* were used as references to identify corresponding orthologs in the other 11 plants including
1168 *T. lepidozooides*. Only orthologs with best-hit scores across all pairs of species were kept. To obtain
1169 high-quality alignment and facilitate downstream analysis, all full orthologs data sets were aligned at

1170 the codon level. Finally, a total of 1970 putative orthologs were identified from 12 plants. We
1171 employed a maximum likelihood approach¹⁴⁵ to detect positively selected genes (PSGs) along
1172 specific lineages. The branch-site model implemented in the program CODEML in PAML¹²⁰ (version
1173 4.9j) was used. The *T. lepidozoides* lineage was defined as foreground branch. For the foreground
1174 branch we compared the selection model and the neutral model using a likelihood ratio test (LRT).
1175 A Chi-square test was conducted for each gene to compute p-values. Multiple testing was corrected
1176 by applying the false discovery rate (FDR) method. For a gene, if the selection model has a
1177 significantly higher likelihood than the neutral model (FDR-adjusted p-value < 0.05), the gene on the
1178 foreground branch may have experienced positive selection.

1179 To identify fast evolving genes (FSGs) in the *T. lepidozoides* lineage, we ran the one-ratio
1180 branch model and two-ratio branch model with CODEML in PAML to estimate the global and local
1181 dN/dS (non - synonymous substitution rates /synonymous substitution rates) ratio, respectively. The
1182 one-ratio model assumes that all branches have been evolving at the same rate (null hypothesis),
1183 and the two-ratio model allows the foreground branch to evolve under a different rate (alternative
1184 hypothesis). *T. lepidozoides* was set to foreground. LRT was employed to compare the one-ratio
1185 branch model to the two-ratio branch model. The LRT p-values were adjusted by FDR correction for
1186 multiple testing. The abnormal dN/dS values (dN/dS < 0.001 or dN/dS >5) were excluded from our
1187 analysis. If a gene had an FDR-adjusted p-value < 0.05 and also a higher dN/dS in the foreground
1188 branch than in the background branch, it was considered as FSGs in the foreground branch. In
1189 addition, we noticed that *H. sendtneri* also lived in the high-altitude area of the Tibetan Plateau. To
1190 detect convergent accelerated evolution of genes between *T. lepidozoides* and *H. sendtneri*, we
1191 defined both of them as foreground branches and repeated above-mentioned analysis. All detected
1192 PSGs and FSGs were submitted to DAVID (<https://david.ncifcrf.gov/home.jsp>) for GO functional
1193 enrichment analyses.

1194

1195 **Identification of transcription factors**

1196 For TFs and TRs, we used the HMMER search method. We downloaded the HMMER model of the
1197 domain structure of each TF from the Pfam website (<https://pfam.xfam.org/>), referring to the TAPscan
1198 TF database⁴³ (<https://plantcode.cup.uni-freiburg.de/tapscan/>). Preliminary candidates were
1199 collected by searching the HMM profile for each species (<1 × 10⁻¹⁰). Then, we filtered genes that
1200 did not match the SwissProt functional annotation (<1 × 10⁻⁵). Finally, we filtered genes containing

1201 an incorrect domain according to the domain rules of the TAPscan TF database. Most TFs/TRs were
1202 confirmed by phylogenetic analysis.

1203

1204 **Gene family classification**

1205 We applied the standard OrthoMCL pipeline¹³⁴ (version 2.0.9)
1206 (<https://legacy.orthomcl.org/orthomcl.b5/>) to identify orthologous gene families among *T.*
1207 *lepidozooides*, *C. reinhardtii*, *K. nitens*, *C. braunii*, *M. polymorpha*, *P. patens*, *S. moellendorffii* and *A.*
1208 *thaliana*. The longest protein from each gene was selected, and the proteins with a length less than
1209 30 amino acids were removed. After this step, pairwise sequence similarities between all input
1210 protein sequences were calculated using BLASTP with an E value cut-off of 1×10^{-5} . Markov
1211 clustering (MCL) of the resulting similarity matrix was used to define the orthologue cluster structure
1212 of the proteins, using an inflation value (-I) of 1.5 (OrthoMCL default). Gene families were aligned by
1213 MAFFT¹⁴⁶ (version 7.453) (<https://mafft.cbrc.jp/alignment/software/>). Following alignment,
1214 Gblocks¹⁴⁷ (version 0.91b) (<http://molevol.cmima.csic.es/castresana/Gblocks.html>) was used to
1215 remove poorly aligned positions and divergent regions. RAXML¹⁴⁸ (version 8.2.12) ([https://cme.h-](https://cme.h-its.org/exelixis/web/software/raxml/)
1216 [its.org/exelixis/web/software/raxml/](https://cme.h-its.org/exelixis/web/software/raxml/)) was used to construct the phylogenetic tree using the
1217 PROTGAMMAIJTTF model and 1,000 bootstrap replicates. The divergence time was estimated
1218 using the MCMCTree program in the PAML package¹²⁰. CAFÉ (version 4.2.1)
1219 (<https://hahnlab.github.io/CAFE/>) was used to estimate gene gain and loss.

1220

1221 **Fossil of *Takakia* sp.**

1222 The studied fossil, collected from the Middle-Late Jurassic Jiulongshan Formation at Daohugou (165
1223 Mya⁶⁶), Ningcheng County, Inner Mongolia, China. The fossil was identified as *Takakia* sp. by Yufei
1224 Wang (Paleobotanist) and Yu Jia (Bryologist) from Institute of Botany, Chinese Academy of Sciences.
1225 The fossil *Takakia*, No. CNU-PLA-NN2023056, is housed at the Key Lab of Insect Evolution and
1226 Environmental Changes, College of Life Sciences, Capital Normal University, Beijing, China (CNUB;
1227 Prof. Dong Ren, Curator). The photographs of fossil were taken with a Nikon SMZ 25 microscope
1228 with an attached Nikon DS-Ri2 digital camera system.

1229

1230 **Evolution of specific morphological traits**

1231 To estimate the evolution of morphogenetic genes including stomata, rhizoids and leaf adaxial-

1232 abaxial polarity in *Takakia*, we carefully checked these gene homologs in 11 bryophytes (including
1233 *T. lepidozoides*, two hornworts, i.e., *Anthoceros agrestis* and *Anthoceros angustus*, one liverwort,
1234 i.e., *Marchantia polymorpha* and seven other mosses, i.e., *Calohypnum plumiforme*, *Ceratodon*
1235 *purpureus*, *Entodon seductrix*, *Fontinalis antipyretica*, *Hypnum curvifolium*, *Physcomitrella patens*,
1236 *Syntrichia caninervis*) (Table S2T) using BLASTP with identity cutoff 50%, e-value cutoff 1×10^{-20} .
1237 Domains were checked in homolog candidates using HMMER. Maximum likelihood tree constructed
1238 using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments of proteins and
1239 trimmed by trimAl. Tree is the consensus of 1,000 replicates. The divergence time was estimated
1240 using the MCMCTree program in the PAML package¹²⁰. We used two fossil¹⁴⁰ calibrations for
1241 MCMCTree analysis (Table S2R), minimum age 405 Mya and maximum age 515.5 Mya calibrations
1242 were applied.

1243

1244 **Gene expression profiling**

1245 Poor quality or technical sequences in Illumina paired-end reads were removed using the
1246 Trimmomatic preprocessing tool¹⁴⁹ (version 0.39)
1247 (<http://www.usadellab.org/cms/?page=trimmomatic>). The qualified paired reads were then aligned
1248 against the *T. lepidozoides* reference transcript sequences using Bowtie 2¹⁵⁰ (version 2.4.1)
1249 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) to find all alignments of a read with no more
1250 than two mismatches. Subsequently, gene and isoform abundances were quantified from paired-end
1251 RNA-seq data using the RSEM software package (version 1.3.3)
1252 (<https://github.com/siddharthab/RSEM>). Differentially expressed transcripts or genes (DEGs) with
1253 biological replicates were identified by running Bioconductor tools edgeR¹⁵¹ (version 3.28.1)
1254 (<https://bioconductor.org/packages/release/bioc/html/edgeR.html>), which implements a range of
1255 statistical methods based on the negative binomial distributions. Orthologs of *T. lepidozoides* were
1256 functionally annotated via the EggNOG database¹⁵² (version 5.0) (<http://eggno5.embl.de/>)
1257 companion tool eggno-mapper¹⁵³ (version 2.0.1) (<http://eggno-mapper.embl.de/>), with “-d euk --
1258 tax_scope 3208”. The inferred GO terms and Kyoto Encyclopedia of Genes and Genomes (KEGG)
1259 pathways were used as the basis for DEG enrichment analyses performed with the R package
1260 clusterProfiler¹⁵⁴ (version 3.14.3)
1261 (<http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html>).

1262

1263 **QUANTIFICATION AND STATISTICAL ANALYSIS**

1264 The statistics of experiments can be found in the figure, table and/or data legends, including the
1265 results represented as mean or median \pm SD, exact value of n, P values with significance levels. All
1266 details of the statistics applied are provided alongside the respective analysis in the Method Details
1267 section.

1268

1269 **SUPPLEMENTARY INFORMATION**

1270 **Figure S1.** Population changes from 2010 to 2021, related to Figure 1

1271 The coverage of commonly found species from 2010 to 2021. The vertical axis represents the
1272 relative coverage of the species in observation quadrants (n = 52), and dotted line marks the
1273 relationship between the species population size and the year. The central mark of each box
1274 indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles
1275 respectively, the whiskers extend to the most extreme data points not considered outliers). (**A – D**)
1276 Most common mosses *Racomitrium himalayanum*, *Dicranum drummondii*, *Hypnum plumaeforme*
1277 and *Polytrichum juniperinum*, (**E – G**) common liverworts *Herbertus sendtneri*, *Plagiochila recurvata*
1278 and *Wettsteinia rotundifolia*, (**H – L**) most common angiosperms *Diapensia wardii*, *Bergenia*
1279 *purpurascens*, *Cassiope selaginoides*, *Saxifraga unguiculata* and *Pedicularis siphonantha*, (**M, N**)
1280 lichens *Stereocaulon alpinum* and *Cladonia rangiferina*. (**O**) The Shannon-Wiener Index change
1281 during 2010 to 2021.

1282

1283 **Figure S2.** The origin of branching in *T. lepidozoides*, observed with the mPS-PI staining technique,
1284 related to Figure 2

1285 The sections (**A, B, C**) from three mutually perpendicular directions at the connection site (the cross
1286 point of white dotted lines) between stem and branch, the numbers mark cells in the parent stem
1287 (yellow) and the branch (white) respectively, both comprise eight rows of cells (1–8th row) in the
1288 section (B), yellow arrows mark the vertical walls between epidermal cells, white arrows mark the
1289 inclined walls between longer cells (potential vessel elements), bar = 50 μ m. The sections (**D, E, F**)
1290 from three mutually perpendicular directions at the connection site (the cross point of white dotted
1291 lines) between stem and branch, numbers mark cells in the parent stem (yellow) and the branch
1292 (white) respectively, which shows the branch originated from cells of the epidermis (such as “yellow
1293 8” and “white 6” marked cells) and sub-epidermis (such as “yellow 7”, 7’, 7 and “white 5” marked

1294 cells) of parent stem, bar = 50 μm . Fluorescence images were obtained using a Leica SP8 laser
1295 scanning confocal microscope, with the excitation wavelength at 561 nm and emission wavelength
1296 from 570 nm to 670 nm.

1297
1298 **Figure S3.** Effects of UV-B irradiation on *A. thaliana*, related to Figure 3

1299 Arabidopsis Col - 0 plants grown for 25 days under normal conditions were treated with or without
1300 UV-B (A) Plants treated without or with 144 kJ m^{-2} UV-B. (B) Plant recovery at 24 hours after UV-B
1301 treatment. (C) Plant recovery at 48 hours after treatment. (D) Plant recovery at 72 hours after UV-B
1302 treatment. Side length of the pots is 7.5 cm.

1303
1304 **Figure S4.** Chromosomes in gametophyte cells of *T. lepidozoides*, related to Figure 4

1305 (A) Hi-C assembly heatmap shows four obvious chromosomes (blue boxes) composing the *T.*
1306 *lepidozoides* genome. (B – E) Four chromosomes shown in a somatic prometaphase cell separated
1307 from a stem tip, bar = 5 μm .

1308
1309 **Figure S5.** Genetic foundation of the lack of stomata and rhizoids, related to Figure 6

1310 Phylogenetic tree of genes relevant for the formation of rhizoids (A). Sequence alignments of SCRM
1311 (B) and FAMA (C). Black boxes mark conserved domains. Phylogenetic tree of genes relevant for
1312 the formation of stomata (D). Maximum likelihood tree using RAXML with the PROTGAMMAJTT
1313 model based upon MAFFT alignments of proteins are the consensus of 1,000 replicates. Numbered
1314 nodes show the percentage of bootstraps which support the node. Bootstraps with less than 50%
1315 support are not shown. Orange font color represents *A. thaliana*, pink font color represents *T.*
1316 *lepidozoides* and blue font color represents *P. patens*. Shade colors as shown on the right.

1317
1318 **Table S1.** Field observation in survey areas, mass spectrum of metabolic compounds, sequencing
1319 library, genome assembly and annotation, positive selection and fast evolution, with index in the first
1320 sheet, related to STAR methods.

1321
1322 **Table S2.** Genomic comparative analyses, key genes operating in morphological innovations and
1323 environmental adaptations, with index in the first sheet, related to STAR methods.

1324

1325 **Data S1.** Field observation in survey areas, morphological characters, physiological and biochemical
1326 properties of *T. lepidozoides*, related to STAR methods.

1327

1328 **Data S2.** Genome features, phylogenetic position, key genes operating in morphological innovations
1329 and environmental adaptations, related to STAR methods.

1330

1331 **Movie S1.** Film of survey areas in the southern part of the Tibetan Plateau. Digital Earth data are
1332 from Google Earth (<https://earth.google.com/>), remote sensing images captured by matrice 300 RTK
1333 (DJI), related to Figure 1.

1334

1335 **Movie S2.** Film of a restructured stem. The raw images were converted to an avi file using Leica
1336 LAS X (3.5.2) at 16 frames per second, related to Figure 2.

1337

1338 **Movie S3.** Time lapse photos of capsule development. Total 155 photos were taken at steps of 0.1
1339 s (TimelapseCams, Wingscapes), related to Figure 1.

1340

1341 **ACKNOWLEDGMENTS**

1342 The authors wish to thank Jingyun Fang (Institute of Ecology, College of Urban and Environmental
1343 Sciences, Laboratory for Earth Surface Processes of the Ministry of Education, Peking University,
1344 Beijing, China), Ming Dong (Key Laboratory of Hangzhou City for Ecosystem Protection and
1345 Restoration, Hangzhou Normal University, Hangzhou, China) and Qingyin Zeng (State Key
1346 Laboratory of Tree Genetics and Breeding, Chinese Academy of Forestry, Beijing, China) for helpful
1347 discussions and comments on the manuscript. The authors wish to thank Congming Lu
1348 (Photosynthesis Research Center, Key Laboratory of Photobiology, Institute of Botany, Chinese
1349 Academy of Sciences, Beijing, China) for the antibodies of photosynthetic protein complexes,
1350 including AtCP47 and AtPsaD. The authors wish to thank Yufei Wang (Institute of Botany, Chinese
1351 Academy of Sciences) for the fossil identification. The authors wish to thank Taiping Gao (College
1352 of Life Sciences, Capital Normal University, Beijing, China) for technical supporting of drawing the
1353 graphical abstract. This work was supported by grants from the National Natural Science Foundation
1354 of China (grant no. 91631109 and 31970658) to Y. H., the key research and development plan of
1355 Science and Technology Department of Tibet Autonomous Region (CGZH2023000196) to Y.H., the

1356 New Interdisciplinary Construction of Bioinformatics and Statistics of Capital Normal University
1357 (2055079) to H. C., Shenzhen Key Laboratory of South Subtropical Plant Diversity open project in
1358 2020 (SSTLAB-2020-02) to Y. H., and the German Research Foundation (DFG) under Germany's
1359 Excellence Strategy (EXC-2189, CIBSS) and CRC-Transregio 141 (project B02) to R. R.. H. R. and
1360 R. R. acknowledge the support of the Freiburg Institute for Advanced Studies (FRIAS) and the
1361 University of Strasbourg Institute of Advanced Study (USIAS) to the METABEVO project.

1362

1363 **AUTHOR CONTRIBUTIONS**

1364 Yikun He conceived the project; Yikun He and Ralf Reski designed the project; Yikun He, Ralf Reski
1365 and Ruoyang Hu wrote the manuscript, supplementary notes, and methods; Xuedong Li, Runjie
1366 Zhang, Ruoyang Hu, Xiaofeng Wu, Feng Zhao, Shuang Xing, Yong Hu, Lei Deng, Fuzhou Duan,
1367 Wenji Zhao, Yu Jia and Meizhi Wang performed the field works; Ruoyang Hu, Yong Hu and Fang
1368 Bao collected and sequenced the plant material; Yikun He and Xianyong Sheng performed three-
1369 dimensional reconstruction of branching and rooting; Qiang Lv and Yuhan Ding performed transgene
1370 experiments; Xiaofeng Wu, Ruoyang Hu, Runjie Zhang, Yikun He, Lihong Xiao and Kai Li performed
1371 ultrastructural observations; Zhukuan Cheng and Hanli You observed chromosomes in somatic cells;
1372 Min Zhang and Huan Yuan performed immunofluorescence labeling of AtPIN1; Zhijia Chen and
1373 Qiang Lv performed immunoblots of photosynthetic proteins; Xiaofeng Wu, Feng Zhao, Ruoyang Hu,
1374 Runjie Zhang, Xiaoqin Wang and Xingru Li detected physiological changes and identified fatty acids;
1375 Ruoyanh Hu, Peiyong Xin and Jinfang Chu performed IAA analyses; Zhijia Chen, Ruoyang Hu,
1376 Chuan Ma and Zhongfeng Li performed metabolic analyses; Chengzhi Liang, Huilong Du, Hongwei
1377 Lu, Qiang Gao, Ruoyang Hu and Nico van Gessel assembled and annotated the genome; Ruoyang
1378 Hu and Qiang Gao conducted the transcriptome sequencing and analysis; Jinzhong Fu and Bin Lv
1379 performed the positive selection and fast evolution analysis; Ruoyang Hu, Liangsheng Zhang, Chien-
1380 Hsun Huang and Hong Ma performed the gene duplication (WGD,TRD, TD, PD, DSD), protein
1381 phosphorylation genes, PPR, phylogenetic position and divergent time, shared genes analysis;
1382 Ruoyang Hu, Xiaoyu Yan, Yong Hu, Ping Wan, Wei Xiao, Eva L. Decker, Hugues Renault, Gertrud
1383 Wiedemann, Nelly A. Horst, Ralf Reski, Hong Ma, Chien-Hsun Huang, Liangsheng Zhang, Fei Chen,
1384 Yong Zhang, Yang Liu and Shanshan Dong performed gene family clustering and comparative
1385 phylogenomics; Fabian B. Haas, Per K.I. Wilhelmsson, Kristian K. Ullrich, Eva Neumann and Stefan
1386 A. Rensing performed contamination analysis and annotated transcription associated proteins;

1387 Stefan A. Rensing supervised gene family clustering analysis; Yu Jia identified the fossil; Stefan A.
1388 Rensing, Hong Ma, Bernard Goffinet, Mitsuyasu Hasebe, Jinzhong Fu, Hengjian Cui, Yong Zhang,
1389 Liangsheng Zhang, Yang Liu, Ruiliang Zhu, Yu Jia, Nico van Gessel and Eva L. Decker reviewed
1390 and modified the manuscript.

1391

1392 **DECLARATION OF INTERESTS**

1393 The authors declare no competing interests.

1394

1395 **REFERENCES**

- 1396 1. Han, J., Morris, S.C., Ou, Q., Shu, D.G., and Huang, H. (2017). Meiofaunal deuterostomes
1397 from the basal Cambrian of Shaanxi (China). *Nature* 542, 228–231. 10.1038/nature21072.
- 1398 2. Reski, R. (2018). Enabling the water-to-land transition. *Nat. Plants* 4, 67–68.
1399 10.1038/s41477-018-0101-5.
- 1400 3. Porada, P., Lenton, T.M., Pohl, A., Weber, B., Mander, L., Donnadieu, Y., Beer, C., Poschl,
1401 U., and Kleidon, A. (2016). High potential for weathering and climate effects of non-vascular
1402 vegetation in the Late Ordovician. *Nat. Commun.* 7, 12113. 10.1038/ncomms12113.
- 1403 4. Lenton, T.M., Crouch, M., Johnson, M., Pires, N., and Dolan, L. (2012). First plants cooled
1404 the Ordovician. *Nat. Geosci.* 5, 86–89. 10.1038/ngeo1390.
- 1405 5. Planavsky, N.J., Reinhard, C.T., Wang, X.L., Thomson, D., McGoldrick, P., Rainbird, R.H.,
1406 Johnson, T., Fischer, W.W., and Lyons, T.W. (2014). Low mid-proterozoic atmospheric
1407 oxygen levels and the delayed rise of animals. *Science* 346, 635–638.
1408 10.1126/science.1258410.
- 1409 6. Lenton, T.M., Dahl, T.W., Daines, S.J., Mills, B.J.W., Ozaki, K., Saltzman, M.R., and Porada,
1410 P. (2016). Earliest land plants created modern levels of atmospheric oxygen. *Proc. Natl. Acad.*
1411 *Sci. USA.* 113, 9704–9709. 10.1073/pnas.1604787113.
- 1412 7. Chater, C.C., Caine, R.S., Tomek, M., Wallace, S., Kamisugi, Y., Cuming, A.C., Lang, D.,
1413 MacAlister, C.A., Casson, S., Bergmann, D.C., et al. (2016). Origin and function of stomata
1414 in the moss *Physcomitrella patens*. *Nat. Plants* 2, 16179. 10.1038/nplants.2016.179.
- 1415 8. Clark, J.W., Harris, B.J., Hetherington, A.J., Hurtado-Castano, N., Brench, R.A., Casson, S.,
1416 Williams, T.A., Gray, J.E., and Hetherington, A.M. (2022). The origin and evolution of stomata.
1417 *Curr. Biol.* 32, R539–R553. 10.1016/j.cub.2022.04.040.

- 1418 9. Rowley, D.B., and Currie, B.S. (2006). Palaeo-altimetry of the late Eocene to Miocene
1419 Lunpola basin, central Tibet. *Nature* 439, 677–681. 10.1038/nature04506.
- 1420 10. Qiu, J. (2008). The third pole. *Nature* 454, 393–396. doi:10.1038/454393a.
- 1421 11. Wang, X.Y., Wang, T., Xu, J.F., Shen, Z.H., Yang, Y.P., Chen, A.P., Wang, S.P., Liang, E.Y.,
1422 and Piao, S.L. (2022). Enhanced habitat loss of the Himalayan endemic flora driven by
1423 warming-forced upslope tree expansion. *Nat. Ecol. Evol.* 6, 890–899 10.1038/s41559-022-
1424 01774-3.
- 1425 12. Zanatta, F., Engler, R., Collart, F., Broennimann, O., Mateo, R.G., Papp, B., Munoz, J.,
1426 Baurain, D., Guisan, A., and Vanderpoorten, A. (2020). Bryophytes are predicted to lag
1427 behind future climate change despite their high dispersal capacities. *Nat. Commun.* 11, 5601.
1428 10.1038/s41467-020-19410-8.
- 1429 13. McDaniel, S.F. (2021). Bryophytes are not early diverging land plants. *New Phytol.* 230,
1430 1300–1304. 10.1111/nph.17241.
- 1431 14. Horst, N.A., Katz, A., Pereman, I., Decker, E.L., Ohad, N., and Reski, R. (2016). A single
1432 homeobox gene triggers phase transition, embryogenesis and asexual reproduction. *Nat.*
1433 *Plants* 2, 15209. 10.1038/nplants.2015.209.
- 1434 15. Cutler, N.A., Belyea, L.R., and Dugmore, A.J. (2008). The spatiotemporal dynamics of a
1435 primary succession. *J. Ecol.* 96, 231–246. 10.1111/j.1365-2745.2007.01344.x.
- 1436 16. Roads, E., Longton, R.E., and Convey, P. (2014). Millennial timescale regeneration in a moss
1437 from Antarctica. *Curr. Biol.* 24, R222–R223. 10.1016/j.cub.2014.01.053.
- 1438 17. Biersma, E.M., Jackson, J.A., Hyvönen, J., Koskinen, S., Linse, K., Griffiths, H., and Convey,
1439 P. (2017). Global biogeographic patterns in bipolar moss species. *R. Soc. Open Sci.* 4,
1440 170147. 10.1098/rsos.170147.
- 1441 18. Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T.,
1442 Perroud, P.F., Lindquist, E.A., Kamisugi, Y., et al. (2008). The *Physcomitrella* genome reveals
1443 evolutionary insights into the conquest of land by plants. *Science* 319, 64–69.
1444 10.1126/science.1150646.
- 1445 19. Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S.Q., Ishizaki, K., Yamaoka, S.,
1446 Nishihama, R., Nakamura, Y., Berger, F., et al. (2017). Insights into land plant evolution
1447 garnered from the *Marchantia polymorpha* genome. *Cell* 171, 287–304.
1448 10.1016/j.cell.2017.09.030.

- 1449 20. Li, F.W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N.,
1450 Barker, M.S., Bennett, T., Blazquez, M.A., et al. (2020). *Anthoceros* genomes illuminate the
1451 origin of land plants and the unique biology of hornworts. *Nat. Plants* 6, 259–272.
1452 10.1038/s41477-020-0618-2.
- 1453 21. Donoghue, P.C.J., Harrison, C.J., Paps, J., and Schneider, H. (2021). The evolutionary
1454 emergence of land plants. *Curr. Biol.* 31, R1281–R1298. 10.1016/j.cub.2021.07.038.
- 1455 22. Zhang, J., Fu, X.X., Li, R.Q., Zhao, X., Liu, Y., Li, M.H., Zwaenepoel, A., Ma, H., Goffinet, B.,
1456 Guan, Y.L., et al. (2020). The hornwort genome and early land plant evolution. *Nat. Plants* 6,
1457 107–118. 10.1038/s41477-019-0588-4.
- 1458 23. Mitten, W. (1860). *Hepaticae Indiae Orientalis: an enumeration of the Hepaticae of the East*
1459 *Indies. Bot. J. Linn. Soc.* 5, 89–108.
- 1460 24. Hattori, S. (1958). Preliminary report on *Takakia lepidozoides*. *Journ. Hattori Bot. Lab.* 19,
1461 133–137.
- 1462 25. Hattori, S., and Mizutani, M. (1958). What is *Takakia lepidozoides*? *Journ. Hattori Bot. Lab.*
1463 20, 295–303.
- 1464 26. Grolle, R. (1963). *Takakia* im Himalaya. *Oesterr. Bot. Z.* 110, 444–447. 10.1007/BF01373680.
- 1465 27. Smith, D.K., and Davison, P.G. (1993). Antheridia and sporophytes in *Takakia ceratophylla*
1466 (Mitt.) Grolle: evidence for reclassification among the mosses. *Journ. Hattori Bot. Lab.* 73,
1467 263–271.
- 1468 28. Schuster, R.M. (1997). On *Takakia* and the phylogenetic relationships of the *Takakiales*. *Nova*
1469 *Hedwig.* 64, 281–310. 10.1127/nova.hedwigia/64/1997/281.
- 1470 29. Liu, Y., Johnson, M.G., Cox, C.J., Medina, R., Devos, N., Vanderpoorten, A., Hedenas, L.,
1471 Bell, N.E., Shevock, J.R., Aguero, B., et al. (2019). Resolution of the ordinal phylogeny of
1472 mosses using targeted exons from organellar and nuclear genomes. *Nat. Commun.* 10, 1485.
1473 10.1038/s41467-019-09454-w.
- 1474 30. Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., and Kent, J. (2000).
1475 Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858. 10.1038/35002501.
- 1476 31. Yao, T.D., Thompson, L., Yang, W., Yu, W.S., Gao, Y., Guo, X.J., Yang, X.X., Duan, K.Q.,
1477 Zhao, H.B., Xu, B.Q., et al. (2012). Different glacier status with atmospheric circulations in
1478 Tibetan Plateau and surroundings. *Nat. Clim. Chang.* 2, 663–667. 10.1038/nclimate1580.
- 1479 32. Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertova, D., Jurgens, G., and

- 1480 Friml, J. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ
1481 formation. *Cell* 115, 591–602. 10.1016/s0092-8674(03)00924-3.
- 1482 33. Ries, G., Heller, W., Puchta, H., Sandermann, H., Seidlitz, H.K., and Hohn, B. (2000).
1483 Elevated UV-B radiation reduces genome stability in plants. *Nature* 406, 98–101.
1484 10.1038/35017595.
- 1485 34. Rensing, S.A., Goffinet, B., Meyberg, R., Wu, S.Z., and Bezanilla, M. (2020). The Moss
1486 *Physcomitrium (Physcomitrella) patens*: A model organism for non-seed plants. *Plant Cell* 32,
1487 1361-1376. 10.1105/tpc.19.00828.
- 1488 35. Resemann, H.C., Herrfurth, C., Feussner, K., Hornung, E., Ostendorf, A.K., Gomann, J.,
1489 Mittag, J., van Gessel, N., de Vries, J., Ludwig-Muller, J., et al. (2021). Convergence of
1490 sphingolipid desaturation across over 500 million years of plant evolution. *Nat. Plants* 7, 219–
1491 232. 10.1038/s41477-020-00844-3.
- 1492 36. Tatuno, S. (1959). Chromosomen von *Takakia lepidozoides* und eine studie zur evolution
1493 der chromosomen der bryophyten. *Cytologia* 24, 138–147. 10.1508/cytologia.24.138.
- 1494 37. Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., dePamphilis, C., Albert,
1495 V.A., Aono, N., Aoyama, T., Ambrose, B.A., et al. (2011). The *Selaginella* genome identifies
1496 genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963.
1497 10.1126/science.1203810.
- 1498 38. Lang, D., Ullrich, K.K., Murat, F., Fuchs, J., Jenkins, J., Haas, F.B., Piednoel, M., Gundlach,
1499 H., Van Bel, M., Meyberg, R., et al. (2018). The *Physcomitrella patens* chromosome-scale
1500 assembly reveals moss genome structure and evolution. *Plant J.* 93, 515–533.
1501 10.1111/tpj.13801.
- 1502 39. Sadamitsu, A., Inoue, Y., Sakakibara, K., Tsubota, H., Yamaguchi, T., Deguchi, H., Nishiyama,
1503 T., and Shimamura, M. (2021). The complete plastid genome sequence of the enigmatic
1504 moss, *Takakia lepidozoides (Takakiopsida, Bryophyta)*: evolutionary perspectives on the
1505 largest collection of genes in mosses and the intensive RNA editing. *Plant Mol. Biol.* 107,
1506 431–449. 10.1007/s11103-021-01214-z.
- 1507 40. Chen, D.Q., Hao, F.S., Mu, H.Q., Ahsan, N., Thelen, J.J., and Stacey, G. (2021). S-acylation
1508 of P2K1 mediates extracellular ATP-induced immune signaling in *Arabidopsis*. *Nat. Commun.*
1509 12, 2750. 10.1038/s41467-021-22854-1.
- 1510 41. Millar, A.H., Heazlewood, J.L., Giglione, C., Holdsworth, M.J., Bachmair, A., and Schulze,

- 1511 W.X. (2019). The scope, functions, and dynamics of posttranslational protein modifications.
1512 In Annual Review of Plant Biology, Vol 70, S.S. Merchant, ed. (Annual Reviews), pp. 119–
1513 151. 10.1146/annurev-arplant-050718-100211.
- 1514 42. Qiu, Y.L., Li, L.B., Wang, B., Chen, Z.D., Knoop, V., Groth-Malonek, M., Dombrowska, O., Lee,
1515 J., Kent, L., Rest, J., et al. (2006). The deepest divergences in land plants inferred from
1516 phylogenomic evidence. Proc. Natl. Acad. Sci. USA. 103, 15511–15516.
1517 10.1073/pnas.0603335103.
- 1518 43. Wilhelmsson, P.K.I., Muhlich, C., Ullrich, K.K., and Rensing, S.A. (2017). Comprehensive
1519 genome-wide classification reveals that many plant-specific transcription factors evolved in
1520 *Streptophyte* algae. Genome Biol. Evol. 9, 3384–3397. 10.1093/gbe/evx258.
- 1521 44. Ariel, F.D., Manavella, P.A., Dezar, C.A., and Chan, R.L. (2007). The true story of the HD-Zip
1522 family. Trends Plant Sci. 12, 419–426. 10.1016/j.tplants.2007.08.003.
- 1523 45. Griffiths, S., Dunford, R.P., Coupland, G., and Laurie, D.A. (2003). The evolution of
1524 CONSTANS-like gene families in barley, rice, and *Arabidopsis*. Plant Physiol. 131, 1855–
1525 1867. 10.1104/pp.102.016188.
- 1526 46. Nakaminami, K., Karlson, D.T., and Imai, R. (2006). Functional conservation of cold shock
1527 domains in bacteria and higher plants. Proc. Natl. Acad. Sci. USA. 103, 10122–10127.
1528 10.1073/pnas.0603168103.
- 1529 47. Sasaki, K., Kim, M.H., and Imai, R. (2013). *Arabidopsis* cold shock domain protein 2 is a
1530 negative regulator of cold acclimation. New Phytol. 198, 95–102. 10.1111/nph.12118.
- 1531 48. Sanchez, R., and Zhou, M.M. (2011). The PHD finger: a versatile epigenome reader. Trends
1532 Biochem. Sci. 36, 364–372. 10.1016/j.tibs.2011.03.005.
- 1533 49. Robles, P., and Quesada, V. (2021). Research progress in the molecular functions of plant
1534 mTERF proteins. Cells 10, 205. 10.3390/cells10020205.
- 1535 50. Babiychuk, E., Vandepoele, K., Wissing, J., Garcia-Diaz, M., De Rycke, R., Akbari, H.,
1536 Joubes, J., Beeckman, T., Jansch, L., Frentzen, M., et al. (2011). Plastid gene expression
1537 and plant development require a plastidic protein of the mitochondrial transcription
1538 termination factor family. Proc. Natl. Acad. Sci. USA. 108, 6674–6679.
1539 10.1073/pnas.1103442108.
- 1540 51. Umezawa, T., Sugiyama, N., Takahashi, F., Anderson, J.C., Ishihama, Y., Peck, S.C., and
1541 Shinozaki, K. (2013). Genetics and phosphoproteomics reveal a protein phosphorylation

- 1542 network in the abscisic acid signaling pathway in *Arabidopsis thaliana*. *Sci. Signal.* *6*, rs8.
1543 10.1126/scisignal.2003509.
- 1544 52. Small, I.D., Schallenberg-Rudinger, M., Takenaka, M., Mireau, H., and Ostersetzer-Biran, O.
1545 (2020). Plant organellar RNA editing: what 30 years of research has revealed. *Plant J.* *101*,
1546 1040–1056. 10.1111/tpj.14578.
- 1547 53. Waltz, F., Nguyen, T.T., Arrive, M., Bochler, A., Chicher, J., Hammann, P., Kuhn, L., Quadrado,
1548 M., Mireau, H., Hashem, Y., and Giege, P. (2019). Small is big in *Arabidopsis* mitochondrial
1549 ribosome. *Nat. Plants* *5*, 106–117. 10.1038/s41477-018-0339-y.
- 1550 54. Wang, C., Blondel, L., Quadrado, M., Dargel-Graffin, C., and Mireau, H. (2022).
1551 Pentatricopeptide repeat protein MTSF3 ensures mitochondrial RNA stability and
1552 embryogenesis. *Plant Physiol.* *190*, 669–681. 10.1093/plphys/kiac309.
- 1553 55. Bray, C.M., and West, C.E. (2005). DNA repair mechanisms in plants: crucial sensors and
1554 effectors for the maintenance of genome integrity. *New Phytol.* *168*, 511–528. 10.1111/j.1469-
1555 8137.2005.01548.x.
- 1556 56. Li, X., and Heyer, W.D. (2008). Homologous recombination in DNA repair and DNA damage
1557 tolerance. *Cell Res.* *18*, 99–113. 10.1038/cr.2008.1.
- 1558 57. Shinohara, A., Ogawa, H., and Ogawa, T. (1992). Rad51 protein involved in repair and
1559 recombination in *S. cerevisiae* is a RecA-like protein. *Cell* *69*, 457–470. 10.1016/0092-
1560 8674(92)90447-k.
- 1561 58. Huen, M.S.Y., Sy, S.M.H., and Chen, J. (2010). BRCA1 and its toolbox for the maintenance
1562 of genome integrity. *Nat. Rev. Mol. Cell Biol.* *11*, 138–148. 10.1038/nrm2831.
- 1563 59. Fraser, C.M., and Chapple, C. (2011). The phenylpropanoid pathway in *Arabidopsis*.
1564 *Arabidopsis Book* *9*, e0152. 10.1199/tab.0152.
- 1565 60. Raes, J., Rohde, A., Christensen, J.H., Van de Peer, Y., and Boerjan, W. (2003). Genome-
1566 wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiol.* *133*, 1051–
1567 1071. 10.1104/pp.103.026484.
- 1568 61. Nelson, D., and Werck-Reichhart, D. (2011). A P450-centric view of plant evolution. *Plant J.*
1569 *66*, 194–211. 10.1111/j.1365-313X.2011.04529.x.
- 1570 62. Renault, H., Alber, A., Horst, N.A., Lopes, A.B., Fich, E.A., Kriegshauser, L., Wiedemann, G.,
1571 Ullmann, P., Herrgott, L., Erhardt, M., et al. (2017). A phenol-enriched cuticle is ancestral to
1572 lignin evolution in land plants. *Nat. Commun.* *8*, 14713. 10.1038/ncomms14713.

- 1573 63. Fukushima, K., and Hasebe, M. (2014). Adaxial-abaxial polarity: the developmental basis of
1574 leaf shape diversity. *Genesis* 52, 1–18. 10.1002/dvg.22728.
- 1575 64. Kerstetter, R.A., Bollman, K., Taylor, R.A., Bomblies, K., and Poethig, R.S. (2001). *KANADI*
1576 regulates organ polarity in *Arabidopsis*. *Nature* 411, 706–709. 10.1038/35079629.
- 1577 65. Burian, A., Paszkiewicz, G., Nguyen, K.T., Meda, S., Raczynska-Szajgin, M., and
1578 Timmermans, M.C.P. (2022). Specification of leaf dorsiventrality via a prepatterned binary
1579 readout of a uniform auxin input. *Nat. Plants* 8, 269–280. 10.1038/s41477-022-01111-3.
- 1580 66. He, H.Y., Wang, X.L., Zhou, Z.H., Zhu, R.X., Jin, F., Wang, F., Ding, X., and Boven, A. (2004).
1581 $^{40}\text{Ar}/^{39}\text{Ar}$ dating of ignimbrite from Inner Mongolia, northeastern China, indicates a post-
1582 Middle Jurassic age for the overlying Daohugou Bed. *Geophys. Res. Lett.* 31, L20609.
1583 10.1029/2004gl020792.
- 1584 67. Khraiwesh, B., Arif, M.A., Seumel, G.I., Ossowski, S., Weigel, D., Reski, R., and Frank, W.
1585 (2010). Transcriptional control of gene expression by MicroRNAs. *Cell* 140, 111–122.
1586 10.1016/j.cell.2009.12.023.
- 1587 68. Crandall-Stotler, B., Stotler, R.E., and Long, D.G. (2009). Phylogeny and classification of the
1588 *Marchantiophyta*. *Edinburgh Journal of Botany* 66, 155–198. 10.1017/S0960428609005393.
- 1589 69. Jones, V.A.S., and Dolan, L. (2012). The evolution of root hairs and rhizoids. *Ann. Bot.* 110,
1590 205–212. 10.1093/aob/mcs136.
- 1591 70. Menand, B., Yi, K.K., Jouannic, S., Hoffmann, L., Ryan, E., Linstead, P., Schaefer, D.G., and
1592 Dolan, L. (2007). An ancient mechanism controls the development of cells with a rooting
1593 function in land plants. *Science* 316, 1477–1480. 10.1126/science.1142618.
- 1594 71. Jang, G., Yi, K., Pires, N.D., Menand, B., and Dolan, L. (2011). RSL genes are sufficient for
1595 rhizoid system development in early diverging land plants. *Development* 138, 2273–2281.
1596 10.1242/dev.060582.
- 1597 72. Proust, H., Honkanen, S., Jones, V.A.S., Morieri, G., Prescott, H., Kelly, S., Ishizaki, K.,
1598 Kohchi, T., and Dolan, L. (2016). RSL class I genes controlled the development of epidermal
1599 structures in the common ancestor of land plants. *Curr. Biol.* 26, 93–99.
1600 10.1016/j.cub.2015.11.042.
- 1601 73. Cheng, S.F., Xian, W.F., Fu, Y., Marin, B., Keller, J., Wu, T., Sun, W.J., Li, X.L., Xu, Y., Zhang,
1602 Y., et al. (2019). Genomes of subaerial *Zygnematophyceae* provide insights into land plant
1603 evolution. *Cell* 179, 1057–1067. 10.1016/j.cell.2019.10.019.

- 1604 74. Harris, B.J., Harrison, C.J., Hetherington, A.M., and Williams, T.A. (2020). Phylogenomic
1605 evidence for the monophyly of Bryophytes and the reductive evolution of stomata. *Curr. Biol.*
1606 *30*, 2001–2012. 10.1016/j.cub.2020.03.048.
- 1607 75. Bowles, A.M.C., Paps, J., and Bechtold, U. (2022). Water-related innovations in land plants
1608 evolved by different patterns of gene cooption and novelty. *New Phytol.* *235*, 732–742.
1609 10.1111/nph.17981.
- 1610 76. Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwarz, H., Hamann, T., Offringa, R., and
1611 Jurgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of
1612 *Arabidopsis*. *Nature* *426*, 147–153. 10.1038/nature02085.
- 1613 77. Reinhardt, D., Pesce, E.R., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., Traas, J.,
1614 Friml, J., and Kuhlemeier, C. (2003). Regulation of phyllotaxis by polar auxin transport.
1615 *Nature* *426*, 255–260. 10.1038/nature02081.
- 1616 78. Bennett, T.A., Liu, M.M., Aoyama, T., Bierfreund, N.M., Braun, M., Coudert, Y., Dennis, R.J.,
1617 O'Connor, D., Wang, X.Y., White, C.D., et al. (2014). Plasma membrane-targeted PIN
1618 proteins drive shoot development in a moss. *Curr. Biol.* *24*, 2776–2785.
1619 10.1016/j.cub.2014.09.054.
- 1620 79. Scotese, C.R., Song, H.J., Mills, B.J.W., and van der Meer, D.G. (2021). Phanerozoic
1621 paleotemperatures: The earth's changing climate during the last 540 million years. *Earth-Sci.*
1622 *Rev.* *215*, 103530. 10.1016/j.earscirev.2021.103503.
- 1623 80. Harris, B.J., Clark, J.W., Schrempf, D., Szollosi, G.J., Donoghue, P.C.J., Hetherington, A.M.,
1624 and Williams, T.A. (2022). Divergent evolutionary trajectories of bryophytes and
1625 tracheophytes from a complex common ancestor of land plants. *Nat. Ecol. Evol.* *6*, 1634–
1626 1643. 10.1038/s41559-022-01885-x.
- 1627 81. Yu, T., Guan, C., Wang, J., Sajjad, M., Ma, L., and Jiao, Y. (2017). Dynamic patterns of gene
1628 expression during leaf initiation. *J. Genet. Genomics* *44*, 599–601. 10.1016/j.jgg.2017.11.001.
- 1629 82. Nakayama, H., Leichty, A.R., and Sinha, N.R. (2022). Molecular mechanisms underlying leaf
1630 development, morphological diversification, and beyond. *Plant Cell* *34*, 2534–2548.
1631 10.1093/plcell/koac118.
- 1632 83. Cance, C., Martin-Arevalillo, R., Boubekour, K., and Dumas, R. (2022). Auxin response
1633 factors are keys to the many auxin doors. *New Phytol.* *235*, 402–419. 10.1111/nph.18159.
- 1634 84. Roosjen, M., Paque, S., and Weijers, D. (2018). Auxin response factors: output control in

- 1635 auxin biology. *J. Exp. Bot.* 69, 179–188. 10.1093/jxb/erx237.
- 1636 85. Ulmasov, T., Hagen, G., and Guilfoyle, T.J. (1997). ARF1, a transcription factor that binds to
1637 auxin response elements. *Science* 276, 1865–1868. 10.1126/science.276.5320.1865.
- 1638 86. Atamian, H.S., Creux, N.M., Brown, E.A., Garner, A.G., Blackman, B.K., and Harmer, S.L.
1639 (2016). Circadian regulation of sunflower heliotropism, floral orientation, and pollinator visits.
1640 *Science* 353, 587–590. 10.1126/science.aaf9793.
- 1641 87. Viaene, T., Landberg, K., Thelander, M., Medvecká, E., Pederson, E., Feraru, E., Cooper,
1642 E.D., Karimi, M., Delwiche, C.F., Ljung, K., et al. (2014). Directional Auxin Transport
1643 Mechanisms in Early Diverging Land Plants. *Curr. Biol.* 24, 2786–2791.
1644 10.1016/j.cub.2014.09.056.
- 1645 88. He, S., Ding, L., Xiong, Z., Spicer, R.A., Farnsworth, A., Valdes, P.J., Wang, C., Cai, F., Wang,
1646 H., Sun, Y., et al. (2022). A distinctive Eocene Asian monsoon and modern biodiversity
1647 resulted from the rise of eastern Tibet. *Sci. Bull.* 67, 2245–2258. 10.1016/j.scib.2022.10.006.
- 1648 89. Peng, M., Shahzad, R., Gul, A., Subthain, H., Shen, S.Q., Lei, L., Zheng, Z.G., Zhou, J.J.,
1649 Lu, D.D., Wang, S.C., et al. (2017). Differentially evolved glucosyltransferases determine
1650 natural variation of rice flavone accumulation and UV-tolerance. *Nat. Commun.* 8, 1975.
1651 10.1038/s41467-017-02168-x.
- 1652 90. Shi, C., and Liu, H.T. (2021). How plants protect themselves from ultraviolet-B radiation stress.
1653 *Plant Physiol.* 187, 1096–1103. 10.1093/plphys/kiab245.
- 1654 91. Kilbride, C., Poole, J., and Hutchings, T.R. (2006). A comparison of Cu, Pb, As, Cd, Zn, Fe,
1655 Ni and Mn determined by acid extraction/ICP-OES and ex situ field portable X-ray
1656 fluorescence analyses. *Environ. Pollut.* 143, 16–23. 10.1016/j.envpol.2005.11.013.
- 1657 92. Gorecka, H., Chojnacka, K., and Gorecki, H. (2006). The application of ICP-MS and ICP-
1658 OES in determination of micronutrients in wood ashes used as soil conditioners. *Talanta* 70,
1659 950–956. 10.1016/j.talanta.2006.05.061.
- 1660 93. Xiao, L.H., Yang, G., Zhang, L.C., Yang, X.H., Zhao, S., Ji, Z.Z., Zhou, Q., Hu, M., Wang, Y.,
1661 Chen, M., et al. (2015). The resurrection genome of *Boea hygrometrica*: A blueprint for
1662 survival of dehydration. *Proc. Natl. Acad. Sci. USA.* 112, 5833–5837.
1663 10.1073/pnas.1505811112.
- 1664 94. Cheng, Z.K., Yan, H.H., Yu, H.X., Tang, S.C., Jiang, J.M., Gu, M.H., and Zhu, L.H. (2001).
1665 Development and applications of a complete set of rice telotrisomics. *Genetics* 157, 361–368.

- 1666 95. Martinez-Garcia, J.F., Monte, E., and Quail, P.H. (1999). A simple, rapid and quantitative
1667 method for preparing *Arabidopsis* protein extracts for immunoblot analysis. *Plant J.* 20, 251–
1668 257. 10.1046/j.1365-313x.1999.00579.x.
- 1669 96. Peng, L.W., Ma, J.F., Chi, W., Guo, J.K., Zhu, S.Y., Lu, Q.T., Lu, C.M., and Zhang, L.X. (2006).
1670 LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in
1671 *Arabidopsis thaliana*. *Plant Cell* 18, 955–969. 10.1105/tpc.105.037689.
- 1672 97. Liu, J., Yang, H.X., Lu, Q.T., Wen, X.G., Chen, F., Peng, L.W., Zhang, L.X., and Lu, C.M.
1673 (2012). PSBP-DOMAIN PROTEIN1, a nuclear-encoded thylakoid lumenal protein, is
1674 essential for photosystem I assembly in *Arabidopsis*. *Plant Cell* 24, 4992–5006.
1675 10.1105/tpc.112.106542.
- 1676 98. Vitha, S., and Osteryoung, K.W. (2011). Immunofluorescence microscopy for localization of
1677 *Arabidopsis* chloroplast proteins. *Methods Mol. Biol.* 774, 33–58. 10.1007/978-1-61779-234-
1678 2_3.
- 1679 99. Vitha, S., McAndrew, R.S., and Osteryoung, K.W. (2001). FtsZ ring formation at the
1680 chloroplast division site in plants. *J. Cell Biol.* 153, 111–119. 10.1083/jcb.153.1.111.
- 1681 100. Cho, H.T., and Cosgrove, D.J. (2002). Regulation of root hair initiation and expansin gene
1682 expression in *Arabidopsis*. *Plant Cell* 14, 3237–3253. 10.1105/tpc.006437.
- 1683 101. Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-
1684 mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. 10.1046/j.1365-
1685 313x.1998.00343.x.
- 1686 102. Ganguly, A., Lee, S.H., Cho, M., Lee, O.R., Yoo, H., and Cho, H.T. (2010). Differential auxin-
1687 transporting activities of PIN-FORMED proteins in *Arabidopsis* root hair cells. *Plant Physiol.*
1688 153, 1046–1061. 10.1104/pp.110.156505.
- 1689 103. Truernit, E., Bauby, H., Dubreucq, B., Grandjean, O., Runions, J., Barthelemy, J., and
1690 Palauqui, J.C. (2008). High-resolution whole-mount imaging of three-dimensional tissue
1691 organization and gene expression enables the study of phloem development and structure
1692 in *Arabidopsis*. *Plant Cell* 20, 1494–1503. 10.1105/tpc.107.056069.
- 1693 104. Xu, B., Ohtani, M., Yamaguchi, M., Toyooka, K., Wakazaki, M., Sato, M., Kubo, M., Nakano,
1694 Y., Sano, R., Hiwatashi, Y., et al. (2014). Contribution of NAC transcription factors to plant
1695 adaptation to land. *Science* 343, 1505–1508. 10.1126/science.1248417.
- 1696 105. Wang, B., Chu, J.F., Yu, T.Y., Xu, Q., Sun, X.H., Yuan, J., Xiong, G.S., Wang, G.D., Wang,

- 1697 Y.H., and Li, J.Y. (2015). Tryptophan-independent auxin biosynthesis contributes to early
1698 embryogenesis in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA.* 112, 4821–4826.
1699 10.1073/pnas.1503998112.
- 1700 106. Estevez, S.L., and Helleur, R. (2005). Fatty acid profiling of lipid classes by silica rod TLC-
1701 thermally assisted hydrolysis and methylation-GC/MS. *J. Anal. Appl. Pyrolysis* 74, 3–10.
1702 10.1016/j.jaap.2004.11.017.
- 1703 107. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017).
1704 Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat
1705 separation. *Genome Res.* 27, 722–736. 10.1101/gr.215087.116.
- 1706 108. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A.,
1707 Zeng, Q.D., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: an Integrated tool for
1708 comprehensive microbial variant detection and genome assembly improvement. *PLoS One*
1709 9, e112963. 10.1371/journal.pone.0112963.
- 1710 109. Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim,
1711 M.S., Machol, I., Lander, E.S., Aiden, A.P., and Aiden, E.L. (2017). De novo assembly of the
1712 *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* 356, 92–95.
1713 10.1126/science.aal3327.
- 1714 110. Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and
1715 Aiden, E.L. (2016). Juicebox provides a visualization system for Hi-C contact maps with
1716 unlimited zoom. *Cell Syst.* 3, 99–101. 10.1016/j.cels.2015.07.012.
- 1717 111. Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-
1718 assembled contigs using SSPACE. *Bioinformatics* 27, 578–579.
1719 10.1093/bioinformatics/btq683.
- 1720 112. English, A.C., Richards, S., Han, Y., Wang, M., Vee, V., Qu, J.X., Qin, X., Muzny, D.M., Reid,
1721 J.G., Worley, K.C., and Gibbs, R.A. (2012). Mind the gap: upgrading genomes with Pacific
1722 Biosciences RS long-read sequencing technology. *PLoS One* 7, e47768.
1723 10.1371/journal.pone.0047768.
- 1724 113. Simao, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015).
1725 BUSCO: assessing genome assembly and annotation completeness with single-copy
1726 orthologs. *Bioinformatics* 31, 3210–3212. 10.1093/bioinformatics/btv351.
- 1727 114. Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length

- 1728 LTR retrotransposons. *Nucleic Acids Res.* 35, W265–W268. 10.1093/nar/gkm286.
- 1729 115. Nussbaumer, T., Martis, M.M., Roessner, S.K., Pfeifer, M., Bader, K.C., Sharma, S., Gundlach,
1730 H., and Spannagl, M. (2013). MIPS PlantsDB: a database framework for comparative plant
1731 genome research. *Nucleic Acids Res.* 41, D1144–D1151. 10.1093/nar/gks1153.
- 1732 116. Lopez, R., Silventoinen, V., Robinson, S., Kibria, A., and Gish, W. (2003). WU-Blast2 server
1733 at the European Bioinformatics Institute. *Nucleic Acids Res.* 31, 3795–3798.
1734 10.1093/nar/gkg573.
- 1735 117. Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic
1736 Acids Res.* 27, 573–580. 10.1093/nar/27.2.573.
- 1737 118. Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible
1738 software for de novo detection of LTR retrotransposons. *BMC Bioinf.* 9, 18. 10.1186/1471-
1739 2105-9-18.
- 1740 119. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H.,
1741 Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version
1742 2.0. *Bioinformatics* 23, 2947–2948. 10.1093/bioinformatics/btm404.
- 1743 120. Yang, Z.H. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.*
1744 24, 1586–1591. 10.1093/molbev/msm088.
- 1745 121. Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic
1746 DNA. *J. Mol. Biol.* 268, 78–94. 10.1006/jmbi.1997.0951.
- 1747 122. Stanke, M., and Waack, S. (2003). Gene prediction with a hidden Markov model and a new
1748 intron submodel. *Bioinformatics* 19, II215–II225. 10.1093/bioinformatics/btg1080.
- 1749 123. Cantarel, B.L., Korf, I., Robb, S.M.C., Parra, G., Ross, E., Moore, B., Holt, C., Alvarado, A.S.,
1750 and Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging
1751 model organism genomes. *Genome Res.* 18, 188–196. 10.1101/gr.6743907.
- 1752 124. Bateman, A., Martin, M.J., O'Donovan, C., Magrane, M., Apweiler, R., Alpi, E., Antunes, R.,
1753 Ar-Ganiska, J., Bely, B., Bingley, M., et al. (2015). UniProt: a hub for protein information.
1754 *Nucleic Acids Res.* 43, D204–D212. 10.1093/nar/gku989.
- 1755 125. Zdobnov, E.M., and Apweiler, R. (2001). InterProScan - an integration platform for the
1756 signature-recognition methods in InterPro. *Bioinformatics* 17, 847–848.
1757 10.1093/bioinformatics/17.9.847.
- 1758 126. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P.,

- 1759 Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for the unification of
1760 biology. *Nature Genet.* 25, 25–29. 10.1038/75556.
- 1761 127. Xue, J.Y., Liu, Y., Li, L.B., Wang, B., and Qiu, Y.L. (2010). The complete mitochondrial
1762 genome sequence of the hornwort *Phaeoceros laevis*: retention of many ancient
1763 pseudogenes and conservative evolution of mitochondrial genomes in hornworts. *Curr.*
1764 *Genet.* 56, 53–61. 10.1007/s00294-009-0279-1.
- 1765 128. Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., and Greiner,
1766 S. (2017). GeSeq - versatile and accurate annotation of organelle genomes. *Nucleic Acids*
1767 *Res.* 45, W6–W11. 10.1093/nar/gkx391.
- 1768 129. Tang, H.B., Bowers, J.E., Wang, X.Y., Ming, R., Alam, M., and Paterson, A.H. (2008). Synteny
1769 and collinearity in plant genomes. *Science* 320, 486–488. 10.1126/science.1153917.
- 1770 130. Sun, P., Jiao, B., Yang, Y., Shan, L., Li, T., Li, X., Xi, Z., Wang, X., and Liu, J. (2022). WGDI:
1771 a user-friendly toolkit for evolutionary analyses of whole-genome duplications and ancestral
1772 karyotypes. *Mol. Plant.* 15, 1841–1851. 10.1016/j.molp.2022.10.018.
- 1773 131. Qiao, X., Li, Q.H., Yin, H., Qi, K.J., Li, L.T., Wang, R.Z., Zhang, S.L., and Paterson, A.H.
1774 (2019). Gene duplication and evolution in recurring polyploidization-diploidization cycles in
1775 plants. *Genome Biol.* 20, 38. 10.1186/s13059-019-1650-2.
- 1776 132. Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs_Calculator 2.0: a toolkit
1777 incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinf.*
1778 8, 77–80. 10.1016/s1672-0229(10)60008-3.
- 1779 133. Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein
1780 sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34,
1781 W609–W612. 10.1093/nar/gkl315.
- 1782 134. Li, L., Stoeckert, C.J., and Roos, D.S. (2003). OrthoMCL: Identification of ortholog groups for
1783 eukaryotic genomes. *Genome Res.* 13, 2178–2189. 10.1101/gr.1224503.
- 1784 135. Ebersberger, I., Strauss, S., and von Haeseler, A. (2009). HaMStR: Profile hidden markov
1785 model based search for orthologs in ESTs. *BMC Evol. Biol.* 9, 157. 10.1186/1471-2148-9-
1786 157.
- 1787 136. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high
1788 throughput. *Nucleic Acids Res.* 32, 1792–1797. 10.1093/nar/gkh340.
- 1789 137. Capella-Gutierrez, S., Silla-Martinez, J.M., and Gabaldon, T. (2009). trimAl: a tool for

- 1790 automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25,
1791 1972–1973. 10.1093/bioinformatics/btp348.
- 1792 138. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and
1793 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.*
1794 32, 268–274. 10.1093/molbev/msu300.
- 1795 139. Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: polynomial time
1796 species tree reconstruction from partially resolved gene trees. *BMC Bioinf.* 19, 153.
1797 10.1186/s12859-018-2129-y.
- 1798 140. Morris, J.L., Puttick, M.N., Clark, J.W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H.,
1799 Yang, Z.H., Schneider, H., and Donoghue, P.C.J. (2018). The timescale of early land plant
1800 evolution. *Proc. Natl. Acad. Sci. USA.* 115, E2274–E2283. 10.1073/pnas.1719588115.
- 1801 141. Hedges, S.B., Tao, Q.Q., Walker, M., and Kumar, S. (2018). Accurate timetrees require
1802 accurate calibrations. *Proc. Natl. Acad. Sci. USA.* 115, E9510–E9511.
1803 10.1073/pnas.1812558115.
- 1804 142. Morris, J.L., Puttick, M.N., Clark, J.W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H.,
1805 Yang, Z.H., Schneider, H., and Donoghue, P.C.J. (2018). Reply to Hedges et al.: Accurate
1806 timetrees do indeed require accurate calibrations. *Proc. Natl. Acad. Sci. USA.* 115, E9512–
1807 E9513. 10.1073/pnas.1812816115.
- 1808 143. Sanderson, M.J. (2002). Estimating absolute rates of molecular evolution and divergence
1809 times: A penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109.
1810 10.1093/oxfordjournals.molbev.a003974.
- 1811 144. Smith, S.A., and O'Meara, B.C. (2012). treePL: divergence time estimation using penalized
1812 likelihood for large phylogenies. *Bioinformatics* 28, 2689–2690.
1813 10.1093/bioinformatics/bts492.
- 1814 145. Nielsen, R., and Yang, Z.H. (1998). Likelihood models for detecting positively selected amino
1815 acid sites and applications to the HIV-1 envelope gene. *Genetics* 148, 929–936.
- 1816 146. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid
1817 multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–
1818 3066. 10.1093/nar/gkf436.
- 1819 147. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use
1820 in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552.

1821 10.1093/oxfordjournals.molbev.a026334.

1822 148. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis
1823 of large phylogenies. *Bioinformatics* 30, 1312–1313. 10.1093/bioinformatics/btu033.

1824 149. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina
1825 sequence data. *Bioinformatics* 30, 2114–2120. 10.1093/bioinformatics/btu170.

1826 150. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat.*
1827 *Methods* 9, 357–359. 10.1038/nmeth.1923.

1828 151. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package
1829 for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–
1830 140. 10.1093/bioinformatics/btp616.

1831 152. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernandez-Plaza, A., Forslund, S.K., Cook, H.,
1832 Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., et al. (2019). eggNOG 5.0: a hierarchical,
1833 functionally and phylogenetically annotated orthology resource based on 5090 organisms
1834 and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. 10.1093/nar/gky1085.

1835 153. Cantalapiedra, C.P., Hernandez-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021).
1836 eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at
1837 the metagenomic scale. *Mol. Biol. Evol.* 38, 5825–5829. 10.1093/molbev/msab293.

1838 154. Yu, G.C., Wang, L.G., Han, Y.Y., and He, Q.Y. (2012). clusterProfiler: an R package for
1839 comparing biological themes among gene clusters. *Omics* 16, 284–287.
1840 10.1089/omi.2011.0118.

1841

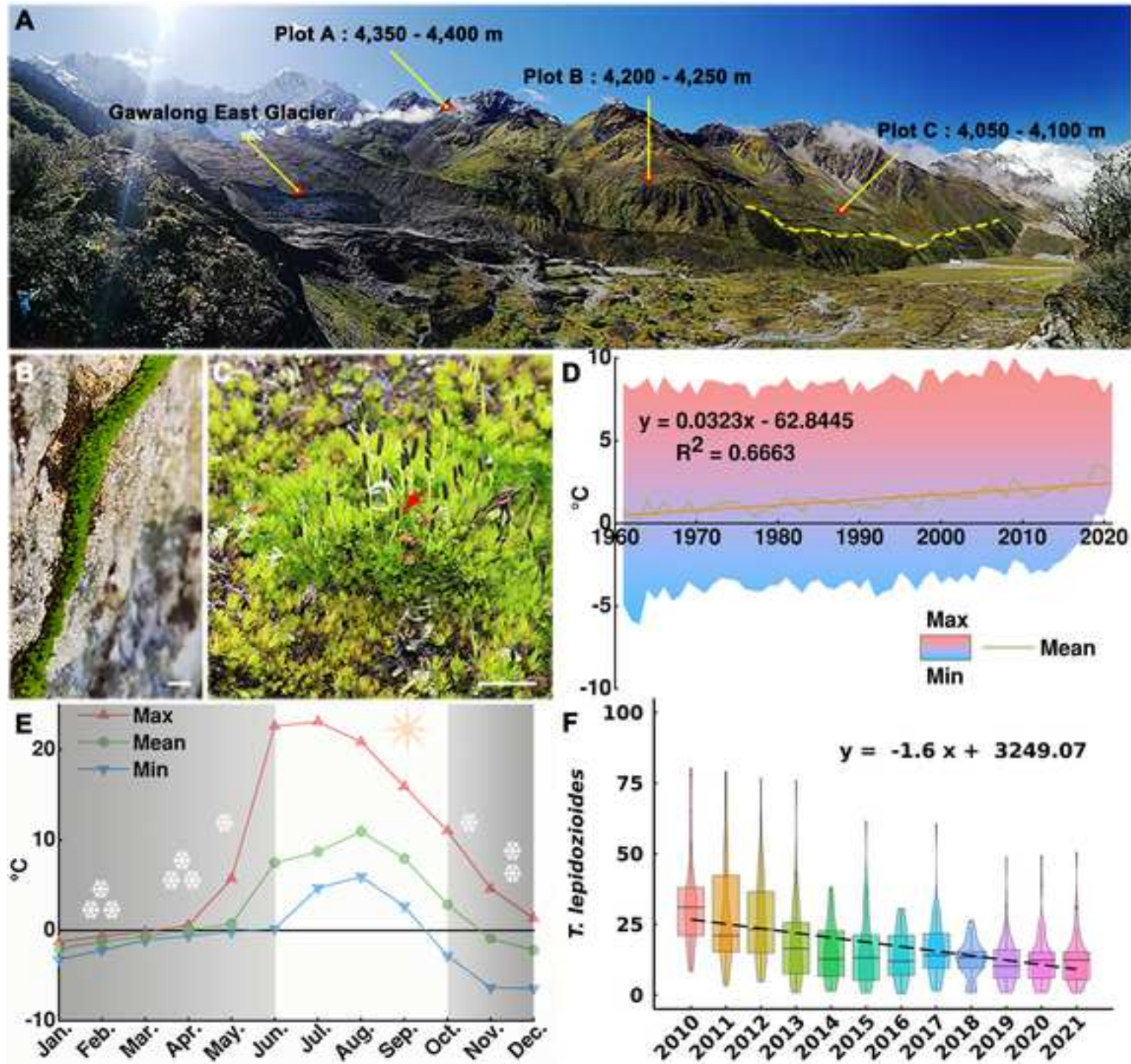
Key resources table

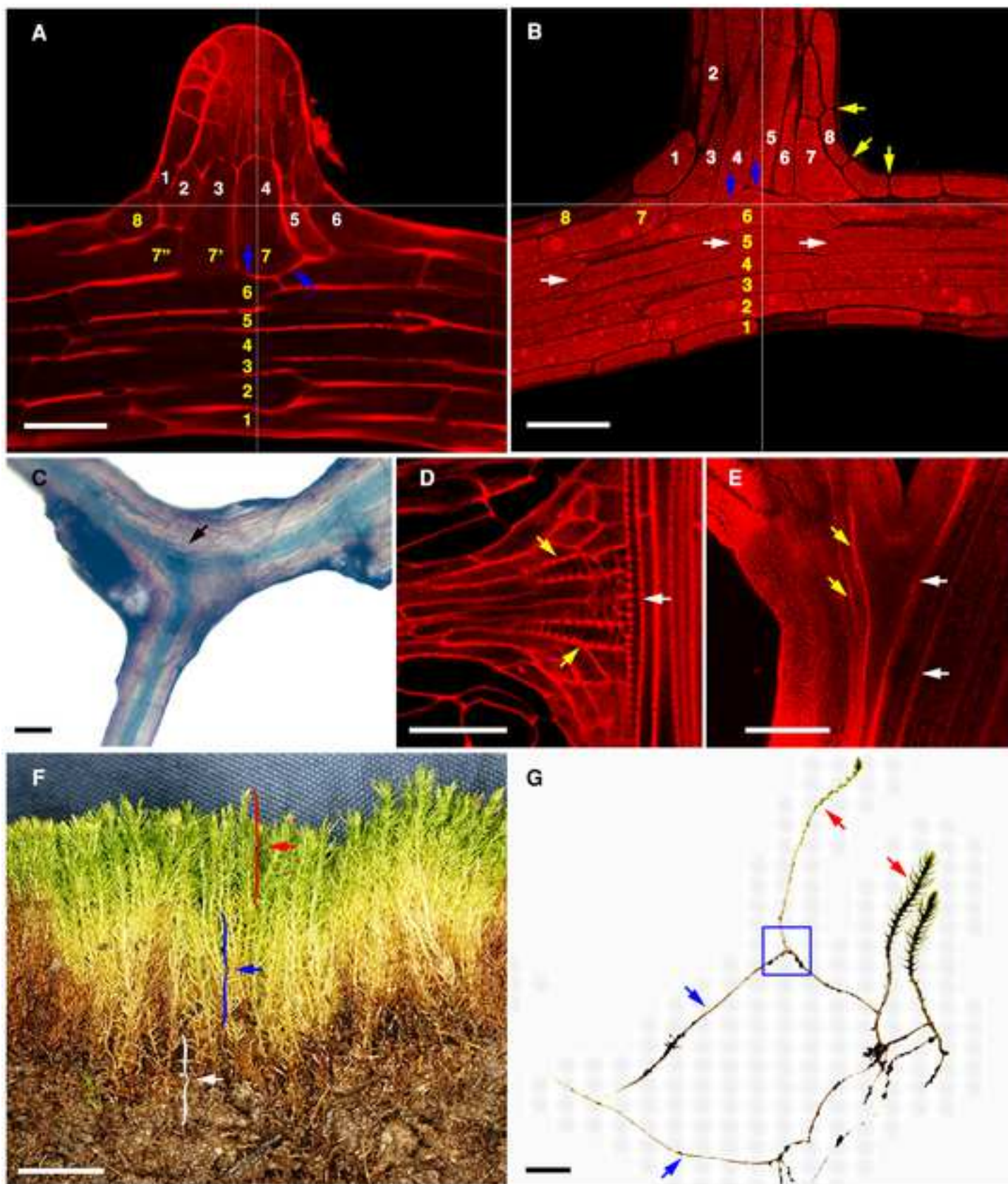
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
AtCP47	97	N/A
AtPsaD	97	N/A
AtPIN1	PhytoAB	PHY1046S
Bacterial and virus strains		
Biological samples		
<i>Takakia lepidozoides</i>	This study	IMSC: 41271
Chemicals, peptides, and recombinant proteins		
Critical commercial assays		
Deposited data		
<i>Takakia lepidozoides</i> : Pacbio CCS, Hi-C, Illumina DNA and RNA sequencing data	This study	NCBI BioProject: PRJNA796444
<i>Takakia lepidozoides</i> : Pacbio CCS, Hi-C, Illumina DNA and RNA sequencing data	This study	NGDC BioProject: PRJCA014819
<i>Takakia lepidozoides</i> genome assembly and gene annotations	This study	https://www.takakia.com
<i>Takakia lepidozoides</i> : Metabolomics raw data	This study	MetaboLights : MTBLS6999
Experimental models: Cell lines		

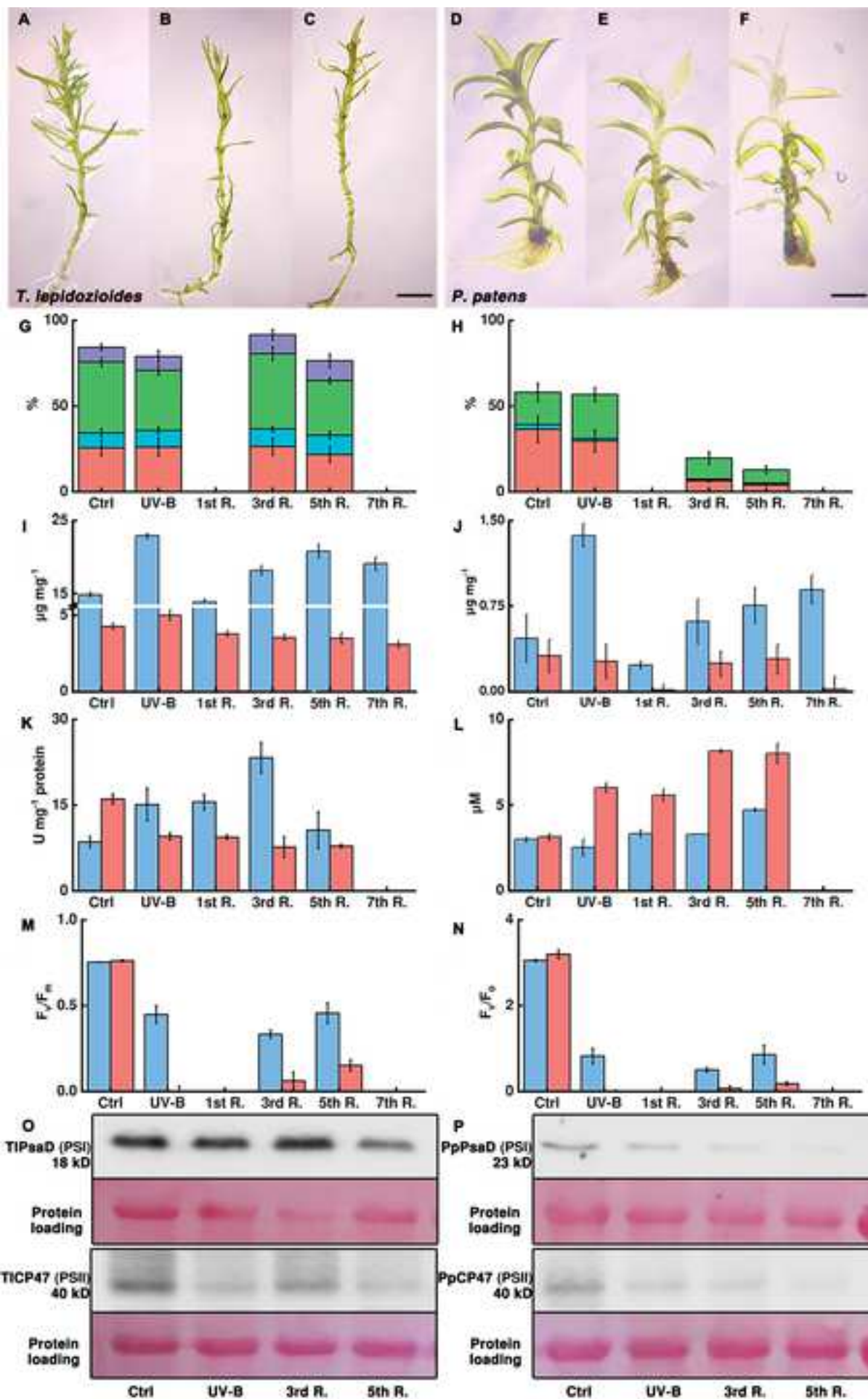
Experimental models: Organisms/strains		
Oligonucleotides		
Primers used for cloning and functional analysis of PIN (see Table S6 A)	This study	N/A
Recombinant DNA		
Software and algorithms		
Vegan version 2.5-7		https://CRAN.R-project.org/package=vegan
MaxEnt version 3.4.1		https://biodiversityinformatics.amnh.org/open_source/maxent/
Canu version 2.0	107	https://github.com/marbl/canu
Arrow version 2.3.1		https://github.com/PacificBiosciences/GenomicConsensus
Pilon version 1.22	108	https://github.com/broadinstitute/pilon
3D-DNA version 190716	109	https://github.com/aidenlab/3d-dna
Juicer version 1.11.08	110	https://github.com/aidenlab/juicer
SSPACE version 2.1.1	111	https://github.com/nsoranzo/sspace_basic
PBJelly version 15.8.24	112	https://sourceforge.net/projects/pb-jelly/
BUSCO version 4.0.6	113	https://busco.ezlab.org/
LTR_FINDER version 1.07	114	https://github.com/xzhu/LTR_Finder
RepeatModeler version 1.0.11		http://www.repeatmasker.org/RepeatModeler/
RepeatMasker version 4.0.8		http://www.repeatmasker.org/

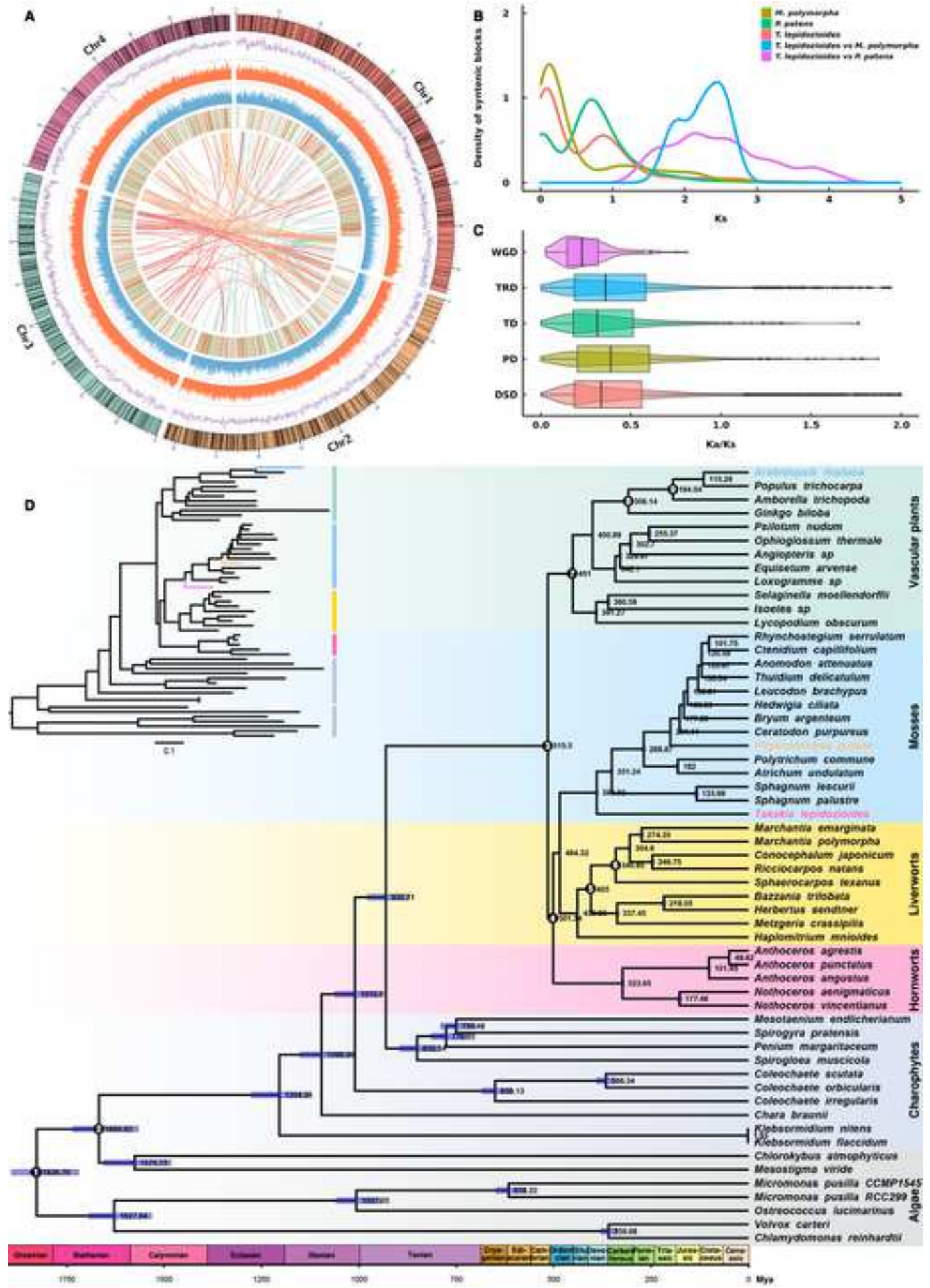
WU-BLAST	116	http://blast.wustl.edu/
Tandem Repeats Finder version 4.04	117	https://tandem.bu.edu/trf/trf.html
LTRharvest version 1.6.1	118	http://genometools.org/tools/gt_ltrharvest.html
ClustalW version 1.81	119	http://www.clustal.org/
PAML version 4.9j	120	http://abacus.gene.ucl.ac.uk/software/paml.html
Genscan version 1.0	121	http://genes.mit.edu/GENSCAN.html
Augustus version 3.3.3	122	http://augustus.gobics.de/
MAKER version 3.01.02	123	http://www.yandell-lab.org/software/maker.html
InterProScan version 5.44-79.0	125	https://github.com/ebi-pf-team/interproscan
GeSeq	128	https://chlorobox.mpi-golm.mpg.de/geseq.html
MCSscan version 1.06	129	https://github.com/tanghaibao/jcvi
WGDI version 0.5.3	130	https://github.com/SunPengChuan/wgdi
DupGen_finder	131	https://github.com/qiao-xin/DupGen_finder
KaKs_Calculator version 2.0	132	https://sourceforge.net/projects/kakscalculator2/
PAL2NAL version 14	133	http://www.bork.embl.de/pal2nal/
OrthoMCL version 2.0.9	134	https://legacy.orthomcl.org/orthomcl.b5/
HaMStR version 1.6.0	135	https://github.com/BI-ONF/HaMStR
MUSCLE version 3.8.31	136	http://www.drive5.com/muscle
trimAl version 1.4	137	http://trimal.cgenomics.org
IQ-tree version 1.6.12	138	http://www.iqtree.org/
ASTRAL version 5.7.1	139	https://github.com/smirarab/ASTRAL
treePL version 1.0	144	https://github.com/blackrim/treePL
TAPscan	43	https://plantcode.uni-marburg.de/tapscan/

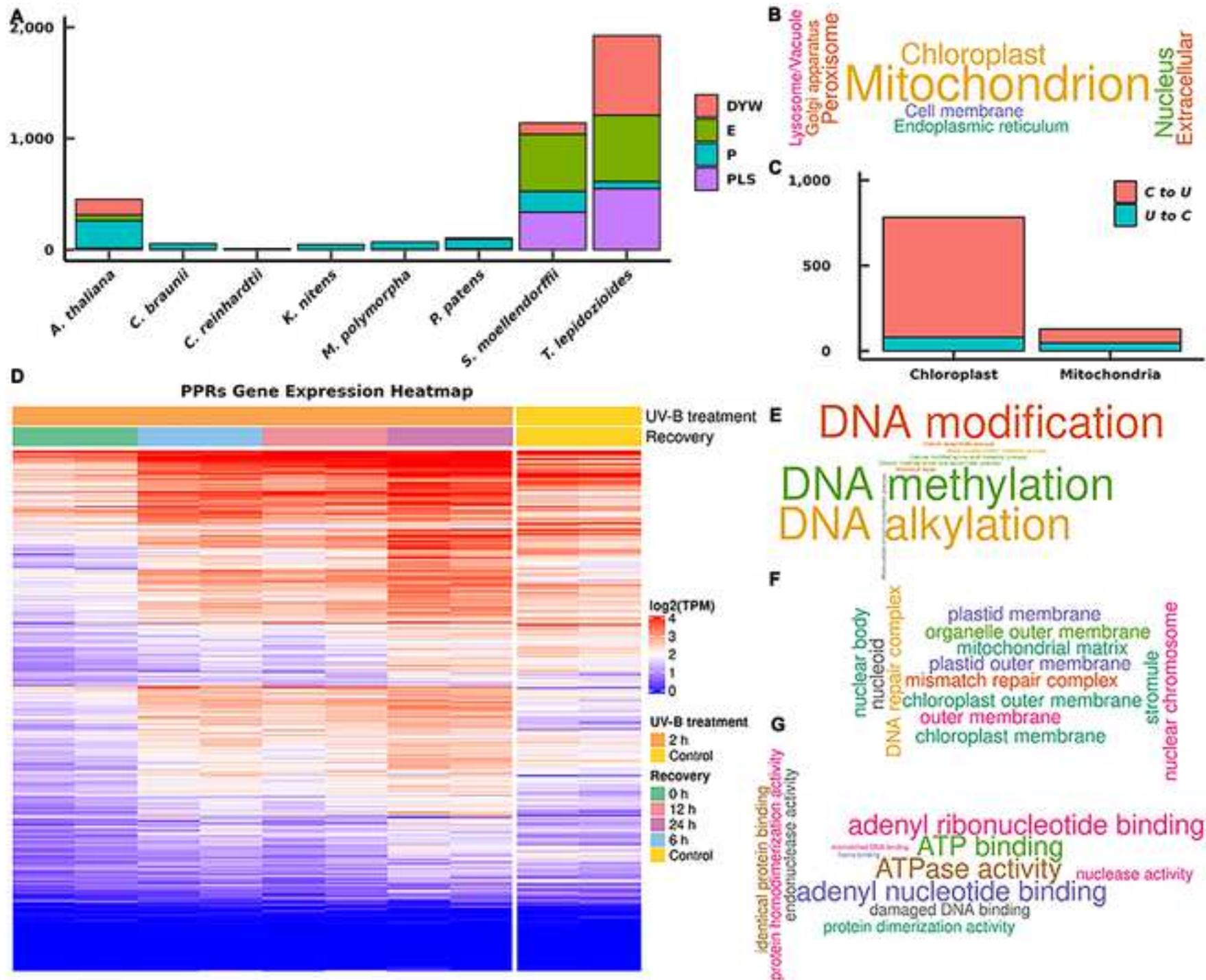
MAFFT version 7.453	146	https://mafft.cbrc.jp/alignment/software/
Gblocks version 0.91b	147	http://molevol.cmima.csic.es/castresana/Gblocks.html
RAxML version 8.2.12	148	https://cme.h-its.org/exelixis/web/software/raxml/
CAFÉ version 4.2.1		https://hahnlab.github.io/CAFE/
Trimmomatic version 0.39	149	http://www.usadellab.org/cms/?page=trimmomatic
Bowtie 2 version 2.4.1	150	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
RSEM version 1.3.3		https://github.com/siddharthab/RSEM
edgeR version 3.28.1	151	https://bioconductor.org/packages/release/bioc/html/edgeR.html
eggNOG-mapper version 2.0.1	153	http://eggNOG-mapper.embl.de/
clusterProfiler version 3.14.3	154	http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html
Other		
WorldClim version 1.4		https://www.worldclim.org/
PGSB Repeat Element Database	115	http://pgsb.helmholtz-muenchen.de/plant/repeat/
Repbase		http://www.girinst.org/rebase
UniProtKB/Swiss-Prot and UniProtKB/TrEMBL	124	https://www.uniprot.org/
Gene Ontology	126	http://geneontology.org/
DAVID		https://david.ncifcrf.gov/home.jsp
Pfam		https://pfam.xfam.org/
EggNOG version 5.0	152	http://eggNOG5.embl.de/
Plants used for phylogenetic analysis (see Table S4 A)	This study	N/A
Calibrations used in age estimation (see Table S6 B)	This study	N/A
Plants used for testing fast evolution and positive selection analysis (see Table S6 C)	This study	N/A
Plants used for estimating the time of morphogenetic genes lost (see Table S6 D)	This study	N/A

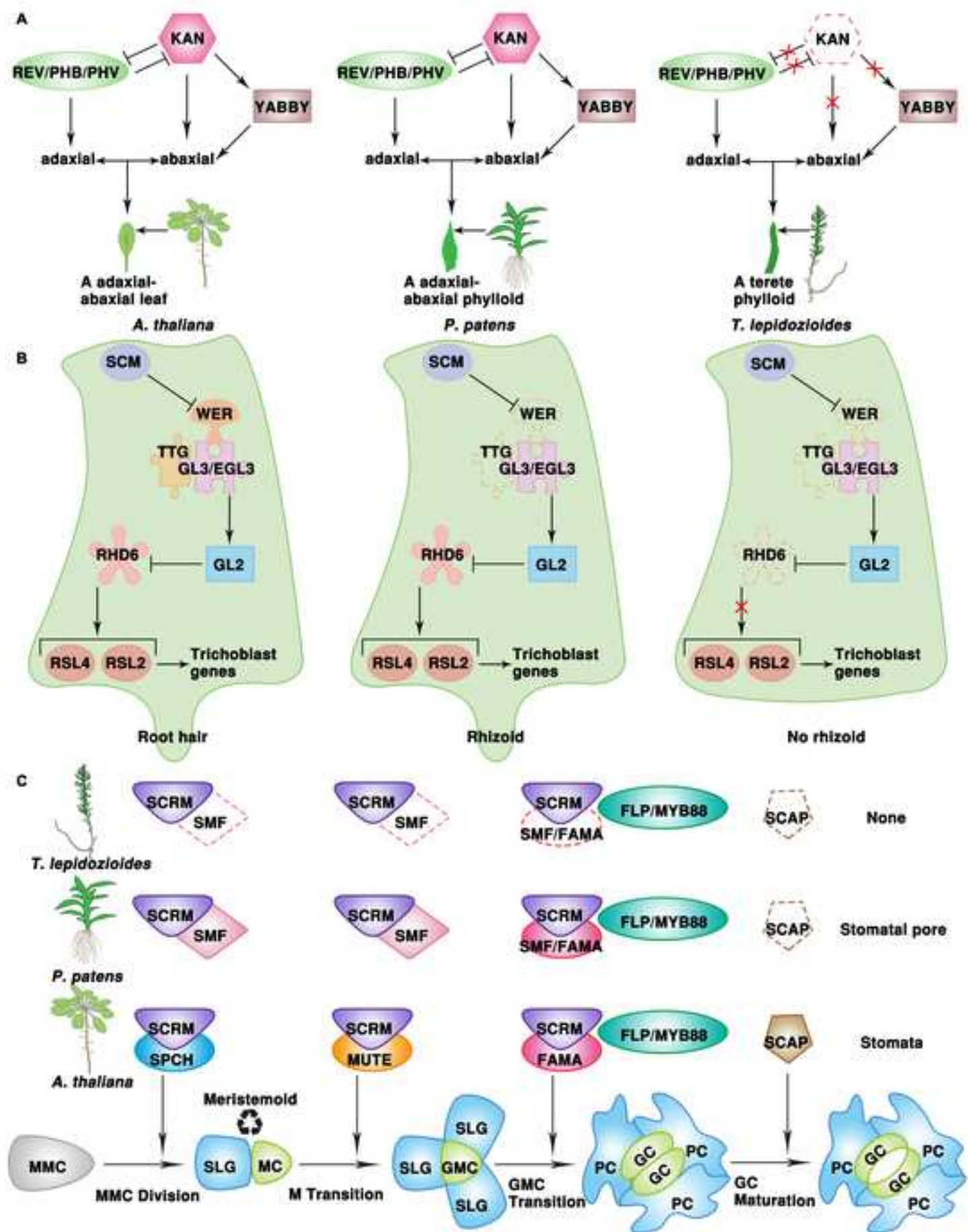






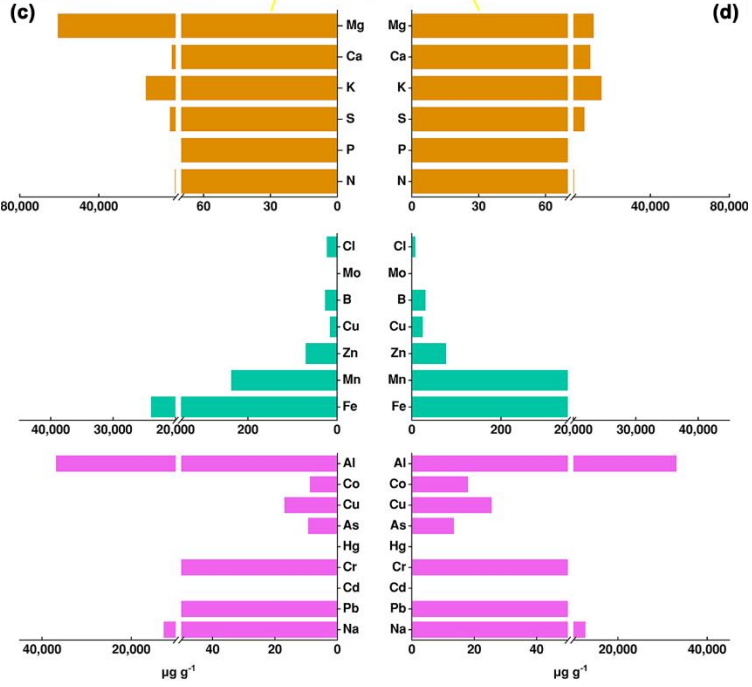
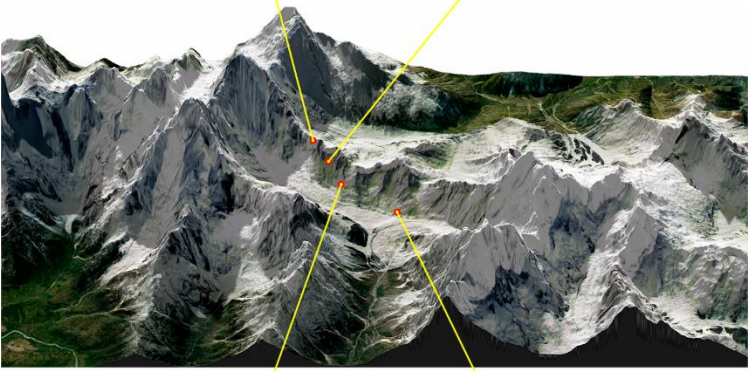
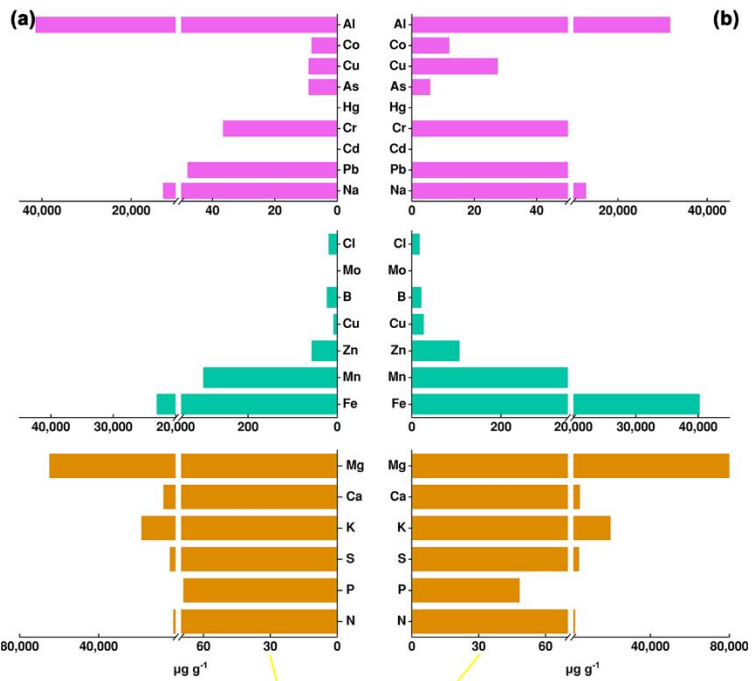






1 **Supplement data 1**

2

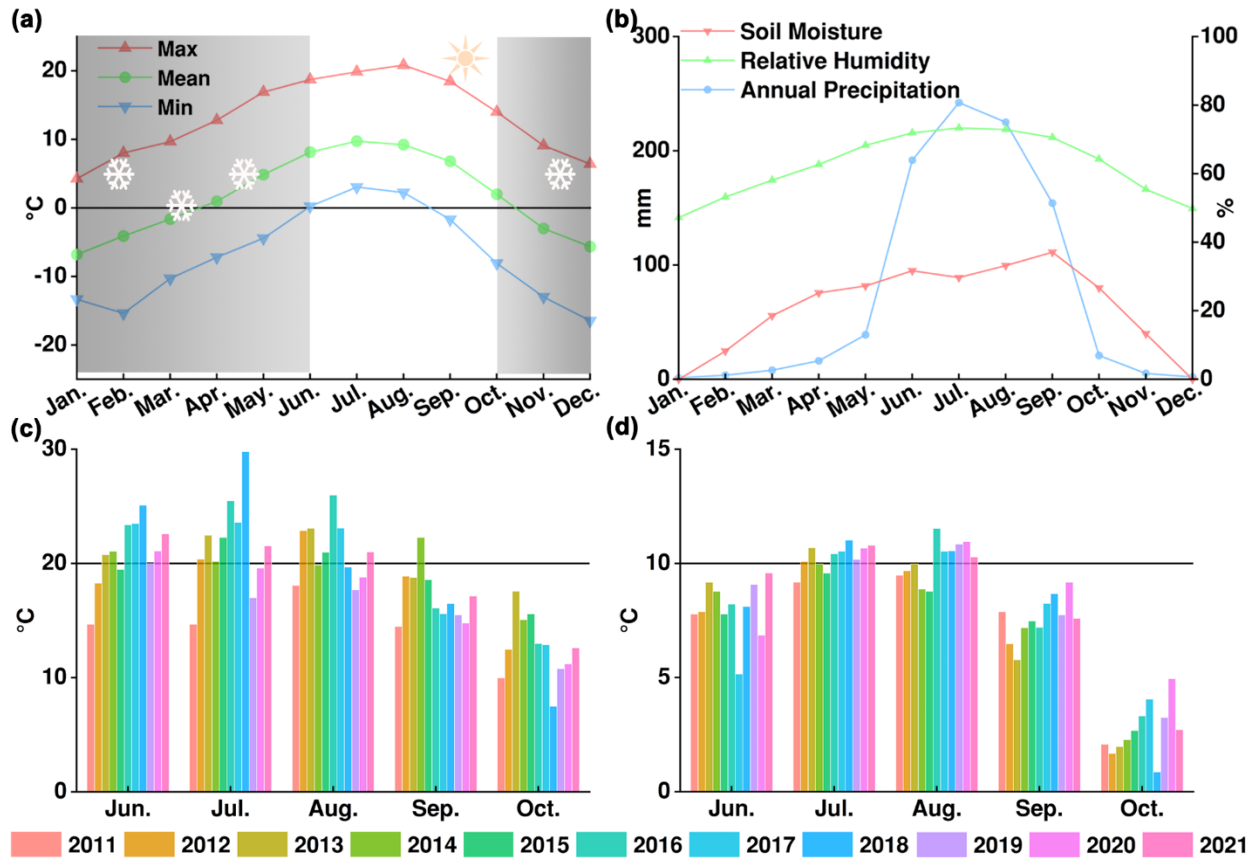


4

5 **Data S1 A. Soil mineral elements in the experimental area**

6 The soil mineral element components in plots with *T. lepidozooides* distribution (**a – c**) and
7 without *T. lepidozooides* distribution (**d**). Red dots show the exact location of experiment plots in
8 the area corresponding to Figure 1A from a different perspective. Pink color represents metal
9 elements, green color represents microelements, and brown color represents macroelements.

10



11

12

13 **Data S1 B. Climate change from 2011 to 2021 in the experimental area**

14 The climate change in the survey area (plot B) from 2011 to 2021. (a) monthly temperature, (b)
 15 monthly relative humidity, precipitation and soil moisture, (c) monthly maximum temperature
 16 during growing season (June to October), (d) monthly average temperature during growing
 17 season (June to October).

18

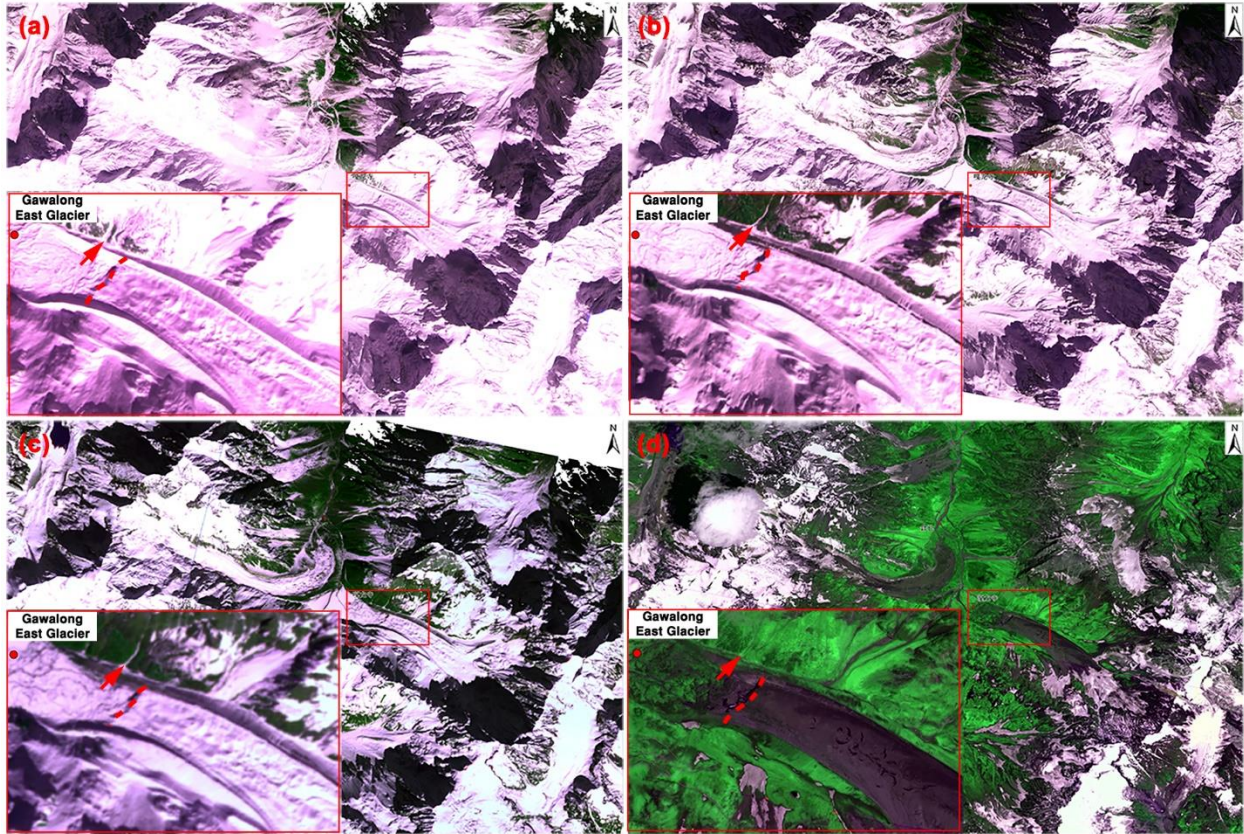


19
20

21 **Data S1 C. Photographs of the edge of the Gawalong East Glacier**

22 Photographs of the edge of the glacier from 2011 to 2021. Years are indicated in red. The
 23 yellow arrows show the mark of water flow (in a fixed position since 2005) near Gawalong East
 24 Glacier. The yellow dotted lines show the edge of the glacier. The data of glacier retreat was
 25 calculated from the measured distance between yellow arrow and dotted line and is compiled in
 26 Table S1B.

27

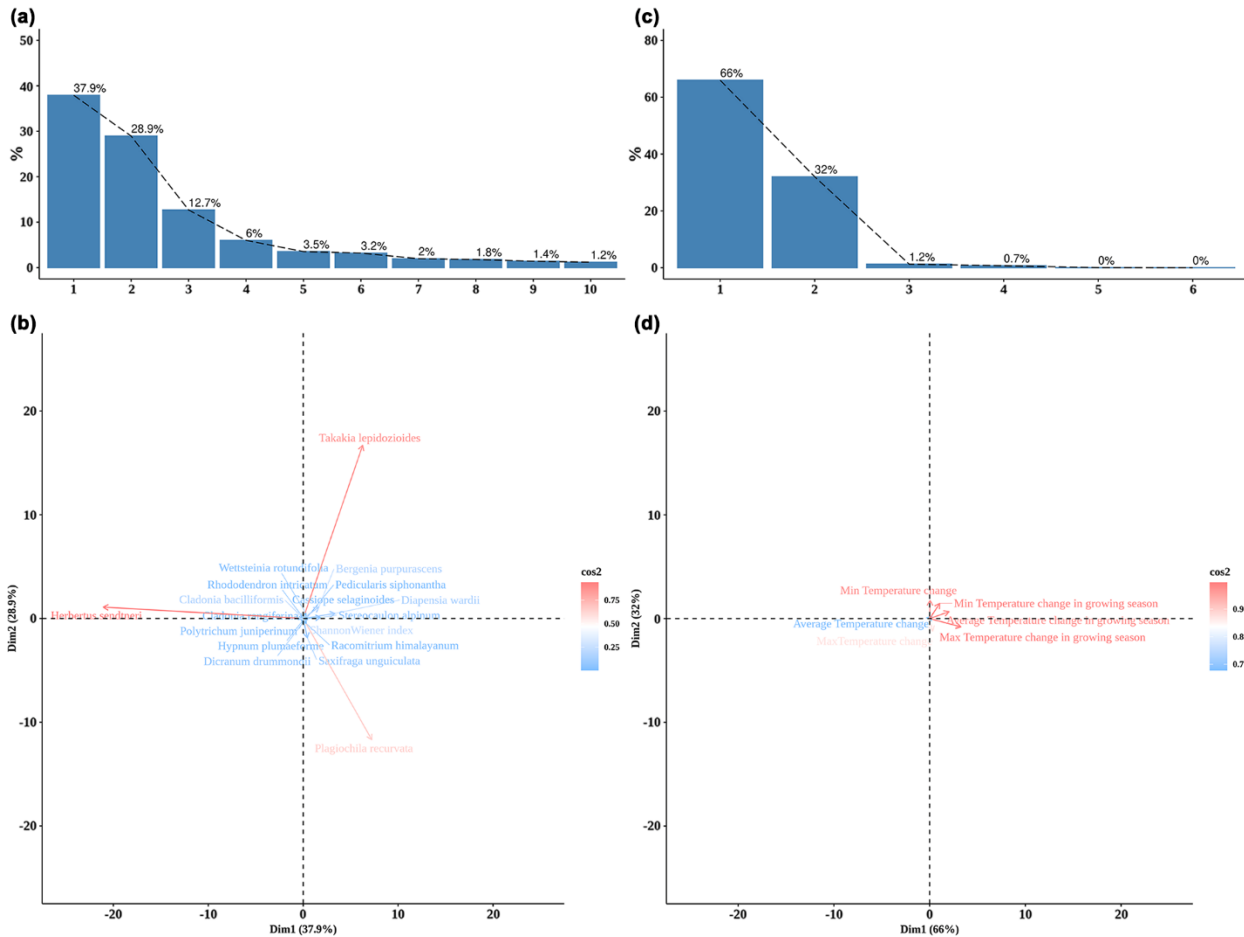


28
29

30 **Data S1 D. Satellite images of the edge of the Gawalong East Glacier**

31 Satellite images of the glacier from Resources-3 (a, b) and Gaofen-1 (c, d). (a) Captured on Jan.
 32 9th, 2013, (b) captured on Feb. 1st, 2014, (c) captured on Nov. 26th, 2013, (d) captured on Jul.
 33 25th, 2015. Red boxes represent magnified views of the glacier. The red arrows show the mark
 34 of water flow (in a fixed position since 2005) near Gawalong East Glacier. The red dotted lines
 35 show the edge of the glacier. The space resolution ratio of both satellites is at 2 m, data from
 36 the same satellite at different year was used to estimate glacier ablation. Detailed data of glacier
 37 ablation estimation is compiled in Table S1B.

38

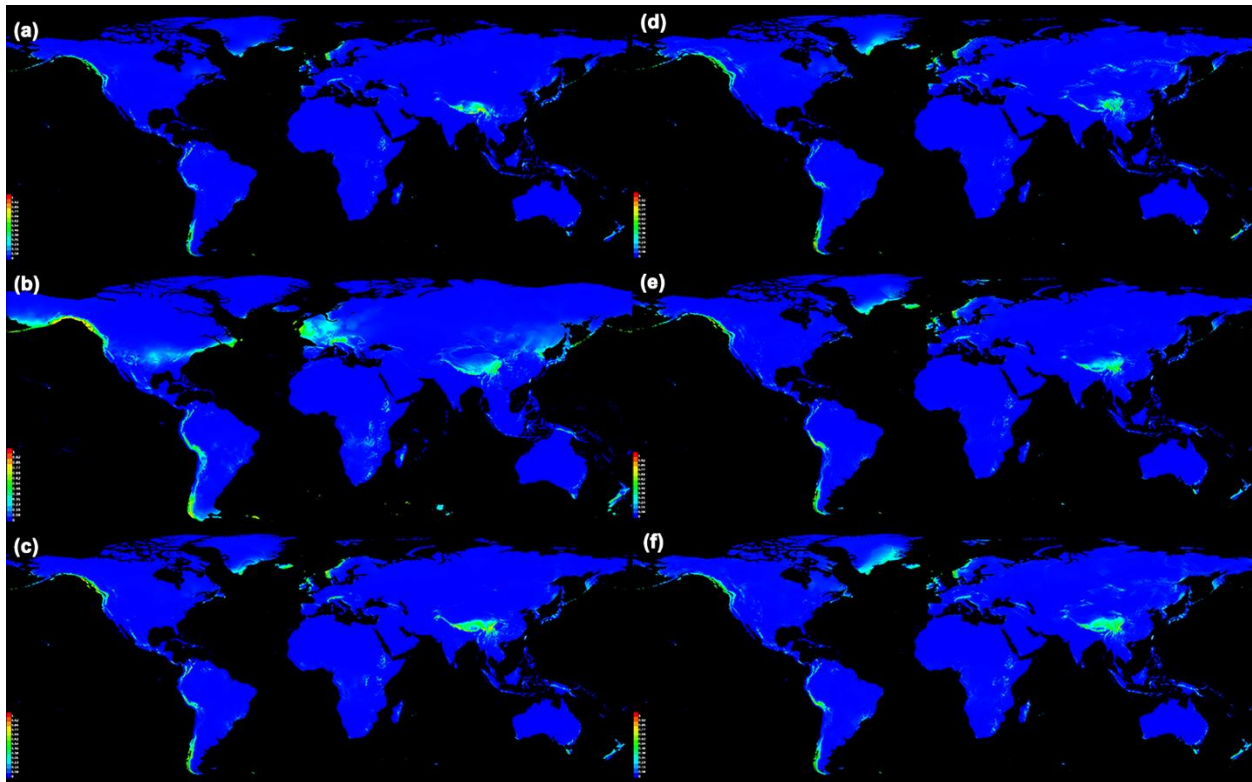


39
40

41 **Data S1 E. Correlation analysis of population changes and climatic changes**

42 Correlations among species population changes and climatic variables changes. (a) Principal
 43 component analysis (PCA) of species population changes, depicted 3 principal component
 44 contribution rate came up to 79.5 %. (b) Redundancy analysis (RDA) of species population
 45 changes, depicted *T. lepidozoioides*, *H. sendtneri* and *P. recurvata* populations changed flexibly.
 46 (c) PCA analysis of climatic variables, depicted 2 principal component contribution rate came up
 47 to 98%. (d) Canonical correspondence analysis (CCA) of climatic variables, depicted
 48 temperature in growing season significantly affected the populations. All multiple regressions
 49 and backward eliminations were performed by vegan package ([https://CRAN.R-](https://CRAN.R-project.org/package=vegan)
 50 [project.org/package=vegan](https://CRAN.R-project.org/package=vegan)) in R.

51



52
53

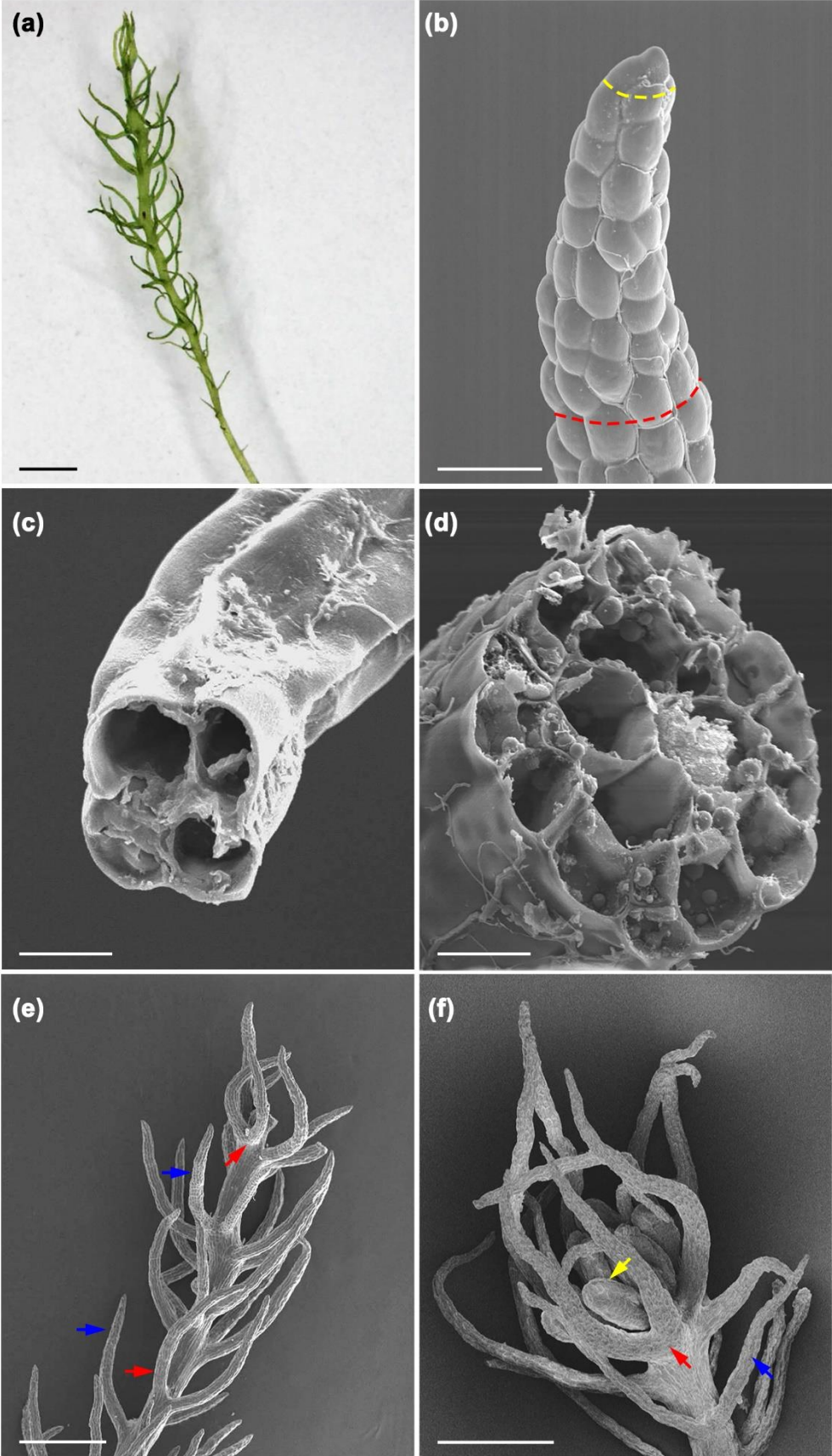
54 **Data S1 F. Potential distribution of *T. lepidozoides* over time**

55 Potential distribution of *T. lepidozoides* (a) in the last interglaciation (~120,000 – 140,000 years ago), (b) in the last glaciation maximum (about 22,000 years ago), (c) in the Holocene mega thermal (about 6,000 years ago), (d) 1970s to 2000s, (e) 2041s to 2060s, (f) 2061s to 2080s.

57 Suitable conditions were classified according to Liu et al. [S1], regions with mean predicted value > 0.68 defined as suitable conditions, sites with value in 0.38 ~ 0.68 were taken as typical candidate distributions and predicted value lower than 0.38 were taken as pseudo absences.

61 Colors indicate predicted probability that conditions are suitable, with warmer colors indicating high probability of suitable conditions for the species, green indicating conditions typical of those where the species is found, and lighter shades of blue indicating low predicted probability of suitable conditions.

65

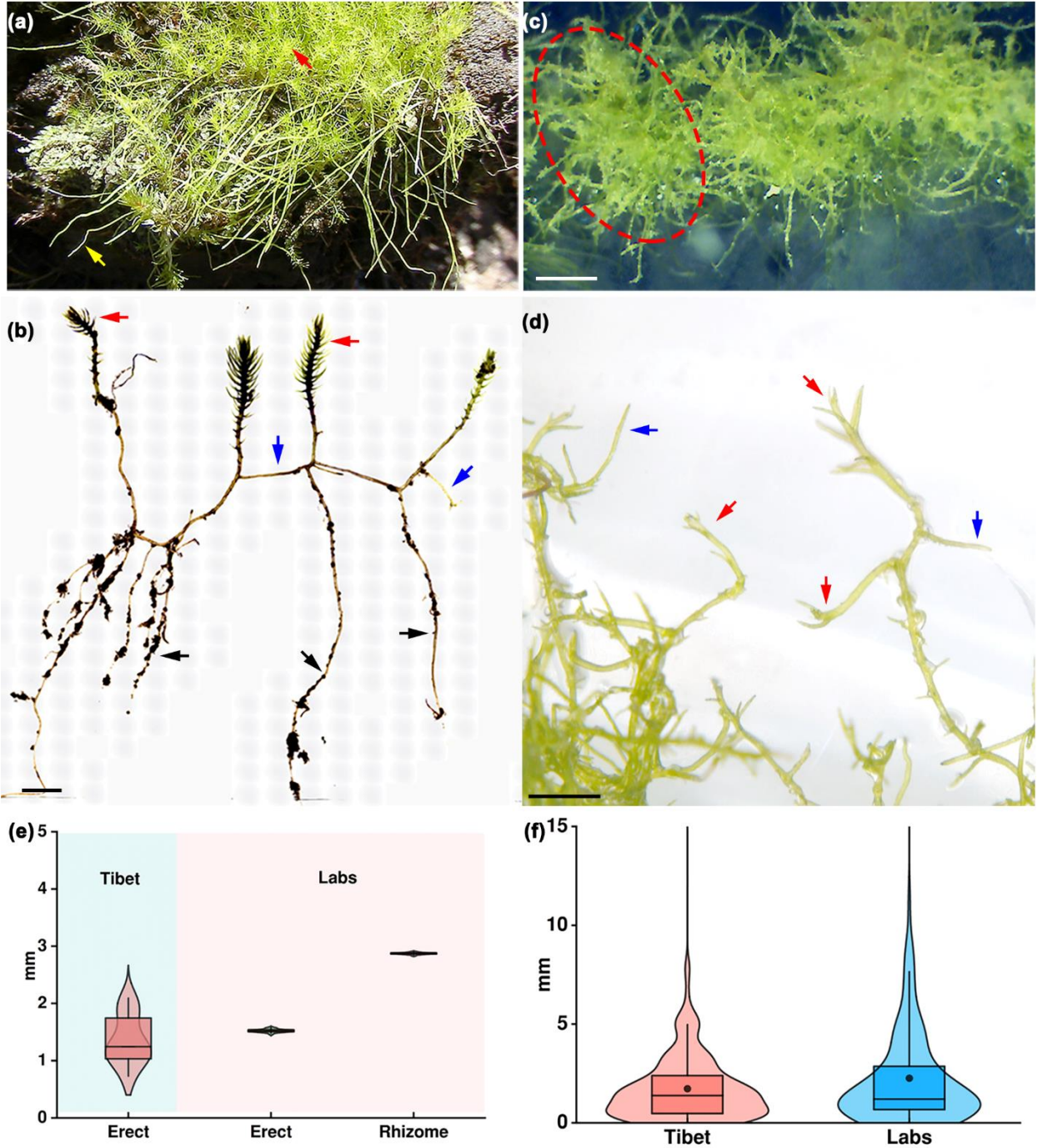


67

68 **Data S1 G. The leaf shape of *T. lepidozoides***

69 Plants are from the population of Plot B in the field. (a) An erect stem with finger-shaped leaves.
70 (b – d) SEM images of a leaf (b), the transection (c) from the tip (yellow circle in b) and the
71 transection (d) from below (red circle in b). (e – f) Arrows mark the single finger-shaped leaf
72 (blue), finger-shaped leaf connected at the base (red) and antheridia (yellow arrow) located at
73 the base of leaves on top of an erect stem. The picture in (a) was made using a stereo
74 microscope (M125 C, Leica), pictures from (b) to (f) were made with a scanning electron
75 microscope (S4800, Hitachi). Bar = 1 mm in (a), 50 μm in (b), 20 μm in (c) and (d), 500 μm in
76 (e) and (f).

77



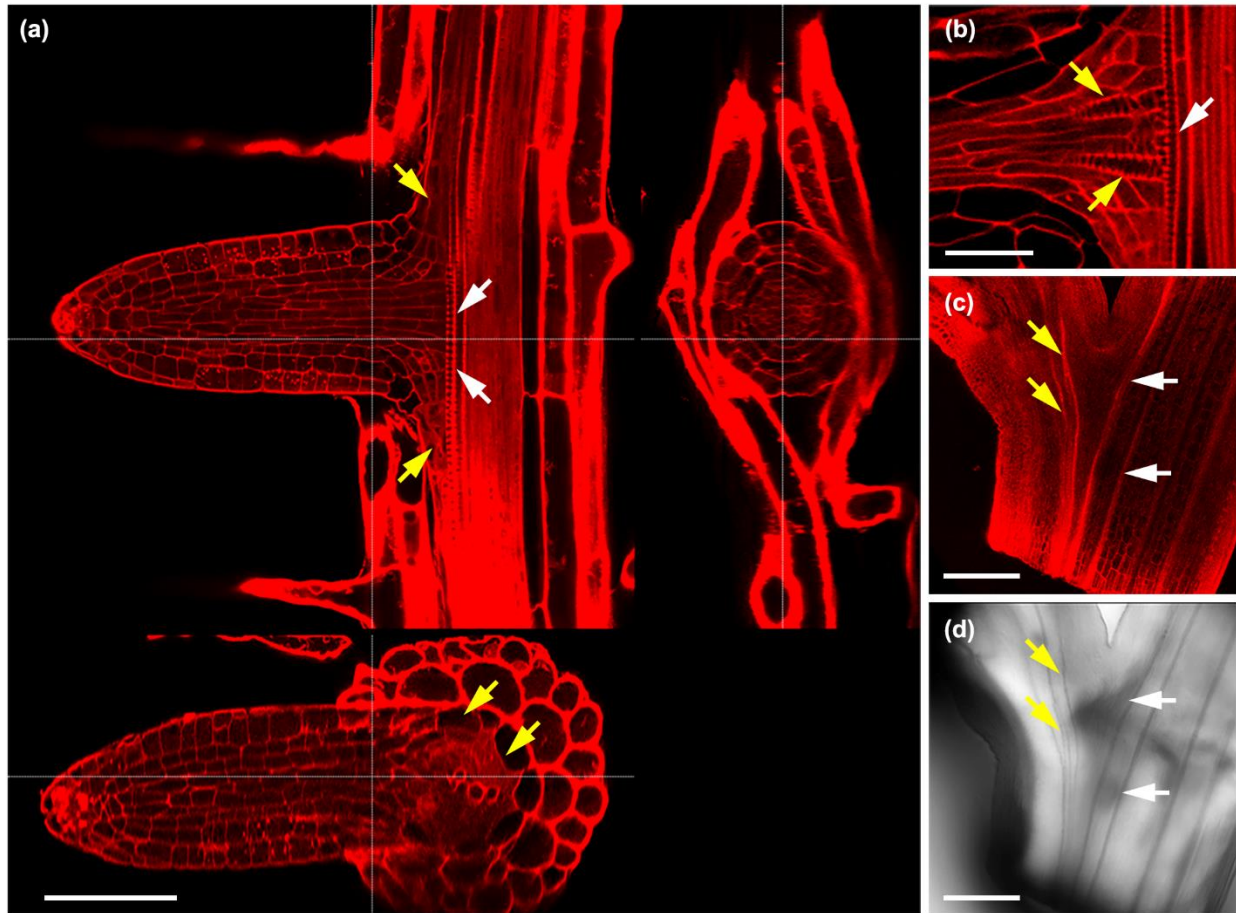
78
79

80 **Data S1 H. Vegetative growth features of *T. lepidozoides***

81 (a) A natural population of *T. lepidozoides* in Plot B, red arrow marks erect stem, yellow arrow
 82 marks rhizome, photograph was made using EOS-1D X Mark II (Canon) with macro lens (MP-
 83 E65 mm, Canon). (b) One part of an individual plant isolated from the population of (a) (bar = 2
 84 mm), arrows show erect stems (red), rhizomes (blue) and aging/faded plant parts (black).

85 Observations were made using an optical microscope (BA600-4, Motic). (c) A population (circle)
86 reproduced from a 3 – 4 mm fragment of erect stem or rhizome cultured on medium for 8
87 months, bar = 2 mm, observation was made using a stereo microscope (Zeiss, AXIO Zoom.
88 V16). (d) The branching pattern of plants in population of (c), red arrows mark erect stems, blue
89 arrows mark rhizomes, bar = 1 mm, observation was made using an optical microscope (Zeiss,
90 AXIO Zoom. V16). (e) The annual growth of stems and rhizomes in the field (n = 42, growth was
91 measured with ImageJ from time-lapse photos) and laboratory (n = 72, measured with ImageJ
92 from microscope). (f) The distance between two adjacent branches in the field (n = 174,
93 measured with ImageJ) and laboratory (n = 135, measured with ImageJ). The central mark of
94 each box indicates the median, and the bottom and top edges of the box indicate the 25th and
95 75th percentiles respectively, the black dots in each box indicates the mean values.

96

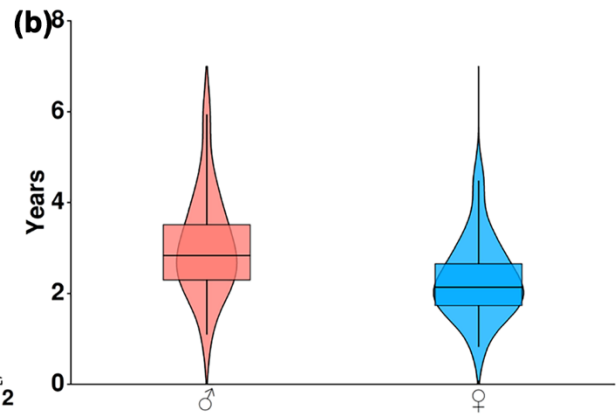
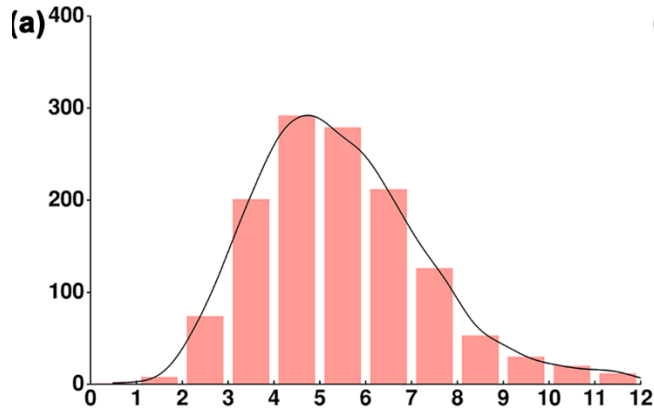


97
98

99 **Data S1 I. The connection mode of vessel elements in vascular bundles between lateral-**
100 **main roots and branches in *A. thaliana*, observed with the mPS-PI staining technique**

101 (a) The sections from three mutually perpendicular directions at the connection site (the cross
102 point of white dotted lines) between main root and lateral root, yellow arrows show pericycle
103 cells, white arrows show vessel elements in vascular bundles of main root, bar = 100 μm. (b)
104 Arrows show vessel elements in main root (white) and lateral root (yellow), bar = 30 μm. (c, d)
105 Arrows show vessel elements in main stem (white) and branch (yellow), PI stain-fluorescent
106 image (c) and bright-field image (d) for the same sample, bar = 500 μm. Fluorescence images
107 were obtained using a Leica SP8 laser scanning confocal microscope, with the excitation
108 wavelength at 561 nm and emission wavelength from 570 nm to 670 nm.

109

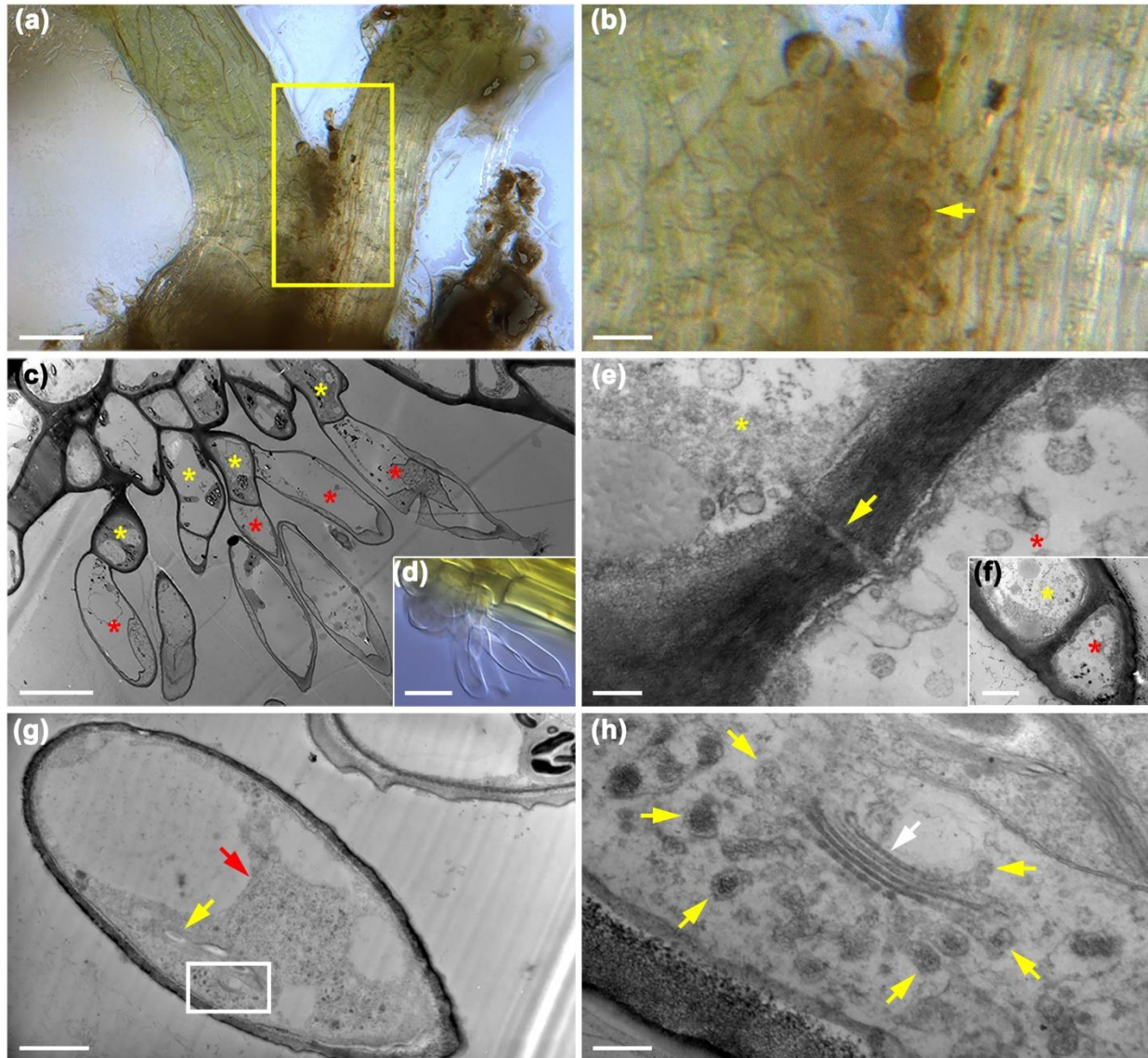


110
111

112 **Data S1 J. Sexual reproduction features of *T. lepidozooides***

113 (a) Distribution blot of the height of erects in the fields. Horizontal axis represents the height of
 114 erects in mm, vertical axis represents the frequency of occurrence (n = 1,315). (b) The timing
 115 that erect stems began sexual reproduction. Male plants achieve sexual maturity in 3.06 ± 1.14
 116 years (median 2.83), while female plants achieve sexual maturity in 2.31 ± 0.86 years (median
 117 2.14) years of vegetative growth (n = 17). The central mark of each box indicates the median,
 118 and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively.

119



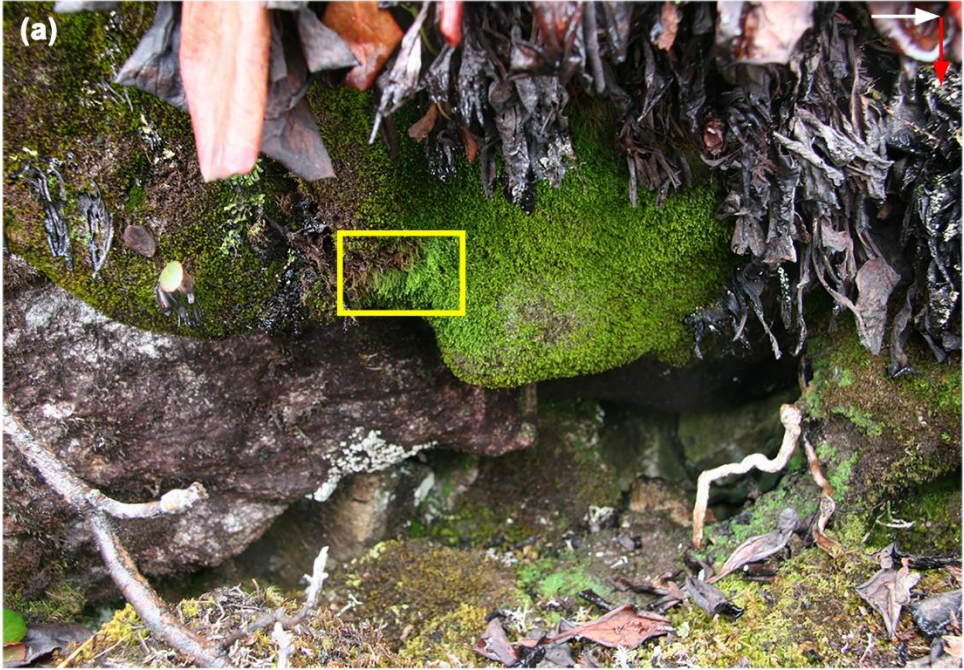
120
121

122 **Data S1 K. The mucilage hairs on the surfaces of stems in *T. lepidozoides***

123 (a) Plants collected from field populations. Mucilage hairs at the cross point between branches,
 124 bar = 100 μm . (b) Magnified views of boxed areas in (a), yellow arrow marks mucilage hairs with
 125 humus, bar = 20 μm . (c) TEM picture of mucilage hairs showing that mucilage hairs are
 126 multicellular. Yellow stars mark the basal cells, red stars mark the tip cells, bar = 10 μm . (d)
 127 Mucilage hairs originate from a stem in cultured populations shown at right corner in (c), bar =
 128 20 μm . (e) The basal cell (yellow star) and tip cell (red star) are connected by plasmodesmata
 129 (yellow arrow), magnified from (f), bar = 200 nm. (f) A mucilage hair with basal cell (yellow star)
 130 and tip cell (red star), bar = 1 μm . (g) A tip cell of mucilage hairs contains a nucleus (red arrow),

131 chloroplast (yellow arrow) and Golgi apparatus (white box), bar = 20 μm . (h) The Golgi
132 apparatus (white arrow) with secretory vesicles (yellow arrows), magnified from the box of (g),
133 bar = 200 nm. (a) and (d) were made using an optical microscope (BA600-4, Motic), (b) was
134 made using a stereo microscope (M125 C, Leica), (c), (e) – (h) were made with a transmission
135 electron microscope (H7500, Hitachi).

136

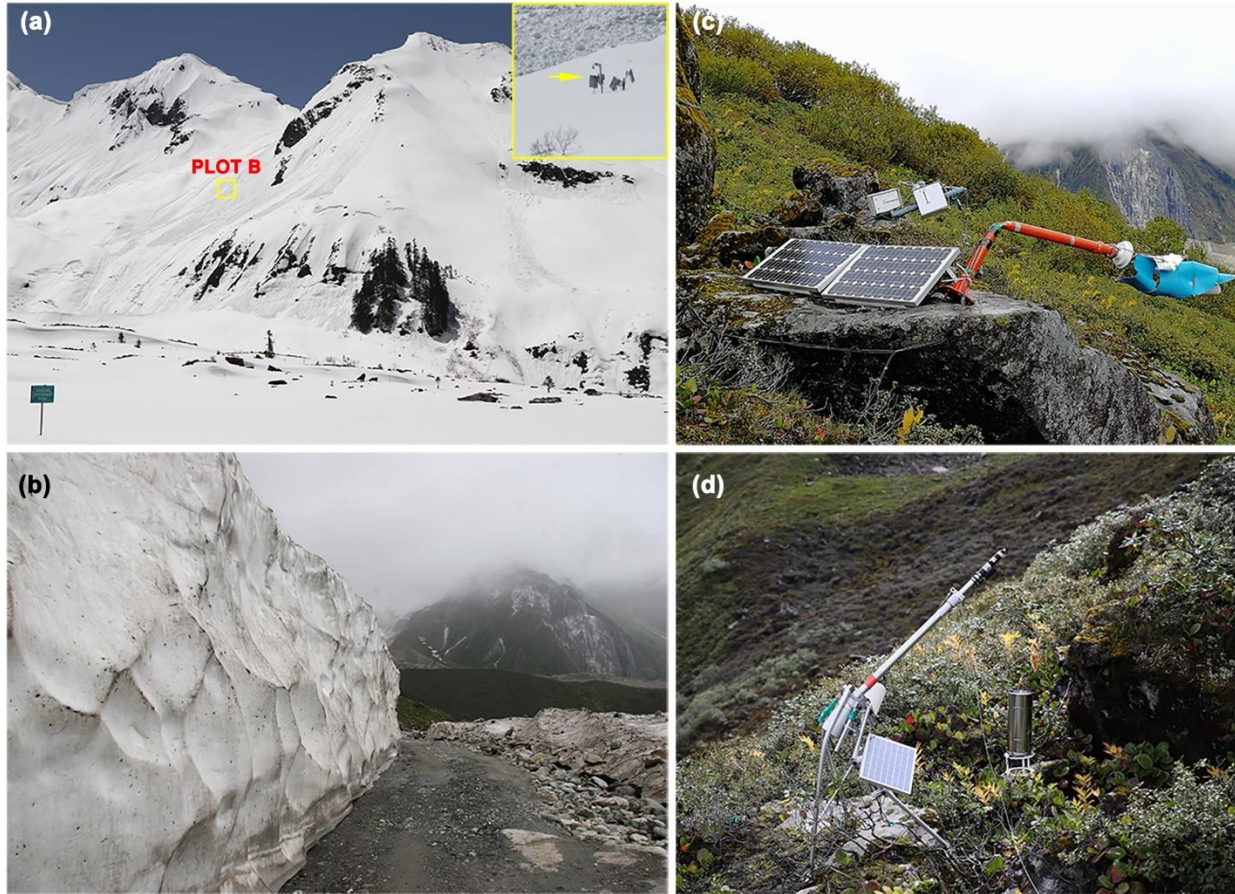


138

139 **Data S1 L. *T. lepidozoides* populations in the field**

140 (a) A population hanging on the top edge of a cave. (b) Magnified views of boxed areas in (a),
141 gametophytes show weak negative geotropism at the apex (yellow arrow), red arrow marks the
142 direction of gravity, and white arrow marks the sunlight direction for most of a day. (c) A
143 population on a steep rock surface. (d) Magnified views of boxed areas in (c), sporophyte shows
144 weak negative geotropism at the apex (yellow arrow), red arrow marks the direction of gravity,
145 and white arrow marks the sunlight direction for most of a day. Photographs were made using
146 EOS-1D X Mark II (Canon) with macro lens (MP-E65 mm, Canon).

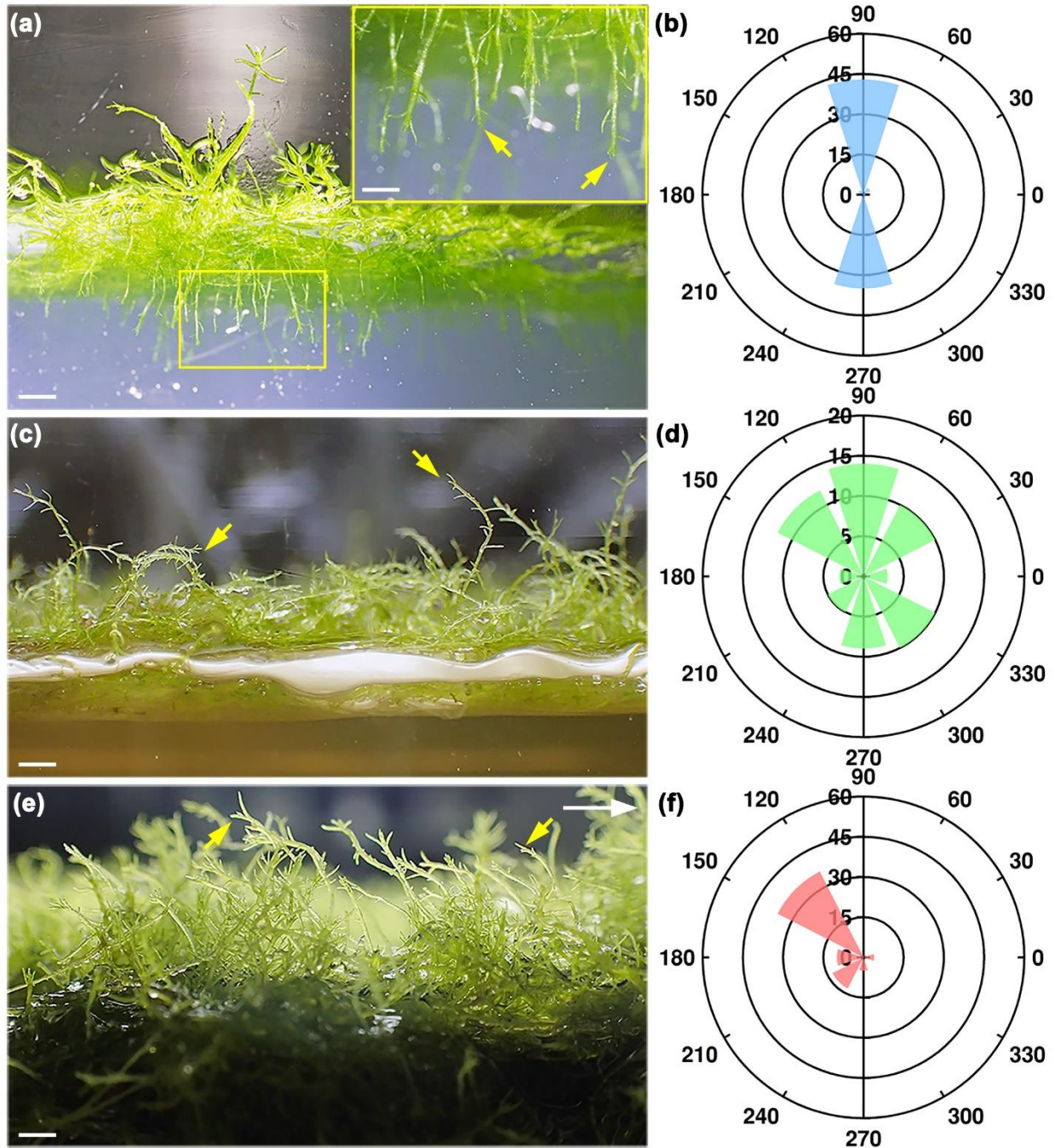
147



148
149

150 **Data S1 M Heavy snowstorm during winter in the experiment sites**

151 Heavy snowstorm during winter (a, b) and meteorological instrument was broken (c, d) by the
 152 heavy snow in the experiment plots. (a) The picture was captured by an unmanned aerial
 153 vehicle (Matrice 300 RTK, DJI) on 5th May, 2020, yellow box marks meteorological instrument in
 154 plot B and magnified at top right (yellow arrow). (b) The picture was captured on 20th June, 2008
 155 in plot B. (c – d) The pictures were captured on 20th August, 2019 in plot B.
 156

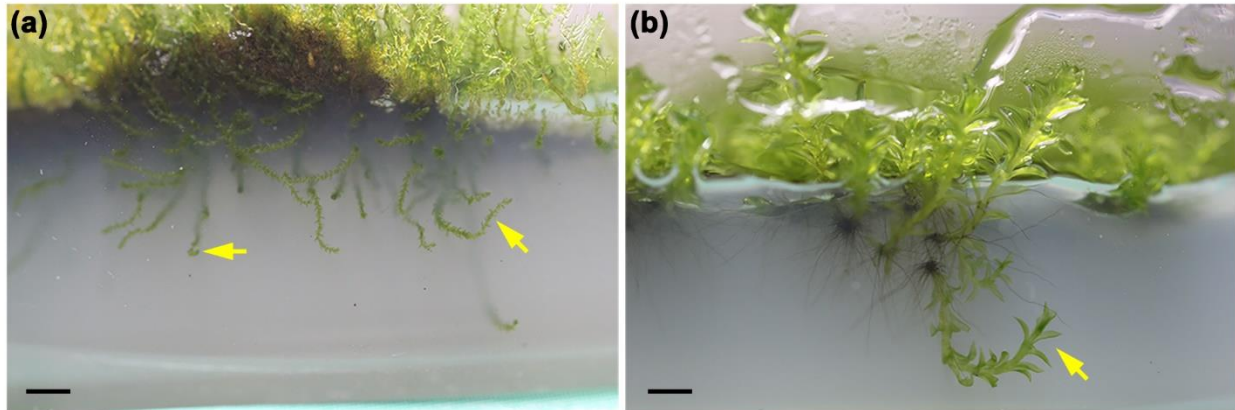


157
158

159 **Data S1 N. The tropism features of *T. lepidozoides* in culture**

160 Plants were grown under normal (a, b), continuous dark (c, d) and continuous unilateral light (e,
161 f) for 8 months. (a) Bar = 2 mm, box magnified at top right with bar = 1 mm, the apex lacks
162 obvious negative geotropism (yellow arrow). (b) The response was quantified by counting the
163 number of tips (n = 70) in all 30° sectors. (c) Bar = 2 mm, the apex lacks obvious tropism

164 (yellow arrow). The response was quantified by counting the number of tips ($n = 66$) in all 30°
165 sectors (**d**). (**e**) Bar = 2 mm, white arrow marks the direction of continuous unilateral light, the
166 apex clearly shows phototropism (yellow arrow). The response was quantified by counting the
167 number of tips ($n = 60$) in all 30° sectors (**f**). Horizontal axis represents sector area, the vertical
168 axis represents the count number.
169



170
171

172 **Data S1 O Negative geotropism of the mosses *Bryum argenteum* and *Physcomitrium***
173 ***patens* in culture**

174 Negative geotropism of (a) *Bryum argenteum* (bar = 1 mm) and (b) *Physcomitrium patens* (bar
175 = 1 mm). *B. argenteum* and *P. patens* shoots firstly grow downward (typical geotropism) without
176 spaces suitable for them to grow upward because the spaces were occupied by other shoots.
177 When the surrounding space becomes available, they return to grow upward (yellow arrow) with
178 apparent negative geotropism. *B. argenteum* was grown on half strength MS medium with a
179 light intensity of $100 \mu\text{mol m}^{-2} \text{s}^{-1}$ delivered in a 16-h/8-h light/dark cycle at 22 °C for 8 weeks; *P.*
180 *patens* was grown on BCD medium with a light intensity of $100 \mu\text{mol m}^{-2} \text{s}^{-1}$ delivered in a 16-
181 h/8-h light/dark cycle at 25 °C for 8 weeks. Observations were made using a stereo microscope
182 (M125 C, Leica).

183

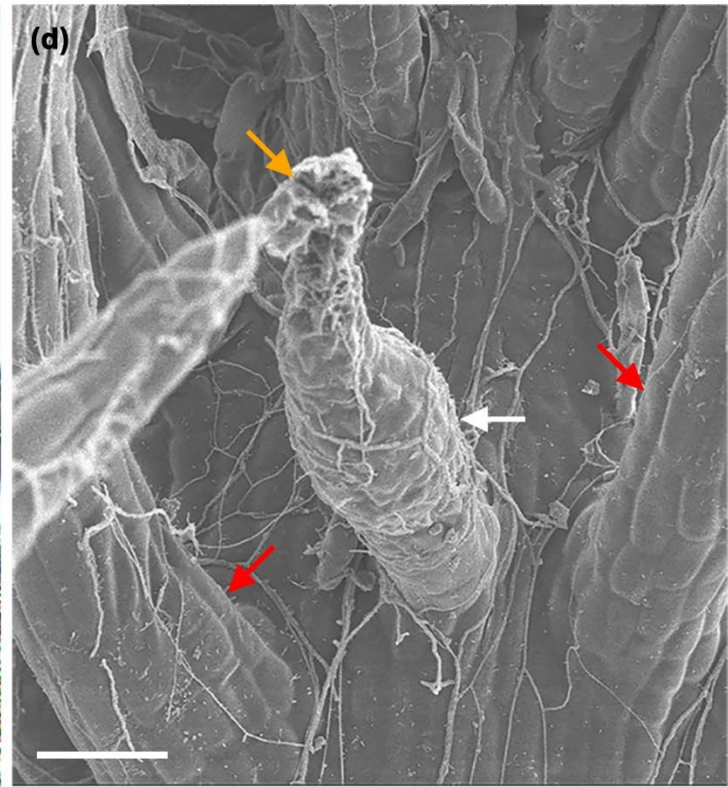
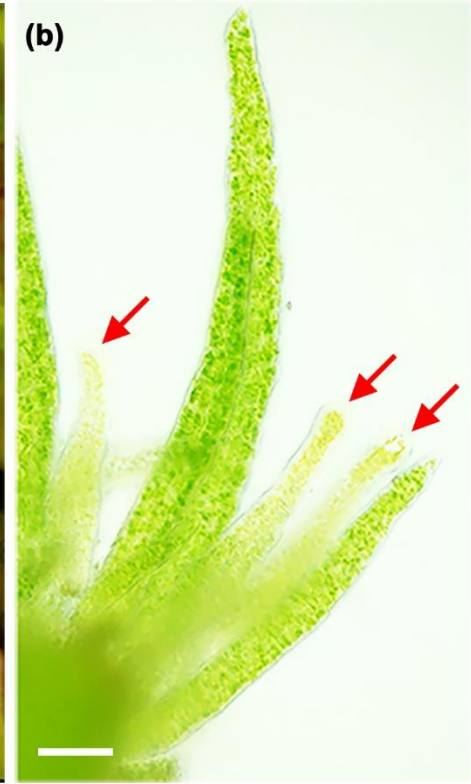
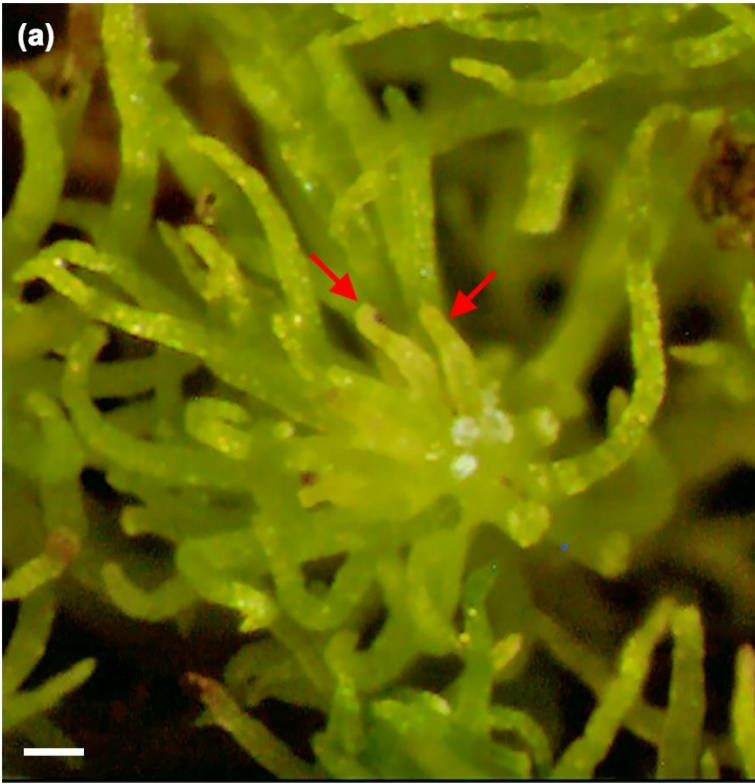


184
185

186 **Data S1 P. Antheridia development in *T. lepidozoides***

187 Development of antheridia in the field. (a) Young antheridia in early July (bar = 2 mm). (b)
188 Mature antheridia in late September, bar = 2 mm. (c) A male erect stem with antheridia (arrow).
189 Red antheridia are a product of the current year and pale green antheridia are a product of the

190 previous year, bar = 200 μm . **(d)** Magnified view of mature antheridia, bar = 200 μm . **(e)**
191 Magnified view of mature and dehisced antheridia. White arrow marks a dehisced antheridium,
192 bar = 200 μm . (a) and (b) were made using EOS-1D X Mark II (Canon) with macro lens (MP-
193 E65 mm, Canon), (c) was made with an optical microscope (BA600-4, Motic), (d) and (e) was
194 made using a stereo microscope (M125 C, Leica).
195

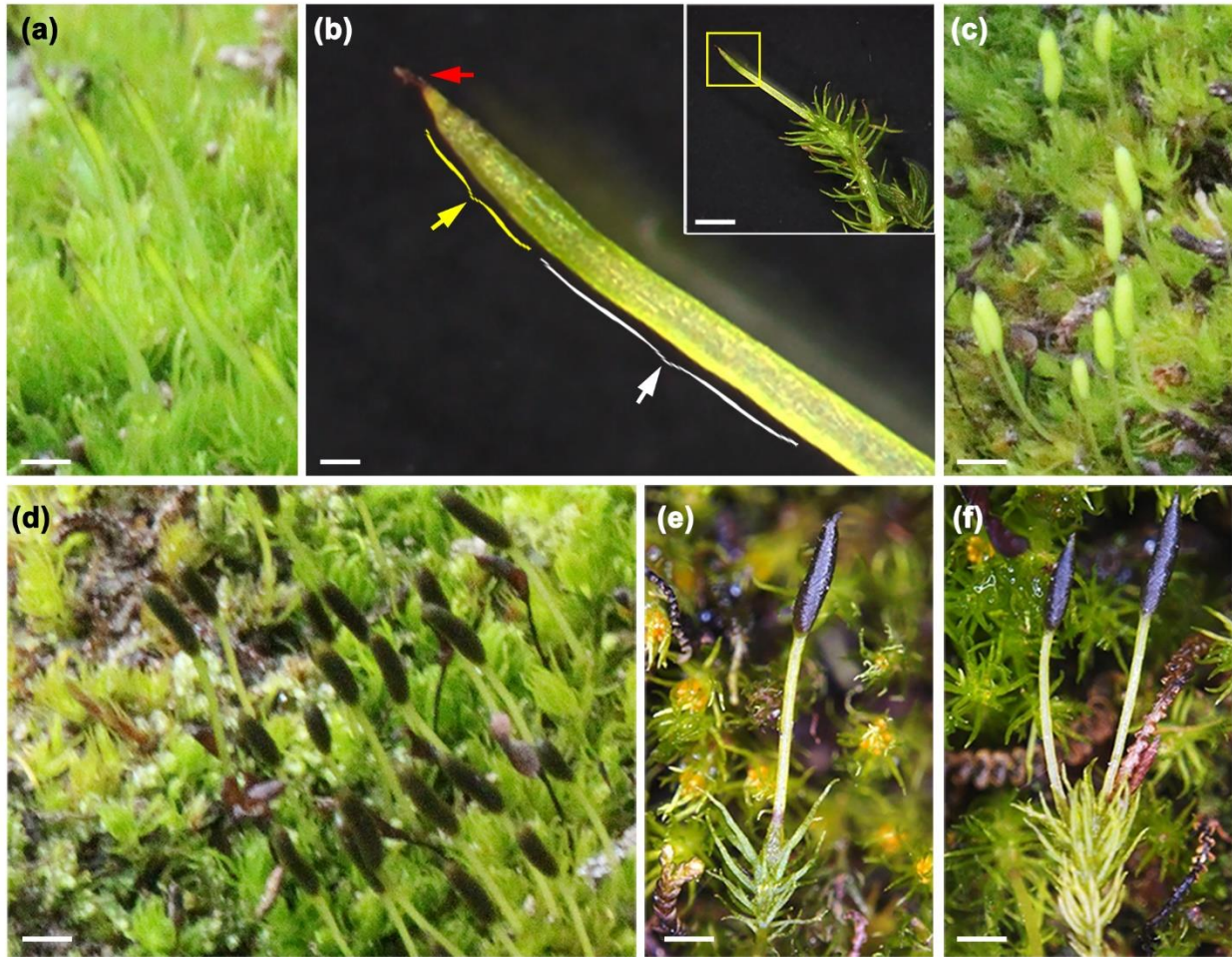


196
197

198 **Data S1 Q. Archegonia development in *T. lepidozoides***

199 Development of archegonia in the field. **(a)** Young archegonia (red arrows) at the base of leaves
200 on top of the erect stem, bar = 100 μm . **(b)** A female erect stem with archegonia (red arrows) at
201 the base of leaves, bar = 100 μm . **(c)** A mature archegonium with an open canal at the top
202 (orange arrow) and immature archegonia (red arrows), bar = 50 μm . **(d)** SEM micrograph of a
203 mature archegonium (white arrow) with an opening (orange arrow). Red arrows mark terete
204 phylloids, bar = 50 μm . (a) was made with a stereo microscope (M125 C, Leica), (b) and (c)
205 were made with an optical microscope (BA600-4, Motic), (d) was made with a scanning electron
206 microscope (S4800, Hitachi).

207

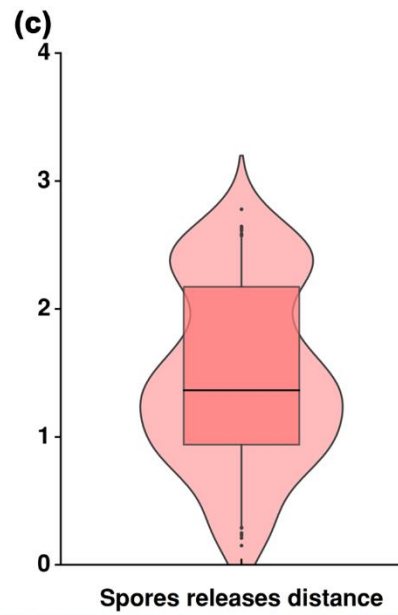
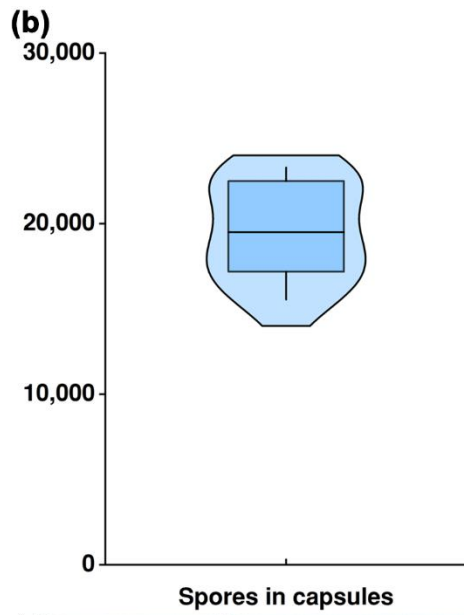
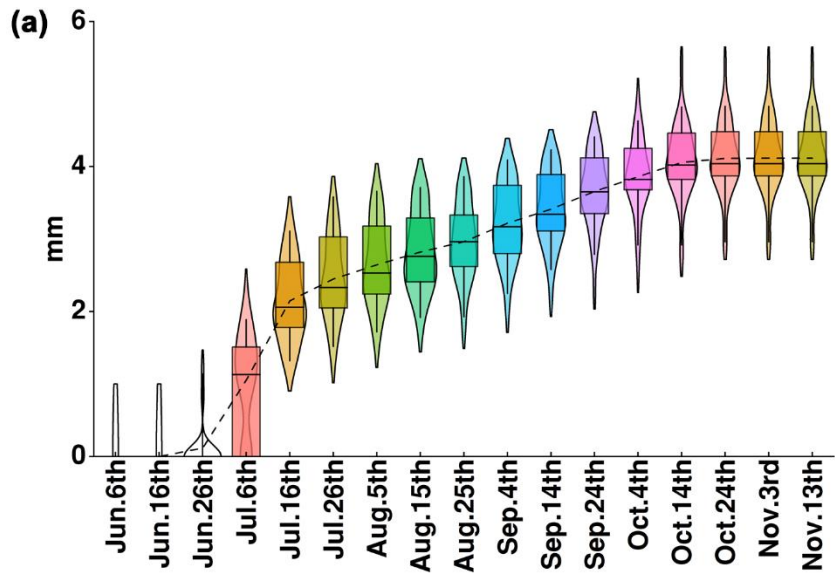


208
209

210 **Data S1 R. Sporophyte development in *T. lepidozoides***

211 Development of sporophytes in the field. (a) Young sporophytes with unexpanded capsules in
 212 early July, bar = 1 mm. (b) A young sporophyte developing (magnified from top right yellow box,
 213 bar = 200 μ m) from a fertilized archegonium with the calyptra (red arrow) on the apiculus of the
 214 unexpanded capsule (yellow arrow), white arrow marks seta, bar = 50 μ m. (c) Sporophytes with
 215 capsule expansion in late July, bar = 1 mm. (d) Sporophytes with mature capsules in mid-
 216 October, bar = 2 mm. (e) Single erect stem with one mature capsule, bar = 1mm. (f) Single
 217 erect stem with two mature capsules, bar = 1mm.

218

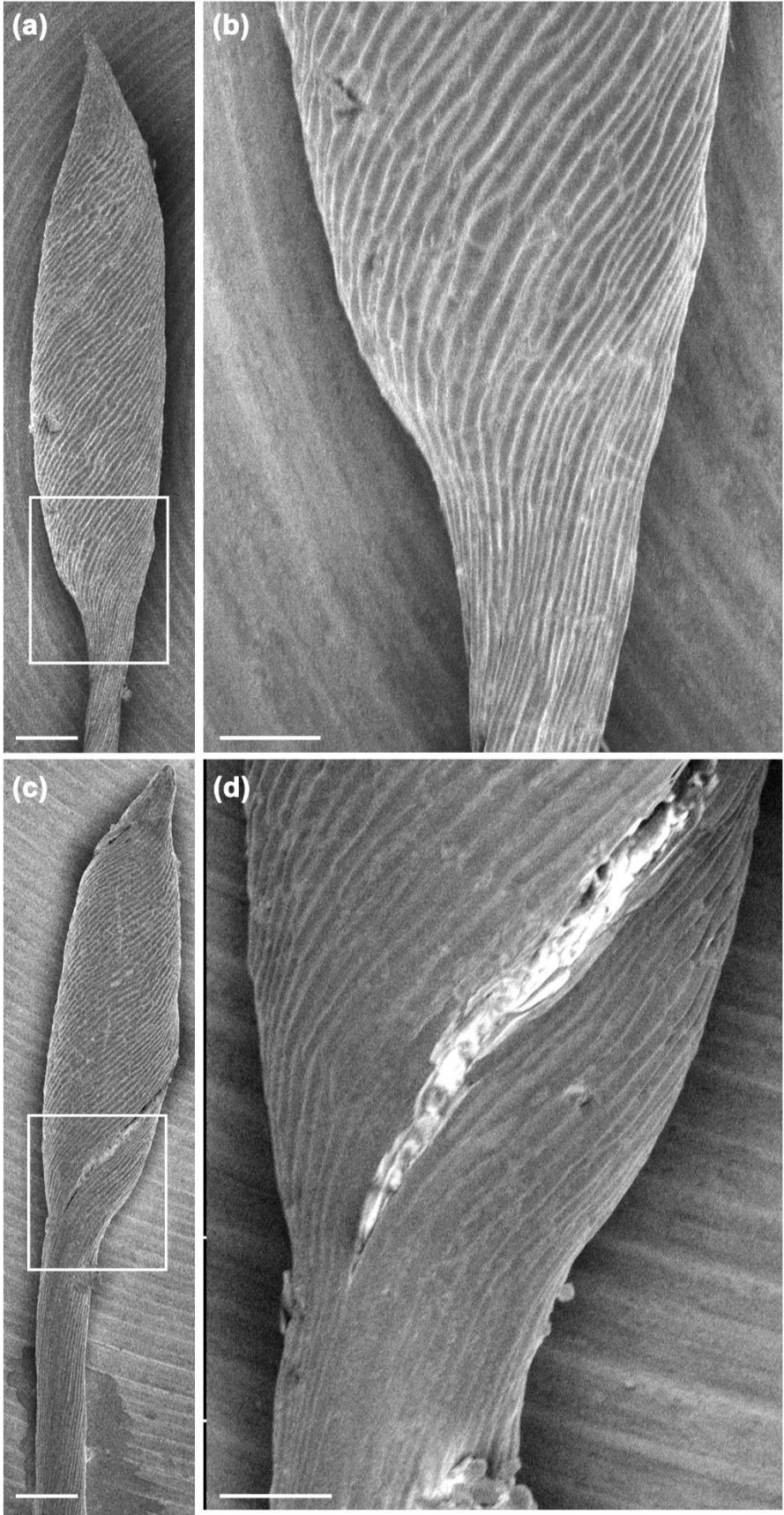


220

221 **Data S1 S. Capsule features of *T. lepidozoides***

222 (a) Analysis of sporophyte growth and development. Growth was measured with ImageJ from
223 time-lapse photos in every 10 days in 2014 (n = 40). The central mark of each box indicates the
224 median, and the bottom and top edges of the box indicate the 25th and 75th percentiles,
225 respectively. (b) Spore numbers in one capsule (n = 8). (c) Spore dispersion distance (in cm)
226 under laboratory conditions in (d), all radius data were from 4 capsules with 30 individual
227 measurements in each. The central mark of each box indicates the median, and the bottom and
228 top edges of the box indicate the 25th and 75th percentiles, respectively. (d) Experimental
229 setting for (c). Erect stem with mature capsule collected from the field and inserted into 1 cm³
230 agar in a petri dish. The dish was placed in a chamber with silica gel to mimic dehiscence. The
231 distance from spore to erect bottom were measured with ImageJ, photograph was made using a
232 stereo microscope (M125 C, Leica), bar = 5 mm.

233

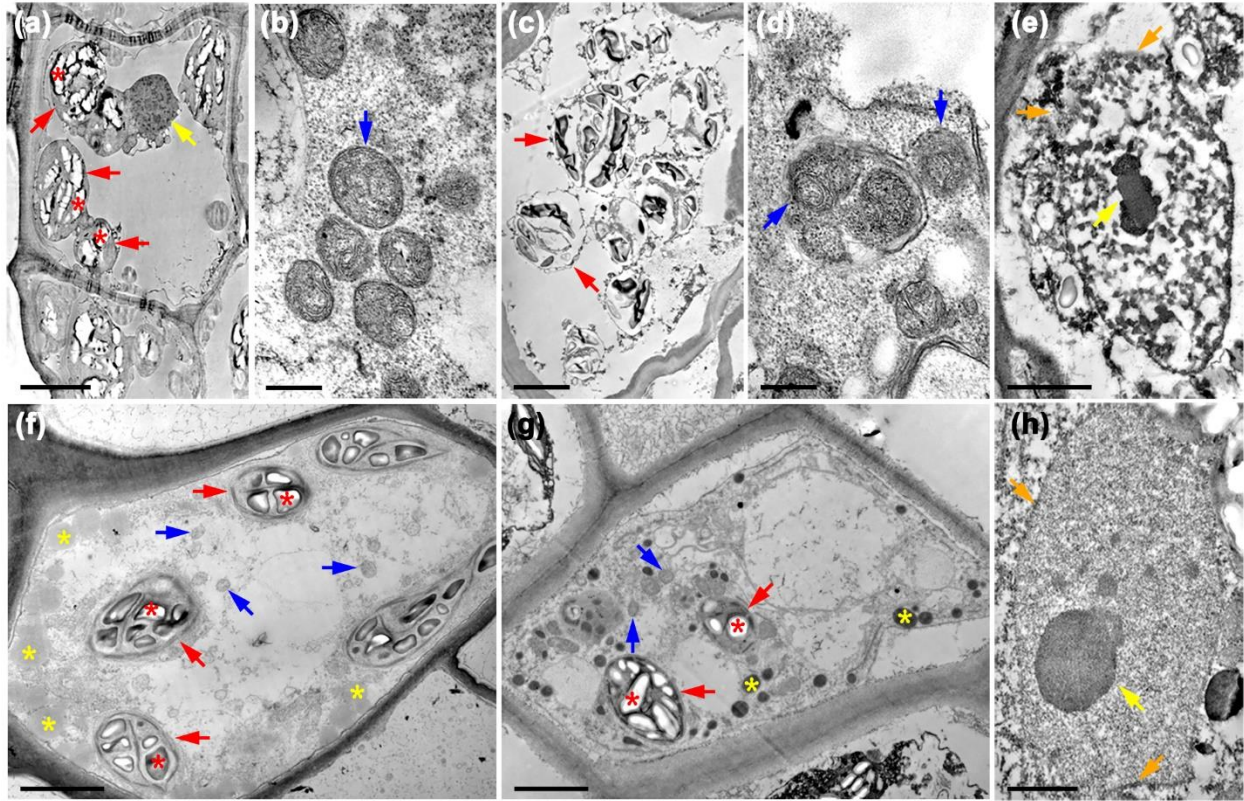


235

236 **Data S1 T. SEM of *T. lepidozoides* capsules**

237 SEM micrographs of *T. lepidozoides* capsules showing that the sporophyte lacks stomata at the
238 capsule base, unlike other moss sporophytes such as *P. patens* [S2]. Young (**a, b**) and mature
239 capsule with spiral split (**c, d**). (b) and (d) are magnified views of boxed areas in (a) and (c)
240 respectively. (a) – (c) bars = 200 μm , (d) bar = 100 μm . Observations were made with a
241 scanning electron microscope (S4800, Hitachi).

242



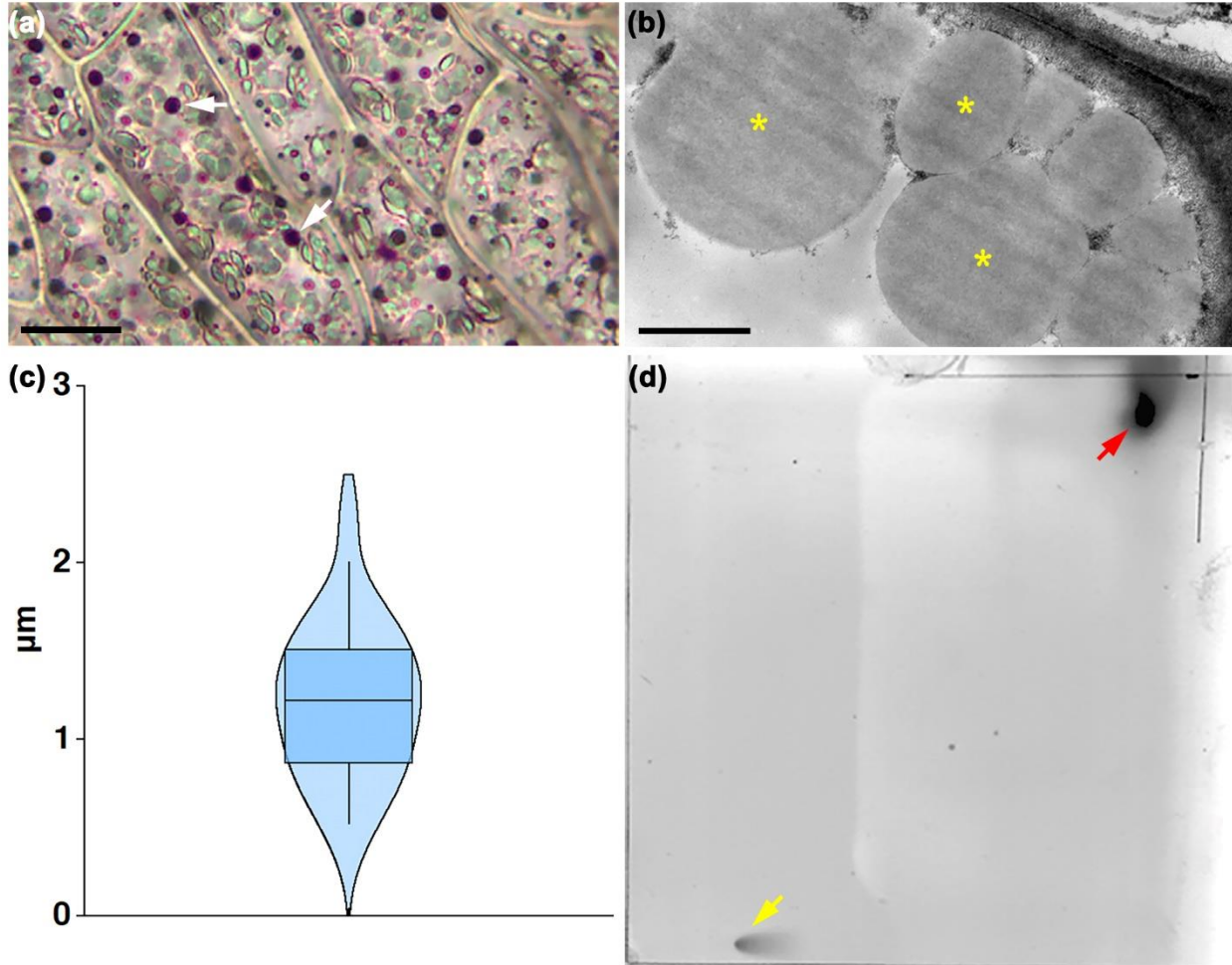
243

244

245 **Data S1 U. TEM of cells before and after UV-B radiation in *P. patens* and *T. lepidozoides***

246 TEM micrographs of *P. patens* (a – e) and *T. lepidozoides* (f – h) before and after UV-B
 247 radiation and recovery. (a, b) Chloroplasts (red arrows) with starch grains (stars) and
 248 mitochondria (blue arrow) in the cell under normal condition. (c – e) Collapsed chloroplasts (red
 249 arrows), expanded mitochondria (blue arrows) and a collapsed nucleus with disintegrated
 250 nuclear membranes (orange arrows) and nucleolus (yellow arrow) surrounded by agglomerated
 251 chromatin in cells recovering for 48 to 72 hours after UV-B radiation. (f) Chloroplasts (red
 252 arrows) with starch grains (red stars), mitochondria (blue arrows) and oil bodies (yellow stars) in
 253 the cell under normal condition. (g, h) Chloroplasts (red arrows) with starch grains (red stars),
 254 mitochondria (blue arrows), oil bodies (yellow stars) and a nucleus with intact nuclear
 255 membrane (orange arrows) and nucleolus (yellow arrow) surrounded by a slight condensed
 256 chromatin in cells recovering for 48 to 72 hours after UV-B radiation. Observations were made
 257 with a transmission electron microscope (H7500, Hitachi), bar = 5 μm in (a) and (c), 3 μm in (g),
 258 2 μm in (e) and (f), 1 μm in (h), 400 nm in (b) and (d).

259

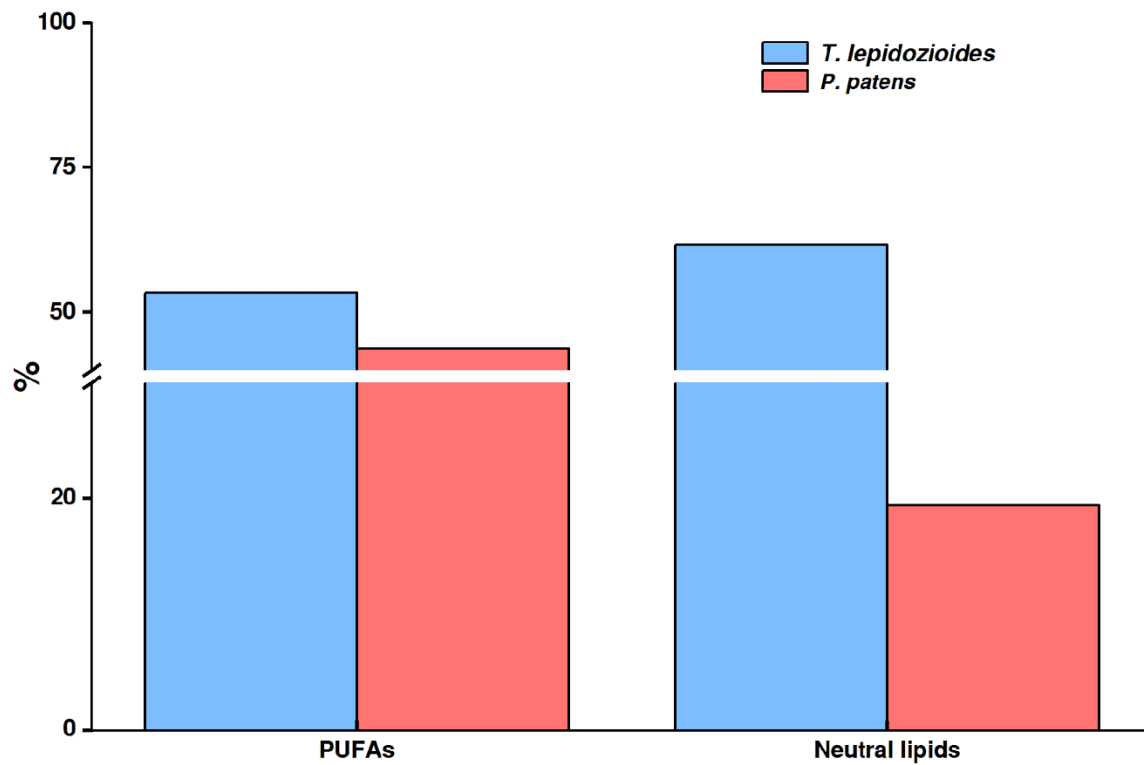


260
261

262 **Data S1 V. The features of oil bodies in *T. lepidozooides***

263 (a) Oil bodies in living cells of *T. lepidozooides* stems after Sudan black staining, white arrows
 264 mark oil bodies, bar = 200 μm . (b) TEM micrograph of oil bodies, yellow stars mark oil bodies,
 265 bar = 1 μm . Observations were made with a transmission electron microscope (H7500,
 266 Hitachi). (c) Radius of oil bodies in living cells (about 1.22 μm) of stems (n = 20). The central
 267 mark of each box indicates the median, and the bottom and top edges of the box indicate the
 268 25th and 75th percentiles respectively. (d) 2D-TLC to identify oil body components. Red arrow
 269 marks nonpolar lipids, yellow arrow marks phosphatidylserine, which is the main content of
 270 plasma membranes [S3].

271



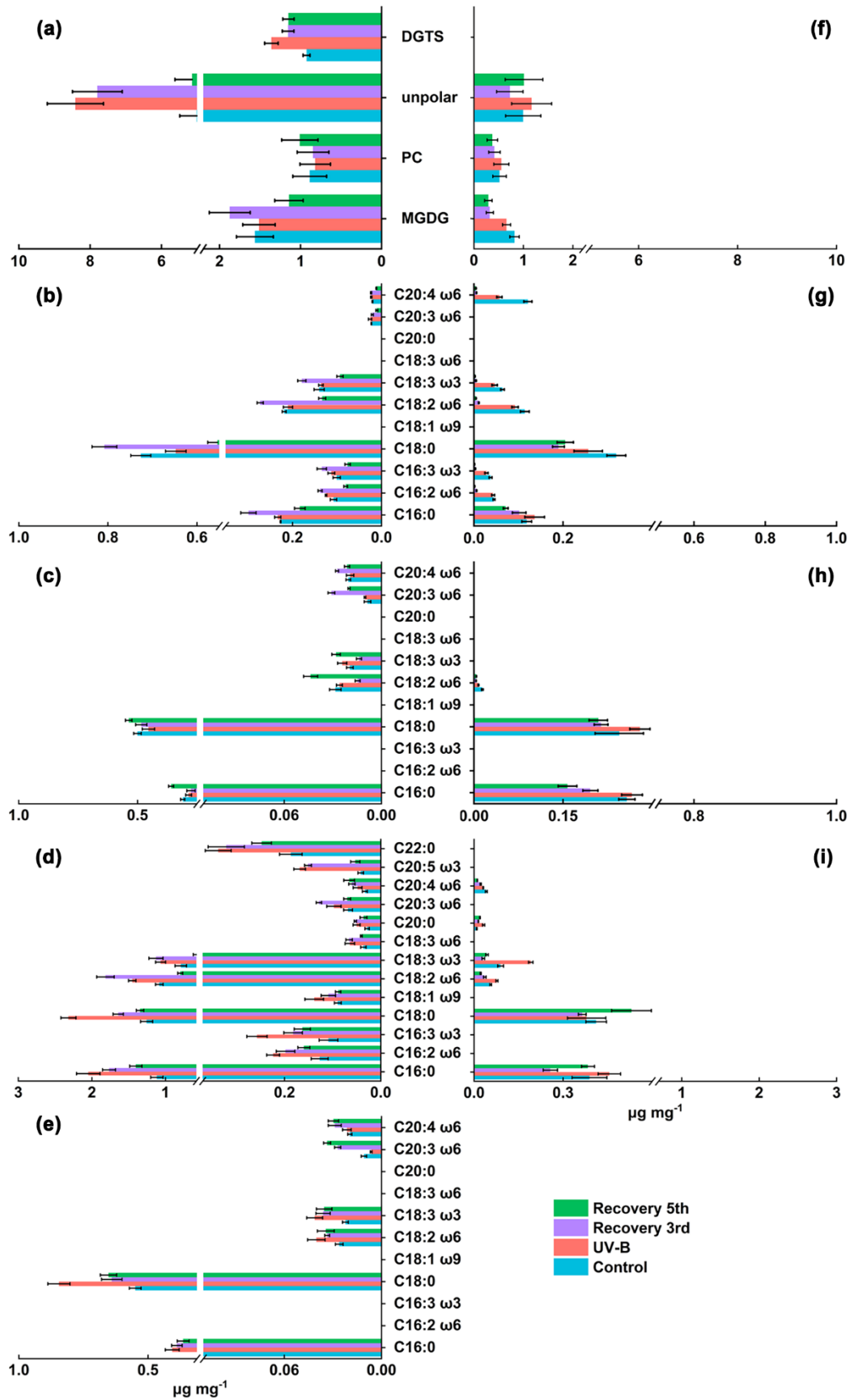
272

273

274 **Data S1 W. The lipid contents of *T. lepidozoides* and *P. patens***

275 *T. lepidozoides* and *P. patens* plants were collected, PUFAs and neutral lipids were qualitatively
 276 and quantitatively analyzed by 2D-TLC and GC-MS. PUFAs: polyunsaturated fatty acids.

277



279

280 **Data S1 X. Lipid composition in *T. lepidozoides* and *P. patens* before and after UV-B**
281 **radiation**

282 Lipid changes before and after UV-B radiation in *T. lepidozoides* (a – e) and *P. patens* (f – i). *T.*
283 *lepidozoides* and *P. patens* gametophytes grown under normal condition for two weeks, UV-B
284 treatment, qualitative and quantitative analysis of lipids. (a, f) Total lipids. (b, g)
285 monogalactoglycerolipid (MGDG). (c, h) phosphatidylcholine (PC). (d, i) Nonpolar lipids. (e)
286 diacylglycerol-N,N,N-trimethylhomoserine (DGTS) in *T. lepidozoides*. There were no detectable
287 DGTS in *P. patens*. Blue color represents plants under normal conditions (control), red color
288 represents plants under 144 kJ m⁻² UV-B treatment for 2 hours (UV-B), violet represents plants
289 recovery on the 3rd day after UV-B, and green represents plants recovery on the 5th day after
290 UV-B.

291

293

294 **Data S1 Y. Non-targeted metabolic profiles of *T. lepidozoides* and *P. patens***

295 The metabolic profiles differed significantly between *T. lepidozoides* and *P. patens* during UV-B
296 radiation and recovery. In total, 85 metabolites were detected and clustered into 3 major clades.
297 Pathway analysis of clade 1 identified 3 pathways that were enriched to flavone and flavonol
298 biosynthesis, flavonoid biosynthesis, and purine metabolism. Clade 2 pathways were mainly
299 enriched in purine metabolism, pyrimidine metabolism, tyrosine metabolism. (a) Negative ion
300 mass spectrum of *P. patens* and *T. lepidozoides*. (b) Positive ion mass spectrum of *P. patens*
301 and *T. lepidozoides*. (c) Heatmap clusters of in total 85 metabolites. Detailed data of the 85
302 metabolites is compiled in Table S6. (d) Pathway enrichment of clade 1 in (c). (e) Pathway
303 enrichment of clade 2 in (c). The horizontal axis “pathway impact” is the cumulative percentage
304 from the matched metabolite nodes.

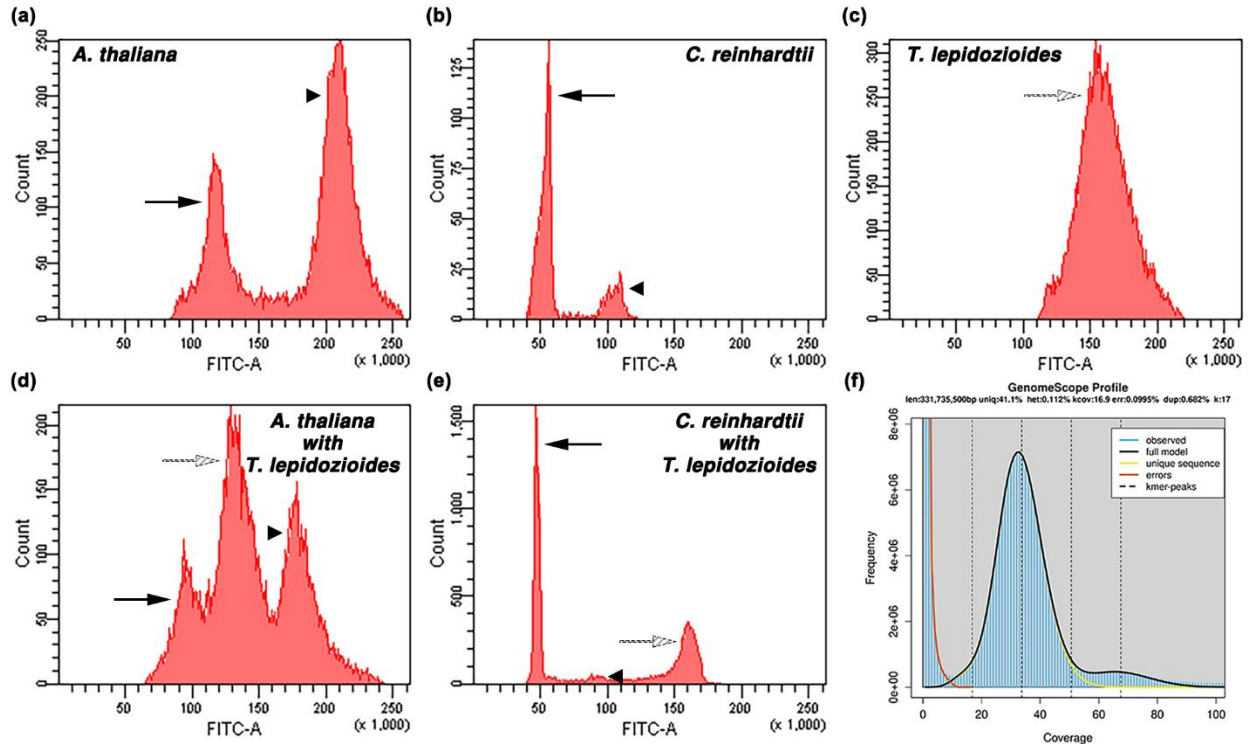
305

306 **SUPPLEMENTAL REFERENCES**

- 307 S1. Liu, C.R., Newell, G., and White, M. (2016). On the selection of thresholds for predicting
308 species occurrence with presence-only data. *Ecol. Evol.* 6, 337–348.
309 10.1002/ece3.1878.
- 310 S2. Chater, C.C., Caine, R.S., Tomek, M., Wallace, S., Kamisugi, Y., Cuming, A.C., Lang,
311 D., MacAlister, C.A., Casson, S., Bergmann, D.C., et al. (2016). Origin and function of
312 stomata in the moss *Physcomitrella patens*. *Nat. Plants* 2, 16179, 16179.
313 10.1038/nplants.2016.179.
- 314 S3. Leventis, P.A., and Grinstein, S. (2010). The distribution and function of
315 phosphatidylserine in cellular membranes. In *Annual Review of Biophysics*, Vol 39, D.C.
316 Rees, K.A. Dill, and J.R. Williamson, eds. (Annual Reviews), pp. 407–427.
317 10.1146/annurev.biophys.093008.131234.
318

1 **Supplement data 2**

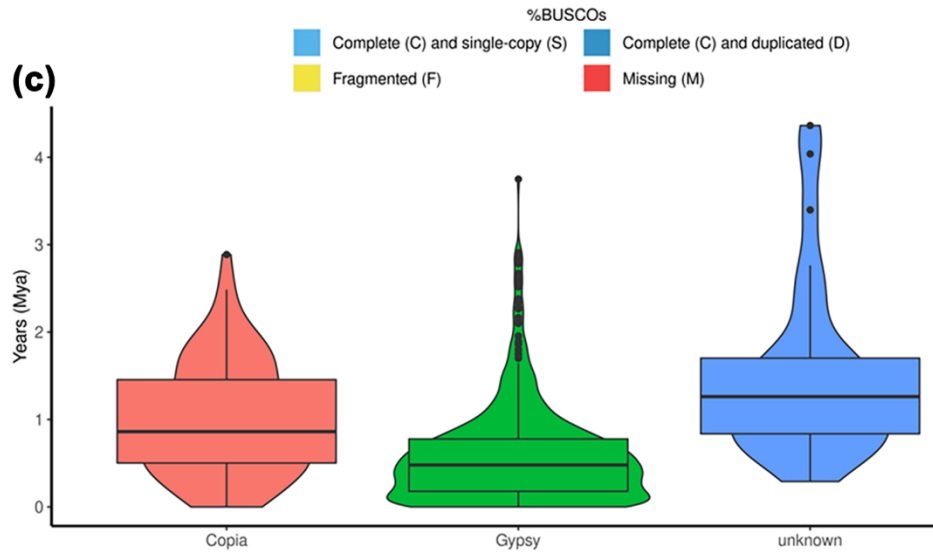
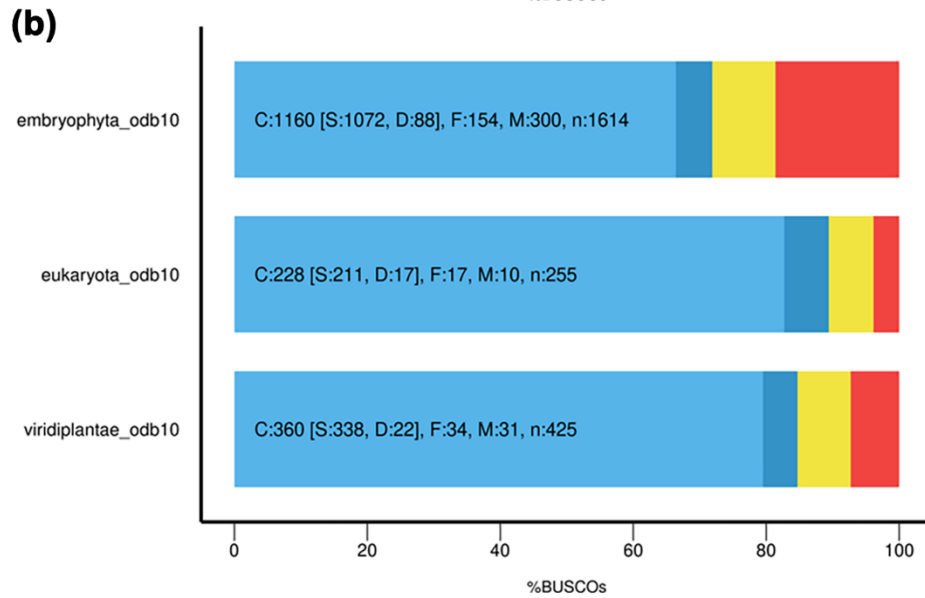
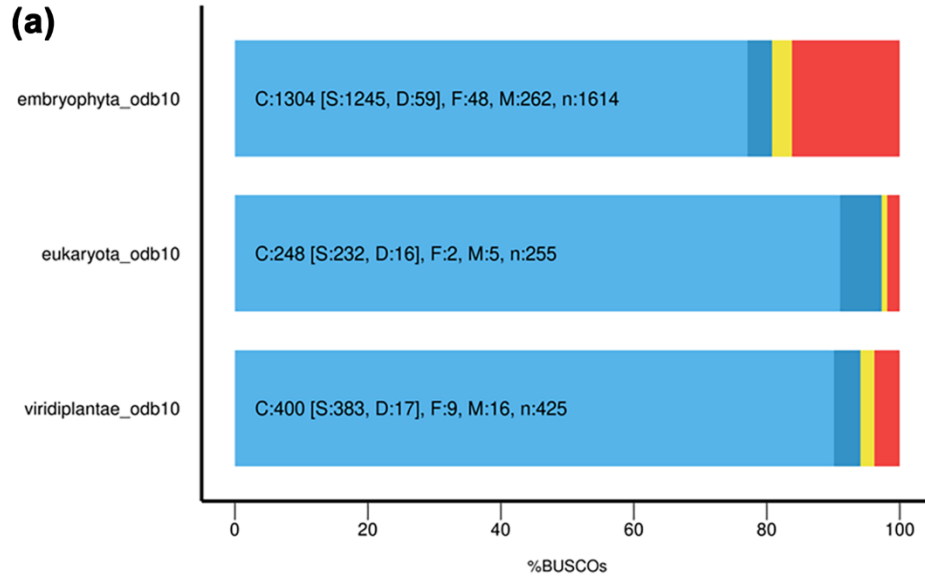
2



3
4

5 Data S2 A. Genome size estimations of *T. lepidozooides*

6 (a – e) Estimate of *T. lepidozooides* genome size based on flow cytometry (FCM). FCM was carried
7 out according to Reski et al. [S1] in a cell sorter (FACSaria, BD) with SYTOX Green stain. (a)
8 Histogram of fluorescence signals of *A. thaliana*, two major peaks at around 100 (arrow) and 200
9 (triangle) with ploidy levels at 2C and 4C, respectively. (b) Histogram of fluorescence signals of
10 *C. reinhardtii*, two major peaks at around 50 (arrow) and 100 (triangle) with ploidy levels at 1C
11 and 2C, respectively. (c) Histogram of fluorescence signals of *T. lepidozooides*, only one peak at
12 around 150 (dotted arrow). (d) Histogram of fluorescence signals of *A. thaliana* and *T.*
13 *lepidozooides* nucleus mixture, the signals of *T. lepidozooides* (dotted arrow) near 150 between
14 two major peaks of *A. thaliana* (arrow and triangle), the estimated DNA contents with a ratio of
15 1.39:1 between *T. lepidozooides* and *A. thaliana* (2C). (e) Histogram of fluorescence signals of *C.*
16 *reinhardtii* and *T. lepidozooides* nucleus mixture, the signals of *T. lepidozooides* (dotted arrow) near
17 150, and two major peaks of *C. reinhardtii* at around 50 (arrow) and 100 (triangle), the estimated
18 DNA contents with a ratio of 3.42:1 between *T. lepidozooides* and *C. reinhardtii* (1C). (f) Estimate
19 of *T. lepidozooides* genome size based on sequencing. Total 20 Gbp Illumina paired-end reads
20 were used in estimation according to Marcais et al. [S2] and Vurture et al. [S3]. The estimated
21 genome size of *T. lepidozooides* is 331 Mbp.



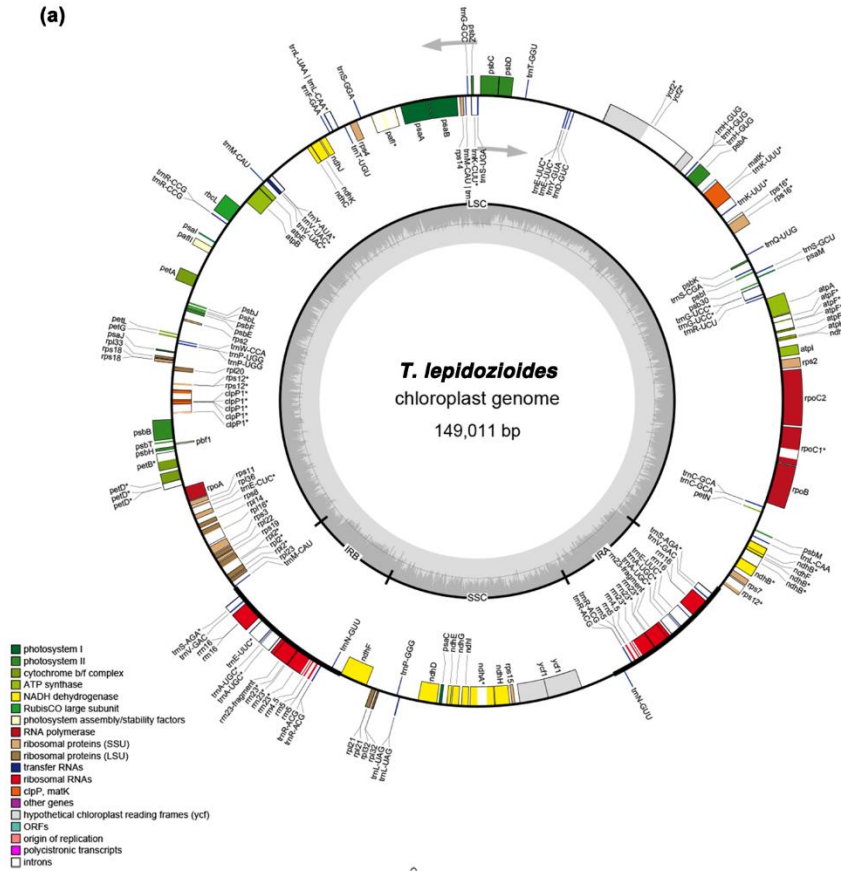
23

24 **Data S2 B. Genome features of *T. lepidozooides***

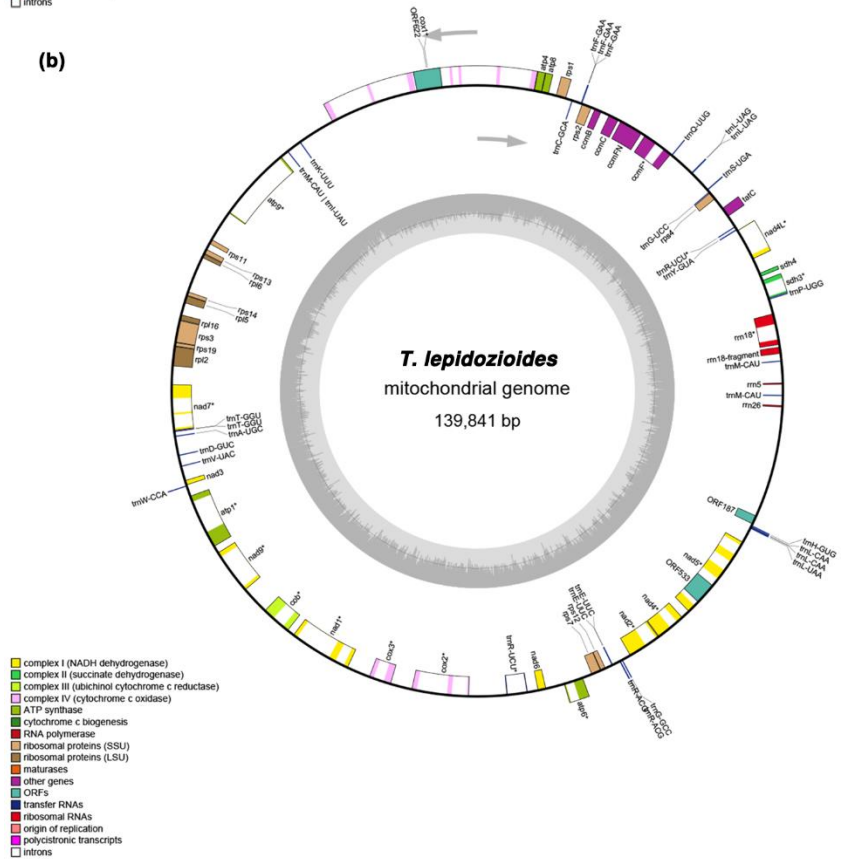
25 BUSCO assessment of assembly (a) and annotation (b). The latest all eukaryotic organism gene
26 set eukaryote odb10, all plants gene set viridiplantae odb10 and all land plants gene set
27 embryophyta odb10, were used as the reference, respectively, and Arabidopsis was used in the
28 -sp option, 97.3% in eukaryota, 94.1% in viridiplantae and 80.8% in embryophyte gene sets are
29 present in the final assembly. The annotation proportions were 89.4%, 84.7% and 71.9%,
30 respectively. (c) Insertion time of *T. lepidozooides* LTR retrotransposons. The central mark of each
31 box indicates the median, and the bottom and top edges of the box indicate the 25th and 75th
32 percentiles, respectively. The whiskers extend to the most extreme data points not considered
33 outliers. The median age of *copia* is 0.85 Mya, the median age of *gypsy* is 0.47 Mya, and most
34 LTRs in *T. lepidozooides* (unknown) have an insertion age of 1 ~ 2 Mya.

35

(a)



(b)

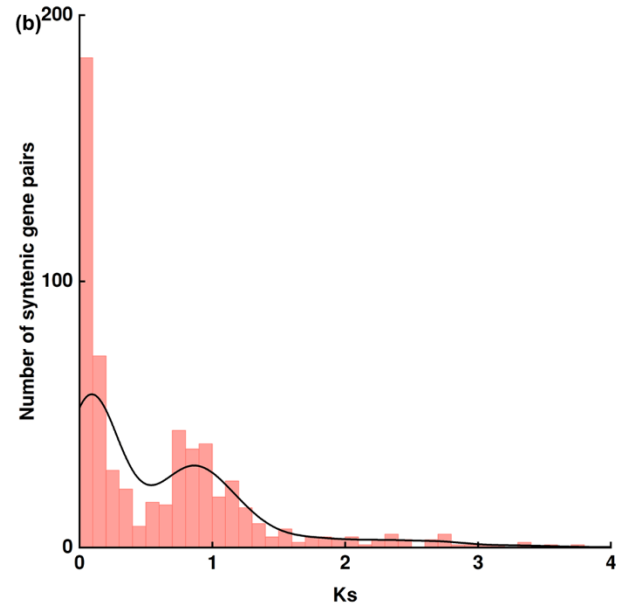
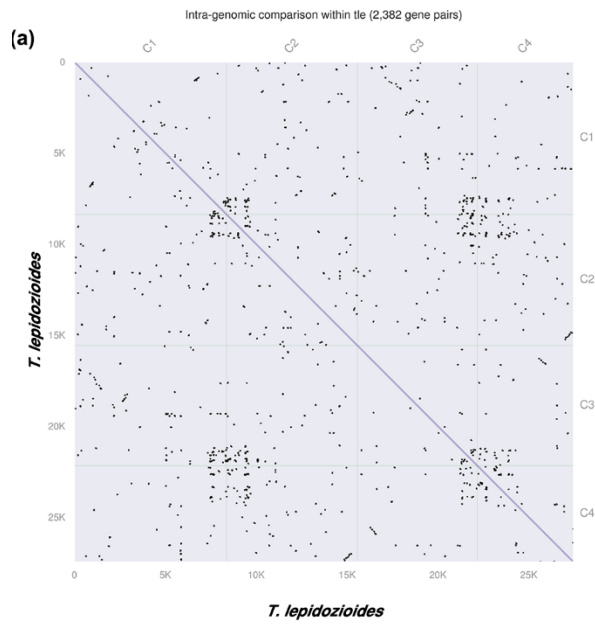


37

38 **Data S2 C. Organellar genomes of *T. lepidozoides***

39 (a) The complete circular *T. lepidozoides* chloroplast genome comprises 149 Kb in four parts:
40 long single copy section (LSC), inverted repeat A (IRA), small single copy section (SSC), and
41 inverted repeat B (IRB). The chloroplast genome encodes 129 genes, including 37 tRNAs. (b)
42 The complete circular *T. lepidozoides* mitochondrial genome comprises 139 Kb. It encodes 70
43 genes, including 24 tRNAs. The figure was generated using OGDRAW program from
44 CHLOROBX (<https://chlorobox.mpimp-golm.mpg.de/index.html>). Color-coded boxes indicate
45 the genes in the genome.

46



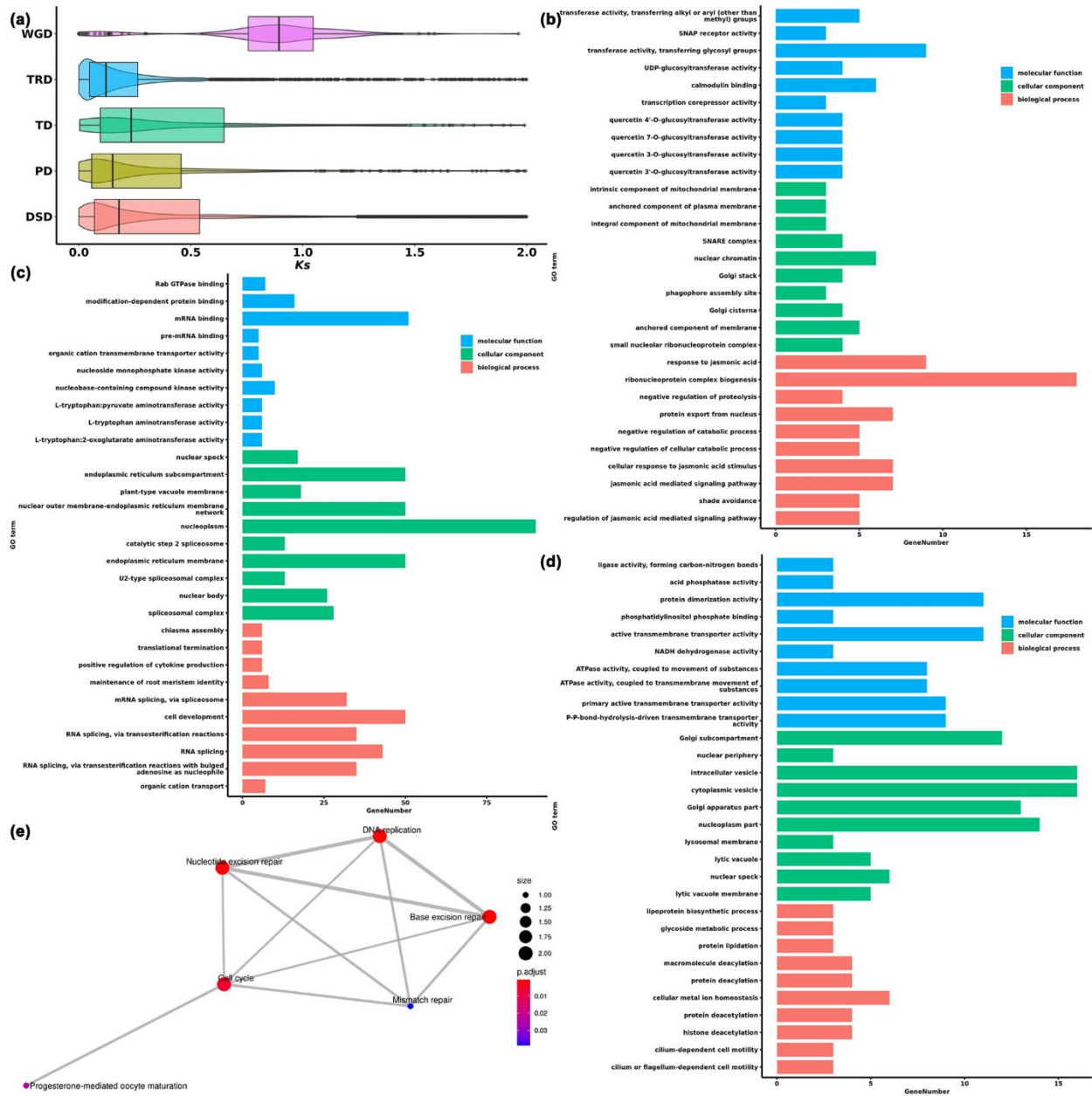
47

48

49 **Data S2 D. Genome characters of *T. lepidozooides***

50 (a) Synteny blocks of *T. lepidozooides* genome. (b) Ks values distribution reveal diploidy events
 51 in *T. lepidozooides*.

52



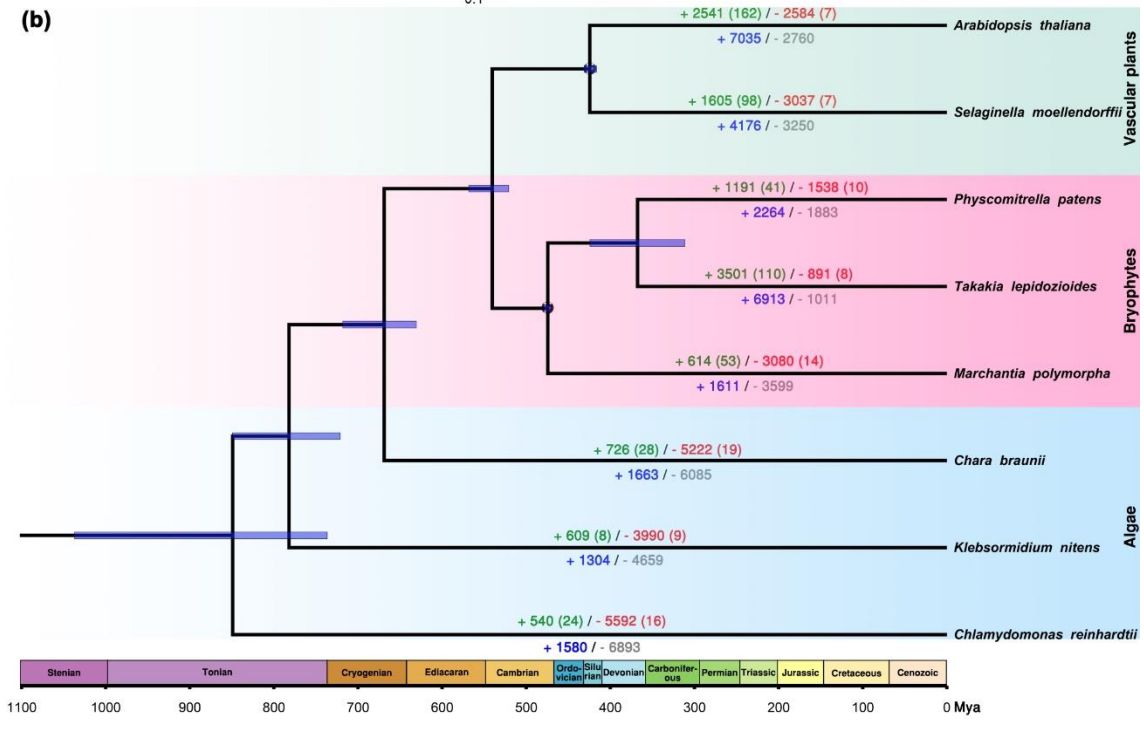
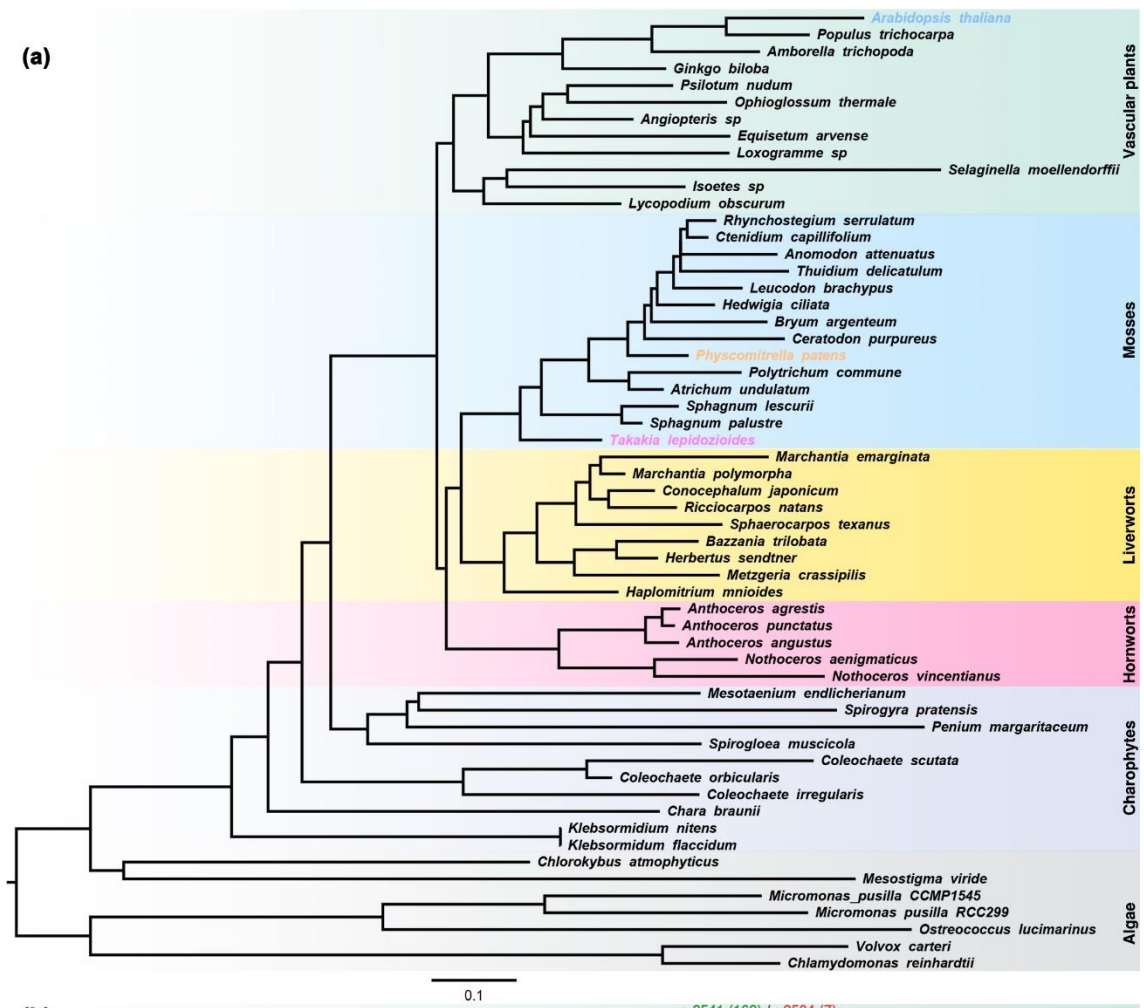
53

54

55 Data S2 E. Evolution and function of duplicated genes in *T. lepidozoides*

56 (a) Ks distribution of the five modes of gene duplications. WGD: whole-genome duplication, TD:
 57 tandem duplication, PD: proximal duplication, TRD: transposed duplication, DSD: dispersed
 58 duplication. In the boxplot representations, points are outliers; the centre line is the median; the
 59 lower and upper hinges correspond to the first and third quartiles (25th and 75th percentiles);
 60 whiskers extend to the minimum (left whiskers) and maximum (right whiskers) estimates located
 61 within the 1.5x interquartile range (IQR) of the first and third quartiles, respectively. (b) Gene

62 ontology (GO) enrichment of PD in *T. lepidozoides* (c) GO enrichment of TRD in *T. lepidozoides*.
63 (d) GO enrichment of WGD in *T. lepidozoides*. (e) Kyoto Encyclopedia of Genes and Genomes
64 (KEGG) enrichment of singleton genes in *T. lepidozoides*.
65



67

68 **Data S2 F. Phylogenetic position and divergence time of *T. lepidozoides***

69 (a) Phylogenetic tree of 56 plant species based on 818 low-copy orthologs amino acid sequence.

70 The tree of 56 plant species inferred from five plant clades reveals *T. lepidozoides* as the sister
71 group to all other mosses, rather than to liverworts or to the LCA of mosses and liverworts. (b) *T.*

72 *lepidozoides* time tree with gene family expansion/contraction analysis compared to

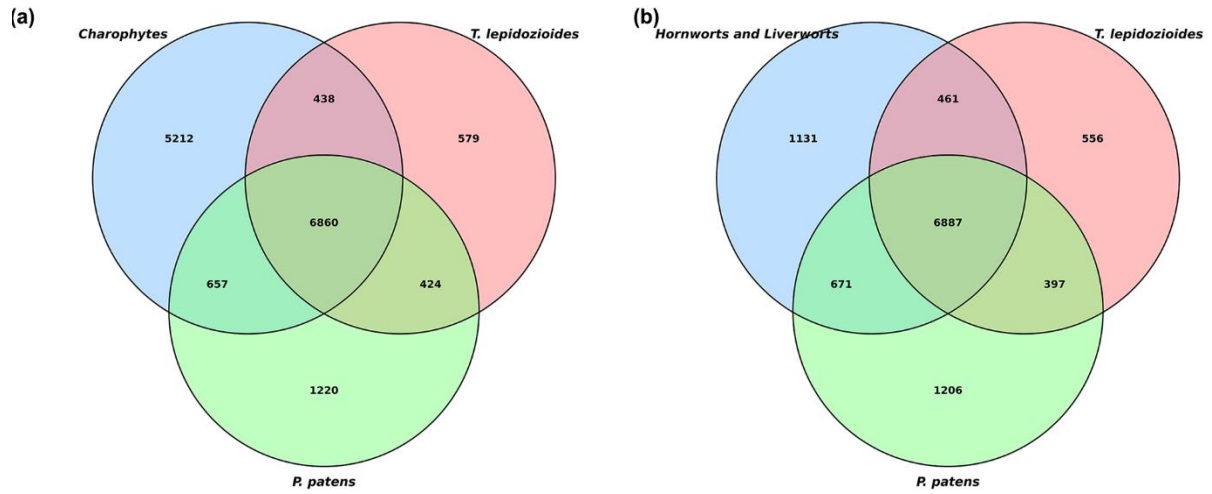
73 representatives of seven other plant lineages, coloured numbers above branches show gene

74 family expansions (+) and contractions (-), coloured numbers under branches show gene gains

75 (+) and losses (-), light purple bars show 95% confidence interval lengths, black dots on nodes

76 mark fossil calibration points.

77



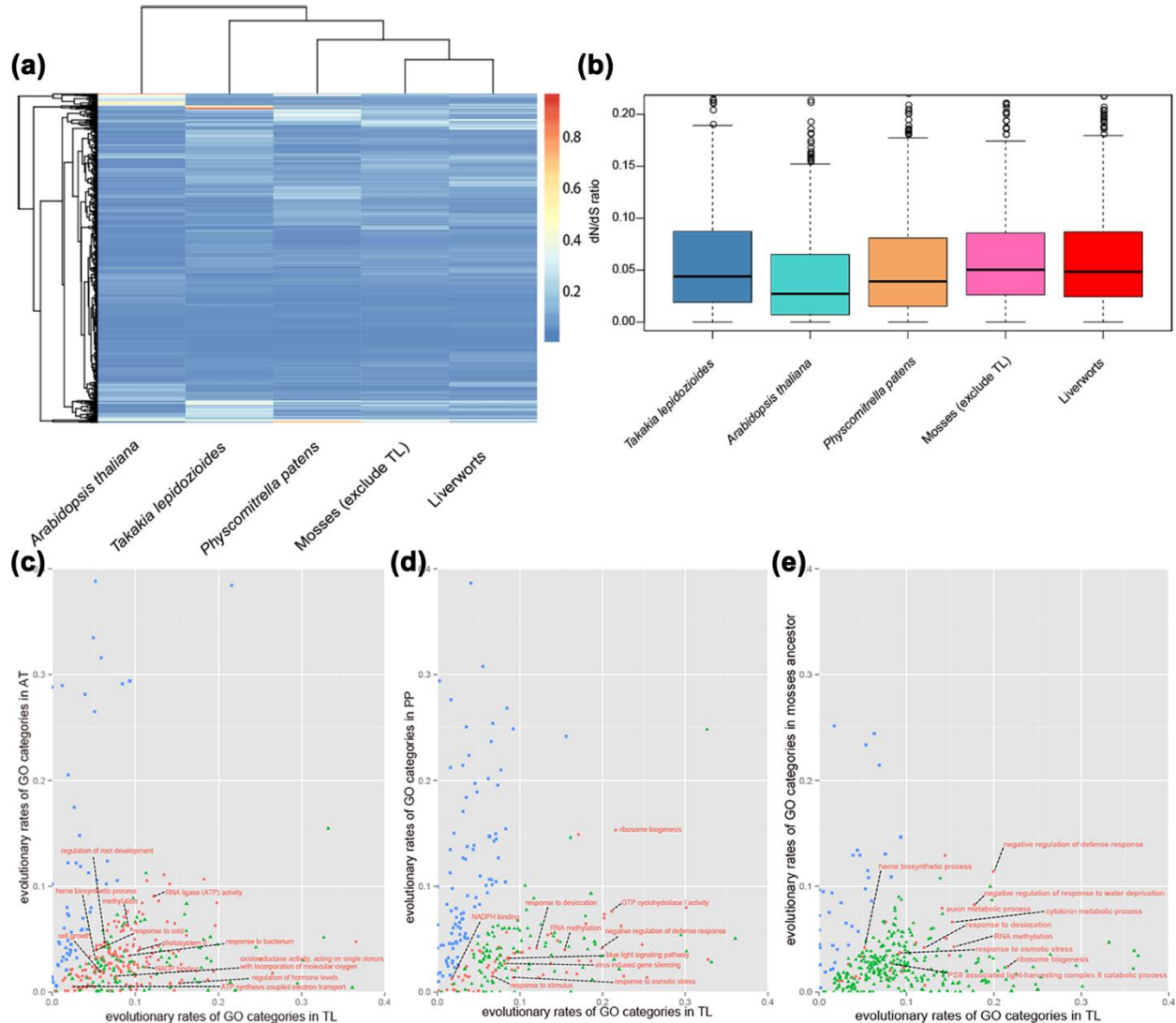
78

79

80 **Data S2 G. Gene families characters of *T. lepidozooides***

81 (a) Gene families of *T. lepidozooides*, *P. patens* and Charophytes (the union of *K. nitens*, *C. braunii*,
 82 *M. endlicherianum* and *S. muscicola* sets). (b) Gene families of *T. lepidozooides*, *P. patens* and
 83 hornworts and liverworts (the union of *A. angustus* and *M. polymorpha* sets).

84



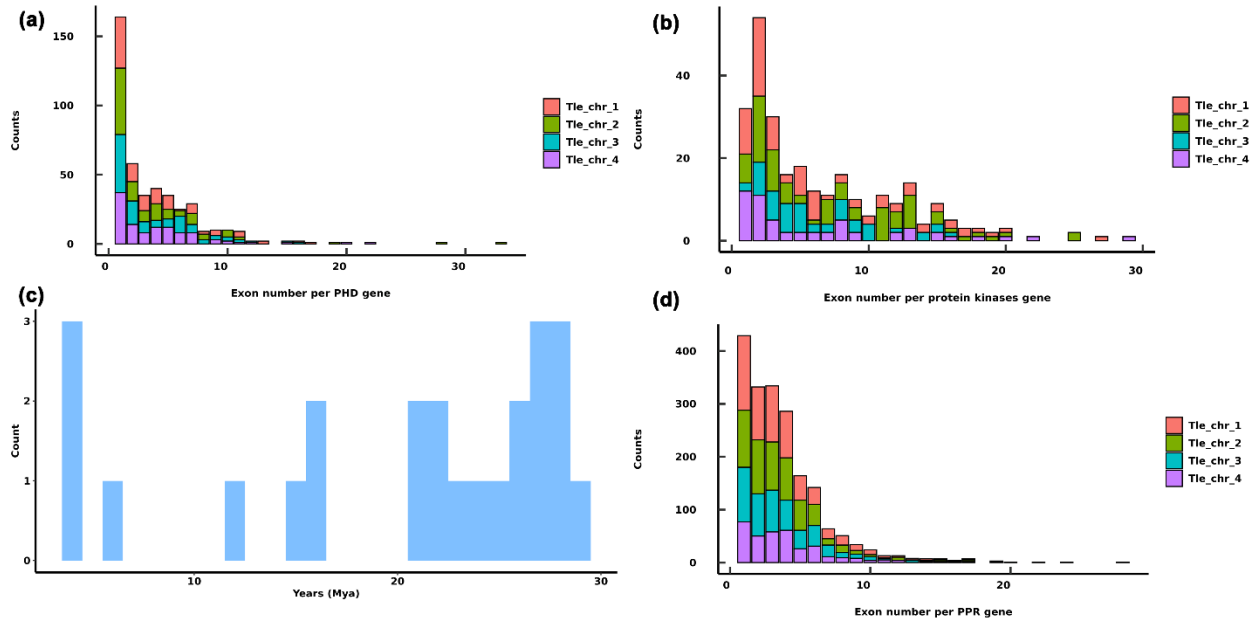
85

86

87 **Data S2 H. Positive selection and fast evolution**

88 (a) Heatmap showing dN/dS ratio for individual genes in *A. thaliana*, *T. lepidozooides*, *P. patens*,
 89 all mosses except *T. lepidozooides*, and all liverworts. Each horizontal line represents a gene. (b)
 90 Boxplot showing dN/dS ratios for individual genes in *T. lepidozooides*, *A. thaliana*, *P. patens*, all
 91 mosses except *T. lepidozooides*, and all liverworts. The upper boundary of the evolution rate range
 92 scale in *T. lepidozooides* is higher than that of other taxon groups surveyed. (c) GO annotation of
 93 significantly fast evolving genes in *T. lepidozooides* and *A. thaliana*. Blue squares represent GO
 94 categories with significantly higher mean dN/dS in *A. thaliana*, red points represent fast evolving
 95 GO categories specific to *T. lepidozooides*, green triangles represent fast evolving GO categories
 96 shared by all mosses; TL, *T. lepidozooides*; AT, *A. thaliana*. (d) GO annotation of significantly fast

97 evolving genes in *T. lepidozoides* and *P. patens*. Blue squares represent GO categories with
98 significantly higher mean dN/dS in *P. patens*, red points represent fast evolving GO categories
99 specific to *T. lepidozoides*, green triangles represent fast evolving GO categories shared by all
100 mosses; TL, *T. lepidozoides*; PP, *P. patens*. (e) GO categories of rapidly evolving genes in *T.*
101 *lepidozoides* and the common ancestor of all mosses including *T. lepidozoides*. Green triangles
102 represent 56 GO categories with significantly higher mean dN/dS in the common ancestor of all
103 mosses, blue squares represent 32 fast evolving GO categories specific to *T. lepidozoides*, red
104 points represent 359 fast evolving GO categories shared by all mosses including *T. lepidozoides*.
105

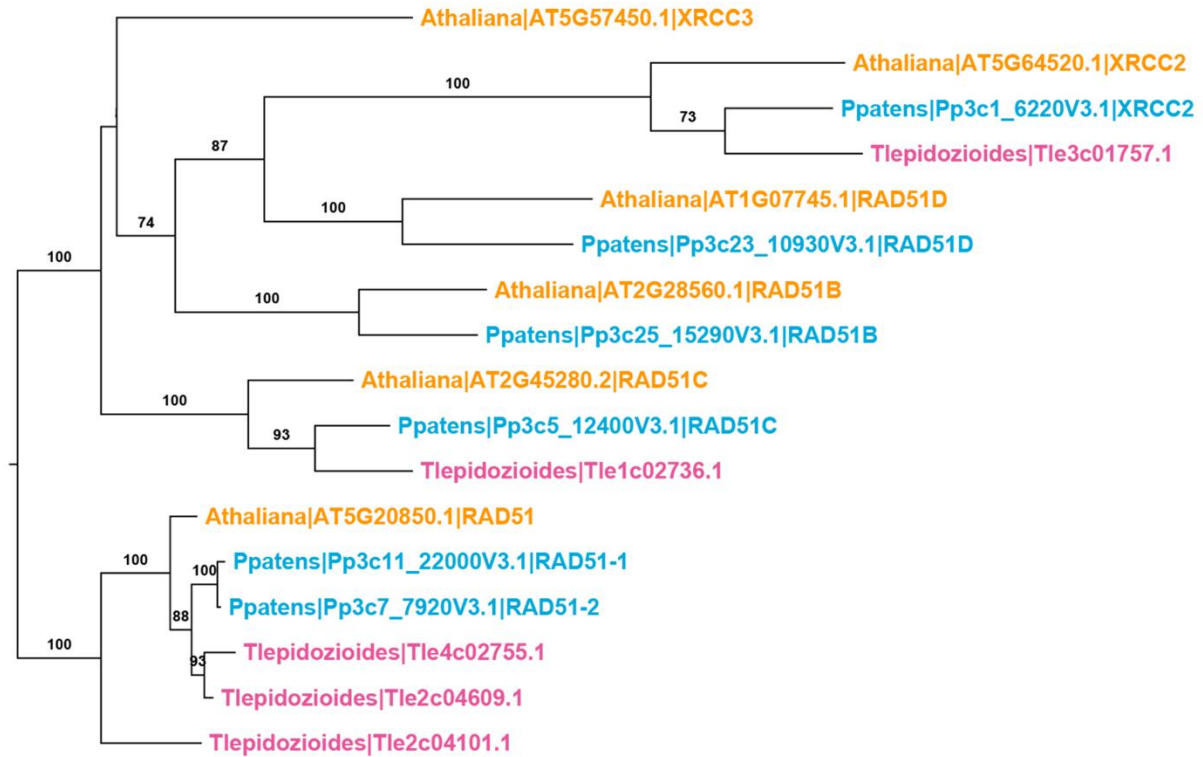


106
107

108 **Data S2 I. Statistics in PHDs, protein kinases and PPRs of *T. lepidozoides***

109 (a) Statistics of exon numbers in each PHD. (b) Statistics of exon numbers in each protein kinase.
110 (c) Numbers of protein kinases which are younger than 30 Mya. (d) Statistics of exon numbers in
111 each PPR.

112



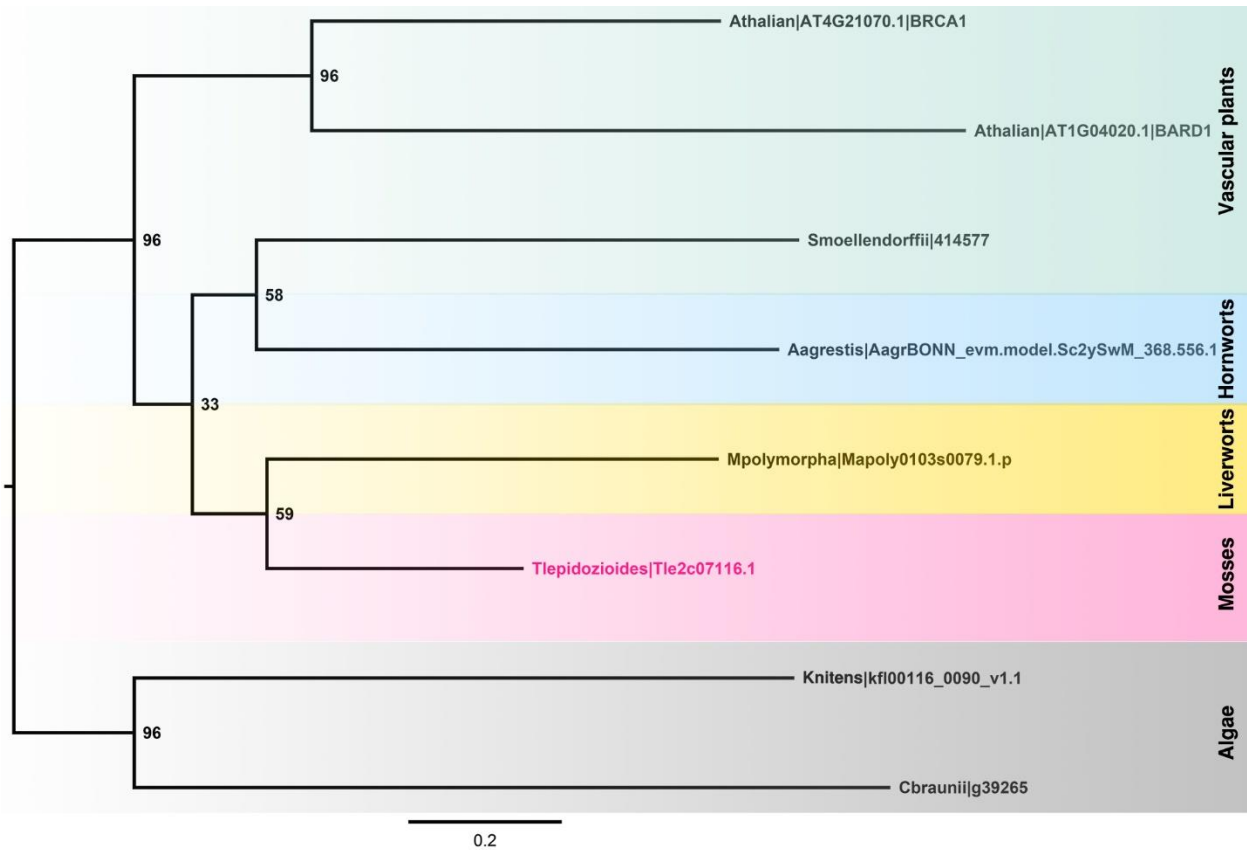
113

114

115 **Data S2 J. Phylogenetic trees of RAD51**

116 Maximum likelihood trees using RAxML with the PROTGAMMAJTT model based upon MAFFT
 117 alignments of proteins are the consensus of 1,000 replicates. Numbered nodes show the
 118 percentage of bootstraps which support the node, bootstraps with less than 50% support are not
 119 shown. Orange font color represents *A. thaliana*, pink font color represents *T. lepidozoides* and
 120 blue font color represents *P. patens*. Shade colors as shown on the right.

121

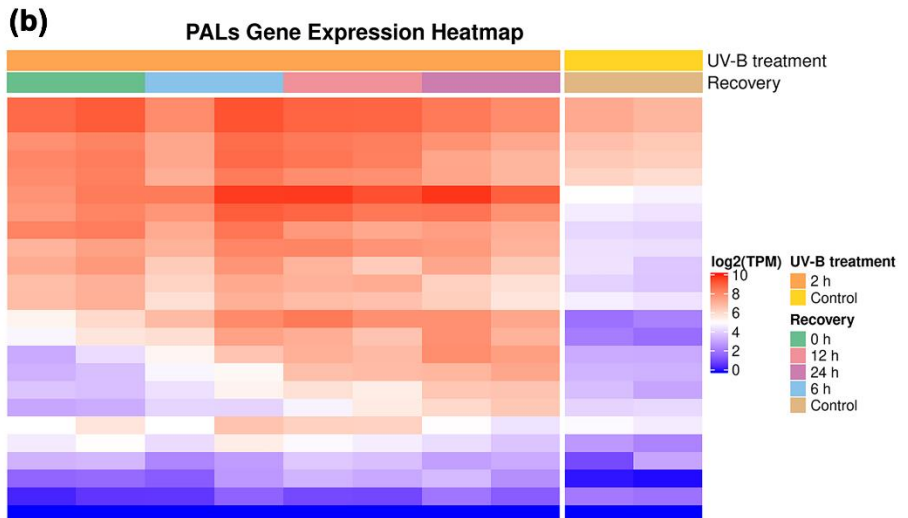
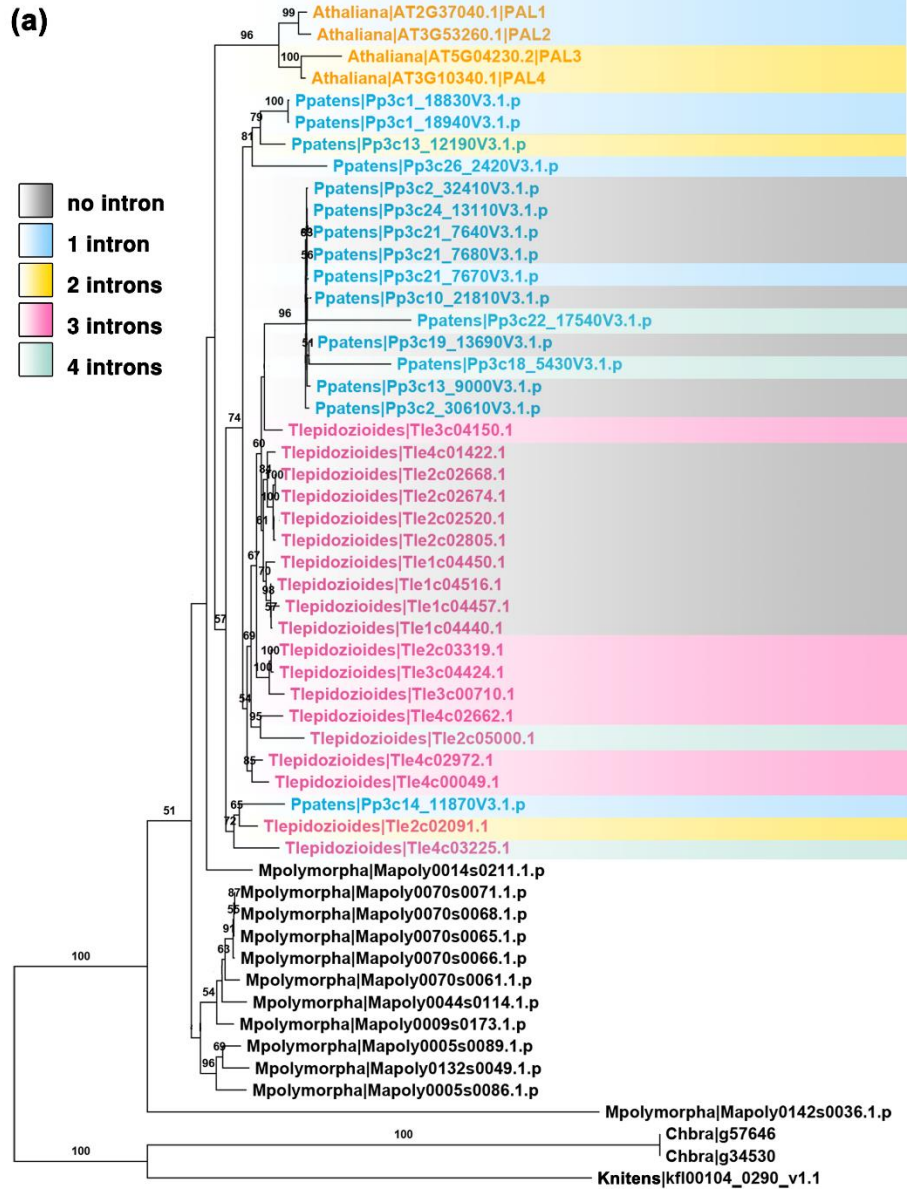


122
123

124 **Data S2 K. Phylogenetic tree of BRCA1-homologs in bryophytes**

125 BRCA1 (AT4G21070) were found using blastp with identity cutoff 50%, e-value cutoff 1×10^{-20} .
 126 Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon MAFFT
 127 alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates.
 128 Numbered nodes show the percentage of bootstraps which supported the node.

129

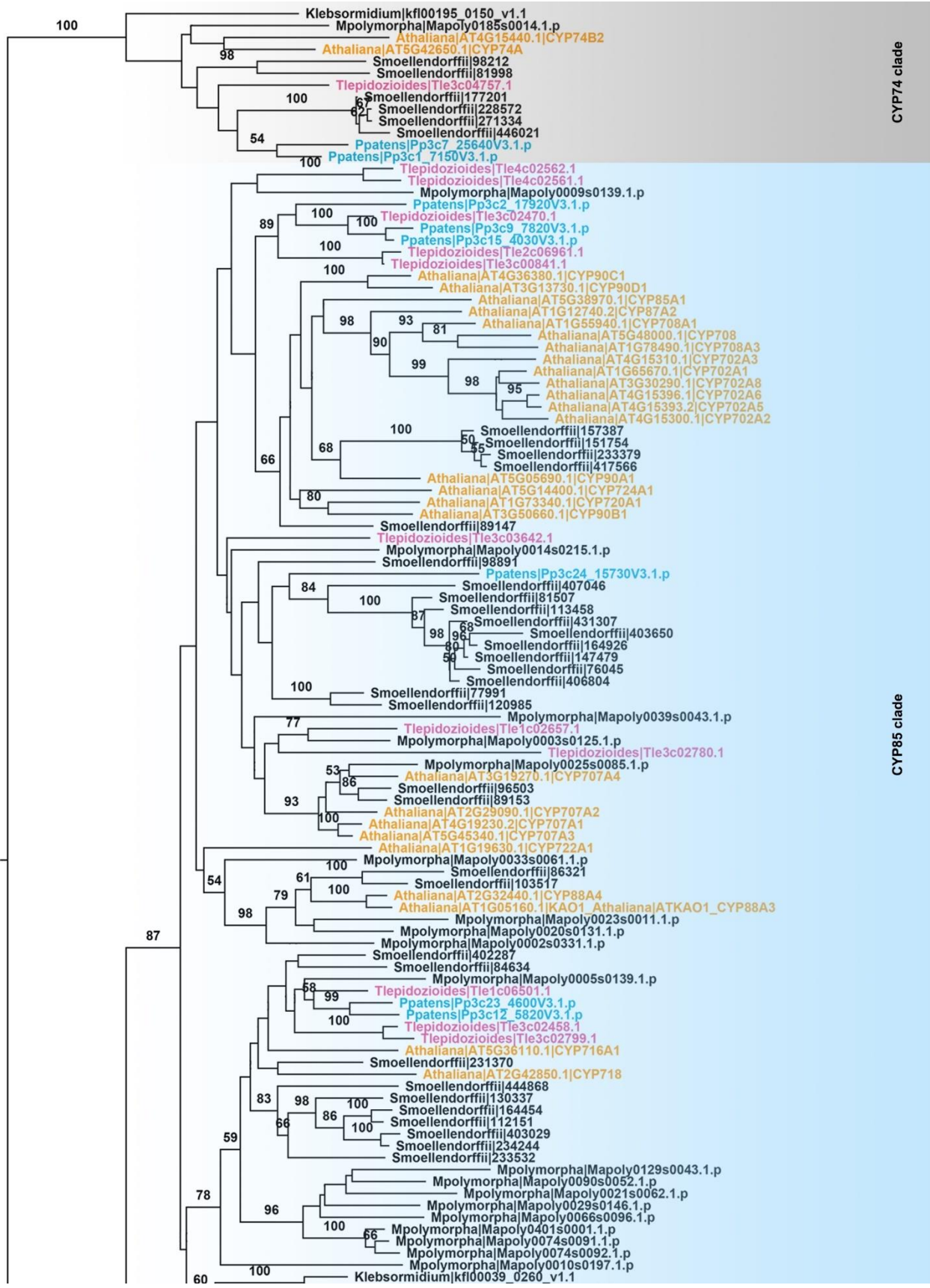


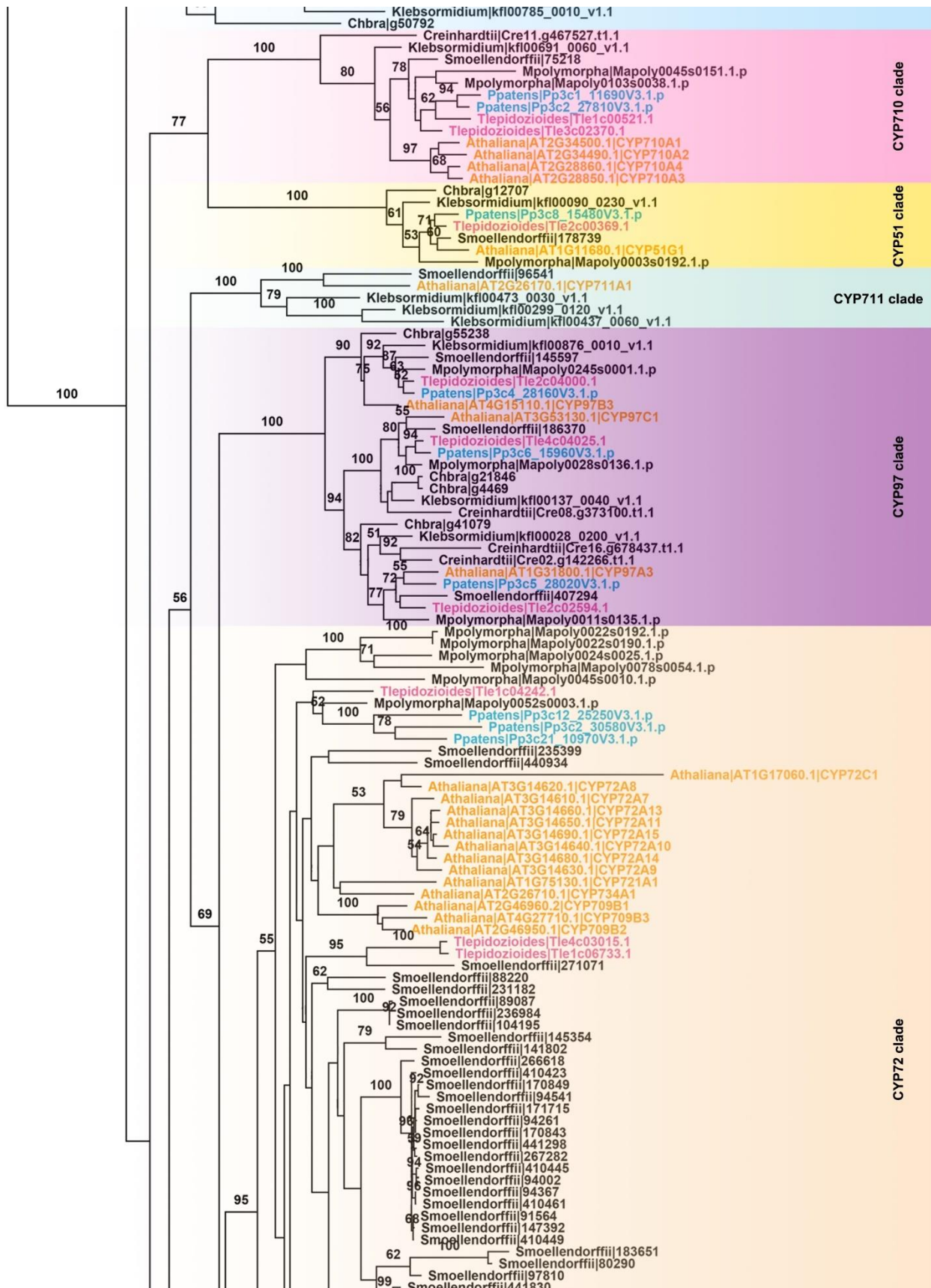
131

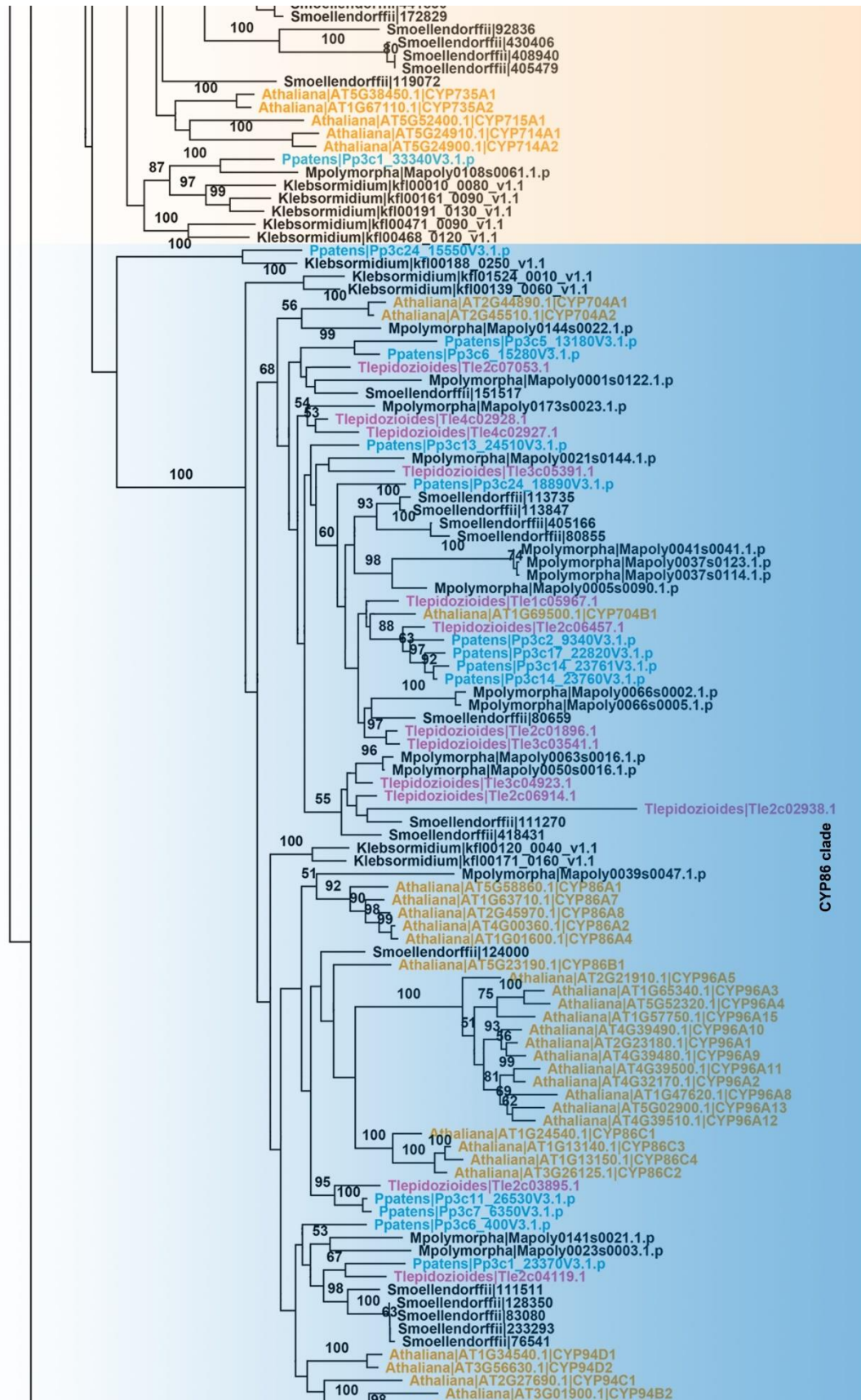
132 **Data S2 L. PAL features in *T. lepidozooides***

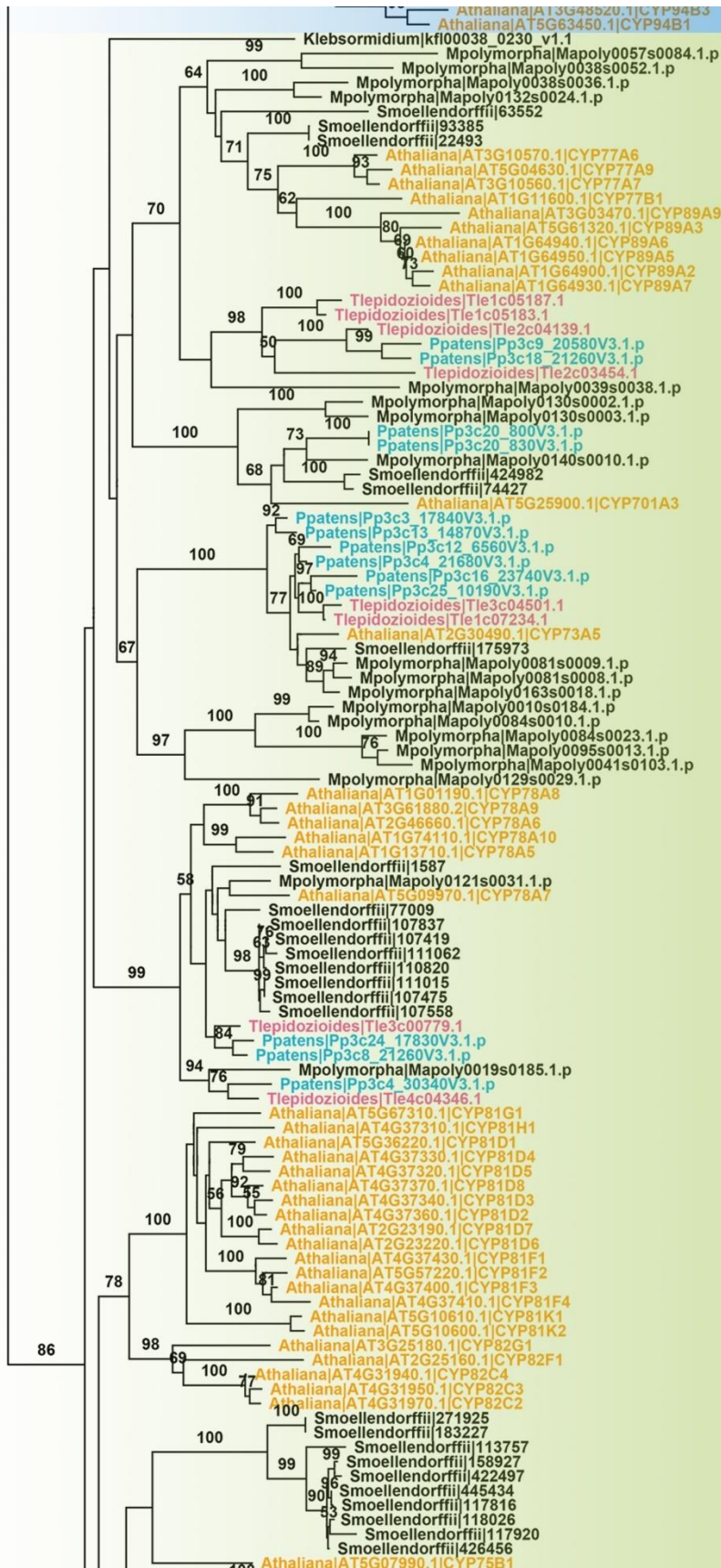
133 (a) Phylogenetic tree of PAL genes. Maximum likelihood tree using RAxML with the
134 PROTGAMMAJTT model based upon MAFFT alignments of proteins and trimmed by trimAl. Tree
135 is the consensus of 1,000 replicates. Numbered nodes show the percentage of bootstraps which
136 support the node, bootstraps with less than 50% support are not shown. Orange font color
137 represents *A. thaliana*, pink font color represents *T. lepidozooides* and blue font color represents
138 *P. patens*. Shade colors as shown on the top left. (b) Expression levels of PALs in response to
139 UV-B radiation and during recovery.

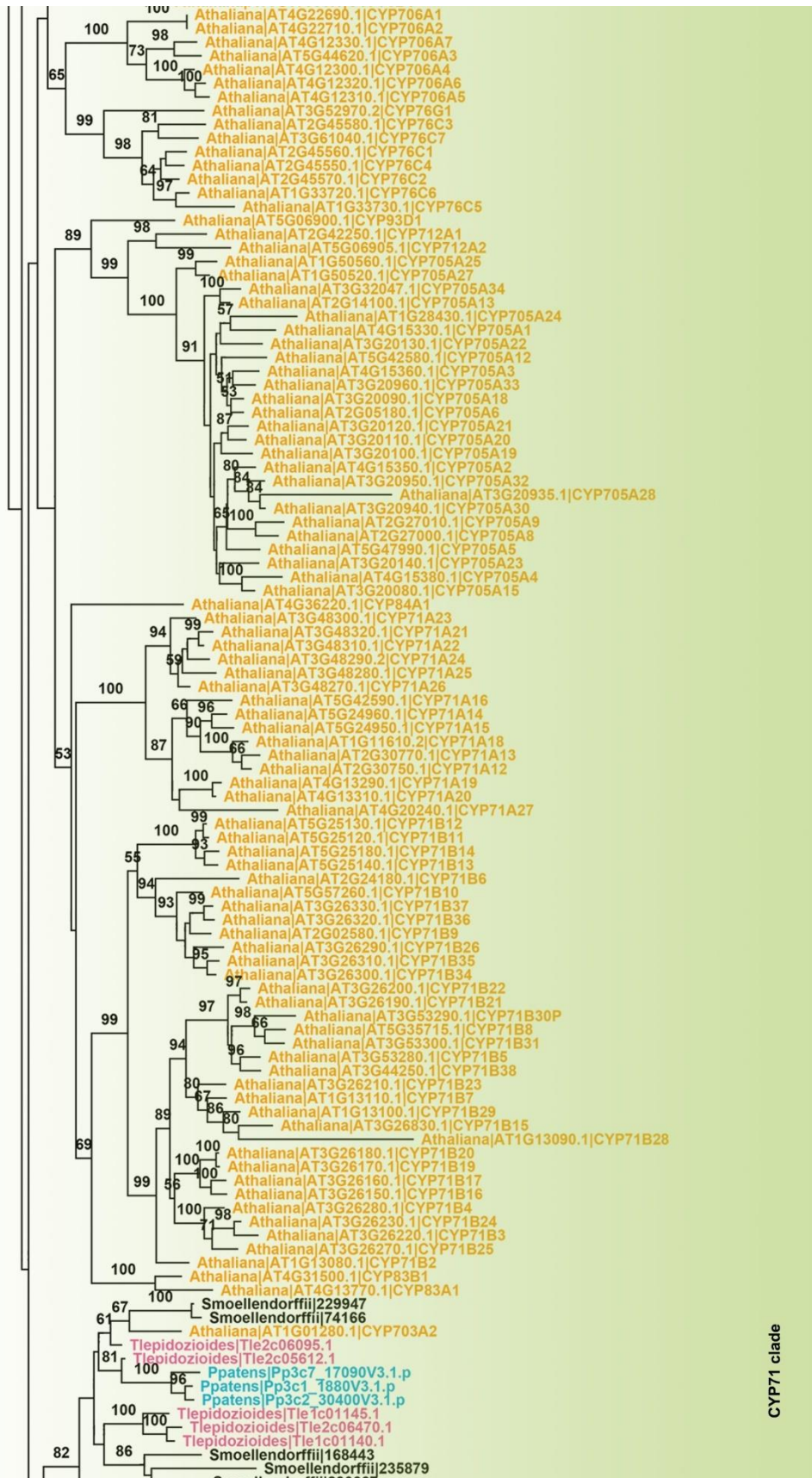
140

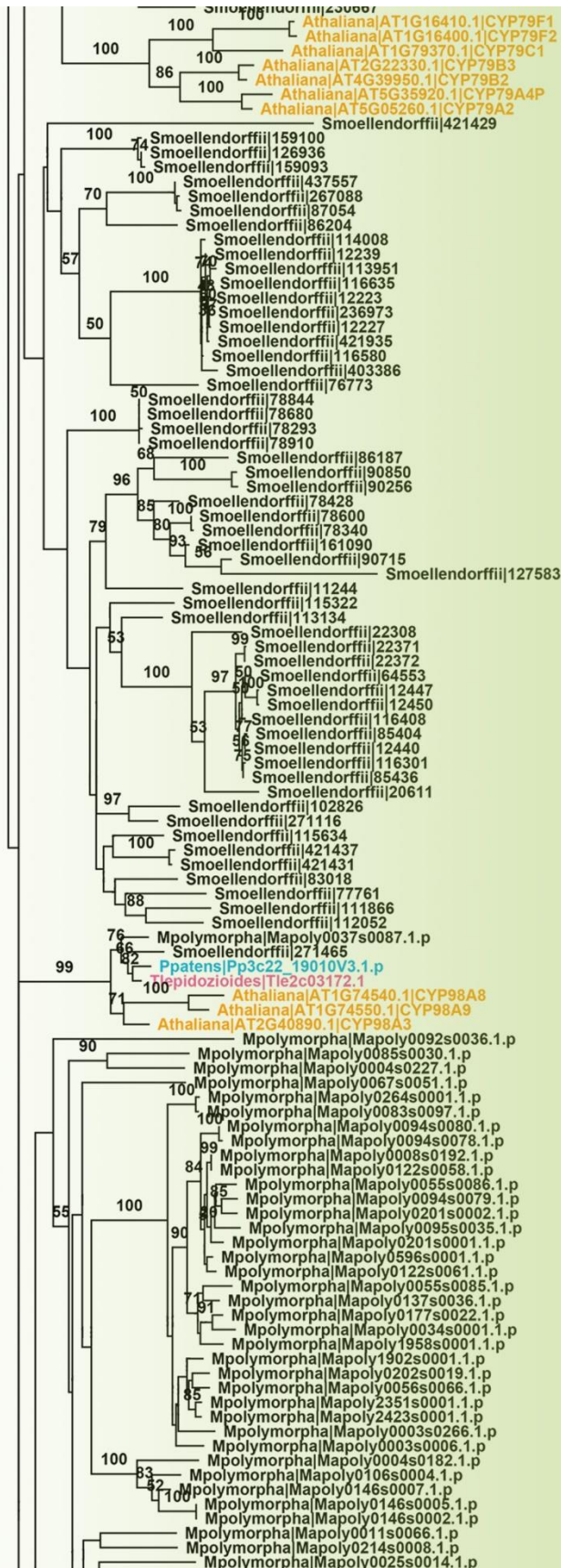


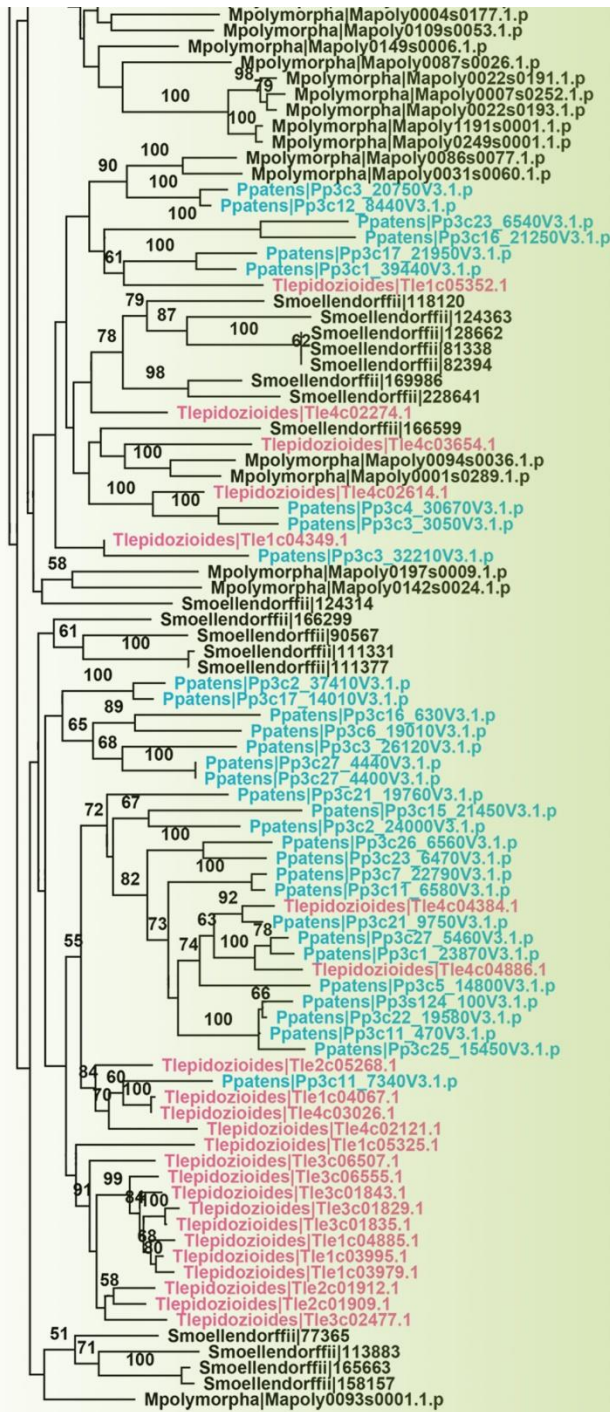












147

148

149 **Data S2 M. Phylogenetic tree of CYP450s**

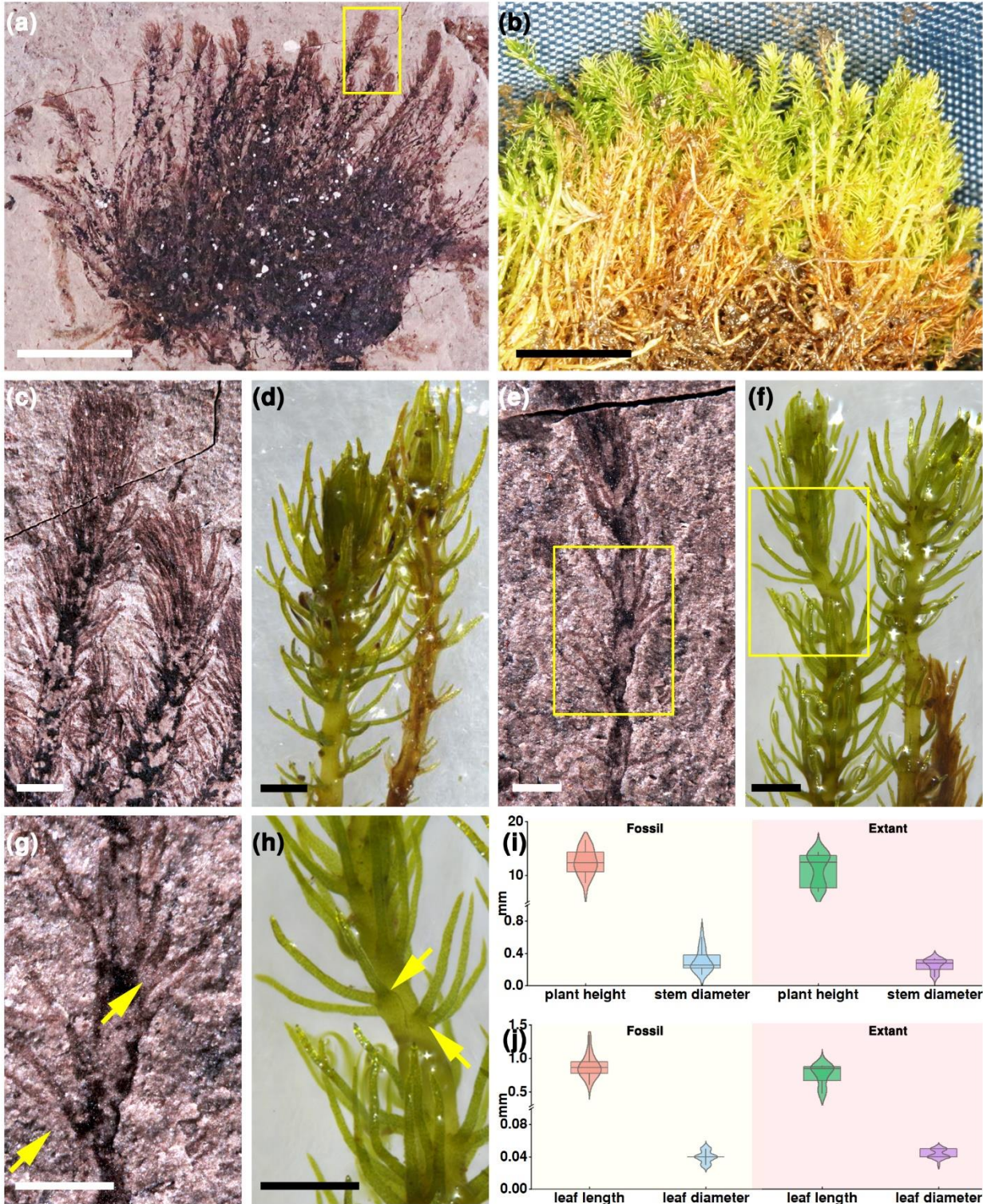
150 Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon MAFFT

151 alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates.

152 Numbered nodes show the percentage of bootstraps which supported the node, bootstraps with

153 less than 50% support are not shown. Orange font color represents *A. thaliana*, pink font color
154 represents *T. lepidozoides* and blue font color represents *P. patens*. Shade colors as shown on
155 the right.

156

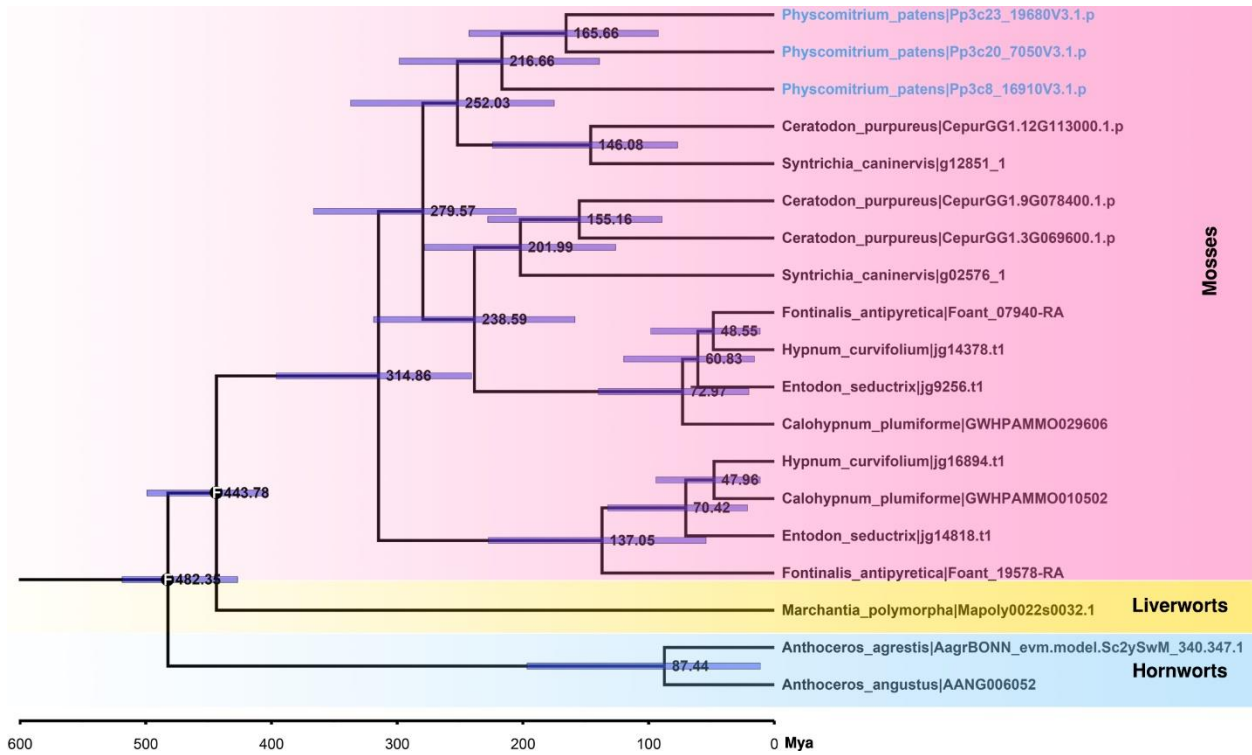


157
158

159 **Data S2 N. Plant morphological characteristics of *Takakia* sp. fossil and *T. lepidozoides***

160 (a) A population of fossil *Takakia* sp. from the Middle–Late Jurassic Daohugou Biota, China, bar
161 = 1 cm. (b) A natural population of *T. lepidozoides* from the experiment site of Tibetan Plateau,
162 bar = 1 cm. (c) Magnified views of boxed areas in (a), bar = 500 μm . (d) Extant plants of *T.*
163 *lepidozoides* in the experiment site of Tibetan Plateau, bar = 500 μm . (e) Details of fossil *Takakia*
164 sp. fossil plant with lobed terete leaves, bar = 500 μm . (f) Details of *T. lepidozoides* plant with
165 lobed terete leaves in the experiment site of Tibetan Plateau, bar = 500 μm . (g) Magnified views
166 of boxed areas in (e), a yellow arrow marks a lobed terete leaf, bar = 500 μm . (h) Magnified views
167 of boxed areas in (f), a yellow arrow marks a lobed terete leaf, bar = 500 μm . (i) The plant type,
168 height 12.48 ± 2.46 mm, stem diameter 0.31 ± 0.14 mm in fossil ($n = 30$); height 11.14 ± 2.90 mm,
169 stem diameter 0.26 ± 0.08 mm in extant plants ($n = 25$). (j) The leaf type, leaf length 0.88 ± 0.17
170 mm, leaf diameter 0.04 ± 0.01 mm in fossil ($n = 30$); leaf length 0.75 ± 0.16 mm, leaf diameter
171 0.04 ± 0.01 mm in extant plants ($n=25$).

172

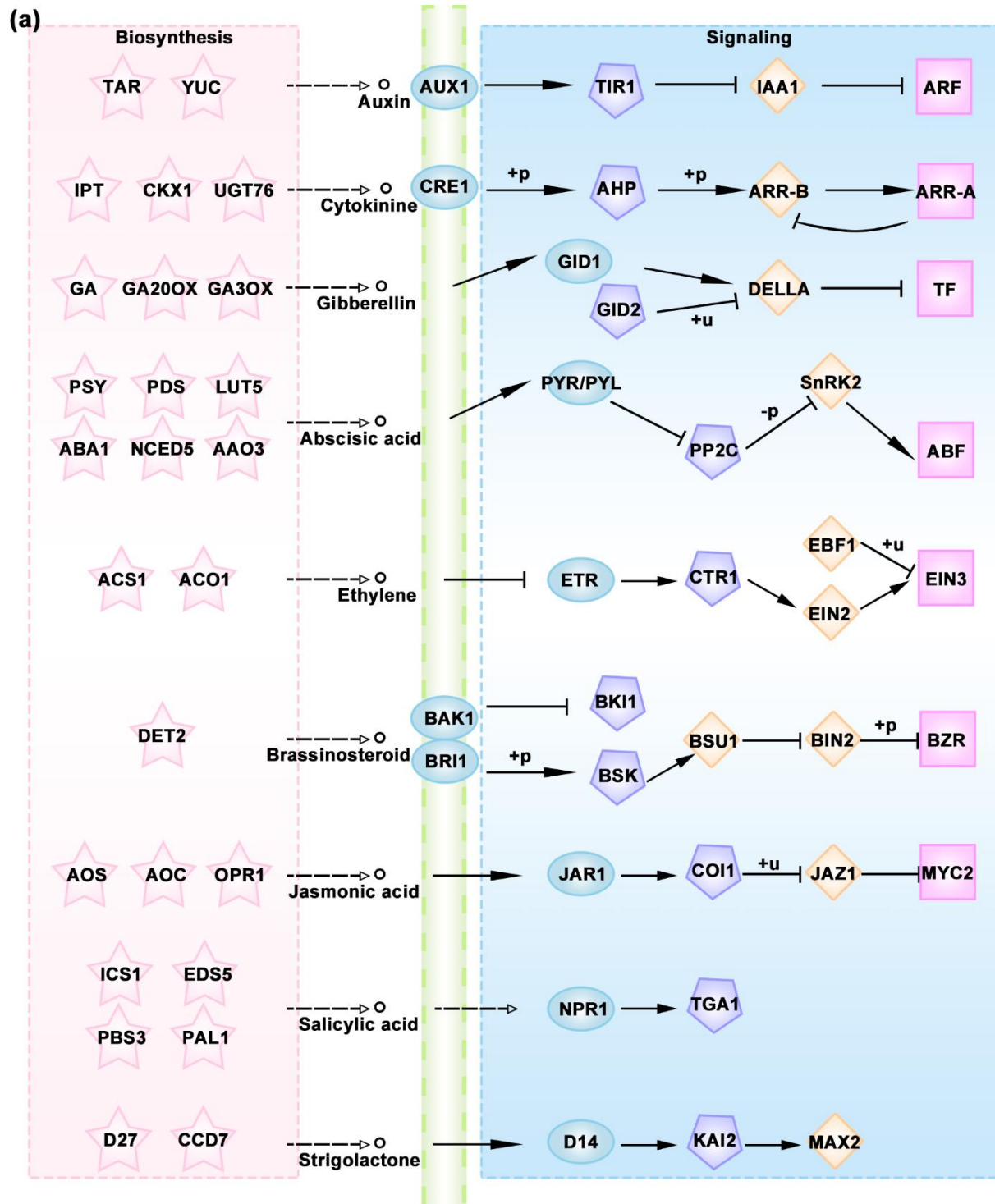


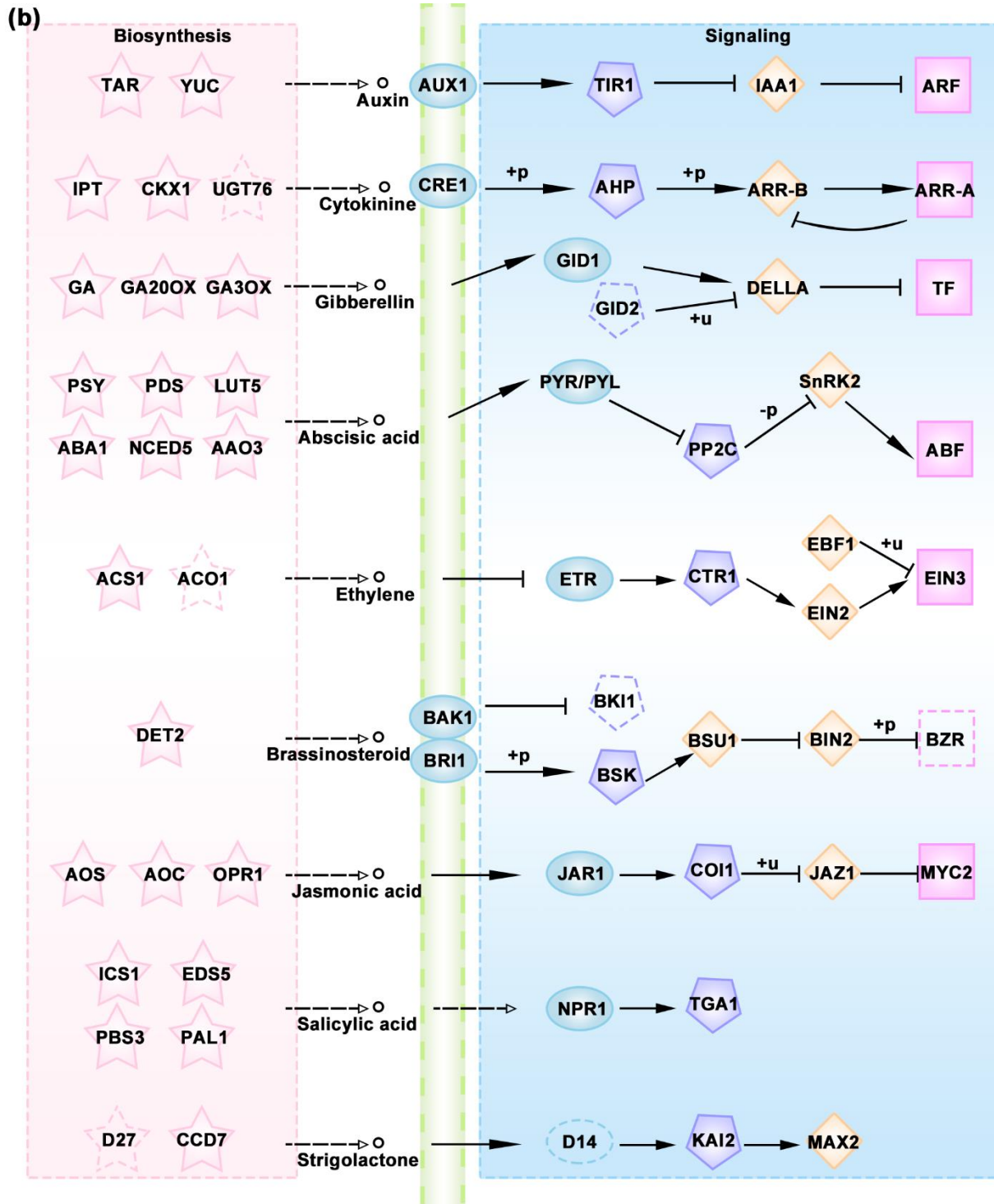
173
174

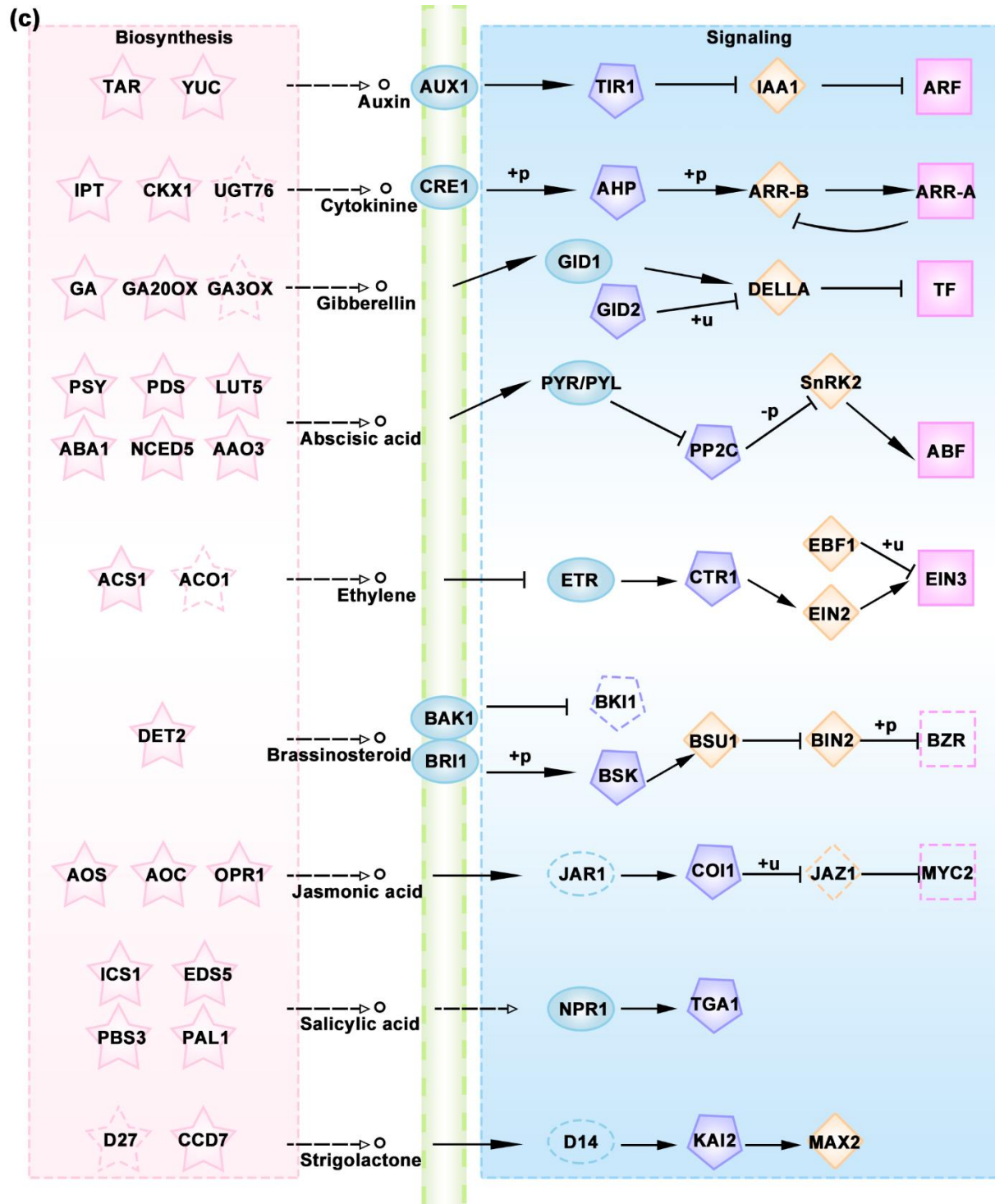
175 **Data S2 O. Phylogenetic timetree of KAN-homologs in bryophytes**

176 KAN (AT5G16560) were found using blastp with identity cutoff 50%, e-value cutoff 1×10^{-20} .
 177 Maximum likelihood tree using RAXML with the PROTGAMMAJTT model based upon MAFFT
 178 alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates. Numbers
 179 on nodes represent divergence time, light purple bars show 95% confidence interval lengths,
 180 nodes with black dots show fossil calibration points.

181







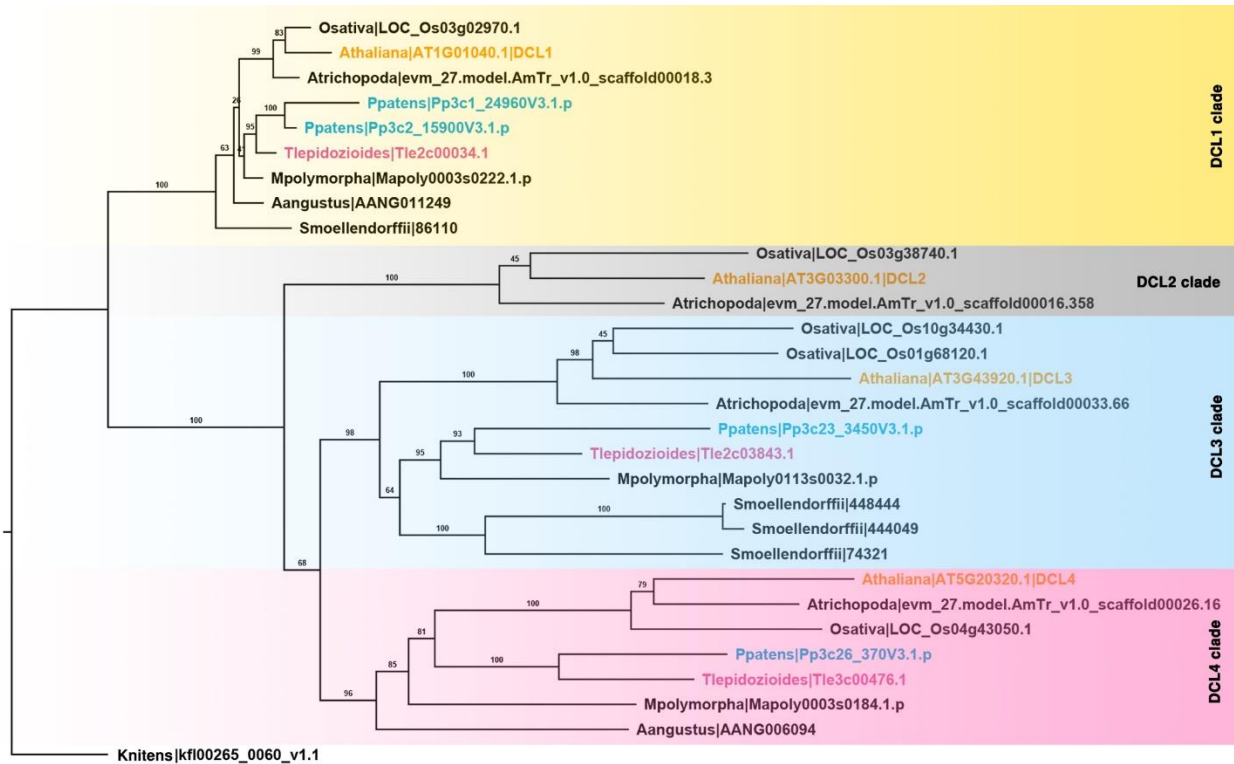
184

185

186 **Data S2 P. Comparison of hormone pathway genes in *A. thaliana*, *P. patens* and *T.***
187 ***lepidozoides***

188 Plant hormones pathways in *A. thaliana* (**a**), *P. patens* (**b**) and *T. lepidozoides* (**c**). Dashed lines
189 with gene name mark the genes absent in the pathway, arrows indicate positive/promoting
190 regulation, and lines with perpendicular end bars indicate negative/repressing regulation. *T.*
191 *lepidozoides* lacks MYC2 and JAZ1, which may be involved in leaf expansion.

192



193

194

195 **Data S2 Q. Phylogenetic trees of Dicer proteins**

196 BLASTP was performed using Dicer of *A. thaliana* sequences to find candidate Dicer proteins.

197 Putative Dicer proteins were selected with “ $E < 1 \times 10^{-20}$, identity > 30%, coverage > 70%”.

198 Maximum likelihood trees using RAxML with the PROTGAMMAJTT model based upon MAFFT

199 alignments of proteins are the consensus of 1,000 replicates. Numbered nodes show the

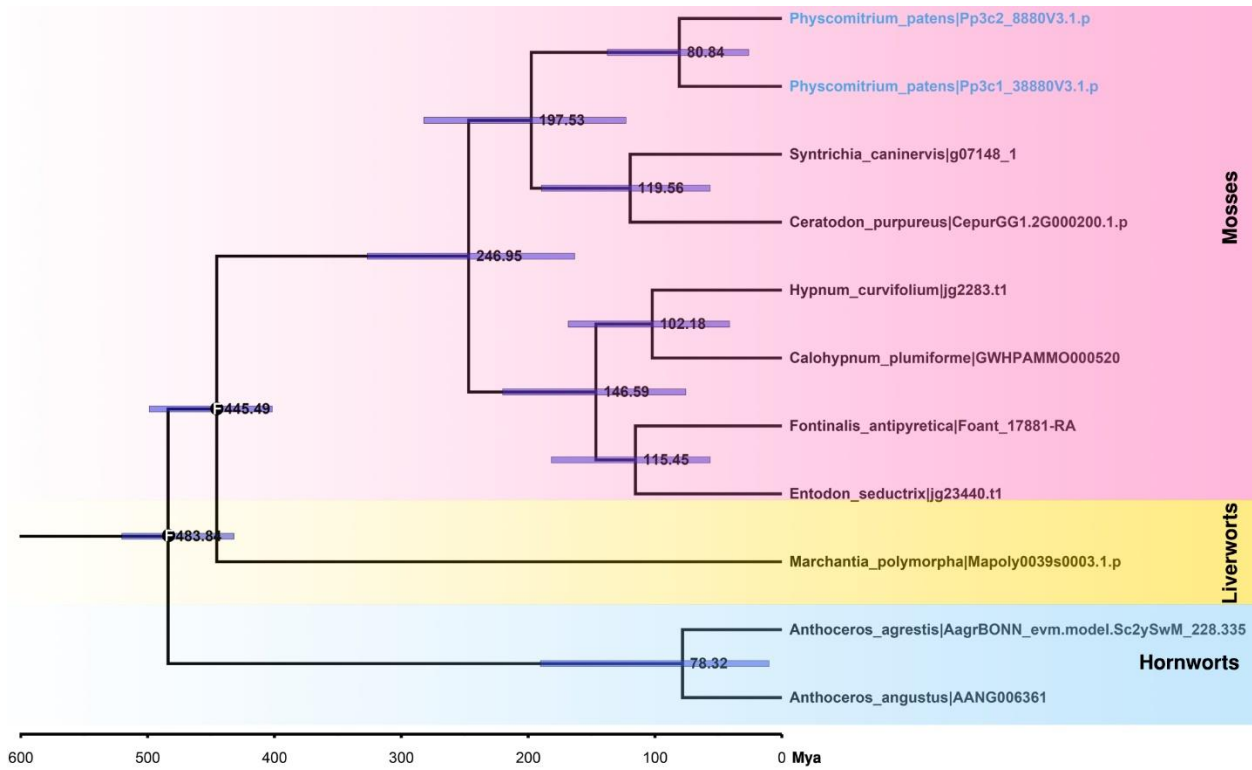
200 percentage of bootstraps which support the node, bootstraps with less than 50% support are not

201 shown. The four clades of DICER-like (DCL) plant proteins are color-coded. Orange font color

202 represents *A. thaliana*, pink font color represents *T. lepidozoides* and blue font color represents

203 *P. patens*. Shade colors as shown on the right.

204



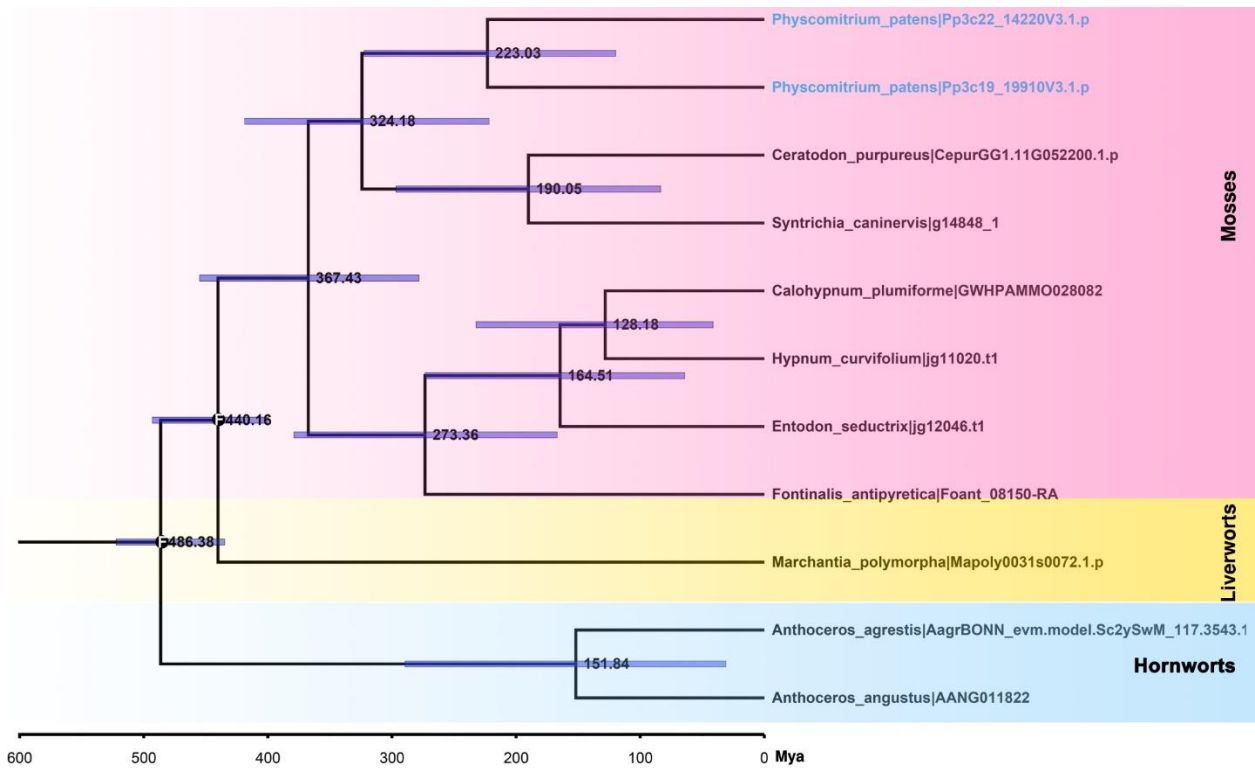
205

206

207 **Data S2 R. Phylogenetic timetree of RHD6-homologs in bryophytes**

208 RHD6 (AT1G66470) homologs were found using blastp with identity cutoff 50%, e-value cutoff
 209 1×10^{-20} . Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon
 210 MAFFT alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates.
 211 Numbers on nodes represent divergence time, light purple bar shows 95% confidence interval
 212 lengths, nodes with black dot show fossil calibration points.

213

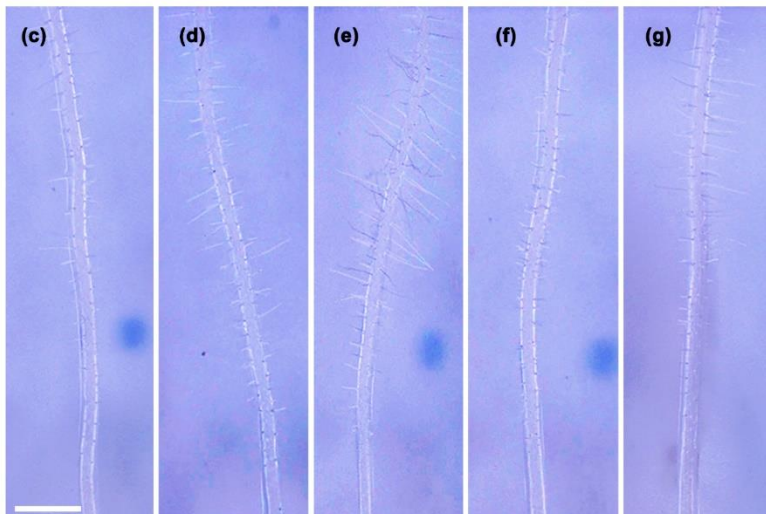
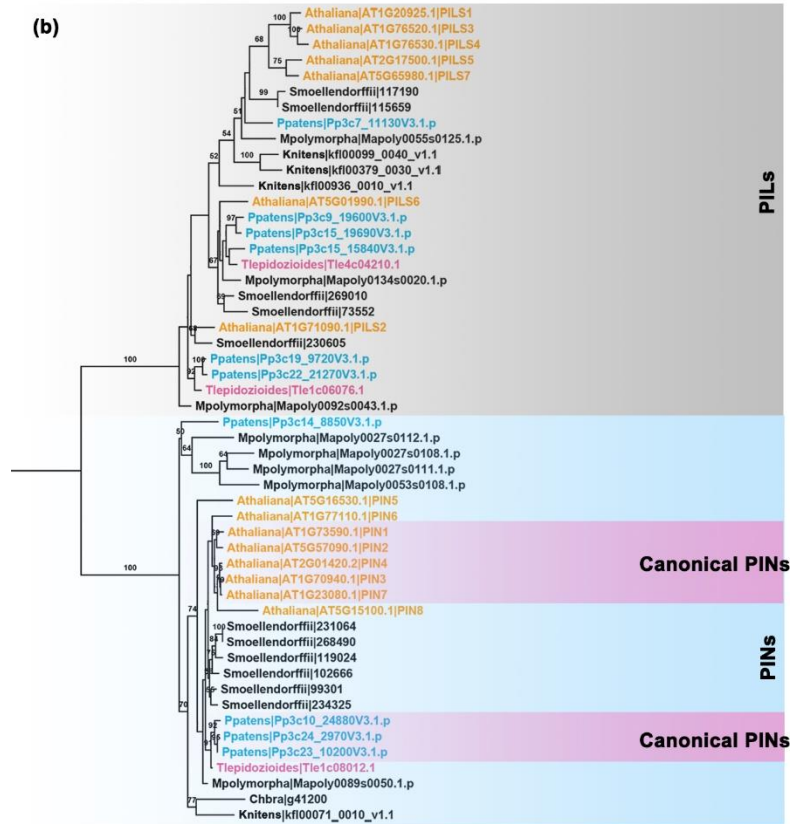
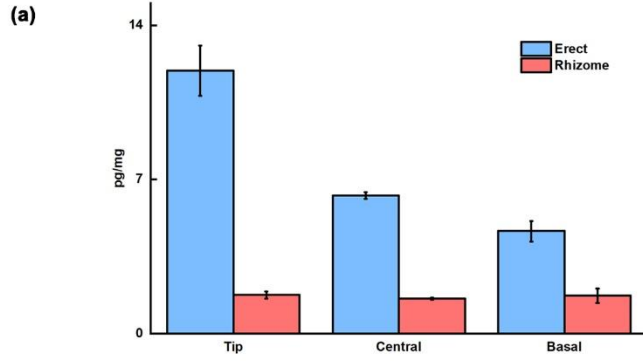


214
215

216 **Data S2 S. Phylogenetic tree of SMF-homologs in bryophytes**

217 FAMA (AT3G24140) homologs were found using blastp with identity cutoff 30%, e-value cutoff
 218 1×10^{-20} . Then HMMER used to check bHLH domain in FAMA-homolog candidates. Maximum
 219 likelihood tree using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments
 220 of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates. Numbers on nodes
 221 represent divergence time, light purple bar show 95% confidence interval length, nodes with black
 222 dot show fossil calibration points.

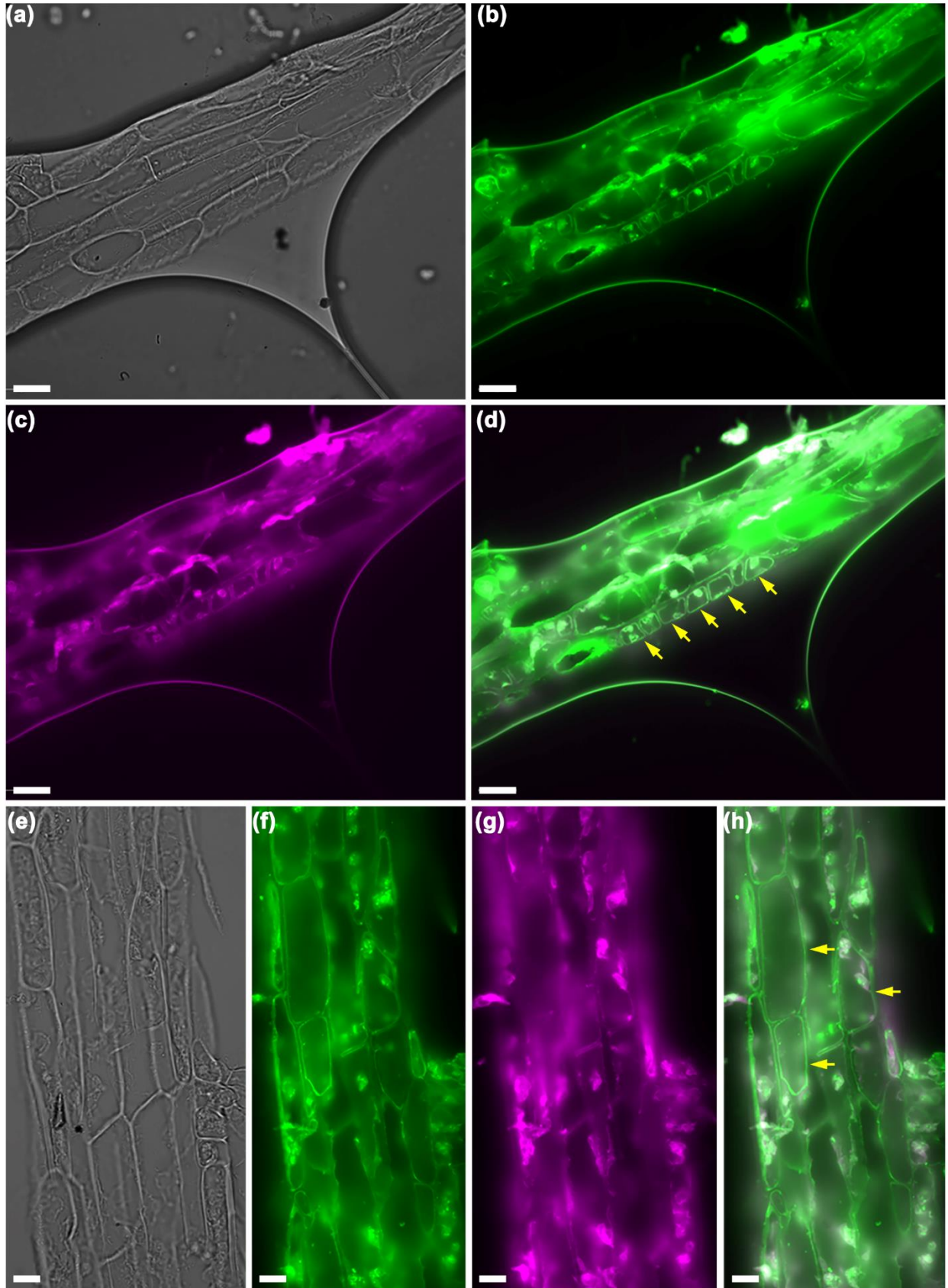
223



225

226 **Data S2 T. IAA content and PINs**

227 (a) Gradient of IAA content from top to base in *T. lepidozooides* stems (mean \pm s.d.; n = 5). Auxin
228 concentrations were highest at the top and lowest at the stem base, and there was no detectable
229 gradient in rhizomes. (b) Phylogenetic tree of PINs. Maximum likelihood tree using RAxML with
230 the PROTGAMMAJTT model based upon MAFFT alignments of proteins are the consensus of
231 1,000 replicates. Numbered nodes show the percentage of bootstraps which support the node.
232 Orange font color represents *A. thaliana*, pink font color represents *T. lepidozooides* and blue font
233 color represents *P. patens*. Shade colors as shown on the right. (c – g) The root phenotypes of
234 *TIPIN1-3* transgenic *A. thaliana* plants, bar = 200 μ m. (c) TIPIN1, (d) TIPIN2, (e) TIPIN3, (f)
235 PpPINB, (g) AtPIN2.
236

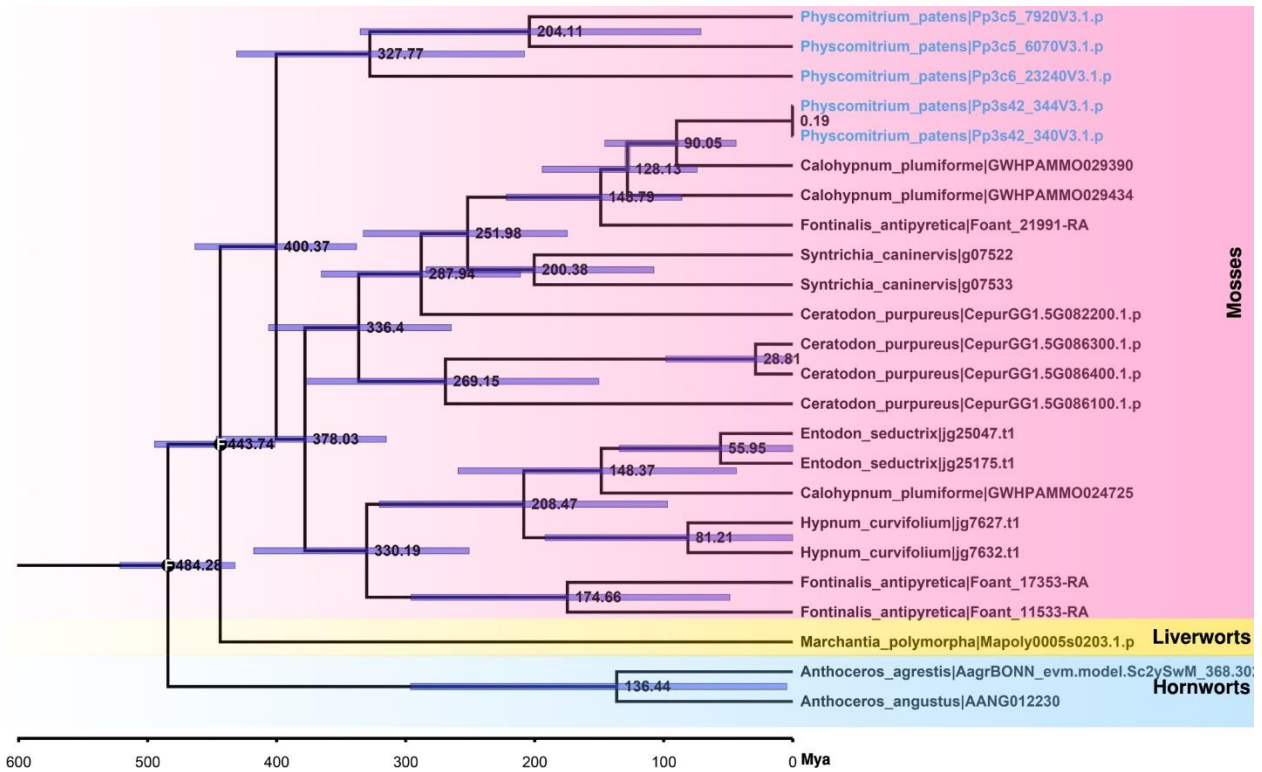


238

239 **Data S2 U. Immune-localization of anti-AtPIN antibodies in cells of *T. lepidozoides***

240 The positive signal (arrows) of anti-AtPIN antibodies immune-localization in leaf cells (**a – d**) and
241 erect-stem cells (**e – h**). Images of bright-field microscopy of leaf cells (**a**) and erect-stem cells
242 (**e**). The immunofluorescence micrographs of the localization of anti-AtPIN1 antibodies for leaf
243 cells (**b**) and erect-stem cells (**f**). The chlorophyll signals for leaf cells (**c**) and erect-stem cells (**g**).
244 The merged views of antibody signals and chlorophyll signals for leaf cells (**d**) and erect-stem
245 cells (**h**). Yellow arrows mark the plasma membrane. Bar = 20 μm .

246



247
248

249 **Data S2 V. Phylogenetic tree of CSD-homologs in bryophytes**

250 CSD domains were found using HMMER. Maximum likelihood tree using RAXML with the
 251 PROTGAMMAJTT model based upon MAFFT alignments of proteins and trimmed by trimAl. Tree
 252 is the consensus of 1,000 replicates. Numbers on the nodes represent divergence time, light
 253 purple bars show 95% confidence interval length, nodes with black dots show fossil calibration
 254 points.

255

256 **SUPPLEMENTAL REFERENCES**

- 257 S1. Reski, R., Faust, M., Wang, X.H., Wehe, M., and Abel, W.O. (1994). Genome analysis of
258 the moss *Physcomitrella patens* (Hedw.) B.S.G. Mol. Gen. Genet. 244, 352-359.
- 259 S2. Marcais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel
260 counting of occurrences of *k*-mers. Bioinformatics 27, 764–770.
261 10.1093/bioinformatics/btr011.
- 262 S3. Vurture, G.W., Sedlazeck, F.J., Nattestad, M., Underwood, C.J., Fang, H., Gurtowski, J.,
263 and Schatz, M.C. (2017). GenomeScope: fast reference-free genome profiling from short
264 reads. Bioinformatics 33, 2202–2204. 10.1093/bioinformatics/btx153.
- 265

