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.

Research Article

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

AUTHORS

Ruoyang Hu[#1,14†], Xuedong Li[#1], Yong Hu[1], Runjie Zhang[1], Qiang Lv[1], Min Zhang[1], Xianyong Sheng[1], Feng Zhao[1], Zhijia Chen[1], Yuhan Ding[1], Huan Yuan[1], Xiaofeng Wu[1], Shuang Xing[1], Xiaoyu Yan[1], Fang Bao[1], Ping Wan[1], Lihong Xiao[1,19†], Xiaoqin Wang[1], Wei Xiao[1], Eva L. Decker[2], Nico van Gessel[2], Hugues Renault[2,20†], Gertrud Wiedemann[2,21†], Nelly A. Horst[2,22†], Fabian B. Haas[3], Per K.I. Wilhelmsson[3], Kristian K. Ullrich[3], Eva Neumann[3], Bin Lv[4,23†], Chengzhi Liang[5], Huilong Du[5,24†], Hongwei Lu[5,25†], Qiang Gao[5,8†], Zhukuan Cheng[5,26†], Hanli You[5,26†], Peiyong Xin[5], Jinfang Chu[5], Chien-Hsun Huang[6], Yang Liu[7,27,28†], Shanshan Dong[27], Liangsheng Zhang[8], Fei Chen[9], Lei Deng[10], Fuzhou Duan[10], Wenji Zhao[10], Kai Li[11], Zhongfeng Li[11], Xingru Li[11], Hengjian Cui[12], Yong Zhang[13], Chuan Ma[14], Ruiliang Zhu[15], Yu Jia[16], Meizhi Wang[16], Mitsuyasu Hasebe[17], Jinzhong Fu[4], Bernard Goffinet[7], Hong Ma[18], Stefan A. Rensing[3], Ralf Reski[2,29*], Yikun He[1*]

# Contributed equally.

† Present addresses.

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

AFFILIATIONS

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

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

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

KEY WORDS

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

SUMMARY

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

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

INTRODUCTION

About 500 million years ago (Mya), when the ancestors of humans were tiny water-borne animals[1], the Earth's rocky land masses were colonized by plants that evolved from freshwater algae. A series of biochemical and morphological innovations helped these first land plants to cope with and diversify in the new and harsher terrestrial environments, triggering a massive change in global carbon cycling through rock weathering and photosynthetic activity[2,3]. This drastically decreased atmospheric CO_2, resulting in an ice age[4], and elevated atmospheric O_2 concentration similar to modern levels, thereby facilitating the evolution of complex animals on land[5,6]. Stomata were one such critical innovations. These microscopic pores on plant surfaces, originated over 400 Mya, facilitating the greening of the land masses by permitting plant-atmosphere gas exchange and plant hydration[7,8].

Starting 65 Mya, the India-Asian collision created the Himalayas with an average altitude exceeding 4,500 m[9], resulting in one of the harshest environments on Earth[10]. A part of it, the Tibetan Plateau, is often referred to as "Roof of the World" where global warming drives upslope tree expansion and habitat loss of the endemic flora[11]. Due to their ecophysiological features and environmental sensitivity, mosses are considered as indicators of climate change[12].

Mosses, together with liverworts and hornworts, compose the bryophytes, the sister lineage to vascular plants[13]. They likely share with the first land plants a gametophytic life style with freely moving sperm cells and a lack of roots, true vasculature, flowers and seeds[14]. Many bryophytes are pioneer plants[15] well adapted to harsh environments[16,17]. Several genome sequences of model bryophytes are available and have shed light on the early evolution of land plants[18-22].

The genus *Takakia* comprises only two species: *T. lepidozioides* and *T. ceratophylla*, and its phylogenetic affinities to either algae, mosses or liverworts[23-27] was long ambiguous, until the discovery of their spore-bearing capsules that are similar to those of mosses[25,28]. *Takakia* is now resolved phylogenetically as the sister group to all other extant mosses[29], from which it differs in leaf morphology and the mode of sporangial dehiscence, while their gametophyte resembles those of some liverworts[23,25], as some biochemical features do. The possession of some algal features further suggests that *Takakia* may share some characteristics with the last common ancestor (LCA) of land

plants and provide critical insights into early land plant evolution.

Both species of *Takakia* co-occur only in the Tibetan Plateau, which may be viewed as the modern distribution center of *Takakia*[28]. In 2005, we found *T. lepidozioides* in the Gawalong glacier area above 4,000 meters elevation, well above the tree line (Figure 1A) where the endemic vegetation is undergoing drastic and profound changes due to local warming under climate change[11], but the fate of specific species is unclear. To comprehensively explore the genetic and phenotypic characteristics of *T. lepidozioides* and correlate its adaptive evolution to environmental changes, we set up sample sites in a ~18 km^2 region, with real-time image and meteorological data recorders to characterize individual plant growth and reproductive development, and population dynamics in a climatic context. Further, we complemented the data of *Takakia* fossil and paleometeorology, and field observations with laboratory experiments on its physiology, its biochemistry and cell biology, and its genome bioinformatics, revealing adaptations to harsh environments, in particular, extreme tolerance to high UV radiation and freezing, resulted from the uplift of our experimental site started 65 Mya.

RESULTS

Global warming impacted population dynamics in the past decade

We studied populations of *T. lepidozioides* in the southern part of the Tibetan Plateau, a top hotspot of global biodiversity conservation[30]. We established 68 sample sites (10×10 cm^2 each) in three areas at altitudinal intervals of about 100 m (Figures 1A–1C; Movie S1). The matrix properties of sample substrates comprised three types: soil, rock covered with a thin layer of soil and bare rock. The composition of mineral elements was analyzed from 20 soil samples (Data S1A).

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

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

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

Vegetative growth for stable network structures

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

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

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

Reproductive strategy

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

Tolerance to extreme UV-B radiation

Exposure to solar UV-B radiation increases with altitudes and hence, at high elevations, mutagenic UV-B[33] is a potent stressor. We exposed *T. lepidozioides,* the model moss *Physcomitrium (Physcomitrella) patens*[34] and the model angiosperm *A. thaliana* to 144 kJ m^{-2} UV-B, about ten times the intensity in our sample site, for two hours. Leaves of *A. thaliana* began wilting within 24 hours after treatment, and within 72 hours plants withered and died (Figure S3). Apical leaves of *P. patens* turned pale within 72 hours, and plants gradually died, whereas *T. lepidozioides* plants appeared to be unaffected (Figures 3A–3F). Cellular damage in *P. patens* included disordered chloroplast lamellae, expanded mitochondria, disintegrated nuclear membranes, agglomerated chromatin, and final cell collapse (Data S1U). By contrast, cells of *T. lepidozioides* exhibited only a slight condensation of heterochromatin but harbored significantly more oil bodies in response to UV-B (Data S1U).

These oil bodies were rich in neutral lipids and membrane lipids (Data S1V). The total lipid content was 1.91 µg / mg fresh weight (fw), including 61.61% neutral lipids, of which 53.41% are polyunsaturated fatty acids (PUFAs). In contrast, the total lipid content of *P. patens* was 1.01 µg / mg fw, including 19.36% neutral lipids, of which 36.29% are PUFAs (Data S1W). Thus, *T. lepidozioides* does not only contain much more overall lipids than *P. patens*, but especially more PUFAs, which are involved in tolerance to various biotic and abiotic stresses[35]. Upon UV-B treatment, the contents of diacylglycerol-*N,N,N*-trimethylhomoserine (DGTS), neutral lipids and PUFAs in neutral lipids increased significantly in *T. lepidozioides* (Figure 3G; Data S1X), compared to no DGTS or stagnation of neutral lipids in *P. patens* (Figure 3H; Data S1X). Monogalactoglycerolipid (MGDG) is the most abundant glycolipid in both mosses. Exposure to UV-B decreased its content slightly in *T. lepidozioides*, but significantly in *P. patens* (Data S1X). Phosphatidylcholine (PC), the main component of nonphotosynthetic membrane lipids, is the major phospholipid in both mosses. Upon radiation, the amount of PC remained stable in *T. lepidozioides* but decreased in *P. patens* (Data S1X).

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

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

Takakia **Genome Features**

Assembly, Annotation and Comparison

The haploid genome of *T. lepidozioides* gametophytes that were initiated from a single spore was sequenced (Table S1G) and assembled. It comprises 325 Mbp, with 98.5% anchored to four chromosomes (Figures 4A and S4A), which coincides with our cytological observation (Figures S4B–S4E), and evidence from a Japanese isolate[36]. These chromosomes (1–4) are 96, 83, 77 and 64 Mbp in size, respectively. The genome sequence represents 87% of the 375 Mbp estimated by flow cytometry (Data S2A; Table S1H) and 98% of the K-mer estimation of 331 Mbp (Data S2A), with a contig N50 length of 1.4 Mbp and a scaffold N50 length of 83 Mbp (Table S1I). The completeness of the genome estimated by BUSCO indicated that 80–97% of the conserved core gene sets are present (Data S2B). As in *P. patens*, we were unable to define specific centromeric regions. Gene annotation combined various strategies of *de novo*, homology, and transcriptome predictions. The final set contains 27,467 high confidence protein-coding gene models with mean gene length of 3,301 bp, of which 85% have an assigned function (Table S1J). Further BUSCO analysis show that 71–89% conserved core gene sets are present (Data S2B).

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

Moreover, we established the organellar genomes of *T. lepidozioides* with the plastome (NC_005087.1) and chondriome (NC_007945.1) of *P. patens* as seeds. The *T. lepidozioides* plastome is a circular molecule of 149,011 bp (Data S2C) and similar to the plastome of a Japanese isolate[39]. The *T. lepidozioides* chondriome is a circular molecule of 139,841 bp (Data S2C).

Whole Genome Evolutionary History

Intragenomic analysis detected 437 syntenic blocks containing 2,382 gene pairs with an average of 5.45 gene pairs / block (Data S2D). Synonymous substitutions per site (Ks) age distributions showing a peak range from 0.7 to 1.1 (median 0.9) unveiled evidence in *T. lepidozioides* for an ancient diploidized state (Figures 4B; Data S2D). Based on an assumed rate of silent-site substitution of 9.4 per silent site per billion years[38], the diploidization occurred 58.51–37.23 Mya (median 47.87 Mya), during the period of the Tibetan Plateau uplifting[9].

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

participated in post translational modifications of proteins, such as protein esterification and deacylation, which may be involved in stress responses[40,41] (Data S2E). According to the rate of silent-site substitution of 9.4 per silent site per billion years[38], the WGD, TD, PD, TRD and DSD occurred 58.51–37.23 Mya (median 47.87 Mya), 34.57–6.38 Mya (median 12.23 Mya), 24.47–3.19 Mya (median 7.98 Mya), 13.83–2.66 Mya (median 6.38 Mya) and 28.72–3.72 Mya (median 9.57 Mya) respectively, and most of duplications occurred during the Tibetan Plateau uplifting. Furthermore, 27 singleton genes occur in *T. lepidozioides*, some of which may participate in DNA repair (Data S2E).

Phylogenetic Position

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

Special Genes

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

Genetic Foundation of Adaptive Features

Positive Selection and Fast Evolution

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

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

Transcription Associated Proteins

Transcription associated proteins (TAPs) comprise transcription factors (TFs) and transcriptional regulators (TRs). Genome-wide, domain-based annotation of TAPs[43] identified 1,091 TAPs, of which 427 are TFs and 664 TRs, which is more than that of the hornwort *A. angustus* (570) and liverwort *M. polymorpha* (682), but less than that of *P. patens* (1,638). However, the proportion of TAPs to protein-coding genes in *T. lepidozioides* (3.97%) is similar to *A. angustus* (3.89%) and *M. polymorpha* (3.53%) but less than in *P. patens* (4.97%) (Table S1R). In general, most TAP families in *T. lepidozioides* have a similarly low number in *A. angustus* and *M. polymorpha* than in *P. patens*. For example, the zf-HD family, which plays a key role in the control of growth and development[44], comprises a single member in *T. lepidozioides*, *A. angustus* and *M. polymorpha*, but 12 in *P. patens*. Similarly, C2C2-CO-like TFs, which act between the circadian clock and genes controlling meristem identity[45], are represented by two genes in *T. lepidozioides* and *A. angustus*, three in *M. polymorpha*, but 16 in *P. patens*. Interestingly, *T. lepidozioides* encodes no nucleic acid-binding cold shock domain protein (CSD) (Table S1R), which is an ancient and conserved family of proteins in various prokaryotes and eukaryotes[46]. In *Arabidopsis*, CSD proteins negatively regulate freezing tolerance[47].

By contrast, with 438 genes, plant homeodomain zinc fingers (PHDs) are the most abundant TRs in *T. lepidozioides*, a diversity vastly outnumbering that in *A. angustus* (32), *M. polymorpha* (55) and *P. patens* (45) (Table S1R). PHDs are structurally conserved modules found in plants as epigenome readers that control gene expression through recruitment of multi-protein complexes of chromatin regulators, and play crucial roles in abiotic stress responses[48]. Among the 438 genes, 118 are in small clusters of 2–5 genes in close proximity (fewer than 10 genes), with 164 of them lacking introns. Ks analysis showed these intronless genes have a peak ranging from 0.19–0.54 (median 0.33), corresponding to their expansion 28.71–10.10 Mya (median 17.50 Mya), suggesting that at least some of them have experienced tandem duplications and retrotranspositions, respectively.

Mitochondrial transcription termination factors (mTERFs) are diversified in *T. lepidozioides* (18) compared to *A. angustus* (16), *M. polymorpha* (15) and *P. patens* (16) (Table S1R). mTERFs are localized in mitochondria or chloroplasts and involved in stimulating transcription and stabilizing post-transcriptionally processed mRNA, terminating transcription, promoting tRNA maturation, intron splicing and chloroplast ribosome assembly and translation[49]. They are important for stress responses and development[50]. The mTERFs in *T. lepidozioides* have a Ks peak range from 1.69–2.57 (median 2.07), suggesting that their diversity occurred 136.70–89.89 Mya (median 110.11 Mya).

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

Protein phosphorylation

Protein phosphorylation is important for regulating protein activity and plays crucial roles in responses to environmental stresses[51]. In *Takakia*, 273 genes encode protein kinases, including 92 that are in small dense clusters of 2–4 genes (Data S2I), suggesting that they were generated by tandem duplications. Moreover, 32 genes lack introns and are scattered throughout the genome, implying that some of them were acquired by retrotranspositions. According to Ks distribution, they diverged 145.07–84.76 Mya (median 107.26 Mya). However, 40 genes are younger than 30 Mya (Data S2I).

PPR proteins and RNA-editing

Nuclear-encoded pentatricopeptide (PPR) proteins localize to mitochondria and chloroplasts, control organellar gene expression, including RNA-editing[52], and thus impact growth and development in response to stresses[53,54]. The *T. lepidozioides* genome encodes 1,926 PPRs, far more than that of *P. patens* (105) or *M. polymorpha* (69) (Figure 5A; Table S2A–S2H). An examination of their chromosomal positions revealed at least ~270 small clusters of two or more tandemly repeated genes (total 789 genes), and 271 of them have only 1–3 introns with a Ks value lower than 0.3, suggesting that they result from tandem duplications at about 15 Mya. About 429 PPR genes lack introns and are dispersed throughout the genome; 95% of them have a Ks value below 0.5 and 213 of them have a Ks value below 0.3, suggesting that they expanded 30–15 Mya. Some of them might have experienced retro-transposon-mediated duplication (Data S2I). While in most bryophytes PPR proteins are equally distributed between plastids and mitochondria, the majority target mitochondria in *T. lepidozioides* (Figure 5B). RNA-editing occurs rarely in mosses (10 sites in *P. patens*). By contrast, we identified 912 editing sites in the *T. lepidozioides* organellar genomes, which is more than in *A. thaliana* (456) (Figure 5C; Table S2I and S2J). PPR-expression is induced by UV-B treatment (Figure 5D). Taken together, the expansion of PPR proteins and RNA-editing sites in *T. lepidozioides* may be linked to adaptation to extreme environments, such as exposure to elevated UV radiation (Figures 5E–5G).

DNA repair

UV-B radiation leads to pyrimidine dimers and (6-4) photoproducts, which can be excised and repaired in nuclear DNA by a group of proteins including ERCC1 and RAD2[55]. While *P. patens* and *A. thaliana* each encode one ERCC1 and one RAD2, *T. lepidozioides* underwent a duplication of each these (Table S2K) about 5.57 Mya and 203.17 Mya, respectively. RAD51 is the eukaryotic homolog of bacterial RecA and a key factor in homologous recombination and DNA damage repair[56,57]. We identified five RAD51 genes in *T. lepidozioides* compared to six in *P. patens* (Data S2J). In *T. lepidozioides* a pair of RAD51 (Tle2c04609.1, Tle4c02755.1) arose at about 20.78 Mya, while paralogs in *P. patens* (Pp3c11_22000V3.1, Pp3c7_7920V3.1) separated at about 44.27 Mya. BRCA1 and BARD1 evolved from a common eukaryotic ancestor and are required for efficient DNA repair in humans and plants[58]. Like all other plants, except *P. patens* which lacks the gene, *T. lepidozioides* encodes one BRCA1 (Table S2L). To trace its evolutionary history, we analysed the BRCA1 and BARD1 homologs in 11 bryophyte species with published genomes (two hornworts, one liverwort, eight mosses), and found that BRCA1 and BARD1 were lost in the ancestor to other mosses after *Takakia* separated from their LCA (Data S2K). Taken together, the genome of *T. lepidozioides* is uniquely equipped with genes enhancing its ability to recover from DNA damage incurred by exposure to high UV radiations associated with its current habitat at high elevations in the Tibetan Plateau.

Phenylalanine ammonia-lyases

T. lepidozioides is rich in phenolic compounds, including flavonoids (Figure 3J; Data S1Y), which are products involved in the phenylpropanoid pathway[59]. Phenylalanine ammonia-lyase (PAL) is the first and key enzyme in phenylpropanoid biosynthesis and is expressed in response to various stresses[60]. *T. lepidozioides* encodes 20 PALs, more than *A. thaliana* (4), *M. polymorpha* (12) or *P. patens* (16). Sixteen of them cluster in a subclade, and nine of them are intronless. From these 20 PALs, 15 expanded 50 Mya and the intronless PALs expended 30 Mya, indicating a recent expansion via segmental or tandem duplications (Data S2L). In *Takakia*, PAL expression increases under UV-B stress and reaches the highest level during the recovery from UV-B exposure (Data S2L).

Cytochromes P450

CYP450s are widespread in eukaryotes. They catalyse various oxygenation / hydroxylation reactions that require molecular oxygen and depend on the electron transfer from NADPH via NADPH-P450 reductase. CYP450s are involved in a variety of pathways producing primary and specialized metabolites[61]. In mosses, they are also important for the formation of the cuticle, a complex of phenol-enriched polymers on the epidermis that enables organogenesis and protects from dehydration[62]. We identified 71 CYP450s, close to the number in *P. patens* (80). Compared to *P. patens*, several clades of CYP450s such as CYP703A2 , CYP704 and CYP77, which are involved in fatty acid and hormone metabolism, are expanded in *T. lepidozioides* (Data S2M). This expansion, dated to 145.21–87.77 Mya (median 136.17 Mya), may explain why its fatty acid content is higher than that of *P. patens* under normal conditions and under UV stress.

Genetic basis of phenotypes and development

Leaf adaxial-abaxial polarity

Most extant land plants have bifacial leaves with adaxial-abaxial (dorsiventral) polarity, which may have been different in the earliest land plants[63]. *T. lepidozioides* leaves have no such polarity (Data S1G). Polarity establishment in angiosperms is controlled by HD-ZIPIII TFs and KAN TFs[64,65]. While the former are present in *T. lepidozioides*, *P. patens* and *A. thaliana*, the latter are missing in *T. lepidozioides* (Table S2M), suggesting that the non-polar leaves resulted from the loss of KAN1 (Figure 6A). The *Takakia* sp. fossil records dated to the Middle-Late Jurassic (165 Mya[66]) help us to trace the evolutionary history of morphological features in *T. lepidozioides*. The fossil plant shape (height 12.48 ± 2.46 mm, stem diameter 0.31 ± 0.14 mm; leaf length 0.88 ± 0.17 mm, leaf diameter 0.04 ± 0.01 mm) is similar to that of extant *T. lepidozioides* (height 11.14 ± 2.90 mm, stem diameter 0.26 ± 0.08 mm; leaf length 0.75 ± 0.16 mm, leaf diameter 0.04 ± 0.01 mm) in the experiment site of Tibet Plateau, and both of them hold similarly lobed terete leaves (Data S2N). Based on data of morphological characteristics of the fossil and extant *Takakia* plants, we suppose that *Takakia* might have lost adaxial-abaxial leaf polarity no later than 165 Mya and after the LCA of all mosses. To better understand the evolutionary history of KAN in mosses, we constructed a timetree using KAN homologs from 11 bryophyte species with two fossil calibrations and found that the KAN diversity within mosses occurred 396.15—240.84 Mya (median 314.86 Mya) (Data S2O), displaying these divergences of KANs did not disturb their function for maintaining the adaxial-abaxial leaf polarity among mosses. The ancestor of *Takakia* lost KAN likely between 389.92 Mya, when it separated

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

In addition, jasmonate signaling is important for leaf expansion[63]. From these factors, *T. lepidozioides* lacks MYC2 and JAZ1 (Data S2P; Table S2P). Another possibility is post-transcriptional regulation of gene expression via microRNAs (miRNAs), as *T. lepidozioides* has only one DICER1 among three DICER proteins (Data S2Q). In *P. patens*, knockout of *DCL1a* or *DCL1b* affects stem and leaf differentiation[67]. Together, these deficiencies may explain the deeply lobed leaf shape in *T. lepidozioides*, which is unique among mosses but found in some liverworts[68].

Rhizoids

Rhizoids are haploid filamentous cells that facilitate interaction with the substrate, including water absorption and nutrient uptake[69]. They share similar regulatory developmental networks to those of root hairs of angiosperms[70]. Unlike other mosses, *T. lepidozioides* lacks rhizoids (Data S1G and S1H). The bHLH TFs RHD6 and RSL1 control root hair development in *A. thaliana* and rhizoid formation in *P. patens*[70], while null mutants in *P. patens* and *M. polymorpha* lack rhizoids[71,72]. The Zygnematophyceae *Spirogloea muscicola* and *Mesotaenium endlicherianum*, the likely sister group to land plants that share the same subaerial/terrestrial habitat with the earliest-diverging bryophytes[73], lack rhizoids and RHD6 homologs. At least one homolog of these TFs (RHD6 clade) occurs in each of the selected land plants, but none in *T. lepidozioides* (Figures 6B and S5A; Table S2N), suggesting that the loss of rhizoids resulted from the loss of these genes. The *Takakia* sp. fossil lack rhizoids (Data S2N), thus we suggest that *Takakia* lost rhizoids no later than the Middle-Late Jurassic after their ancestor separated from the LCA of all mosses. We constructed a timetree using RHD6 homologs from 11 bryophyte species with two fossil calibrations and found that the RHD6 diversity within mosses occurred 326.81—163.40 Mya (median 246.95 Mya) (Data S2R), the lower boundary close to the date of fossil 165 Mya.

Stomata

Stomata are microscopic pores on plant surfaces that facilitate gas exchange, hydration and in *P. patens* the release of spores[7]. They were one of the first developmental innovations of land plants, at about 400 Mya[2,74,75]. However, *T. lepidozioides* lacks stomata (Data S1T). PpSCRM1 and PpSMF1, orthologs of bHLH TFs, regulate stomata development in *A. thaliana*, and control stomata formation

in *P. patens*, given that their individual knock-out resulted in the lack of stomata[7]. We identified two SCRM-homologs but no SMF-homologs in *T. lepidozioides* (Figures 6C, S5B–S5D; Table S2O), suggesting that the lack of stomata is due to a loss of SMF. The *Takakia* sp. fossil is uninformative for stomata. To better understand the evolutionary history of SMF in mosses, we constructed a timetree using SMF homologs from 11 bryophyte species with two fossil calibrations and found that the SMF diversity within mosses occurred 455.07—278.06 Mya (median 367.43 Mya) (Data S2S), the lower boundary earlier than the date of fossil 165 Mya.

Tropisms

Plant tropisms rely on auxin gradients[76,77]. In contrast to vascular plants, auxin concentrations in *T. lepidozioides* were highest at the top and lowest at the base of erect shoots, whereas the concentration remained constant along the rhizomes (Data S2T). The *T. lepidozioides* genome encodes three PIN proteins possibly related to auxin efflux (Data S2T), but only *TlPIN1* (Tle1c08012.1) has the classical domain structure. Ectopic expression of these genes in *A. thaliana* revealed that *TlPIN1* represses root hair growth greatly, *TlPIN2* (Tle4c04210.1) mildly, and *TlPIN3* (Tle1c06076.1) not at all (Data S2T). AtPIN1 antibodies detected TlPIN proteins on plasma membranes in cells from leaves and stems with no polar localization (Data S2U), which is different from leaf cells of *P. patens*[78] and root cells of *A. thaliana*[76,77]. The functional divergence of TlPIN1-3 occurred 134.59–101.08 Mya. The unusual nonpolar localization of TlPIN1 in *Takakia* may account for its unusual weak negative geotropism, which facilitates the adaptive growth of populations on rugged and steep surfaces in the abscense of rhizoids.

DISCUSSION

Evolution of specific morphological traits predated the uplift of the Himalayas

Takakia has been an enigmatic genus, as it combines a mixture of morphological traits of algae, liverworts and mosses. Our analysis of the *T. lepidozioides* genome combined with the fossil evidence suggests that its extant shape evolved earlier than 165 Mya, after their ancestor separated from the LCA of all other mosses, in which there were at least two periods of severe high temperatures according to paleometeorological data[79]. This predates the uplift of the Himalayas, which started 65 Mya. Based on fossils and phylogenomics data, it is believed that extant vascular plants and bryophytes originated from a more complex ancestral land plant during the Cambrian,

515–494 Mya[80]. Thus, we suggest that adaptive evolution of *Takakia* included the loss of stomata, rhizoids, and leaf polarity, rather than it representing an ancestral state.

The determination of leaf shape is controlled by complex molecular regulatory networks, which are involved in leaf origination, polarity fixation and blade expansion[81,82]. Our analysis of 11 bryophyte genomes revealed that *T. lepidozioides* encodes gene homologs related to leaf origination and blade expansion but lost the leaf polarity-determining TF KAN1 earlier than 165 Mya (Data S2O), suggesting that its ancestor had dorsiventral leaves, and that loss of polarity may be a special adaptive morphological trait. Likewise, all 11 bryophyte genomes, except *T. lepidozioides*, encode RHD6, which controls rhizoids development in bryophytes[70], and diverged 326.81—163.40 Mya (median 246.95 Mya) (Data S2R). Therefore, we suggest that rhizoid loss is a consequence of the loss of RHD6 in the unique ancestor of *Takakia*. Fossils and phylogenomic data indicate that stomata were present in the LCA of land plants, prior to the divergence of bryophytes and tracheophytes, and subsequently lost multiple times during the diversification of bryophytes[8]. Our analysis of 11 bryophyte genomes reveals that all of them, except *T. lepidozioides*, had genetic traces related to the TF SMF, which is associated with stomata development in embryophytes[7], during 455.07—278.06 Mya (median 367.43 Mya) (Data S2S). Therefore, we suggest that the loss of stomata resulted from the loss of SMF in *T. lepidozioides*, rather than this species never encoded SMF.

Instead of rhizoids that attach plants to the soil matrix, *T. lepidozioides* evolved a morphology unique among mosses. Plant tropism and branching are closely related to auxin signalling[76,77]. Aux/IAAs interact with ARFs leading to the repression of ARF-controlled genes, which are involved in tropisms, braching and root formation[83-85]. The timing history of Aux/IAA expansion (218.62–142.02 Mya), RHD6 loss (no later than 165 Mya) and PIN divergence (134.59–101.08 Mya) might reveal the reasons why *Takakia* plants display a weak negative geotropism, continued branching and randomness of branching positions. The expansion of Aux/IAA genes might have enhanced the function of Aux/IAA, strengthing the inhibitory effect of ARF-regulated gene expression, which could lead to weaker effects of auxin sigalling. The timing of RHD6 loss and Aux/IAA divergence may overlap and follow with the loss of polar PIN localization. Thus, there is no IAA distribution gradient (Data S2T) in *Takakia* in contrast to the situation in seed plants[86] and *P. patens*[87]. If such a causal relationship between the three gene families exists could be tested experimentally.

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

While some morphological features characteristic of *Takakia* evolved before 65 Mya, the duplication and divergence of genes with implications in stress protection occurred mainly during the uplift of the Himalayas, namely between 50 Mya to present. This period is characterized by a sharp increase in harmful UV-B radiation and a sharp drop in temperature, both caused by the uplift of the region[88].

The evolutionary deeply conserved RAD51 is a key factor in DNA repair and recombination in prokaryotes and eukaryotes[56,57]. Its enhanced expression under elevated UV-B levels results in increasing frequencies of homologous recombination to maintain genome stability in *A. thaliana* and potato plants[33]. In *Takakia*, two of the five RAD51 genes arose from duplications about 20.78 Mya when the experimental site was being uplifted to about 4,000 m elevation[88], suggesting that this expansion enhances their protective effect on genome stability in responses to elevated UV-B radiation.

UV-B radiation leads to pyrimidine dimers and (6-4) photoproducts, which can be excised and repaired in nuclear DNA by a group of proteins including RAD2 and ERCC1[55]. Compared with one RAD2 and ERCC1 gene in all surveyed plants, two paralogs of each occurred in *Takakia* (Table S2K) from duplication dated to about 203.17 Mya and 5.57 Mya, respectively. While the driving force for RAD2 expansion remains enigmatic, the expansion of ERCC1 may have been driven by severe UV-B radiation at high altitude. BRCA1 and BARD1 evolved from a common eukaryotic ancestor and are required for efficient DNA repair in humans and plants[58]. While all surveyed moss genomes lack BRCA1 and BARD1, the *T. lepidozioides* genome encodes one BRCA1, suggesting that the common ancestor of the other mosses lost BRCA1 and BARD1 after their separation from the lineage that leads to *Takakia* (Data S2K).

One strategy of plants to resist UV-B damage is to produce abundant metabolites that effectively absorb radiation or eliminate ROS, such as accumulate flavones to increase UV-tolerance[89], phenols and unsaturated fatty acids to enhance resistance to UV-B[90]. *T. lepidozioides* is rich in phenolic compounds, including flavonoids (Figure 3J; Data S1Y), which are products involved in the phenylpropanoid pathway[59]. PAL is the first and key enzyme in phenylpropanoid biosynthesis and is expressed in response to various stresses[60]. *T. lepidozioides* encodes 20 PALs, more than all surveyed plants, 15 of them expanded within 50 Mya and intronless PALs expanded 30 Mya, indicating a recent expansion via segmental or tandem duplications (Data S2L) likely in response to the changing environment during the uplifting of the Himalayas. Expression of these PALs increases

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

The uplift of the Himalayas has caused the annual average temperature in the area of our experimental site to drop from 27 °C to about 5 °C[88], forcing *Takakia* plants to withstand freezing. CSD proteins are evolutionary deeply conserved in prokaryotes and eukaryotes[46]. In *Arabidopsis*, CSD proteins negatively regulate freezing tolerance as the null mutants showed higher freezing tolerance than WT when cold-acclimated. The increase in freezing tolerance was associated with up-regulation of CBF transcription factors and their downstream genes[47]. We constructed a timetree using CSD homologs from 11 bryophyte species with two fossil calibrations and found that the CSD diversity within mosses occurred 463.43—338.08 Mya (median 400.37 Mya) (Data S2V), the lower boundary earlier than the date of fossil 165 Mya. So we supposed the ancestor of *Takakia* may lost CSD after their ancestor separated from other mosses LCA, suggesting that CSD loss may help *Takakia* plants to withstand harsh conditions, such as being covered by snow for eight months at -4 °C.

Limitations of the study

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

Figure titles and legends

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

(**A**) The geomorphologic view of the survey area: exact location of experimental plots (red dots) in the area of ~18 km^2 near Gawalong East Glacier, Lat 29°45′ N, Long 95°42′ E, altitude 3,800–4,400 meters, at Bomi County, Tibet, China, yellow dotted line marks the treeline position in survey area. (**B**) A population of *T. lepidozioides* growing on the surface of a cliff in area of Plot C (bar = 1 cm). (**C**) *Takakia* plants with capsules (red arrow) in one population of Plot B (bar = 5 mm). (**D**) Temperature change in the survey area from 1960 to 2021, with T˚ rising at a rate of 0.32°C per decade. (**E**) Temperature change per year in the sample plot from 2010 to 2021 (mean ± s.d.; n = 6 investigated plots), the gray shadow area shows the period during which *Takakia* plants were covered by snow. (**F**) The coverage of *T. lepidozioides* population decreased continually at a rate of 1.6% per year from 2010 to 2021. The vertical axis represents the relative coverage of *T. lepidozioides* in observation quadrants (n = 52), and the dotted line marks the relationship between the *T. lepidozioides* population size and year. The central mark of each box indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles respectively. The whiskers extend to the most extreme data points that are not considered as outliers.

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

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

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

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

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

Figure 4. Organization and characteristics of the *T. lepidozioides* genome.

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

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

Figure 5. PPR features in *T. lepidozioides*.

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

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

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

STAR Methods

RESOURCE AVAILABILITY

Lead contact

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

Materials availability

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

Data and code availability

Whole-genome assembly and raw reads of *T. lepidozioides* have been deposited in the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) under BioProject accession number PRJNA796444, the pseudomolecules genome with GenBank accession number JAKLZA000000000, the raw reads of DNA and RNA sequencing with Sequence Read Archive accession number SRP354614 (Illumina PCR-free reads SRR17573914—SRR17573930 and PacBio reads SRR17573931—SRR17573932). The genome assembly and raw data was also deposited in National Genomics Data Center (NGDC, https://ngdc.cncb.ac.cn/?lang=en) under BioProject accession number PRJCA014819, the genome with Genome Warehouse accession number GWHBRAK00000000 and the raw sequences data with Genome Sequence Archive accession number CRA009836. Metabolomics raw data of *T. lepidozioides* have been deposited in MetaboLights (https://www.ebi.ac.uk/metabolights/) under accession number MTBLS6999. The Takakia genome assembly sequences and gene annotations are available at https://www.takakia.com. This study does not report original code. All other data and materials are available from the corresponding authors upon reasonable request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Natural samples of *T. lepidozioides* with mature capsules was collected from a shady, moist rock on Galongla Mountain, Bomi County, Tibet, China (29°45′ N, 95°42′ E) in October 2005 by Xuedong Li and Meizhi Wang. The *T. lepidozioides* tissue culture plantlets were grown in the greenhouses at

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

METHOD DETAILS

Field study

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

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

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

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

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

Plant materials

Samples of *T. lepidozioides* with mature capsules were collected from a shady, moist rock on Galongla Mountain, Tibet, China. Healthy capsules were selected and surface-sterilized as follows: (1) five 3-min rinses with sterilized water, (2) five rounds of sterilization for 5 s with 75 % ethanol and washing for 1 min with sterilized water, (3) sterilization with 0.05 % $HgCl_2$ solution for 2 min and (4) five 1-min washes with sterilized water. The capsules were transferred into sterilized water to prepare a spore suspension, which was inoculated onto the surface of Beneke's medium with some modification (200 mg L^{-1} NH_4NO_3, 100 mg L^{-1} KH_2PO_4, 100 mg L^{-1} $MgSO_4 \cdot 7H_2O$, 100 mg L^{-1} $CaCl_2 \cdot 2H_2O$, 15 mg L^{-1} $FeSO_4 \cdot 7H_2O$) containing 0.5 % (w/v) glucose (pH 5.8), and were cultivated under the following conditions: 14/10 °C day/night temperature, a 12-h/12-h light - dark photoperiod, illumination at 100 µmol photons m^{-2} s^{-1}, and approximately 85 % relative humidity (RH). Plants were transferred to the same medium and incubated under the same light and temperature conditions mentioned above every 2 weeks. The *Physcomitrella patens* wild-type "Gransden 2004" was grown on BCD medium (1.8 mM KH_2PO_4, 1 mM $MgSO_4 \cdot 7H_2O$, 45 µM $FeSO_4 \cdot 6H_2O$, 10 mM KNO_3, 0.22 µM $CuSO_4 \cdot 5H_2O$, 0.19 µM $ZnSO_4 \cdot 7H_2O$, 10 µM H_3BO_3, 0.1 µM $Na_2MoO_4 \cdot 2H_2O$, 2 µM $MnCl_2 \cdot 4H_2O$, 0.23 µM $CoCl_2 \cdot 6H_2O$, 0.17 µM KI) containing 0.5 % (w/v) glucose (pH 6.5) with a light intensity of 100 µmol m^2 s^{-1} delivered in a 16-h/8-h light/dark cycle at 25 °C. Plants were transferred to the same medium and incubated under the same light and temperature conditions every 2 weeks, and the 2-weeks-old gametophytes were used as the material for future experiments, unless specifically mentioned. Seeds of *Arabidopsis thaliana* wild-type accession Col-0 lines were germinated on MS

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

UV-B treatment

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

Cytological observations

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

The chromosome preparation was conducted according to a previous study[94]. Shoot tips of vigorously growing plants were harvested and fixed with Cannoy Solution (ethanol:acetic acid = 3:1) overnight. After washing with distilled water, the shoots were digested in a mixture of 2% cellulose (Sigma-Aldrich) and 1% pectinase (Sigma-Aldrich) at 37 °C for 90 min. After that, they were quickly rinsed in distilled water and fixed with the same fixation solution. The samples were spread on precooled slides and flame dried. Chromosomes were counterstained with 4',6-diamidino-2-

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

Immunoblot of photosynthetic proteins

Thylakoid membranes proteins were extracted as described[95,96]. Plants leaves were collected and ground under liquid nitrogen, then homogenized in an ice-cold isolation buffer containing 125 mm Tris–HCl pH 8.8, 1% (w/v) SDS, 10% (v/v) glycerol, 50 mm $Na_2S_2O_5$ with a chilled mortar and pestle and filtrated through two layers of cheesecloth. The filtrate was centrifuged at 5,000g for 10 min. The thylakoid pellets were washed with isolation buffer and recentrifuged. The thylakiod membranes were washed in 330 mM sorbitol and 50 mM BisTris-HCl, pH 7.0, and suspended in resuspension buffer (20% glycerol and 25 mM BisTris-HCl, pH 7.0) at 1.0 mg chlorophyll/mL. An equal volume of resuspension buffer containing 2% (w/v) DM was added to the thylakiod suspension in a dropwise manner. After incubation at 4°C for 5 min, insoluble material was removed by centrifugation at 12,000g for 10 min. The supernatant was combined with one-tenth volume of 5% Serva blue G in 100 mM BisTris-HCl, pH 7.0, 0.5 M 6-amino-n-caproic acid, and 30% (w/v) glycerol and applied to 0.75-mm-thick 6 to 12% acrylamide gradient gels in a Hoefer Mighty Small vertical electrophoresis unit connected to a cooling circulator. For two-dimensional analysis, excised BN-PAGE lanes were soaked in SDS sample buffer and 5% β-mercaptoethanol for 30 min and layered onto 1-mm-thick 15% SDS polyacrylamide gels containing 6 M urea. After SDS-PAGE, proteins were transferred to nitrocellulose membranes (Amersham Biosciences). The membranes were incubated with specific antibodies and visualized using the enhanced chemiluminescence method. The antibodies of AtCP47 and AtPsaD were from Liu et al. [97].

Immunofluorescence labeling

T. lepidozioides plants were fixed in FAA (3% w/v formaldehyde, 5% v/v acetic acid, 50% v/v ethanol) at room temperature overnight, embedded in STEEDMAN's wax after dehydration, and sectioned as described[98]. Rehydrated sections were then subjected to an antigen retrieval procedure[99]. Slides were first incubated in blocking buffer (2% nonfat milk in PBST) for 1 h, then with anti-AtPIN1 antibody (PhytoAB, PHY1046S) at a dilution of 1:100 in blocking buffer with 2% normal goat serum for 5 h. Slides were washed with PBST buffer three times, each for 10 min, and then incubated for 2

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

Transgene constructs

The *Arabidopsis EXPA7* (AT1G12560) promoter[100] and the 35S promoter were used to drive all PIN related transgenes. The binary vector *pCAMBIA1300-GFP-NOS* was used as the cloning vector. The *ProEXPA7:GFP-NOS* vector was generated by replacing the 35S promoter region with the genomic *EXPA7* promoter fragment by PCR from genomic DNA of *Arabidopsis* in the binary vector *pCAMBIA1300-GFP-NOS* using the primer sets in Table S2Q. The *Pro35S:AtPIN2* (AT5G57090), *ProEXPA7:AtPIN2, PpPINB* (Pp3c24_2970V3.1), *TlPIN1* (Tle1c08012.1), *TlPIN2* (Tle4c04210.1), and *TlPIN3* (Tle1c06076.1)-*GFP* constructs were generated by connecting the *Pro35S* and *ProEXPA7* promoter to *PIN* coding regions by PCR from cDNA of corresponding species in the binary vector *pCAMBIA1300-GFP-NOS*. Plants were transformed with the floral dip method[101]. Root hairs of transgenic plants were observed as described[102]. For estimation of root hair length, digital photographs of roots were taken under a stereomicroscope (Zeiss, AXIO Zoom. V16) at 40x to 50x magnifications. The hairs in the hair maturation region were observed.

Observation of branching origin

To analyze the characteristics of the origin of the branches (stem and root) and the connection of the conducting elements at the cross-point between stems, the mPS-PI staining technique[103], in which polysaccharides (especially cell walls) were specifically labeled with propidium iodide, with some modifications was used to reconstruct 3D images of plant organs. In brief, *T. lepidozioides* and *A. thaliana* organs were obtained and fixed in fixative (50% methanol and 10% acetic acid) at 4 °C overnight. After washing in ddH_2O for three times, samples were incubated in 1% periodic acid at room temperature for 40 min. Next, samples were rinsed with ddH_2O and incubated in Schiff's reagent (100 mM sodium metabisulphite and 0.15 N HCl) containing 100 mg/mL propidium iodide for 1 to 2 h. Finally, samples were transferred to glass bottom dishes and cleaned with a chloral hydrate solution (4 g chloral hydrate, 1 mL glycerol, and 2 mL water) for additional 1—2 d.

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

Assay of macromolecule transport

The Evans blue staining technique[104] with some modifications was used to monitor how macromolecules are transported in stems of *T. lepidozioides*. Plants from Tibet were collected and incubated in Evans blue staining buffer (3% w/v Evans Blue, 50 mM NaH_2PO_4) at room temperature for 3 min, followed by washing four times with deionized water. Digital photographs of stems were taken under a stereomicroscope (BA600-4, Motic).

Measurement of free IAA

IAA measurements were carried out according to Wang et al. [105] with some modifications. In brief, the tip, middle and base fragments were isolated separately from erects and rhizomes of *T. lepidozioides* plants under normal condition for 3 months subculture and collected, 200 mg (fresh weight) per sample were weighted, homogenized under liquid nitrogen, and extracted at -20 °C for 24 h with 2 mL of cold methanol containing antioxidant and $^2H_2$-IAA (internal standard, CDN Isotopes). Purification was performed with Oasis MAX solid-phase extract cartridge (150 mg/6 cc; Waters). Measurement of free IAA was carried out using a UPLC-MS/MS system consisting of a UPLC system (ACQUITY UPLC; Waters) and a triple quadruple tandem mass spectrometer (Quattro Premier XE; Waters).

Physiological experiments

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

Two-dimension thin layer chromatography (2D-TLC) and gas chromatography/mass spectrometry (GC-MS) were employed to detect and identify total fatty acids[106]. In brief, plants were grinded in chlorohydrocarbon and a methyl alcohol mixture (1:2) at 4 °C, then mixed with 1M KCl to

a final volume ratio at 1:1:0.9, centrifuged for 10 min at 2,500 g. Subsequently, the organic phase was removed with N_2. 20 × 20 cm × 250 μm glass Si gel plates (Silica Gel GF Uniplates, Sigma) were activated for 1 h at 100 °C and stored in a desiccator to increase reproducibility. Crude extract material of lipids applied to form a single small spot (2–4 mm diameter) in a corner 1 cm from either of two sides of a Si gel plate, then TLC plates were placed in a solvent chamber ($CHCl_3$:CH_3OH:H_2O, 65:25:4, v/v/v), and the extract spot was allowed to develop along the left-hand side of the plate. TLC plates were then removed from the solvent chamber immediately before the solvent front reached the top of the plates. These plates were allowed to air-dry and subsequently placed in another solvent chamber ($CHCl_3$:CH_3OH:$NH_3 \cdot H_2O$, 65:35:5, v/v/v) so that the developed lane of extract was placed horizontally just above the solvent. Each plate was allowed to develop until the solvent front reached 1 cm from the top of the plate and then removed from the chamber and air-dried. Sample spots were collected and mixed with 10 μg N-heptadecanoic acid, 5 μg butylated hydroxytoluene, 1.5 mL sulfuric acid: methyl ester (5:95), 85 °C for 1 h and then with 1 mL H_2O, 2 mL N-pentane, centrifuged for 10 min at 3,000 g. Subsequently, the water phase was removed. GC-MS (Trace GC 2000-Trace DSQ, Thermo Finnigan) was performed using a quartz capillary column. The heating program used an initial temperature of 50 °C, increasing at a rate of 10 °C/min to 150 °C and maintain 2 min, then increasing to 280 °C at a rate of 3 °C/min and maintain 10 min, then maintained at 250 °C until the running was completed. Individual compounds were identified by comparing the MS fragmentation patterns with those provided by the National Institute of Standards and Technology (https://chemdata.nist.gov/).

To classify the oil body components, the oil bodies were separated according to Zhang et al. (2010). In brief, plants were collected and washed three times with 50 ml 0.001% PBS Triton-X100; then washed with 50 ml buffer A (25 mm Tricine, pH 7.6, 250 mm sucrose, and 0.2 mm phenylmethylsulfonyl fluoride); homogenized using a polytron (Cole-Parmer® Labgen™ 125 and 700 Tissue Homogenizers) in 10 ml buffer A, and centrifuged at 1,000 × g for 30 s; supernatant was homogenized again by nitrogen cavitation (Ashcroft Duralife Pressure Gauge) at 500 psi for 15 min on ice, and then centrifuged at 1,000 × g for 10 min; nine milliliters of this postnuclear supernatant was loaded into an SW40 tube, and 3 ml buffer B (20 mm HEPES, pH 7.4, 100 mm KCl, and 2 mm $MgCl_2$) was overlaid on top. The tube was then centrifuged at 12,628 × g for 1 h at 4 °C. The oil body fraction was carefully collected from the top layer and washed three times with 200 μl buffer B. The oil body components were identified using 2D-TLC.

Non-targeted metabolome profile

Plants were collected and milled with 80% methyl alcohol and dried with N_2. 25 mg powder per sample were solved with 50% acetonitrile. Chromatographic separation was performed on a Thermo Ultimate 3000 (Thermo Fisher Scientific) with 100 mm x 2.1 mm x 1.8 μm C18 column (ZORBAX SB-Aq, Agilent Technology). The chromatographic conditions were set as follows: flow rate of 0.3 mL/min, column temperature of 40 °C, and injection volume of 4 μL. The mobile phase comprised water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B), by using the following 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% B; 11.0–11.5 min, 5% B; 11.5–14.5 min, 5% B.

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

Genome and transcriptome sequencing

Total DNA for *de novo* genome sequencing was extracted from young gametophytes produced from a single spore. For transcriptome sequencing we sampled tissues from *T. lepidozioides* including gametophytes in different growth stages and under different growth conditions, including protonemata, phylloids (leaves) at 7, 14, 21, 28, 35 and 42 days as well as phylloids under 144 kJ 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 for 6 h.

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

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

Genome assembly

The contig-level assembly was performed on full PacBio long reads using the Canu package[107] (version 2.0) (https://github.com/marbl/canu). We then polished the draft assembly using Arrow (version 2.3.1) (https://github.com/PacificBiosciences/GenomicConsensus). To increase the accuracy of the assembly, Illumina short reads were recruited for further polishing with the Pilon program[108] (version 1.22) (https://github.com/broadinstitute/pilon). The paired-end reads from Hi-C were uniquely mapped onto the draft assembly contigs. We used the 3D-DNA pipeline[109] (version 190716) (https://github.com/aidenlab/3d-dna) and Juicer[110] (version 1.11.08) (https://github.com/aidenlab/juicer) to group them into scaffolds and chromosomes. Additionally, the 10-Kb and 20-Kb mate pair reads were used to connect sequence contigs into scaffolds using SSPACE[111] (version 2.1.1) (https://github.com/nsoranzo/sspace_basic) with default parameters. To minimize errors introduced in the scaffolding step, we only connected the sequence contigs supported by linkage evidence from mate pair reads and HiC data. The SMRT reads were also applied to fill the gap in scaffolds using PBJelly[112] (version 15.8.24) (https://sourceforge.net/projects/pb-jelly/) with default parameters to generate longer sequence contigs. The scaffolds clustering into 5 groups initially. As cytological evidence strongly supports 4 chromosomes in *T. lepidozioides*, 2 shortest pseudochromosomes were linked together after error correction and manual inspection finally. The genome assembly quality was measured using BUSCO[113] (version 4.0.6) (https://busco.ezlab.org/).

Analysis of repetitive elements

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

LTR_FINDER[114] (version 1.07) (https://github.com/xzhub/LTR_Finder), RepeatModeler (version 1.0.11) (http://www.repeatmasker.org/RepeatModeler/) with default parameters. The library was aligned to PGSB Repeat Element Database[115] (http://pgsb.helmholtz-muenchen.de/plant/recat/) to classify the type of each repeat family. For identification of the repeats throughout the genome, RepeatMasker (version 4.0.8) (http://www.repeatmasker.org/) was applied with both the *ab initio* repeat databases and Repbase (http://www.girinst.org/repbase) using the WU-BLAST[116] (http://blast.wustl.edu/) search engine. Overlapping transposable elements belonging to the same repeat class were collated and combined. In addition, we annotated the tandem repeats using Tandem Repeats Finder[117] (version 4.04) (https://tandem.bu.edu/trf/trf.html). Solo-LTR were identified using LTRharvest[118] (version 1.6.1) (http://genometools.org/tools/gt_ltrharvest.html). A customized script was implemented to identify intact LTR/Copia and LTR/Gypsy retrotransposons. ClustalW[119] (version 1.81) (http://www.clustal.org/) was applied to align 5′ and 3′ solo-LTRs to intact LTR elements. The evolutionary distance of the two LTR sequences was estimated using the Kimura two-parameter method embedded in BASEML in PAML[120] (version 4.9j) (http://abacus.gene.ucl.ac.uk/software/paml.html). A substitution rate of 1.3×10^{-8} mutations per site per year was used to convert evolutionary distance between 5′ and 3′ solo-LTRs to insertion age of retrotransposons.

Gene prediction and functional annotation

Genscan[121] (version 1.0) (http://genes.mit.edu/GENSCAN.html) and Augustus[122] (version 3.3.3) (http://augustus.gobics.de/) were used to carry out *de novo* predictions with gene model parameters trained from *Chlamydomonas reinhardtii*, *Marchantia polymorpha*, *P. patens*, *Selaginella moellendorffii* and *A. thaliana*. Furthermore, gene models were *de novo* predicted using MAKER[123] (version 3.01.02) (http://www.yandell-lab.org/software/maker.html). We then evaluated the genes by comparing MAKER results with the corresponding transcript evidence to select gene models that were the most consistent. The gene prediction quality was measured using BUSCO[113] (version 4.0.6) (https://busco.ezlab.org/). Protein-coding genes were retrieved and functionally annotated by blast searches against databases including UniProtKB/Swiss-Prot and UniProtKB/TrEMBL[124] (https://www.uniprot.org/). They were also subjected to Gene Ontology (GO) annotation and protein family annotation by the InterProScan pipeline[125] (version 5.44-79.0) (https://github.com/ebi-pf-team/interproscan). GO terms for each gene were assigned by blast searches against the GO

database[126] (http://geneontology.org/).

Organellar genomes

The organellar raw reads of *T. lepidozioides* were identified using BLASTn against the PacBio subreads with plastome (NC_005087.1) and chondriome (NC_007945.1) of *P. patens* as references, respectively. The organellar assembly were performed on raw reads using the Canu package[107]. The organellar genome sequences were annotated as previously described[127]. Protein-coding genes, rRNA genes, exact gene and exon/intron boundaries were annotated using GeSeq[128] (https://chlorobox.mpimp-golm.mpg.de/geseq.html).

Identification of genome synteny

We identified homologs using BLAST with default parameters, and scanned syntenic blocks consisting of homologous genes within the *T. lepidozioides* genome using MCScan[129] (version 1.0.6) (https://github.com/tanghaibao/jcvi).

We identified WGD events in *T. lepidozioides* genome and two bryophytes (*M. polymorpha* and *P. patens*) using the WGDI toolkit[130] (version 0.5.3) (https://github.com/SunPengChuan/wgdi). In brief, delineation of paralogous gene families and one-to-one orthologs starts from all-versus-all BLASTP similarity searches with "-evalue 1e-5 -num_alignments 10". Collinear genes were first identified with the parameter "-icl" of WGDI within each genome and between genomes, and the collinear genes dot plots were used to count the syntenic ratios between different species to confirm the polyploidy level of each species, then gene pairs located on the same chromosome or on the same scaffold were removed, tandem gene pairs were also removed. Frequencies of synonymous substitutions per synonymous site (Ks) between colinear genes were estimated using the Nei-Gojobori approach as implemented in PAML[120]. The median Ks values of each block were selected to perform the Ks peak fitting by WGDI with the parameter "-pf". Ks peaks of *T. lepidozioides* vs. two species (*M. polymorpha* and *P. patens*) syntenic orthologs allows calculating Ks per year (r) following $r = Ks/(2 \times (\text{divergence time}))$[38].

We used DupGen_finder[131] (https://github.com/qiao-xin/DupGen_finder) with default parameters to search for genome-wide duplications. First, all-versus-all local BLASTP was performed using protein sequences with "-evalue 1e-5 -num_alignments 10". Second, we identified different modes of gene duplication as whole-genome duplicates (WGD), tandem duplicates (TD),

proximal duplicates (less than 10 gene distance on the same chromosome: PD), transposed duplicates (transposed gene duplications: TRD), or dispersed duplicates (other duplicates than WGD, TD, PD and TRD: DSD). Third, the Ka (number of substitutions per nonsynonymous site), Ks, and Ka/Ks values were estimated for gene pairs generated by different modes of duplication based on the YN model in KaKs_Calculator[132] (version 2.0, June 2009) (https://sourceforge.net/projects/kakscalculator2/), after conversion of amino acid alignments into the corresponding codon alignments with PAL2NAL[133] (version 14) (http://www.bork.embl.de/pal2nal/).

Phylogenomic reconstruction

To explore the phylogenetic position of *T. lepidozioides*, we constructed data sets from a total of 56 plant taxa (Table S1L), including nine genomes from Ensembl plants (http://plants.ensembl.org/index.html) and JGI phytozome (http://phytozome.jgi.doe.gov), 20 transcriptomes from the onekp project (http://mirrors.iplantcollaborative.org/browse/iplant/home/shared/onekp_pilot), 13 transcriptomes from the NCBI sequence read archive (https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?), genes of *Klebsormidium nitens* were obtained from the *K. nitens* genome project (http://www.plantmorphogenesis.bio.titech.ac.jp/~algae_genome_project/klebsormidium/index.html), those of *Chara braunii* from ORCAE (https://bioinformatics.psb.ugent.be/orcae/overview/Chbra), those of *Spirogloea muscicola* and *Mesotaenium endlicherianum* from Figshare (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 University of Zurich (https://www.hornworts.uzh.ch/en/download.html), while the remaining data are from our newly sequenced transcriptomes. Putative orthologous genes were identified using the best reciprocal blast hits (BRBH) method. Eight genomes (*Anthoceros agrestis*, *Klebsormidium nitens*, *Mesotaenium endlicherianum*, *Physcomitrella patens*, *Chara braunii*, *Marchantia polymorpha*, *Penium margaritaceum*, *Takakia lepidozioides*) were selected for inferring low-copy orthologous genes by OrthoMCL pipeline[134] (version 2.0.9) (https://legacy.orthomcl.org/orthomcl.b5/), then 818 single-copy orthologous genes (OGs) were selected and set as queries to search homologous sequences from the 56 species (Table S1L) by HaMStR[135] (version 1.6.0) (https://github.com/BIONF/HaMStR) with an e-value of 1×10^{-5}. Protein sequences of each OGs were

aligned by MUSCLE[136] (version 3.8.31) (http://www.drive5.com/muscle) and poorly aligned regions were trimmed (-automated1) using trimAl[137] (version 1.4) (http://trimal.cgenomics.org). Each OGs were reconstructed by IQ-tree[138] (version 1.6.12) (http://www.iqtree.org/) with ultrafast bootstrap parameter and suggested models, and then the 818 trees analyzed using ASTRAL[139] (version 5.7.1) (https://github.com/smirarab/ASTRAL) to infer species trees based on coalescent methods.

Divergence time estimation

We used a total of nine fossil[140-142] calibrations for TreePL analysis (Table S2R). Considering the arguments of the maximum age of five fossils, we set two calibrations: in set1, for the #3-7 fossils only minimum ages were used for calibration, whereas in set2, minimum and maximum age calibrations were applied to all fossils. The concatenation of the alignments of the 818 OGs was used to calculate the branch length by IQ-tree, with the 818-gene coalescent tree as the constraint tree. Divergence times were estimated using the penalized likelihood method[143] by treePL[144] (version 1.0) (https://github.com/blackrim/treePL). The cross-validation tested a set of smoothing parameters and then an optimal smoothing value of 0.1 was selected.

Testing positive selection and fast evolution

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

the codon level. Finally, a total of 1970 putative orthologs were identified from 12 plants. We employed a maximum likelihood approach[145] to detect positively selected genes (PSGs) along specific lineages. The branch-site model implemented in the program CODEML in PAML[120] (version 4.9j) was used. The *T. lepidozioides* lineage was defined as foreground branch. For the foreground branch we compared the selection model and the neutral model using a likelihood ratio test (LRT). A Chi-square test was conducted for each gene to compute p-values. Multiple testing was corrected by applying the false discovery rate (FDR) method. For a gene, if the selection model has a significantly higher likelihood than the neutral model (FDR-adjusted p-value < 0.05), the gene on the foreground branch may have experienced positive selection.

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

Identification of transcription factors

For TFs and TRs, we used the HMMER search method. We downloaded the HMMER model of the domain structure of each TF from the Pfam website (https://pfam.xfam.org/), referring to the TAPscan TF database[43] (https://plantcode.cup.uni-freiburg.de/tapscan/). Preliminary candidates were collected by searching the HMM profile for each species ($<1 \times 10^{-10}$). Then, we filtered genes that did not match the SwissProt functional annotation ($<1 \times 10^{-5}$). Finally, we filtered genes containing

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

Gene family classification

We applied the standard OrthoMCL pipeline[134] (version 2.0.9) (https://legacy.orthomcl.org/orthomcl.b5/) to identify orthologous gene families among *T. lepidozioides*, *C. reinhardtii*, *K. nitens*, *C. braunii*, *M. polymorpha*, *P. patens*, *S. moellendorffii* and *A. thaliana*. The longest protein from each gene was selected, and the proteins with a length less than 30 amino acids were removed. After this step, pairwise sequence similarities between all input protein sequences were calculated using BLASTP with an E value cut-off of 1×10^{-5}. Markov clustering (MCL) of the resulting similarity matrix was used to define the orthologue cluster structure of the proteins, using an inflation value (-I) of 1.5 (OrthoMCL default). Gene families were aligned by MAFFT[146] (version 7.453) (https://mafft.cbrc.jp/alignment/software/). Following alignment, Gblocks[147] (version 0.91b) (http://molevol.cmima.csic.es/castresana/Gblocks.html) was used to remove poorly aligned positions and divergent regions. RAxML[148] (version 8.2.12) (https://cme.h-its.org/exelixis/web/software/raxml/) was used to construct the phylogenetic tree using the PROTGAMMAIJTTF model and 1,000 bootstrap replicates. The divergence time was estimated using the MCMCTree program in the PAML package[120]. CAFÉ (version 4.2.1) (https://hahnlab.github.io/CAFE/) was used to estimate gene gain and loss.

Fossil of *Takakia* sp.

The studied fossil, collected from the Middle-Late Jurassic Jiulongshan Formation at Daohugou (165 Mya[66]), Ningcheng County, Inner Mongolia, China. The fossil was identified as *Takakia* sp. by Yufei Wang (Paleobotanist) and Yu Jia (Bryologist) from Institute of Botany, Chinese Academy of Sciences. The fossil *Takakia*, No. CNU-PLA-NN2023056, is housed at the Key Lab of Insect Evolution and Environmental Changes, College of Life Sciences, Capital Normal University, Beijing, China (CNUB; Prof. Dong Ren, Curator). The photographs of fossil were taken with a Nikon SMZ 25 microscope with an attached Nikon DS-Ri2 digital camera system.

Evolution of specific morphological traits

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

abaxial polarity in *Takakia*, we carefully checked these gene homologs in 11 bryophytes (including *T. lepidozioides*, two hornworts, i.e., *Anthoceros agrestis* and *Anthoceros angustus*, one liverwort, i.e., *Marchantia polymorpha* and seven other mosses, i.e., *Calohypnum plumiforme*, *Ceratodon purpureus*, *Entodon seductrix*, *Fontinalis antipyretica*, *Hypnum curvifolium*, *Physcomitrella patens*, *Syntrichia caninervis*) (Table S2T) using BLASTP with identity cutoff 50%, e-value cutoff 1×10^{-20}. Domains were check in homolog candidates using HMMER. Maximum likelihood tree constructed using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates. The divergence time was estimated using the MCMCTree program in the PAML package[120]. We used two fossil[140] calibrations for MCMCTree analysis (Table S2R), minimum age 405 Mya and maximum age 515.5 Mya calibrations were applied.

Gene expression profiling

Poor quality or technical sequences in Illumina paired-end reads were removed using the Trimmomatic preprocessing tool[149] (version 0.39) (http://www.usadellab.org/cms/?page=trimmomatic). The qualified paired reads were then aligned against the *T. lepidozioides* reference transcript sequences using Bowtie 2[150] (version 2.4.1) (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) to find all alignments of a read with no more than two mismatches. Subsequently, gene and isoform abundances were quantified from paired-end RNA-seq data using the RSEM software package (version 1.3.3) (https://github.com/siddharthab/RSEM). Differentially expressed transcripts or genes (DEGs) with biological replicates were identified by running Bioconductor tools edgeR[151] (version 3.28.1) (https://bioconductor.org/packages/release/bioc/html/edgeR.html), which implements a range of statistical methods based on the negative binomial distributions. Orthologs of *T. lepidozioides* were functionally annotated via the EggNOG database[152] (version 5.0) (http://eggnog5.embl.de/) companion tool eggnog-mapper[153] (version 2.0.1) (http://eggnog-mapper.embl.de/), with “-d euk --tax_scope 3208”. The inferred GO terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were used as the basis for DEG enrichment analyses performed with the R package clusterProfiler[154] (version 3.14.3) (http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html).

QUANTIFICATION AND STATISTICAL ANALYSIS

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

SUPPLEMENTARY INFORMATION

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

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

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

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

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

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

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

Figure S4. Chromosomes in gametophyte cells of *T. lepidozioides*, related to Figure 4

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

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

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

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

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

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

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

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

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

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

ACKNOWLEDGMENTS

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

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

AUTHOR CONTRIBUTIONS

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

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

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

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

9. Rowley, D.B., and Currie, B.S. (2006). Palaeo-altimetry of the late Eocene to Miocene Lunpola basin, central Tibet. Nature *439*, 677–681. 10.1038/nature04506.

10. Qiu, J. (2008). The third pole. Nature *454*, 393–396. doi:10.1038/454393a.

11. Wang, X.Y., Wang, T., Xu, J.F., Shen, Z.H., Yang, Y.P., Chen, A.P., Wang, S.P., Liang, E.Y., and Piao, S.L. (2022). Enhanced habitat loss of the Himalayan endemic flora driven by warming-forced upslope tree expansion. Nat. Ecol. Evol. *6*, 890–899 10.1038/s41559-022-01774-3.

12. Zanatta, F., Engler, R., Collart, F., Broennimann, O., Mateo, R.G., Papp, B., Munoz, J., Baurain, D., Guisan, A., and Vanderpoorten, A. (2020). Bryophytes are predicted to lag behind future climate change despite their high dispersal capacities. Nat. Commun. *11*, 5601. 10.1038/s41467-020-19410-8.

13. McDaniel, S.F. (2021). Bryophytes are not early diverging land plants. New Phytol. *230*, 1300–1304. 10.1111/nph.17241.

14. Horst, N.A., Katz, A., Pereman, I., Decker, E.L., Ohad, N., and Reski, R. (2016). A single homeobox gene triggers phase transition, embryogenesis and asexual reproduction. Nat. Plants *2*, 15209. 10.1038/nplants.2015.209.

15. Cutler, N.A., Belyea, L.R., and Dugmore, A.J. (2008). The spatiotemporal dynamics of a primary succession. J. Ecol. *96*, 231–246. 10.1111/j.1365-2745.2007.01344.x.

16. Roads, E., Longton, R.E., and Convey, P. (2014). Millennial timescale regeneration in a moss from Antarctica. Curr. Biol. *24*, R222–R223. 10.1016/j.cub.2014.01.053.

17. Biersma, E.M., Jackson, J.A., Hyvönen, J., Koskinen, S., Linse, K., Griffiths, H., and Convey, P. (2017). Global biogeographic patterns in bipolar moss species. R. Soc. Open Sci. *4*, 170147. 10.1098/rsos.170147.

18. Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.F., Lindquist, E.A., Kamisugi, Y., et al. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. Science *319*, 64–69. 10.1126/science.1150646.

19. Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S.Q., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., et al. (2017). Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. Cell *171*, 287–304. 10.1016/j.cell.2017.09.030.

20. Li, F.W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N., Barker, M.S., Bennett, T., Blazquez, M.A., et al. (2020). *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. Nat. Plants *6*, 259–272. 10.1038/s41477-020-0618-2.

21. Donoghue, P.C.J., Harrison, C.J., Paps, J., and Schneider, H. (2021). The evolutionary emergence of land plants. Curr. Biol. *31*, R1281–R1298. 10.1016/j.cub.2021.07.038.

22. Zhang, J., Fu, X.X., Li, R.Q., Zhao, X., Liu, Y., Li, M.H., Zwaenepoel, A., Ma, H., Goffinet, B., Guan, Y.L., et al. (2020). The hornwort genome and early land plant evolution. Nat. Plants *6*, 107–118. 10.1038/s41477-019-0588-4.

23. Mitten, W. (1860). Hepaticae Indiae Orientalis: an enumeration of the Hepaticae of the East Indies. Bot. J. Linn. Soc. *5*, 89–108.

24. Hattori, S. (1958). Preliminary report on *Takakia lepidozioides*. Journ. Hattori Bot. Lab. *19*, 133–137.

25. Hattori, S., and Mizutani, M. (1958). What is *Takakia lepidozioides*? Journ. Hattori Bot. Lab. *20*, 295–303.

26. Grolle, R. (1963). *Takakia* im Himalaya. Oesterr. Bot. Z. *110*, 444–447. 10.1007/BF01373680.

27. Smith, D.K., and Davison, P.G. (1993). Antheridia and sporophytes in *Takakia ceratophylla* (Mitt.) Grolle: evidence for reclassification among the mosses. Journ. Hattori Bot. Lab. *73*, 263–271.

28. Schuster, R.M. (1997). On *Takakia* and the phylogenetic relationships of the *Takakiales*. Nova Hedwig. *64*, 281–310. 10.1127/nova.hedwigia/64/1997/281.

29. Liu, Y., Johnson, M.G., Cox, C.J., Medina, R., Devos, N., Vanderpoorten, A., Hedenas, L., Bell, N.E., Shevock, J.R., Aguero, B., et al. (2019). Resolution of the ordinal phylogeny of mosses using targeted exons from organellar and nuclear genomes. Nat. Commun. *10*, 1485. 10.1038/s41467-019-09454-w.

30. Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. Nature *403*, 853–858. 10.1038/35002501.

31. Yao, T.D., Thompson, L., Yang, W., Yu, W.S., Gao, Y., Guo, X.J., Yang, X.X., Duan, K.Q., Zhao, H.B., Xu, B.Q., et al. (2012). Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. Nat. Clim. Chang. *2*, 663–667. 10.1038/nclimate1580.

32. Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertova, D., Jurgens, G., and

Friml, J. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell *115*, 591–602. 10.1016/s0092-8674(03)00924-3.

33. Ries, G., Heller, W., Puchta, H., Sandermann, H., Seidlitz, H.K., and Hohn, B. (2000). Elevated UV-B radiation reduces genome stability in plants. Nature *406*, 98–101. 10.1038/35017595.

34. Rensing, S.A., Goffinet, B., Meyberg, R., Wu, S.Z., and Bezanilla, M. (2020). The Moss *Physcomitrium* (*Physcomitrella*) *patens*: A model organism for non-seed plants. Plant Cell *32*, 1361-1376. 10.1105/tpc.19.00828.

35. Resemann, H.C., Herrfurth, C., Feussner, K., Hornung, E., Ostendorf, A.K., Gomann, J., Mittag, J., van Gessel, N., de Vries, J., Ludwig-Muller, J., et al. (2021). Convergence of sphingolipid desaturation across over 500 million years of plant evolution. Nat. Plants *7*, 219–232. 10.1038/s41477-020-00844-3.

36. Tatuno, S. (1959). Chromosomen von *Takakia lepidozioides* und eine studie zur evolution der chromosomen der bryophyten. Cytologia *24*, 138–147. 10.1508/cytologia.24.138.

37. Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., dePamphilis, C., Albert, V.A., Aono, N., Aoyama, T., Ambrose, B.A., et al. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. Science *332*, 960–963. 10.1126/science.1203810.

38. Lang, D., Ullrich, K.K., Murat, F., Fuchs, J., Jenkins, J., Haas, F.B., Piednoel, M., Gundlach, H., Van Bel, M., Meyberg, R., et al. (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. Plant J. *93*, 515–533. 10.1111/tpj.13801.

39. Sadamitsu, A., Inoue, Y., Sakakibara, K., Tsubota, H., Yamaguchi, T., Deguchi, H., Nishiyama, T., and Shimamura, M. (2021). The complete plastid genome sequence of the enigmatic moss, *Takakia lepidozioides* (*Takakiopsida*, *Bryophyta*): evolutionary perspectives on the largest collection of genes in mosses and the intensive RNA editing. Plant Mol. Biol. *107*, 431–449. 10.1007/s11103-021-01214-z.

40. Chen, D.Q., Hao, F.S., Mu, H.Q., Ahsan, N., Thelen, J.J., and Stacey, G. (2021). S-acylation of P2K1 mediates extracellular ATP-induced immune signaling in *Arabidopsis*. Nat. Commun. *12*, 2750. 10.1038/s41467-021-22854-1.

41. Millar, A.H., Heazlewood, J.L., Giglione, C., Holdsworth, M.J., Bachmair, A., and Schulze,

W.X. (2019). The scope, functions, and dynamics of posttranslational protein modifications. In Annual Review of Plant Biology, Vol 70, S.S. Merchant, ed. (Annual Reviews), pp. 119–151. 10.1146/annurev-arplant-050718-100211.

42. Qiu, Y.L., Li, L.B., Wang, B., Chen, Z.D., Knoop, V., Groth-Malonek, M., Dombrovska, O., Lee, J., Kent, L., Rest, J., et al. (2006). The deepest divergences in land plants inferred from phylogenomic evidence. Proc. Natl. Acad. Sci. USA. *103*, 15511–15516. 10.1073/pnas.0603335103.

43. Wilhelmsson, P.K.I., Muhlich, C., Ullrich, K.K., and Rensing, S.A. (2017). Comprehensive genome-wide classification reveals that many plant-specific transcription factors eolved in *Streptophyte* algae. Genome Biol. Evol. *9*, 3384–3397. 10.1093/gbe/evx258.

44. Ariel, F.D., Manavella, P.A., Dezar, C.A., and Chan, R.L. (2007). The true story of the HD-Zip family. Trends Plant Sci. *12*, 419–426. 10.1016/j.tplants.2007.08.003.

45. Griffiths, S., Dunford, R.P., Coupland, G., and Laurie, D.A. (2003). The evolution of CONSTANS-like gene families in barley, rice, and *Arabidopsis*. Plant Physiol. *131*, 1855–1867. 10.1104/pp.102.016188.

46. Nakaminami, K., Karlson, D.T., and Imai, R. (2006). Functional conservation of cold shock domains in bacteria and higher plants. Proc. Natl. Acad. Sci. USA. *103*, 10122–10127. 10.1073/pnas.0603168103.

47. Sasaki, K., Kim, M.H., and Imai, R. (2013). *Arabidopsis* cold shock domain protein 2 is a negative regulator of cold acclimation. New Phytol. *198*, 95–102. 10.1111/nph.12118.

48. Sanchez, R., and Zhou, M.M. (2011). The PHD finger: a versatile epigenome reader. Trends Biochem. Sci. *36*, 364–372. 10.1016/j.tibs.2011.03.005.

49. Robles, P., and Quesada, V. (2021). Research progress in the molecular functions of plant mTERF proteins. Cells *10*, 205. 10.3390/cells10020205.

50. Babiychuk, E., Vandepoele, K., Wissing, J., Garcia-Diaz, M., De Rycke, R., Akbari, H., Joubes, J., Beeckman, T., Jansch, L., Frentzen, M., et al. (2011). Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family. Proc. Natl. Acad. Sci. USA. *108*, 6674–6679. 10.1073/pnas.1103442108.

51. Umezawa, T., Sugiyama, N., Takahashi, F., Anderson, J.C., Ishihama, Y., Peck, S.C., and Shinozaki, K. (2013). Genetics and phosphoproteomics reveal a protein phosphorylation

network in the abscisic acid signaling pathway in *Arabidopsis thaliana*. Sci. Signal. *6*, rs8. 10.1126/scisignal.2003509.

52. Small, I.D., Schallenberg-Rudinger, M., Takenaka, M., Mireau, H., and Ostersetzer-Biran, O. (2020). Plant organellar RNA editing: what 30 years of research has revealed. Plant J. *101*, 1040–1056. 10.1111/tpj.14578.

53. Waltz, F., Nguyen, T.T., Arrive, M., Bochler, A., Chicher, J., Hammann, P., Kuhn, L., Quadrado, M., Mireau, H., Hashem, Y., and Giege, P. (2019). Small is big in *Arabidopsis* mitochondrial ribosome. Nat. Plants *5*, 106–117. 10.1038/s41477-018-0339-y.

54. Wang, C., Blondel, L., Quadrado, M., Dargel-Graffin, C., and Mireau, H. (2022). Pentatricopeptide repeat protein MTSF3 ensures mitochondrial RNA stability and embryogenesis. Plant Physiol. *190*, 669–681. 10.1093/plphys/kiac309.

55. Bray, C.M., and West, C.E. (2005). DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity. New Phytol. *168*, 511–528. 10.1111/j.1469-8137.2005.01548.x.

56. Li, X., and Heyer, W.D. (2008). Homologous recombination in DNA repair and DNA damage tolerance. Cell Res. *18*, 99–113. 10.1038/cr.2008.1.

57. Shinohara, A., Ogawa, H., and Ogawa, T. (1992). Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. Cell *69*, 457–470. 10.1016/0092-8674(92)90447-k.

58. Huen, M.S.Y., Sy, S.M.H., and Chen, J. (2010). BRCA1 and its toolbox for the maintenance of genome integrity. Nat. Rev. Mol. Cell Biol. *11*, 138–148. 10.1038/nrm2831.

59. Fraser, C.M., and Chapple, C. (2011). The phenylpropanoid pathway in *Arabidopsis*. Arabidopsis Book *9*, e0152. 10.1199/tab.0152.

60. Raes, J., Rohde, A., Christensen, J.H., Van de Peer, Y., and Boerjan, W. (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. Plant Physiol. *133*, 1051–1071. 10.1104/pp.103.026484.

61. Nelson, D., and Werck-Reichhart, D. (2011). A P450-centric view of plant evolution. Plant J. *66*, 194–211. 10.1111/j.1365-313X.2011.04529.x.

62. Renault, H., Alber, A., Horst, N.A., Lopes, A.B., Fich, E.A., Kriegshauser, L., Wiedemann, G., Ullmann, P., Herrgott, L., Erhardt, M., et al. (2017). A phenol-enriched cuticle is ancestral to lignin evolution in land plants. Nat. Commun. *8*, 14713. 10.1038/ncomms14713.

63. Fukushima, K., and Hasebe, M. (2014). Adaxial-abaxial polarity: the developmental basis of leaf shape diversity. Genesis *52*, 1–18. 10.1002/dvg.22728.

64. Kerstetter, R.A., Bollman, K., Taylor, R.A., Bomblies, K., and Poethig, R.S. (2001). *KANADI* regulates organ polarity in *Arabidopsis*. Nature *411*, 706–709. 10.1038/35079629.

65. Burian, A., Paszkiewicz, G., Nguyen, K.T., Meda, S., Raczynska-Szajgin, M., and Timmermans, M.C.P. (2022). Specification of leaf dorsiventrality via a prepatterned binary readout of a uniform auxin input. Nat. Plants *8*, 269–280. 10.1038/s41477-022-01111-3.

66. He, H.Y., Wang, X.L., Zhou, Z.H., Zhu, R.X., Jin, F., Wang, F., Ding, X., and Boven, A. (2004). $^{40}Ar/^{39}Ar$ dating of ignimbrite from Inner Mongolia, northeastern China, indicates a post-Middle Jurassic age for the overlying Daohugou Bed. Geophys. Res. Lett. *31*, L20609. 10.1029/2004gl020792.

67. Khraiwesh, B., Arif, M.A., Seumel, G.I., Ossowski, S., Weigel, D., Reski, R., and Frank, W. (2010). Transcriptional control of gene expression by MicroRNAs. Cell *140*, 111–122. 10.1016/j.cell.2009.12.023.

68. Crandall-Stotler, B., Stotler, R.E., and Long, D.G. (2009). Phylogeny and classification of the *Marchantiophyta*. Edinburgh Journal of Botany *66*, 155–198. 10.1017/S0960428609005393.

69. Jones, V.A.S., and Dolan, L. (2012). The evolution of root hairs and rhizoids. Ann. Bot. *110*, 205–212. 10.1093/aob/mcs136.

70. Menand, B., Yi, K.K., Jouannic, S., Hoffmann, L., Ryan, E., Linstead, P., Schaefer, D.G., and Dolan, L. (2007). An ancient mechanism controls the development of cells with a rooting function in land plants. Science *316*, 1477–1480. 10.1126/science.1142618.

71. Jang, G., Yi, K., Pires, N.D., Menand, B., and Dolan, L. (2011). RSL genes are sufficient for rhizoid system development in early diverging land plants. Development *138*, 2273–2281. 10.1242/dev.060582.

72. Proust, H., Honkanen, S., Jones, V.A.S., Morieri, G., Prescott, H., Kelly, S., Ishizaki, K., Kohchi, T., and Dolan, L. (2016). RSL class I genes controlled the development of epidermal structures in the common ancestor of land plants. Curr. Biol. *26*, 93–99. 10.1016/j.cub.2015.11.042.

73. Cheng, S.F., Xian, W.F., Fu, Y., Marin, B., Keller, J., Wu, T., Sun, W.J., Li, X.L., Xu, Y., Zhang, Y., et al. (2019). Genomes of subaerial *Zygnematophyceae* provide insights into land plant evolution. Cell *179*, 1057–1067. 10.1016/j.cell.2019.10.019.

74. Harris, B.J., Harrison, C.J., Hetherington, A.M., and Williams, T.A. (2020). Phylogenomic evidence for the monophyly of Bryophytes and the reductive evolution of stomata. Curr. Biol. *30*, 2001–2012. 10.1016/j.cub.2020.03.048.

75. Bowles, A.M.C., Paps, J., and Bechtold, U. (2022). Water-related innovations in land plants evolved by different patterns of gene cooption and novelty. New Phytol. *235*, 732–742. 10.1111/nph.17981.

76. Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwarz, H., Hamann, T., Offringa, R., and Jurgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. Nature *426*, 147–153. 10.1038/nature02085.

77. Reinhardt, D., Pesce, E.R., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., Traas, J., Friml, J., and Kuhlemeier, C. (2003). Regulation of phyllotaxis by polar auxin transport. Nature *426*, 255–260. 10.1038/nature02081.

78. Bennett, T.A., Liu, M.M., Aoyama, T., Bierfreund, N.M., Braun, M., Coudert, Y., Dennis, R.J., O'Connor, D., Wang, X.Y., White, C.D., et al. (2014). Plasma membrane-targeted PIN proteins drive shoot development in a moss. Curr. Biol. *24*, 2776–2785. 10.1016/j.cub.2014.09.054.

79. Scotese, C.R., Song, H.J., Mills, B.J.W., and van der Meer, D.G. (2021). Phanerozoic paleotemperatures: The earth's changing climate during the last 540 million years. Earth-Sci. Rev. *215*, 103530. 10.1016/j.earscirev.2021.103503.

80. Harris, B.J., Clark, J.W., Schrempf, D., Szollosi, G.J., Donoghue, P.C.J., Hetherington, A.M., and Williams, T.A. (2022). Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants. Nat. Ecol. Evol. *6*, 1634–1643. 10.1038/s41559-022-01885-x.

81. Yu, T., Guan, C., Wang, J., Sajjad, M., Ma, L., and Jiao, Y. (2017). Dynamic patterns of gene expression during leaf initiation. J. Genet. Genomics *44*, 599–601. 10.1016/j.jgg.2017.11.001.

82. Nakayama, H., Leichty, A.R., and Sinha, N.R. (2022). Molecular mechanisms underlying leaf development, morphological diversification, and beyond. Plant Cell *34*, 2534–2548. 10.1093/plcell/koac118.

83. Cance, C., Martin-Arevalillo, R., Boubekeur, K., and Dumas, R. (2022). Auxin response factors are keys to the many auxin doors. New Phytol. *235*, 402–419. 10.1111/nph.18159.

84. Roosjen, M., Paque, S., and Weijers, D. (2018). Auxin response factors: output control in

auxin biology. J. Exp. Bot. *69*, 179–188. 10.1093/jxb/erx237.

85. Ulmasov, T., Hagen, G., and Guilfoyle, T.J. (1997). ARF1, a transcription factor that binds to auxin response elements. Science *276*, 1865–1868. 10.1126/science.276.5320.1865.

86. Atamian, H.S., Creux, N.M., Brown, E.A., Garner, A.G., Blackman, B.K., and Harmer, S.L. (2016). Circadian regulation of sunflower heliotropism, floral orientation, and pollinator visits. Science *353*, 587–590. 10.1126/science.aaf9793.

87. Viaene, T., Landberg, K., Thelander, M., Medvecka, E., Pederson, E., Feraru, E., Cooper, E.D., Karimi, M., Delwiche, C.F., Ljung, K., et al. (2014). Directional Auxin Transport Mechanisms in Early Diverging Land Plants. Curr. Biol. *24*, 2786–2791. 10.1016/j.cub.2014.09.056.

88. He, S., Ding, L., Xiong, Z., Spicer, R.A., Farnsworth, A., Valdes, P.J., Wang, C., Cai, F., Wang, H., Sun, Y., et al. (2022). A distinctive Eocene Asian monsoon and modern biodiversity resulted from the rise of eastern Tibet. Sci. Bull. *67*, 2245–2258. 10.1016/j.scib.2022.10.006.

89. Peng, M., Shahzad, R., Gul, A., Subthain, H., Shen, S.Q., Lei, L., Zheng, Z.G., Zhou, J.J., Lu, D.D., Wang, S.C., et al. (2017). Differentially evolved glucosyltransferases determine natural variation of rice flavone accumulation and UV-tolerance. Nat. Commun. *8*, 1975. 10.1038/s41467-017-02168-x.

90. Shi, C., and Liu, H.T. (2021). How plants protect themselves from ultraviolet-B radiation stress. Plant Physiol. *187*, 1096–1103. 10.1093/plphys/kiab245.

91. Kilbride, C., Poole, J., and Hutchings, T.R. (2006). A comparison of Cu, Pb, As, Cd, Zn, Fe, Ni and Mn determined by acid extraction/ICP-OES and ex situ field portable X-ray fluorescence analyses. Environ. Pollut. *143*, 16–23. 10.1016/j.envpol.2005.11.013.

92. Gorecka, H., Chojnacka, K., and Gorecki, H. (2006). The application of ICP-MS and ICP-OES in determination of micronutrients in wood ashes used as soil conditioners. Talanta *70*, 950–956. 10.1016/j.talanta.2006.05.061.

93. Xiao, L.H., Yang, G., Zhang, L.C., Yang, X.H., Zhao, S., Ji, Z.Z., Zhou, Q., Hu, M., Wang, Y., Chen, M., et al. (2015). The resurrection genome of *Boea hygrometrica*: A blueprint for survival of dehydration. Proc. Natl. Acad. Sci. USA. *112*, 5833–5837. 10.1073/pnas.1505811112.

94. Cheng, Z.K., Yan, H.H., Yu, H.X., Tang, S.C., Jiang, J.M., Gu, M.H., and Zhu, L.H. (2001). Development and applications of a complete set of rice telotrisomics. Genetics *157*, 361–368.

95. Martinez-Garcia, J.F., Monte, E., and Quail, P.H. (1999). A simple, rapid and quantitative method for preparing *Arabidopsis* protein extracts for immunoblot analysis. Plant J. *20*, 251–257. 10.1046/j.1365-313x.1999.00579.x.

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). LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in *Arabidopsis thaliana*. Plant Cell *18*, 955–969. 10.1105/tpc.105.037689.

97. Liu, J., Yang, H.X., Lu, Q.T., Wen, X.G., Chen, F., Peng, L.W., Zhang, L.X., and Lu, C.M. (2012). PSBP-DOMAIN PROTEIN1, a nuclear-encoded thylakoid lumenal protein, is essential for photosystem I assembly in *Arabidopsis*. Plant Cell *24*, 4992–5006. 10.1105/tpc.112.106542.

98. Vitha, S., and Osteryoung, K.W. (2011). Immunofluorescence microscopy for localization of *Arabidopsis* chloroplast proteins. Methods Mol. Biol. *774*, 33–58. 10.1007/978-1-61779-234-2_3.

99. Vitha, S., McAndrew, R.S., and Osteryoung, K.W. (2001). FtsZ ring formation at the chloroplast division site in plants. J. Cell Biol. *153*, 111–119. 10.1083/jcb.153.1.111.

100. Cho, H.T., and Cosgrove, D.J. (2002). Regulation of root hair initiation and expansin gene expression in *Arabidopsis*. Plant Cell *14*, 3237–3253. 10.1105/tpc.006437.

101. Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. *16*, 735–743. 10.1046/j.1365-313x.1998.00343.x.

102. Ganguly, A., Lee, S.H., Cho, M., Lee, O.R., Yoo, H., and Cho, H.T. (2010). Differential auxin-transporting activities of PIN-FORMED proteins in *Arabidopsis* root hair cells. Plant Physiol. *153*, 1046–1061. 10.1104/pp.110.156505.

103. Truernit, E., Bauby, H., Dubreucq, B., Grandjean, O., Runions, J., Barthelemy, J., and Palauqui, J.C. (2008). High-resolution whole-mount imaging of three-dimensional tissue organization and gene expression enables the study of phloem development and structure in *Arabidopsis*. Plant Cell *20*, 1494–1503. 10.1105/tpc.107.056069.

104. Xu, B., Ohtani, M., Yamaguchi, M., Toyooka, K., Wakazaki, M., Sato, M., Kubo, M., Nakano, Y., Sano, R., Hiwatashi, Y., et al. (2014). Contribution of NAC transcription factors to plant adaptation to land. Science *343*, 1505–1508. 10.1126/science.1248417.

105. Wang, B., Chu, J.F., Yu, T.Y., Xu, Q., Sun, X.H., Yuan, J., Xiong, G.S., Wang, G.D., Wang,

Y.H., and Li, J.Y. (2015). Tryptophan-independent auxin biosynthesis contributes to early embryogenesis in *Arabidopsis*. Proc. Natl. Acad. Sci. USA. *112*, 4821–4826. 10.1073/pnas.1503998112.

106. Estevez, S.L., and Helleur, R. (2005). Fatty acid profiling of lipid classes by silica rod TLC-thermally assisted hydrolysis and methylation-GC/MS. J. Anal. Appl. Pyrolysis *74*, 3–10. 10.1016/j.jaap.2004.11.017.

107. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res. *27*, 722–736. 10.1101/gr.215087.116.

108. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q.D., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: an Integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One *9*, e112963. 10.1371/journal.pone.0112963.

109. Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., and Aiden, E.L. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science *356*, 92–95. 10.1126/science.aal3327.

110. Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. *3*, 99–101. 10.1016/j.cels.2015.07.012.

111. Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. Bioinformatics *27*, 578–579. 10.1093/bioinformatics/btq683.

112. English, A.C., Richards, S., Han, Y., Wang, M., Vee, V., Qu, J.X., Qin, X., Muzny, D.M., Reid, J.G., Worley, K.C., and Gibbs, R.A. (2012). Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. PLoS One *7*, e47768. 10.1371/journal.pone.0047768.

113. Simao, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics *31*, 3210–3212. 10.1093/bioinformatics/btv351.

114. Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length

LTR retrotransposons. Nucleic Acids Res. *35*, W265–W268. 10.1093/nar/gkm286.

115. Nussbaumer, T., Martis, M.M., Roessner, S.K., Pfeifer, M., Bader, K.C., Sharma, S., Gundlach, H., and Spannagl, M. (2013). MIPS PlantsDB: a database framework for comparative plant genome research. Nucleic Acids Res. *41*, D1144–D1151. 10.1093/nar/gks1153.

116. Lopez, R., Silventoinen, V., Robinson, S., Kibria, A., and Gish, W. (2003). WU-Blast2 server at the European Bioinformatics Institute. Nucleic Acids Res. *31*, 3795–3798. 10.1093/nar/gkg573.

117. Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. *27*, 573–580. 10.1093/nar/27.2.573.

118. Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinf. *9*, 18. 10.1186/1471-2105-9-18.

119. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947–2948. 10.1093/bioinformatics/btm404.

120. Yang, Z.H. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*, 1586–1591. 10.1093/molbev/msm088.

121. Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. *268*, 78–94. 10.1006/jmbi.1997.0951.

122. Stanke, M., and Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics *19*, II215–II225. 10.1093/bioinformatics/btg1080.

123. Cantarel, B.L., Korf, I., Robb, S.M.C., Parra, G., Ross, E., Moore, B., Holt, C., Alvarado, A.S., and Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. *18*, 188–196. 10.1101/gr.6743907.

124. Bateman, A., Martin, M.J., O'Donovan, C., Magrane, M., Apweiler, R., Alpi, E., Antunes, R., Ar-Ganiska, J., Bely, B., Bingley, M., et al. (2015). UniProt: a hub for protein information. Nucleic Acids Res. *43*, D204–D212. 10.1093/nar/gku989.

125. Zdobnov, E.M., and Apweiler, R. (2001). InterProScan - an integration platform for the signature-recognition methods in InterPro. Bioinformatics *17*, 847–848. 10.1093/bioinformatics/17.9.847.

126. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P.,

Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for the unification of biology. Nature Genet. *25*, 25–29. 10.1038/75556.

127. Xue, J.Y., Liu, Y., Li, L.B., Wang, B., and Qiu, Y.L. (2010). The complete mitochondrial genome sequence of the hornwort *Phaeoceros laevis*: retention of many ancient pseudogenes and conservative evolution of mitochondrial genomes in hornworts. Curr. Genet. *56*, 53–61. 10.1007/s00294-009-0279-1.

128. Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., and Greiner, S. (2017). GeSeq - versatile and accurate annotation of organelle genomes. Nucleic Acids Res. *45*, W6–W11. 10.1093/nar/gkx391.

129. Tang, H.B., Bowers, J.E., Wang, X.Y., Ming, R., Alam, M., and Paterson, A.H. (2008). Synteny and collinearity in plant genomes. Science *320*, 486–488. 10.1126/science.1153917.

130. Sun, P., Jiao, B., Yang, Y., Shan, L., Li, T., Li, X., Xi, Z., Wang, X., and Liu, J. (2022). WGDI: a user-friendly toolkit for evolutionary analyses of whole-genome duplications and ancestral karyotypes. Mol. Plant. *15*, 1841–1851. 10.1016/j.molp.2022.10.018.

131. Qiao, X., Li, Q.H., Yin, H., Qi, K.J., Li, L.T., Wang, R.Z., Zhang, S.L., and Paterson, A.H. (2019). Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. Genome Biol. *20*, 38. 10.1186/s13059-019-1650-2.

132. Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genom. Proteom. Bioinf. *8*, 77–80. 10.1016/s1672-0229(10)60008-3.

133. Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. *34*, W609–W612. 10.1093/nar/gkl315.

134. Li, L., Stoeckert, C.J., and Roos, D.S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. Genome Res. *13*, 2178–2189. 10.1101/gr.1224503.

135. Ebersberger, I., Strauss, S., and von Haeseler, A. (2009). HaMStR: Profile hidden markov model based search for orthologs in ESTs. BMC Evol. Biol. *9*, 157. 10.1186/1471-2148-9-157.

136. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797. 10.1093/nar/gkh340.

137. Capella-Gutierrez, S., Silla-Martinez, J.M., and Gabaldon, T. (2009). trimAl: a tool for

automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973. 10.1093/bioinformatics/btp348.

138. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274. 10.1093/molbev/msu300.

139. Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinf. *19*, 153. 10.1186/s12859-018-2129-y.

140. Morris, J.L., Puttick, M.N., Clark, J.W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H., Yang, Z.H., Schneider, H., and Donoghue, P.C.J. (2018). The timescale of early land plant evolution. Proc. Natl. Acad. Sci. USA. *115*, E2274–E2283. 10.1073/pnas.1719588115.

141. Hedges, S.B., Tao, Q.Q., Walker, M., and Kumar, S. (2018). Accurate timetrees require accurate calibrations. Proc. Natl. Acad. Sci. USA. *115*, E9510–E9511. 10.1073/pnas.1812558115.

142. Morris, J.L., Puttick, M.N., Clark, J.W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H., Yang, Z.H., Schneider, H., and Donoghue, P.C.J. (2018). Reply to Hedges et al.: Accurate timetrees do indeed require accurate calibrations. Proc. Natl. Acad. Sci. USA. *115*, E9512–E9513. 10.1073/pnas.1812816115.

143. Sanderson, M.J. (2002). Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol. *19*, 101–109. 10.1093/oxfordjournals.molbev.a003974.

144. Smith, S.A., and O'Meara, B.C. (2012). treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics *28*, 2689–2690. 10.1093/bioinformatics/bts492.

145. Nielsen, R., and Yang, Z.H. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics *148*, 929–936.

146. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. *30*, 3059–3066. 10.1093/nar/gkf436.

147. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. *17*, 540–552.

10.1093/oxfordjournals.molbev.a026334.

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

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

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

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

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

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

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

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		
Takakia lepidozioides	This study	IMSC: 41271
Chemicals, peptides, and recombinant proteins		
Critical commercial assays		
Deposited data		
Takakia lepidozioides: Pacbio CCS, Hi-C, Illumina DNA and RNA sequencing data	This study	NCBI BioProject: PRJNA796444
Takakia lepidozioides: Pacbio CCS, Hi-C, Illumina DNA and RNA sequencing data	This study	NGDC BioProject: PRJCA014819
Takakia lepidozioides genome assembly and gene annotations	This study	https://www.takakia.com
Takakia lepidozioides: 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/xzhub/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.mpimp-golm.mpg.de/geseq.html
MCScan 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/BIONF/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.online.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/recat/
Repbase		http://www.girinst.org/repbase
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

Figure 1

Click here to access/download;Figure;submit-Figure1-17.4cm.tif

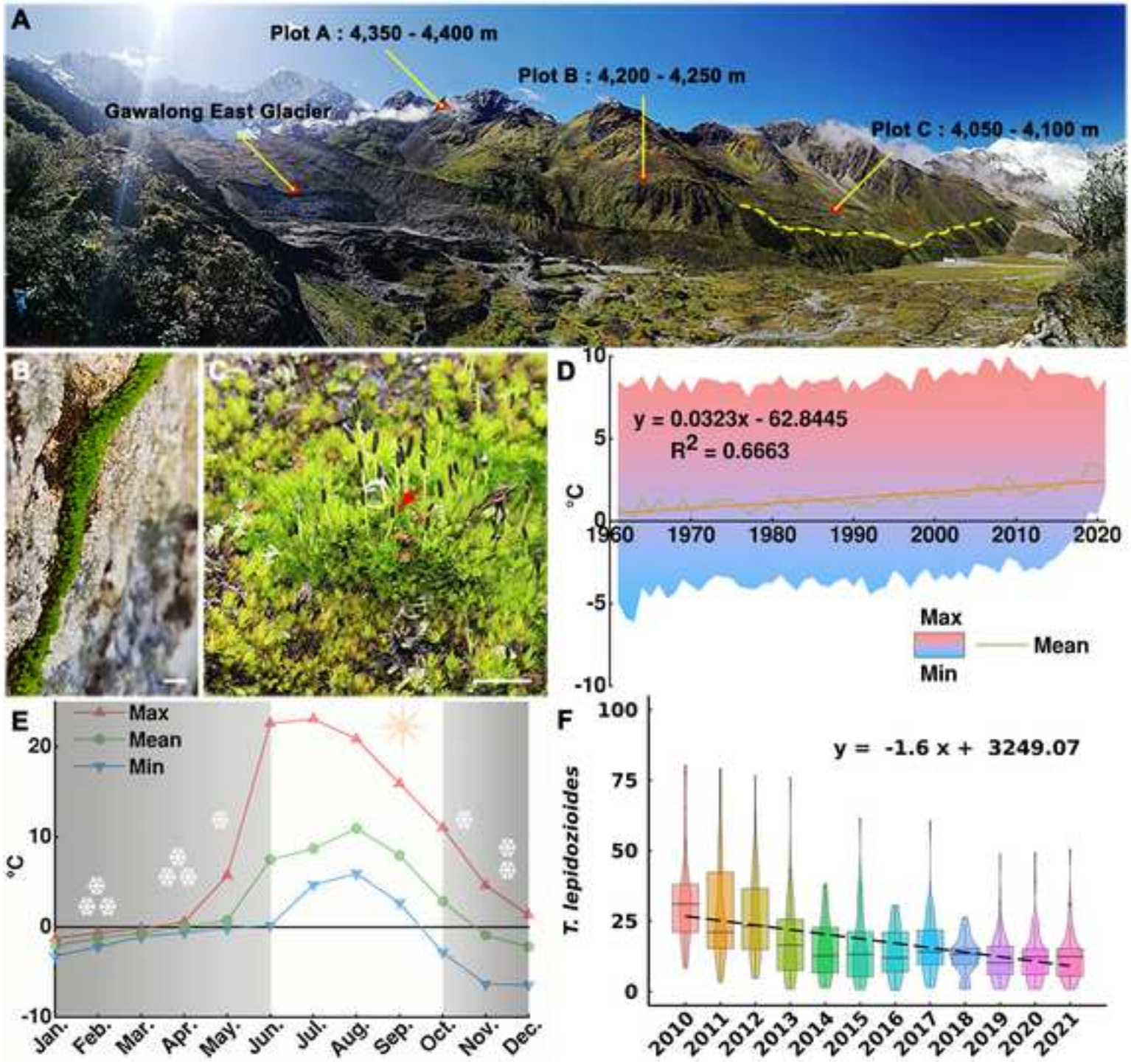


Figure 2

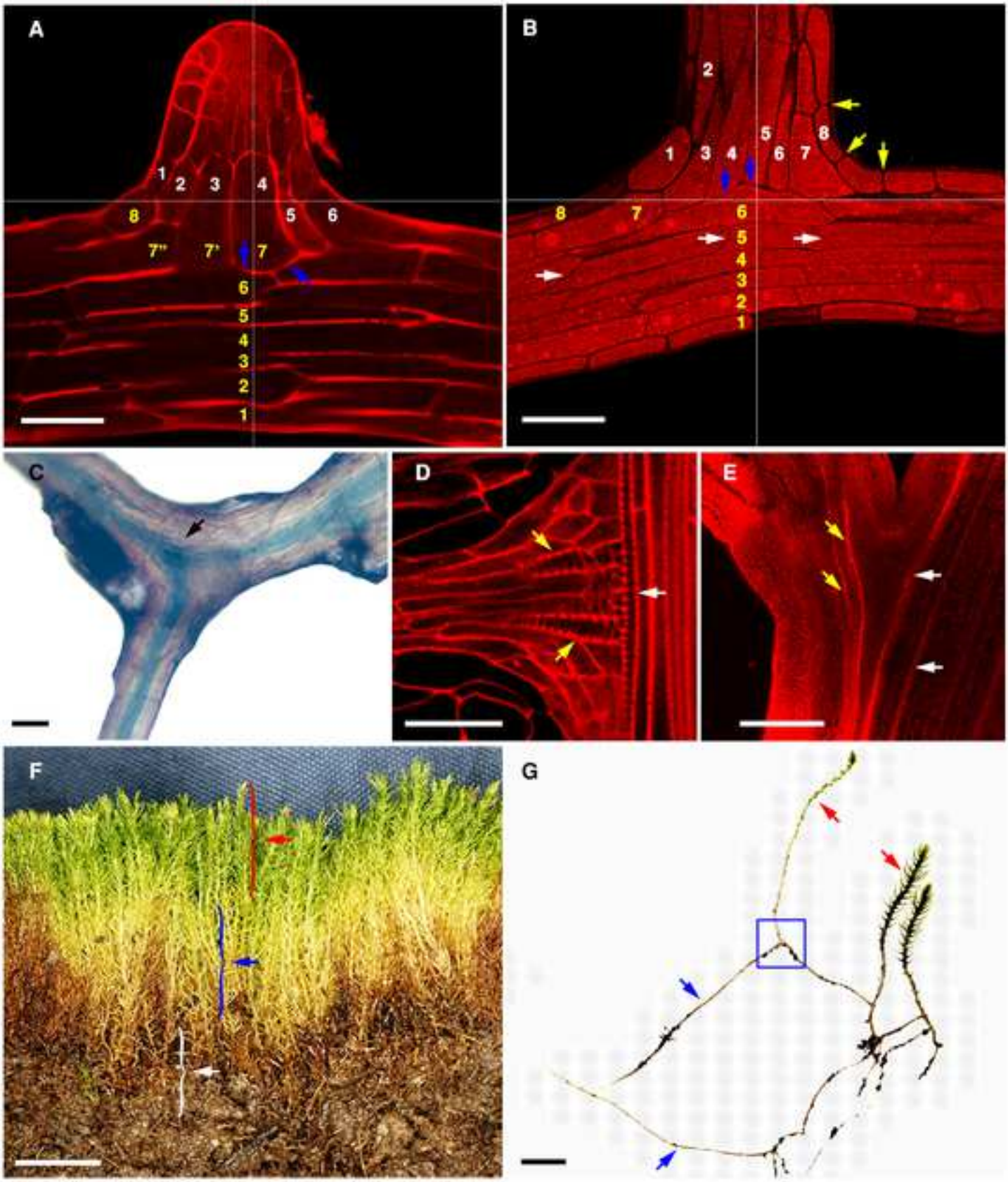


Figure 3

Click here to access/download;Figure;submit-Figure3-24.7cm.tif

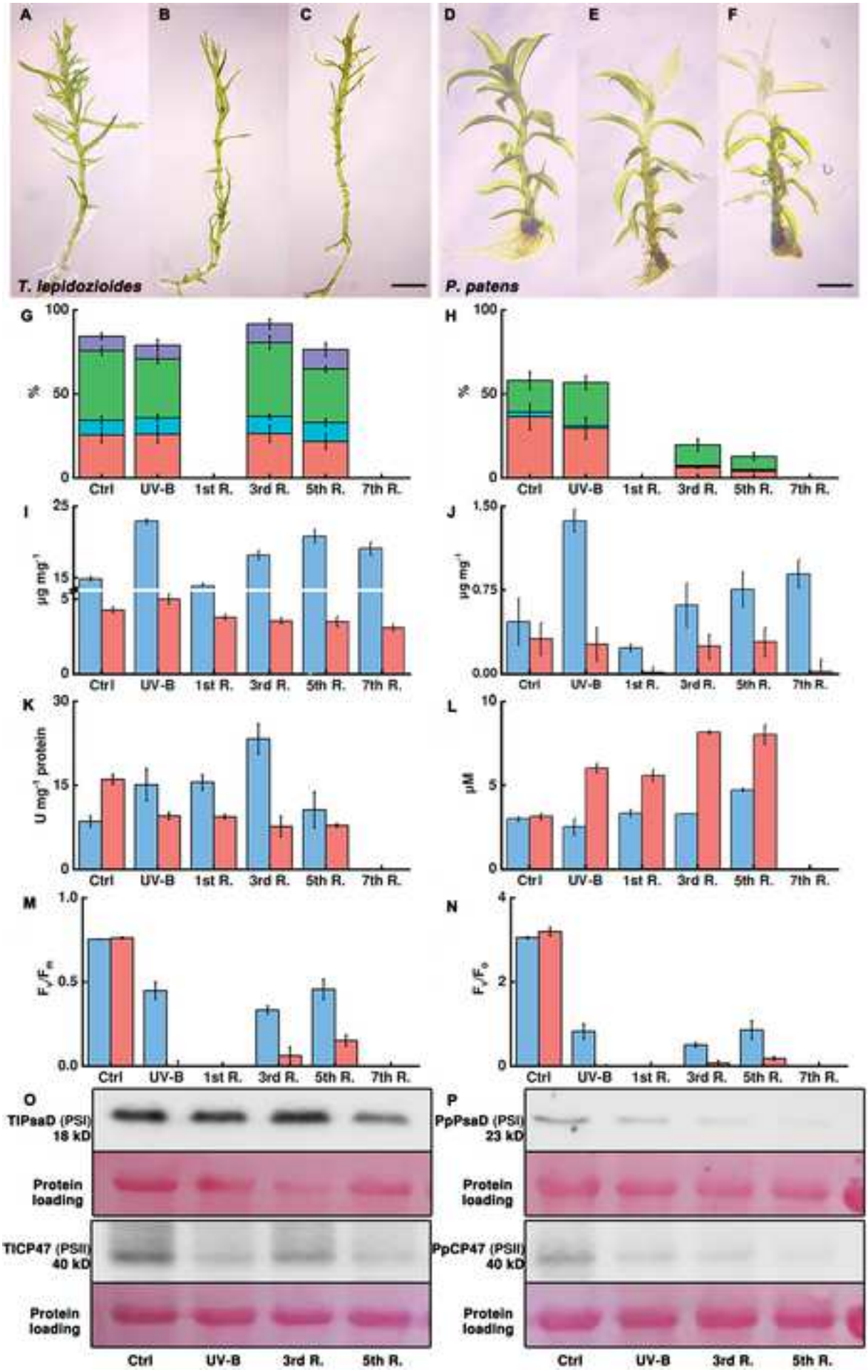


Figure 4

Click here to access/download;Figure;submit-Figure4-17.4cm.tif

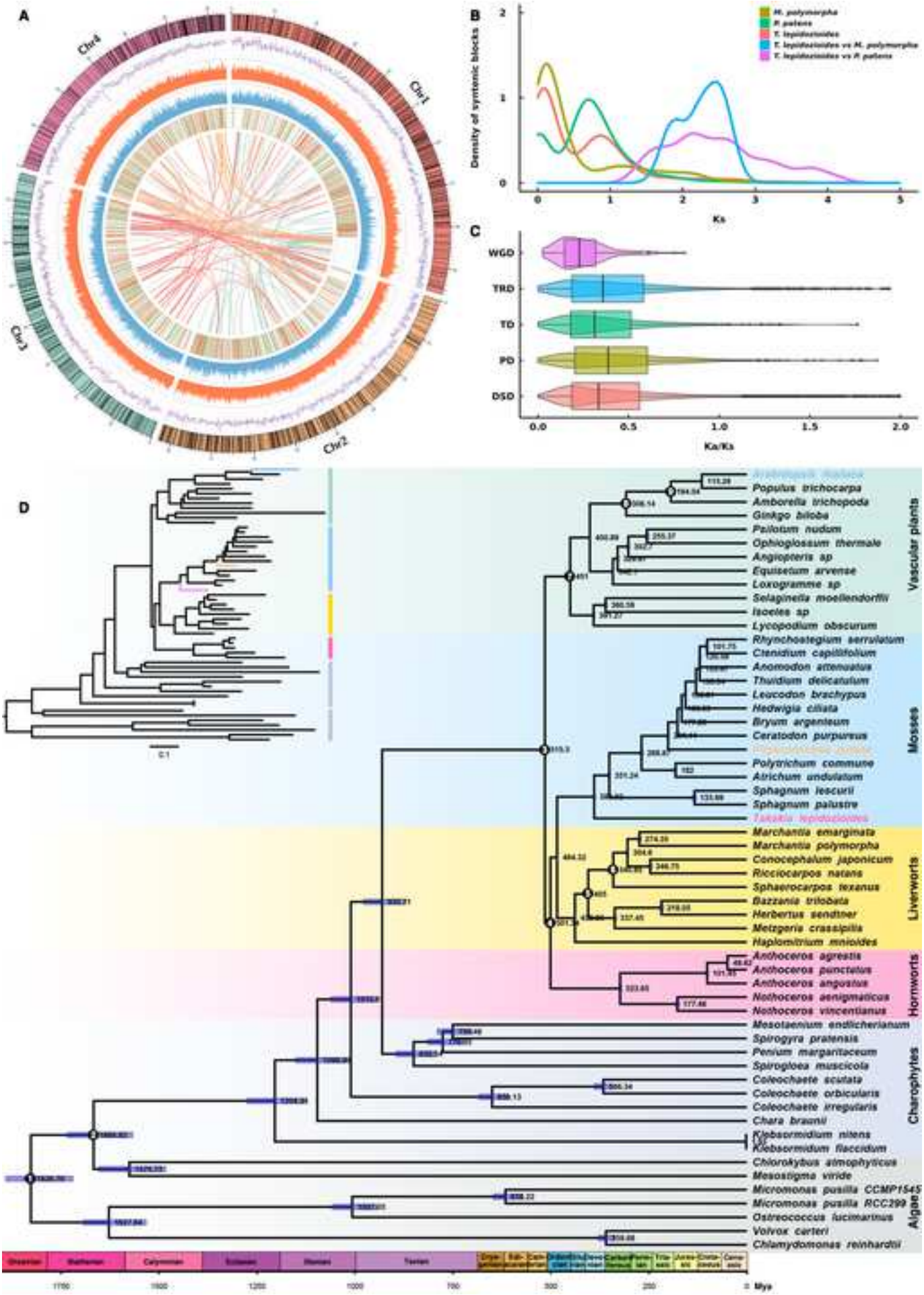


Figure 5

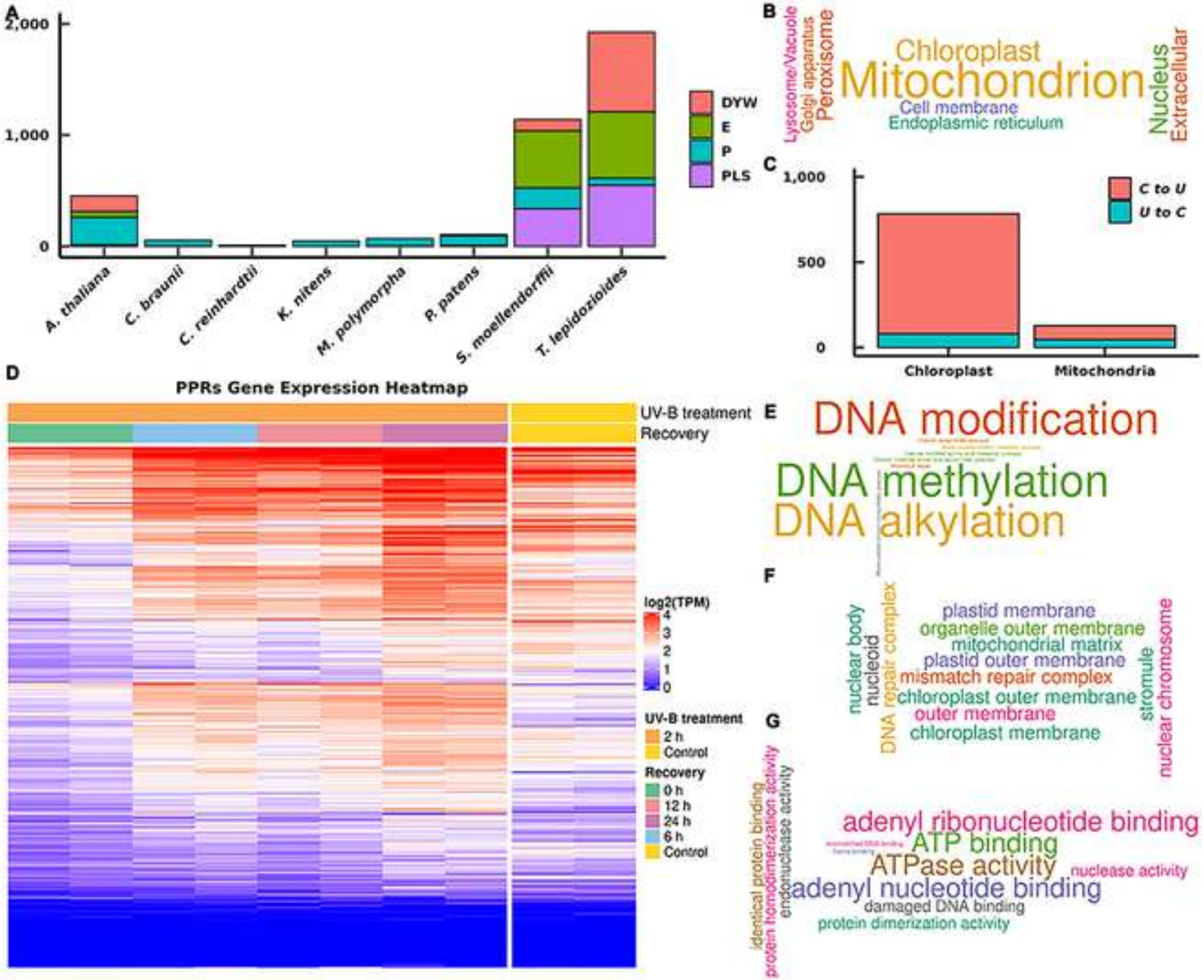
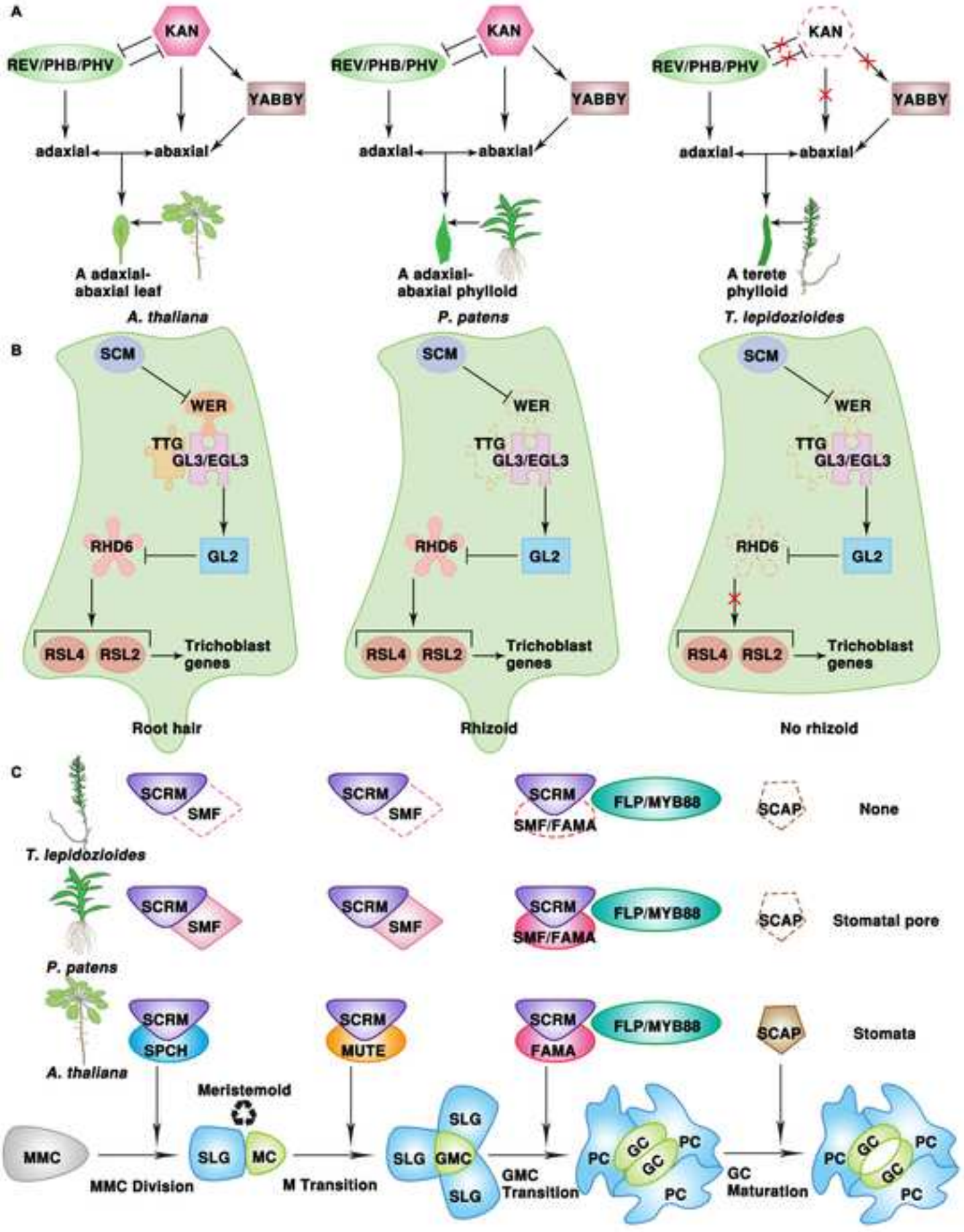
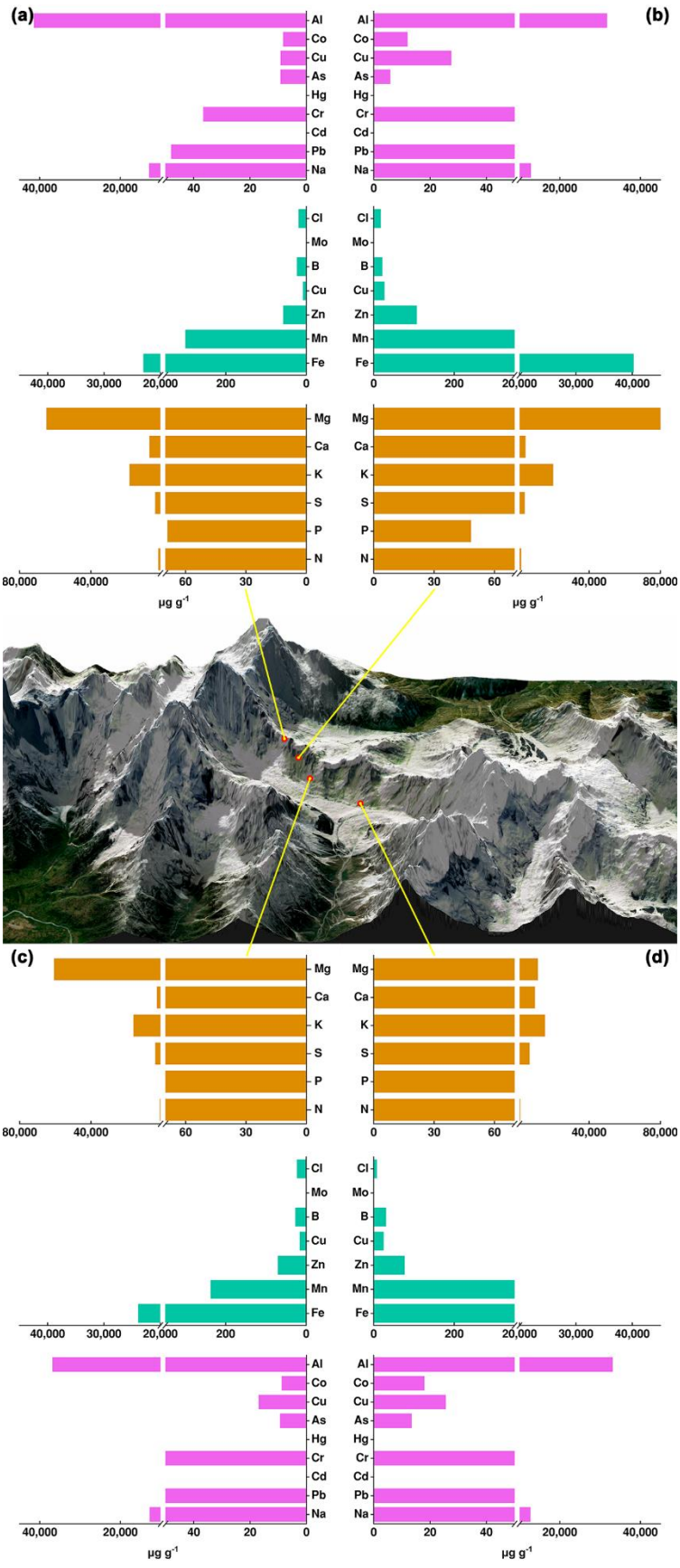


Figure 6



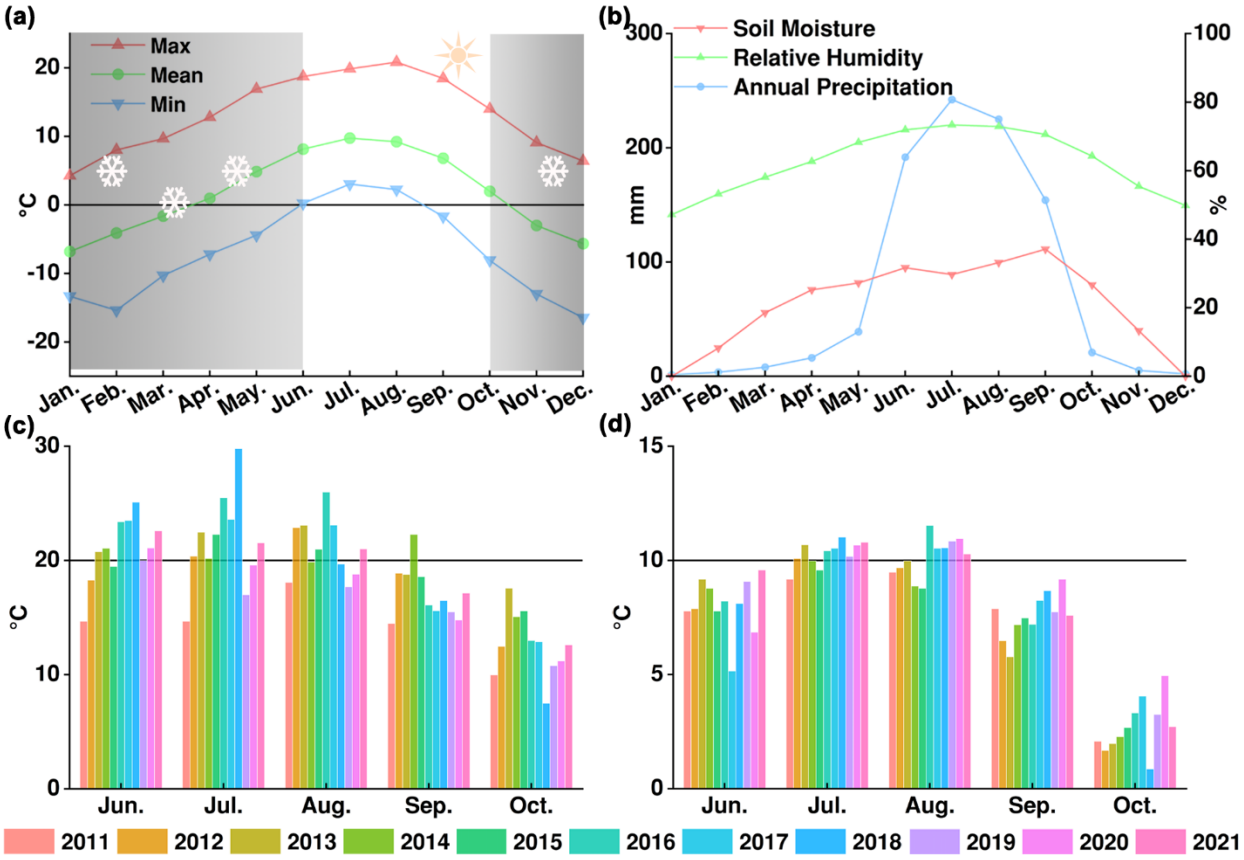
Supplement data 1


(a)
(b)
(c)
(d)
Al
Co
Cu
As
Hg
Cr
Cd
Pb
Na
Cl
Mo
B
Cu
Zn
Mn
Fe
Mg
Ca
K
S
P
N
40,000
20,000
40
20
0
200
30,000
60
30
80,000
μg g-1

3

Data S1 A. Soil mineral elements in the experimental area

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

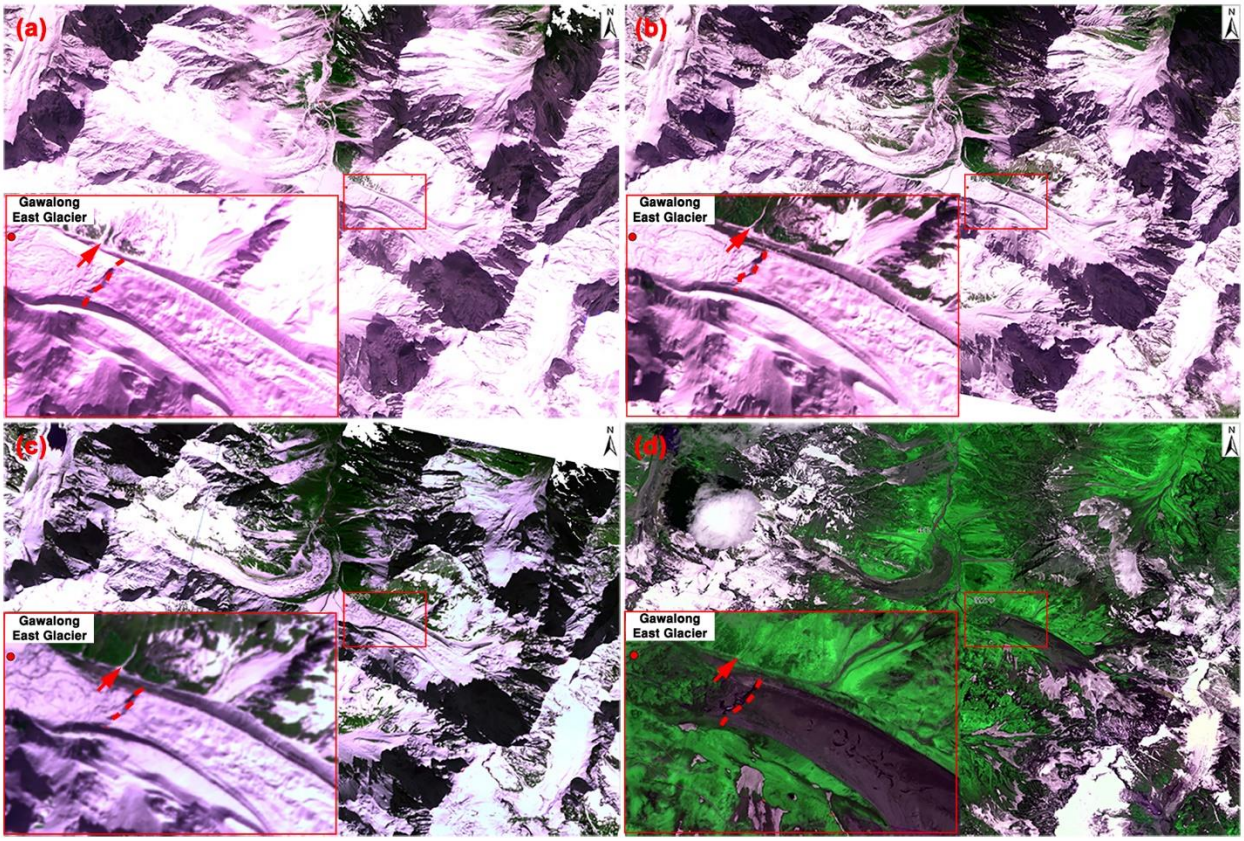


Data S1 B. Climate change from 2011 to 2021 in the experimental area

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

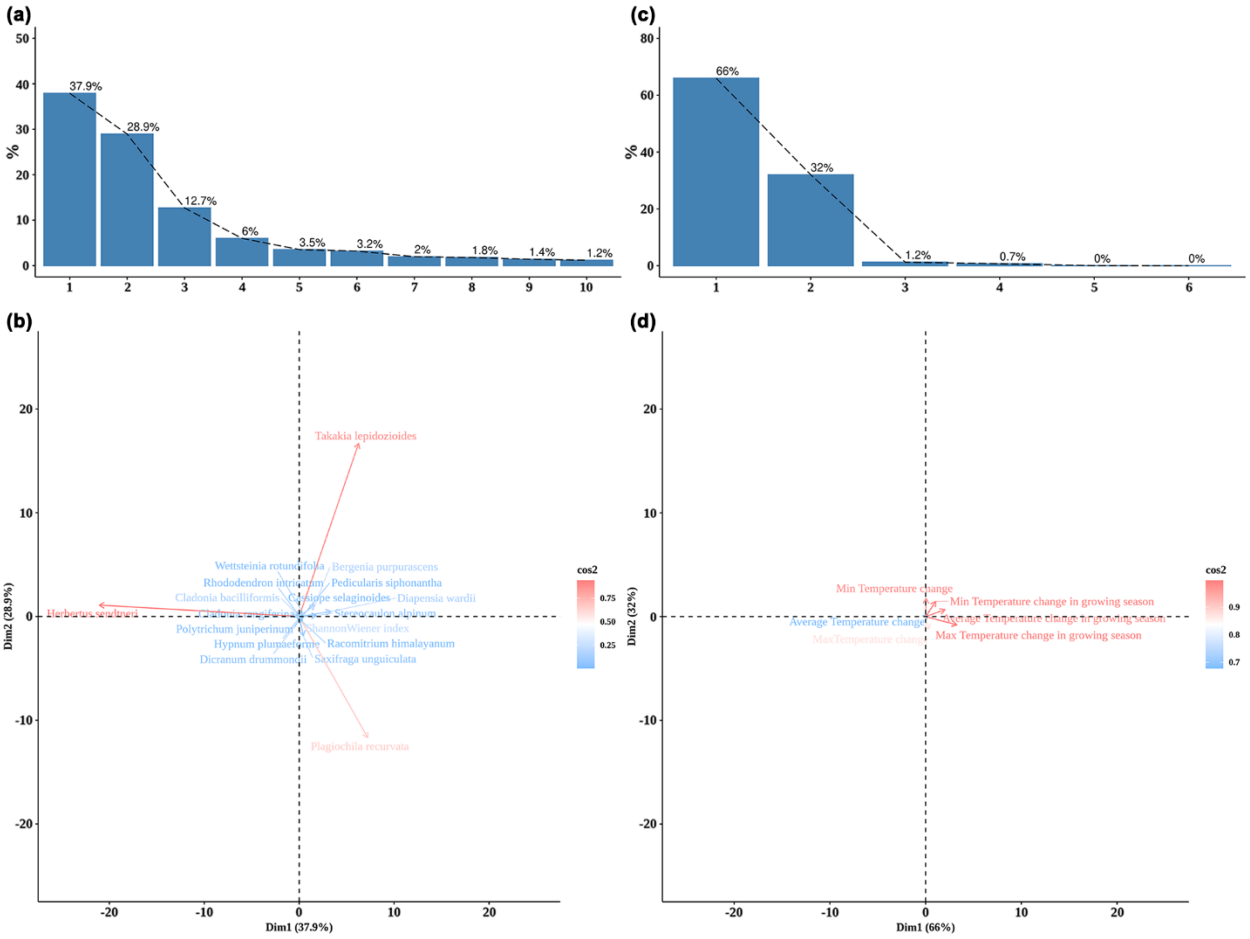
Data S1 C. Photographs of the edge of the Gawalong East Glacier

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



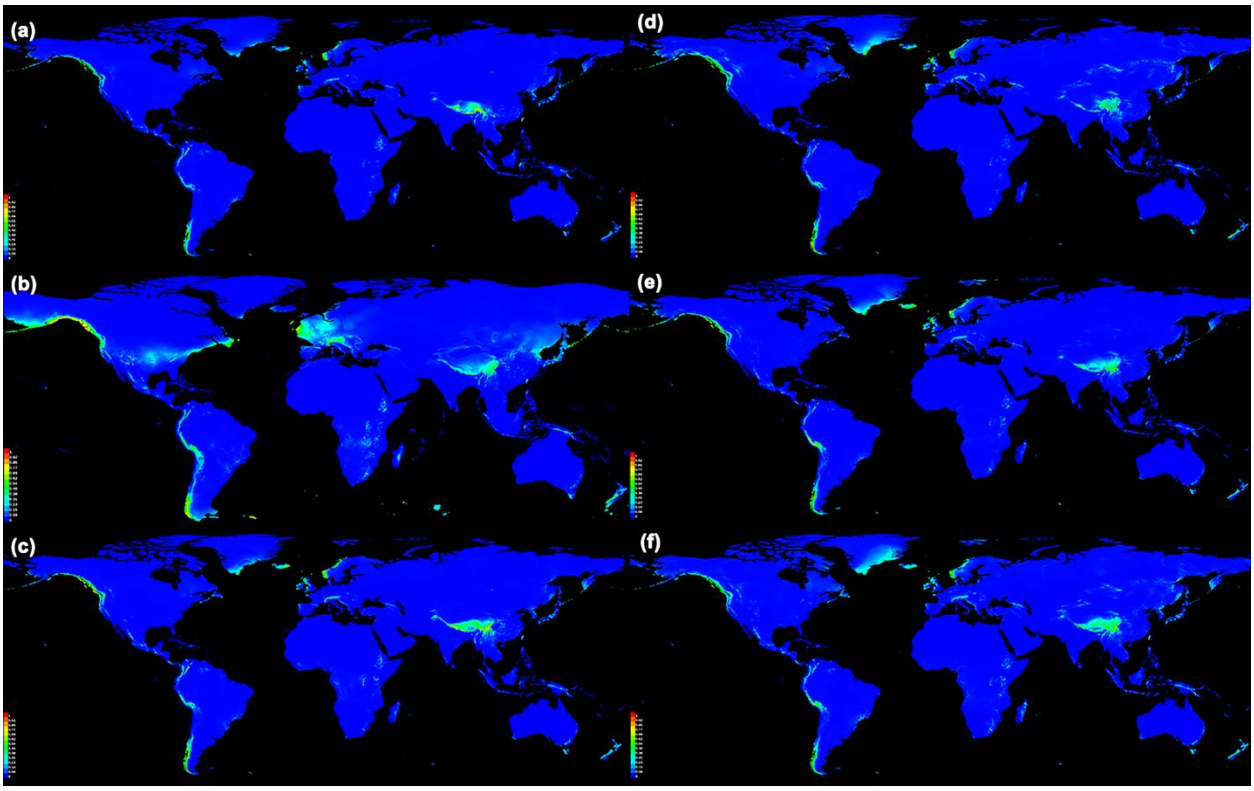
Data S1 D. Satellite images of the edge of the Gawalong East Glacier

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



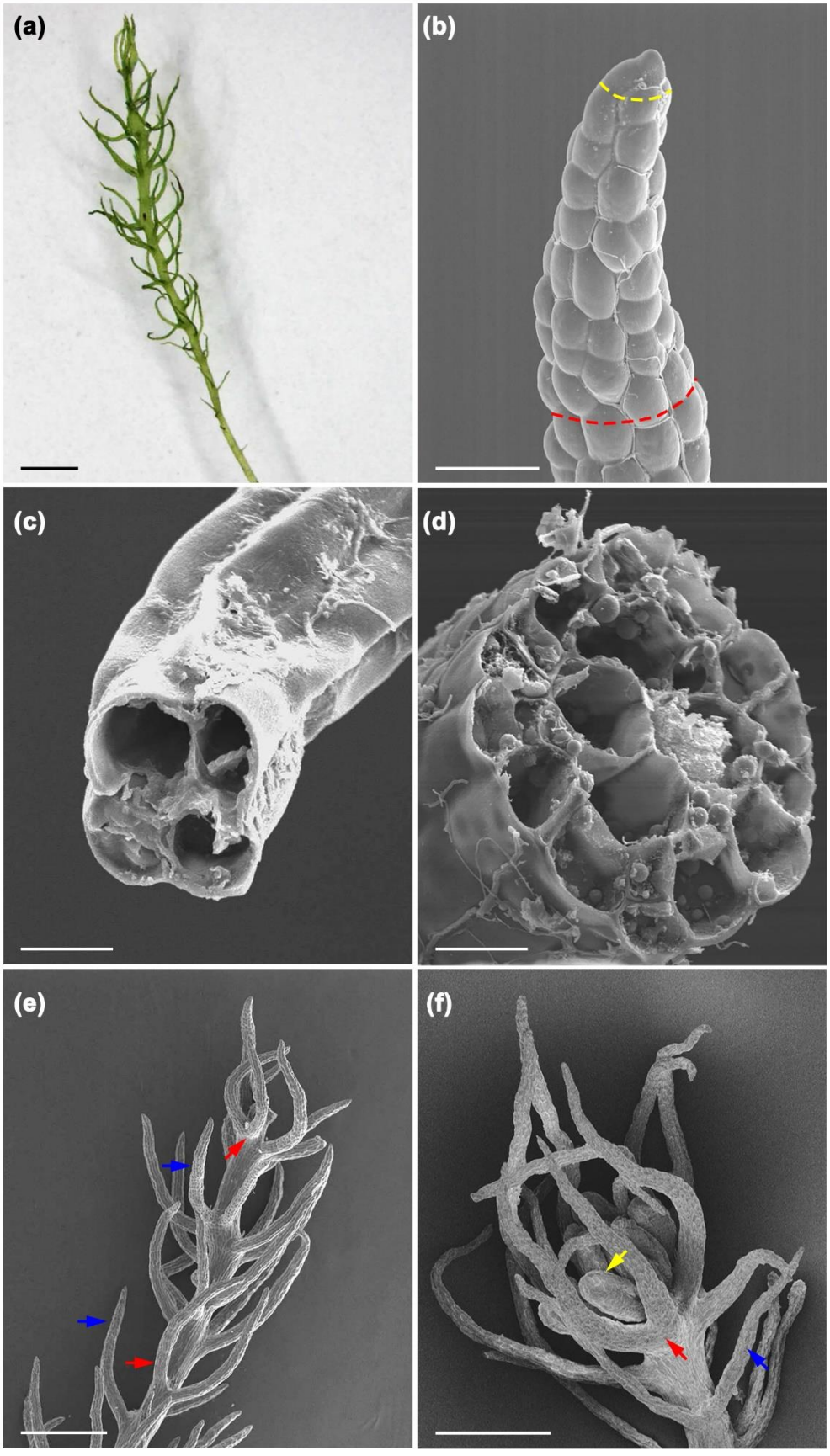
Data S1 E. Correlation analysis of population changes and climatic changes

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



Data S1 F. Potential distribution of *T. lepidozioides* over time

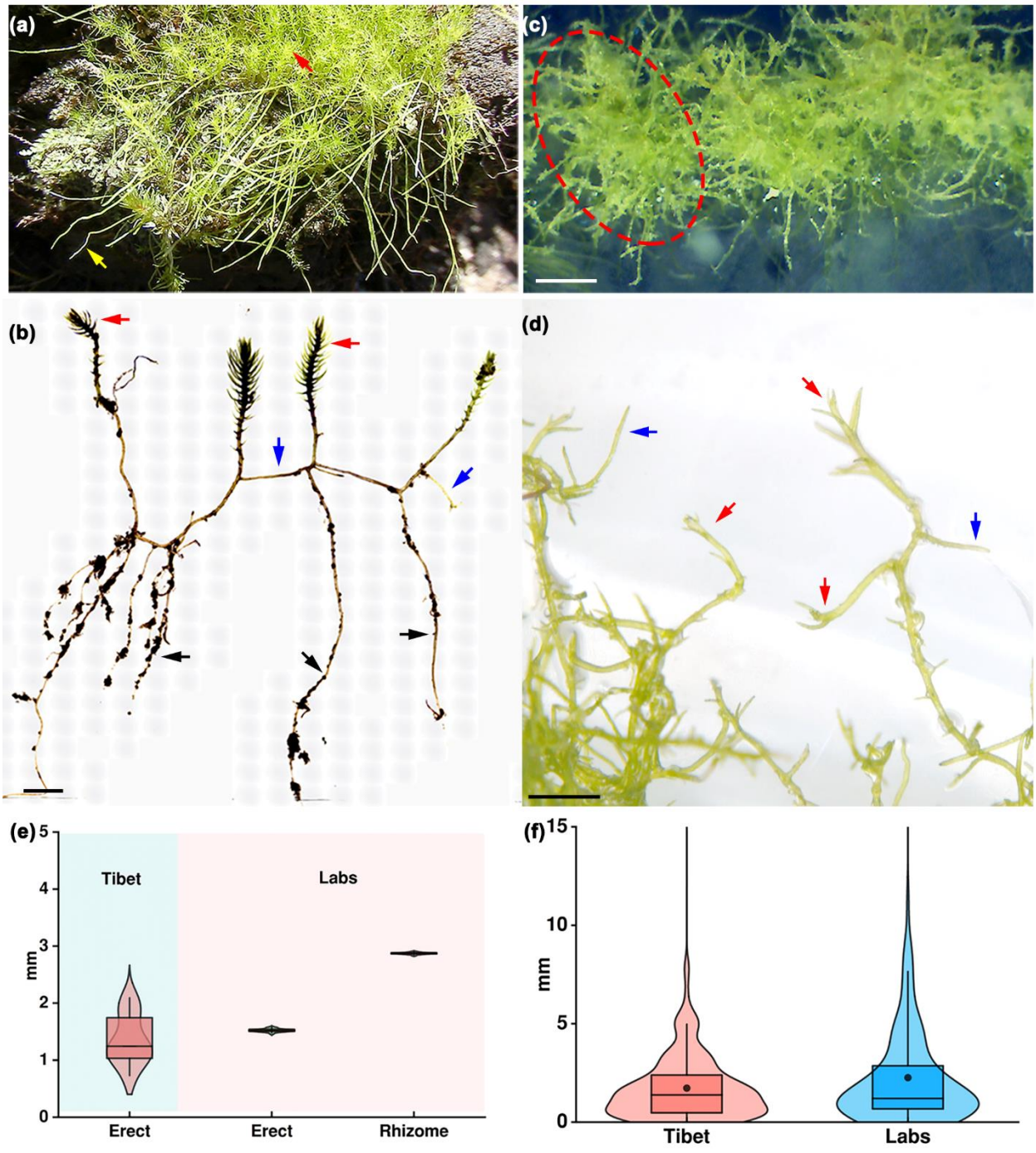
Potential distribution of *T. lepidozioides* (**a**) in the last interglaciation (~120,000 – 140,000 years ago), (**b**) in the last glaciation maximum (about 22,000 years ago), (c) in the Holocene mega thermal (about 6,000 years ago), (**d**) 1970s to 2000s, (**e**) 2041s to 2060s, (**f**) 2061s to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



66

Data S1 G. The leaf shape of *T. lepidozioides*

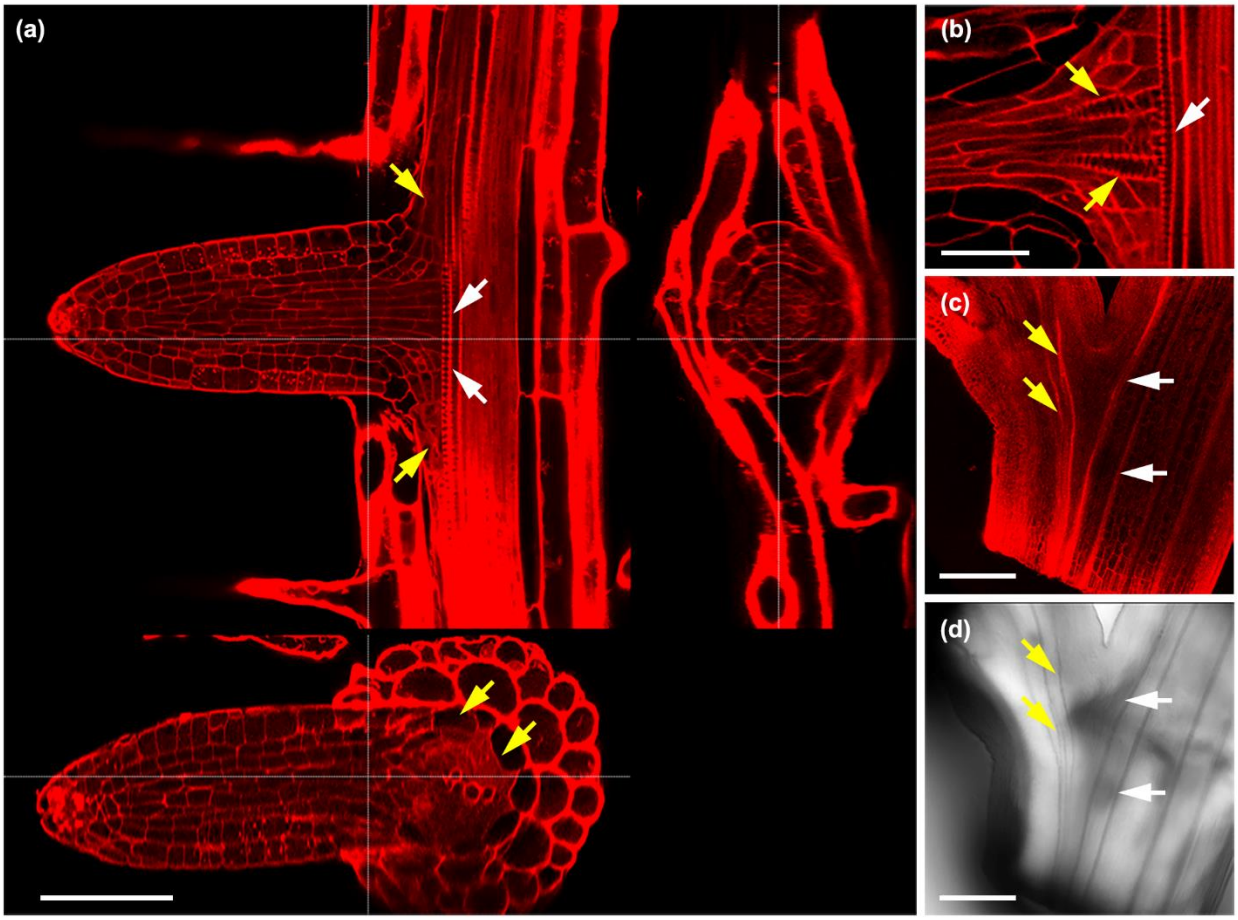
Plants are from the population of Plot B in the field. (**a**) An erect stem with finger-shaped leaves. (**b – d**) SEM images of a leaf (**b**), the transection (**c**) from the tip (yellow circle in b) and the transection (**d**) from below (red circle in b). (**e – f**) Arrows mark the single finger-shaped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



Data S1 H. Vegetative growth features of *T. lepidozioides*

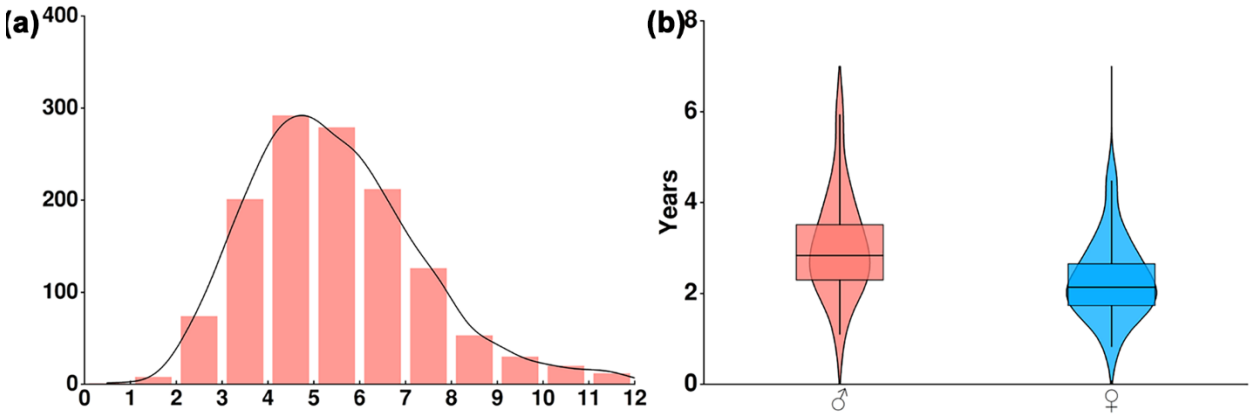
(**a**) A natural population of *T. lepidozioides* in Plot B, red arrow marks erect stem, yellow arrow marks rhizome, photograph was made using EOS-1D X Mark II (Canon) with macro lens (MP-E65 mm, Canon). (**b**) One part of an individual plant isolated from the population of (a) (bar = 2 mm), arrows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

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



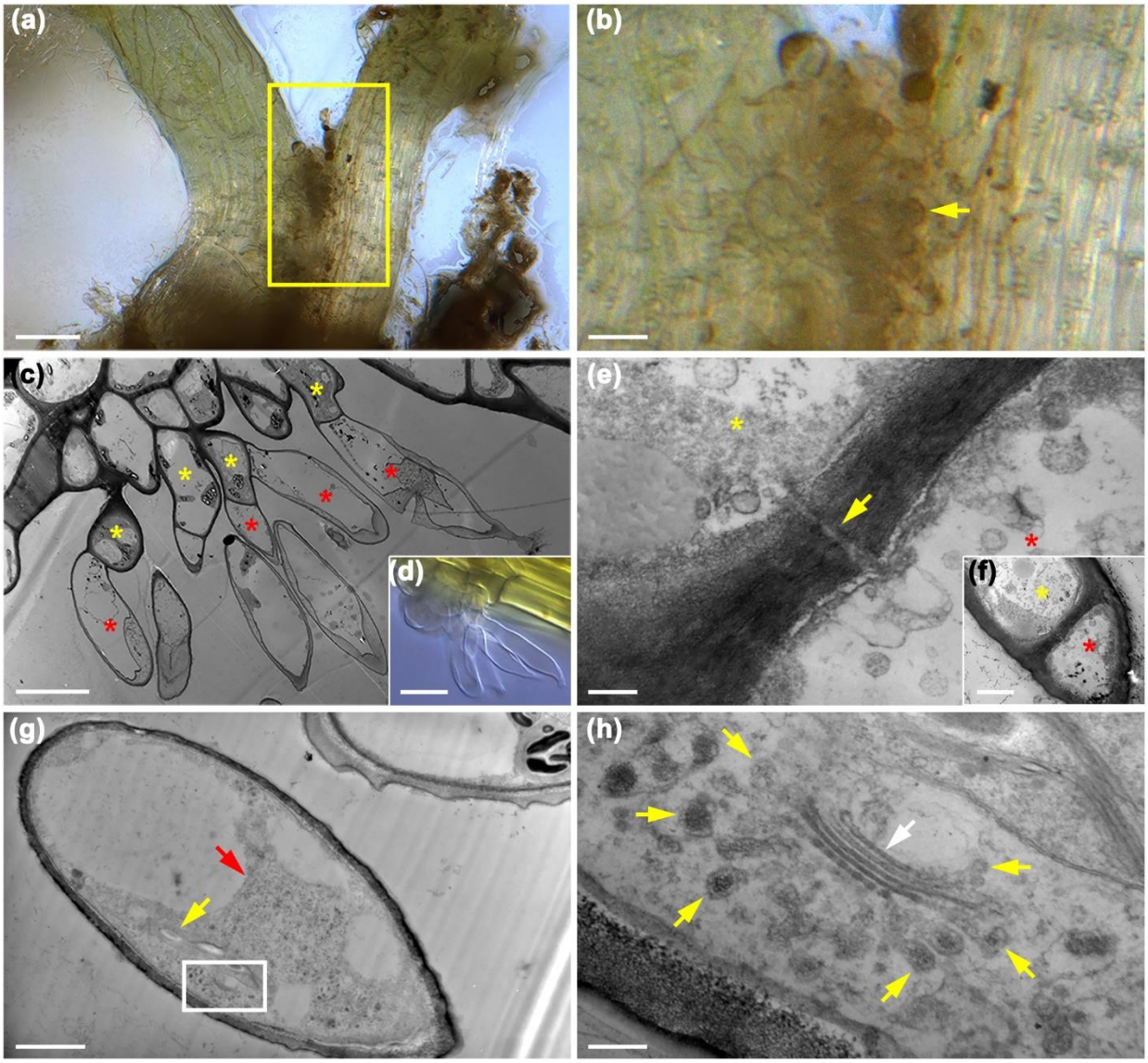
Data S1 I. The connection mode of vessel elements in vascular bundles between lateral-main roots and branches in *A. thaliana*, observed with the mPS-PI staining technique

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



Data S1 J. Sexual reproduction features of *T. lepidozioides*

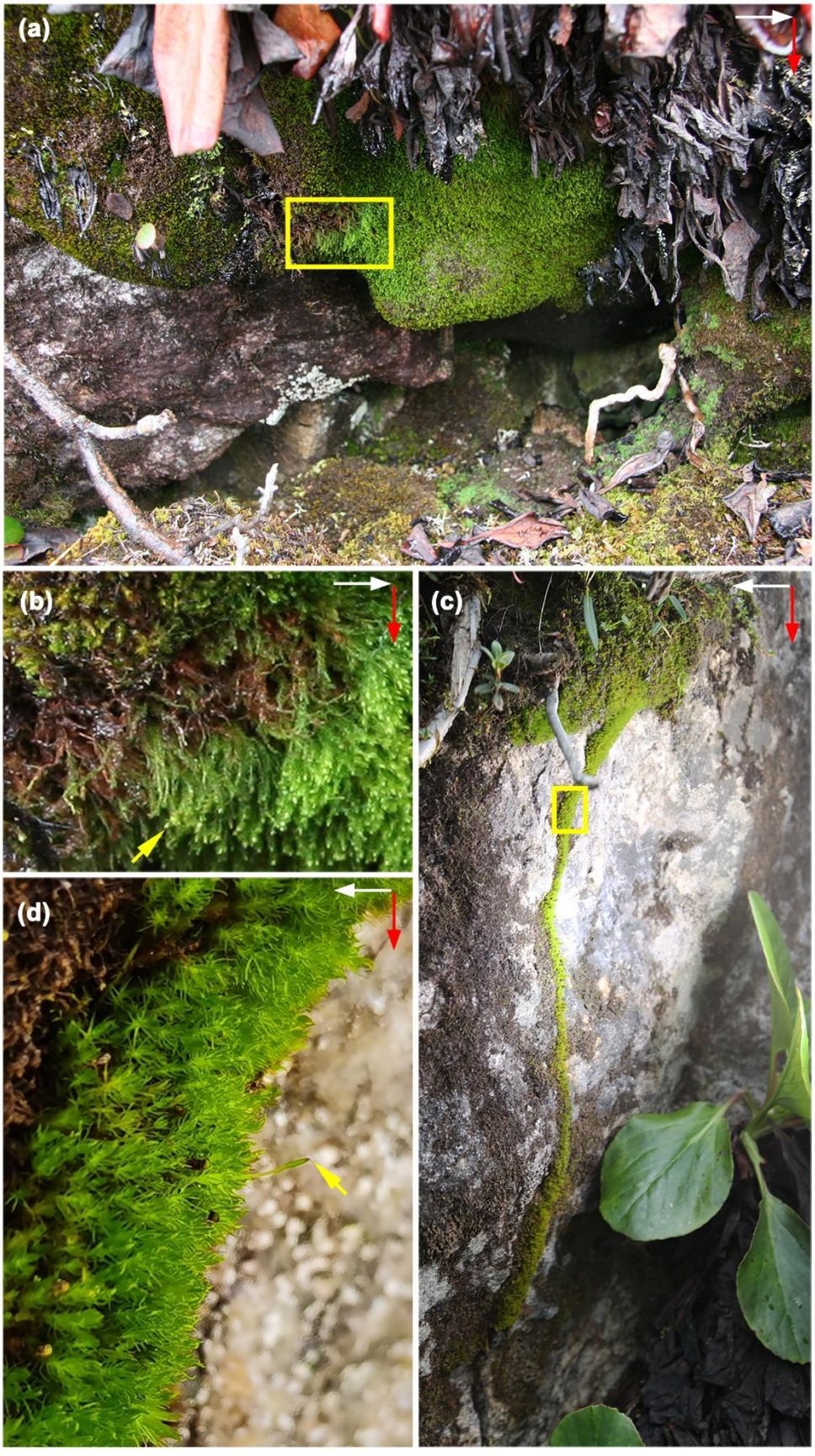
(**a**) Distribution blot of the height of erects in the fields. Horizontal axis represents the height of erects in mm, vertical axis represents the frequency of occurrence (n = 1,315). (**b**) The timing that erect stems began sexual reproduction. Male plants achieve sexual maturity in 3.06 ± 1.14 years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



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

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

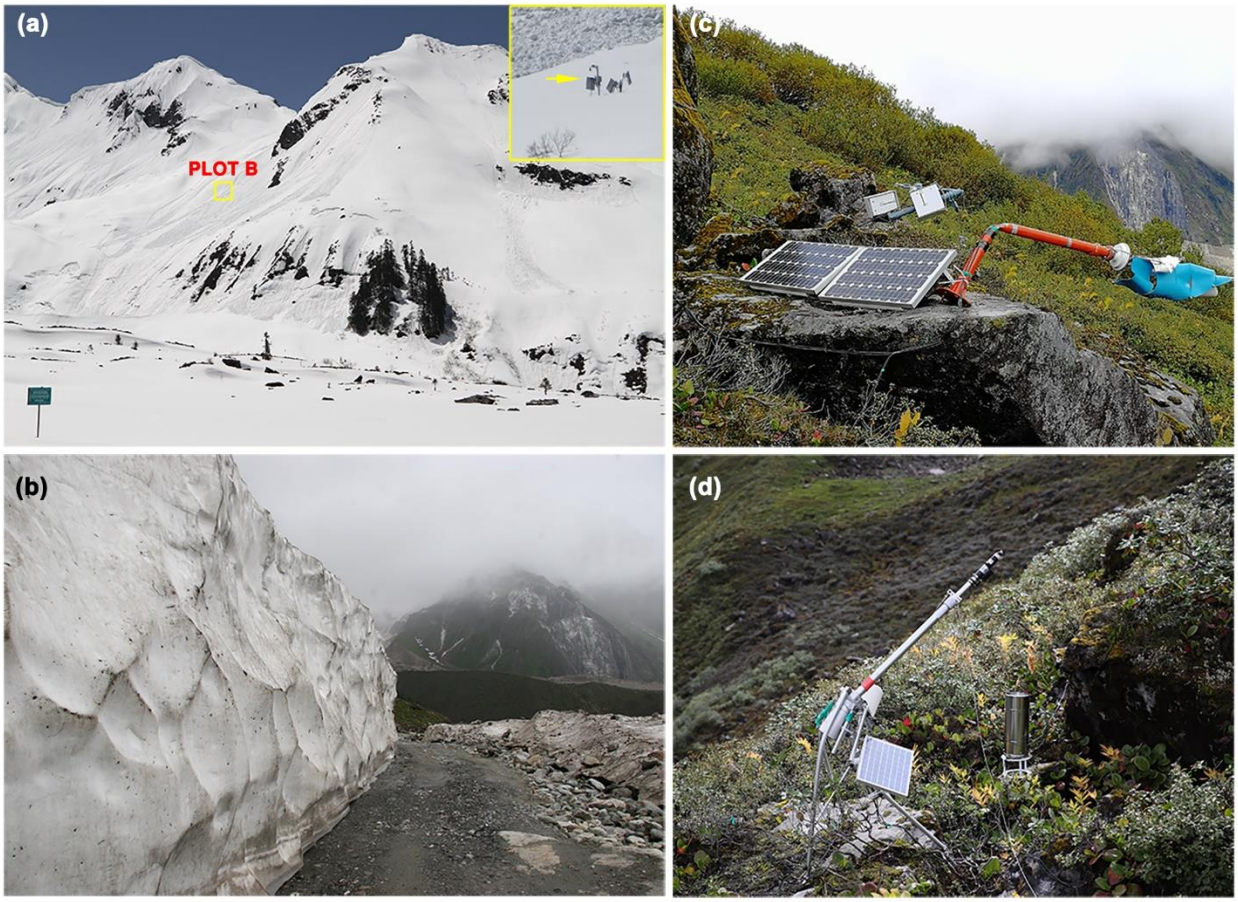
chloroplast (yellow arrow) and Golgi apparatus (white box), bar = 20 µm. (**h**) The Golgi apparatus (white arrow) with secretory vesicles (yellow arrows), magnified from the box of (g), bar = 200 nm. (a) and (d) were made using an optical microscope (BA600-4, Motic), (b) was made using a stereo microscope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



137

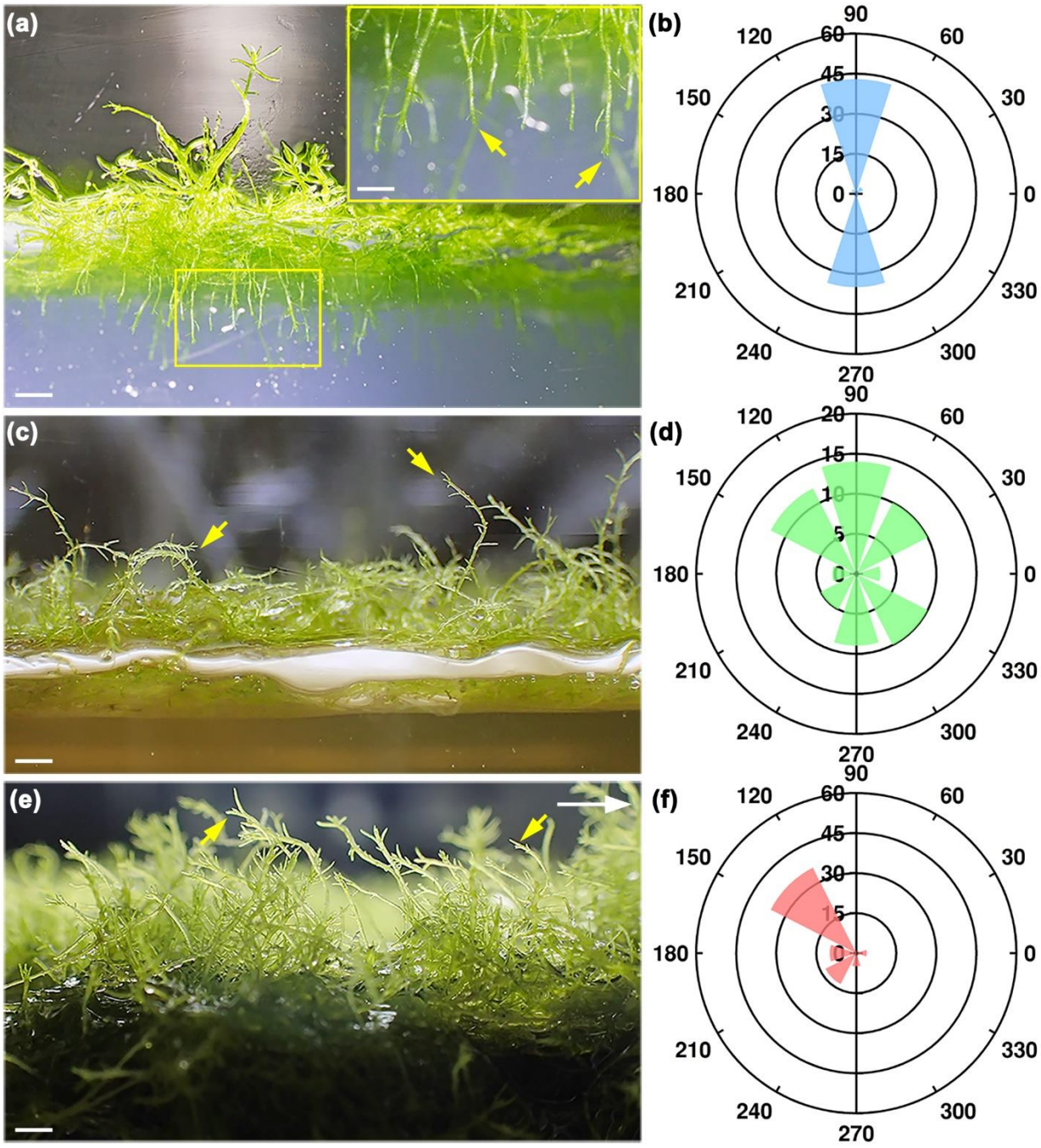
Data S1 L. *T. lepidozioides* populations in the field

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



Data S1 M Heavy snowstorm during winter in the experiment sites

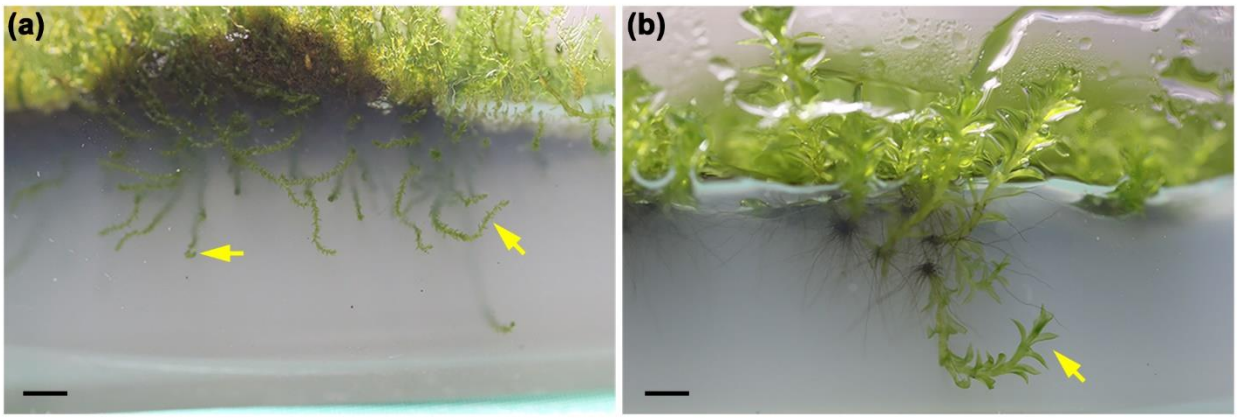
Heavy snowstorm during winter (**a, b**) and meteorological instrument was broken (**c, d**) by the heavy snow in the experiment plots. (**a**) The picture was captured by an unmanned aerial vehicle (Matrice 300 RTK, DJI) on 5th May, 2020, yellow box marks meteorological instrument in plot B and magnified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



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

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

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



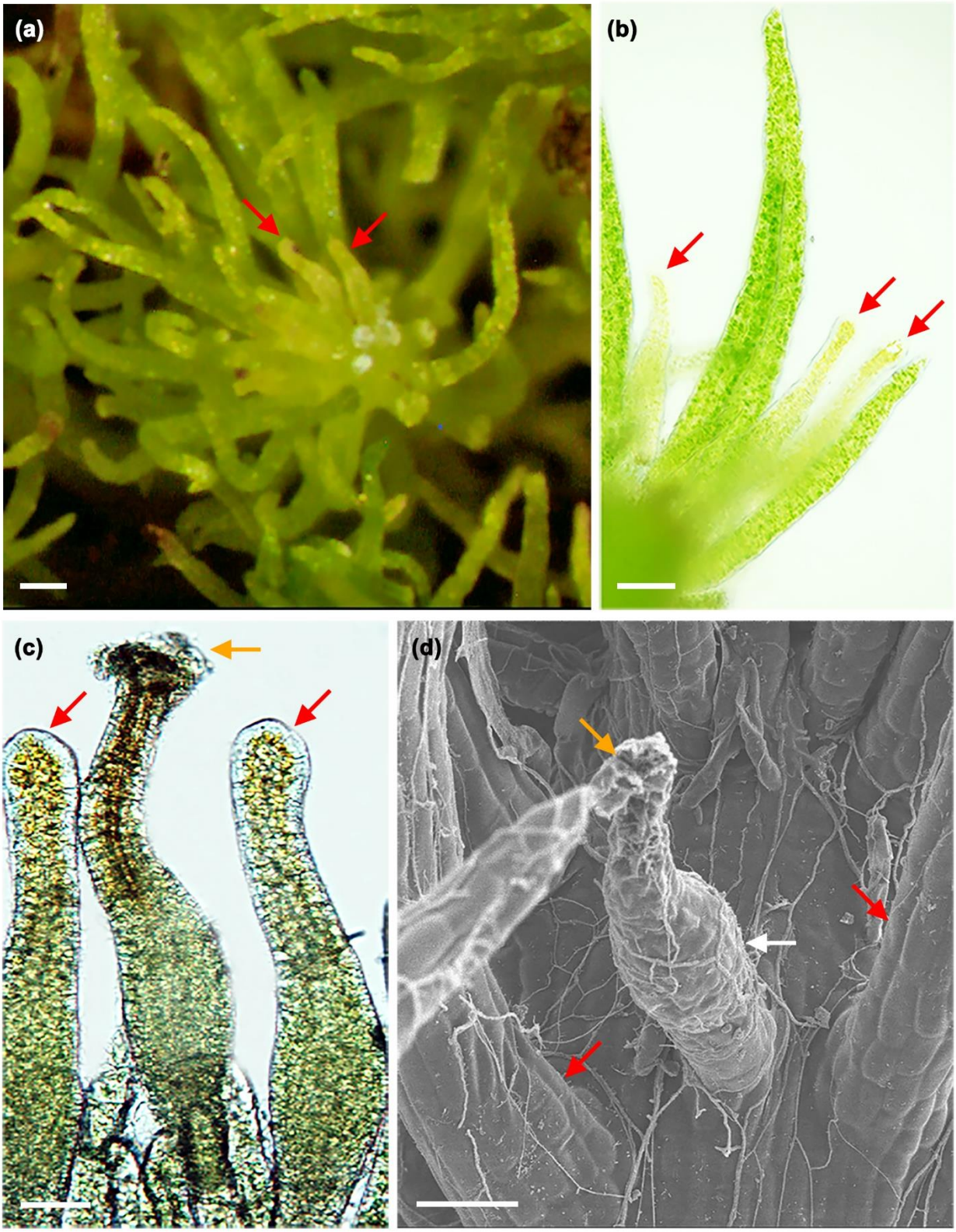
Data S1 O Negative geotropism of the mosses *Bryum argenteum* and *Physcomitrium patens* in culture

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

Data S1 P. Antheridia development in *T. lepidozioides*

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

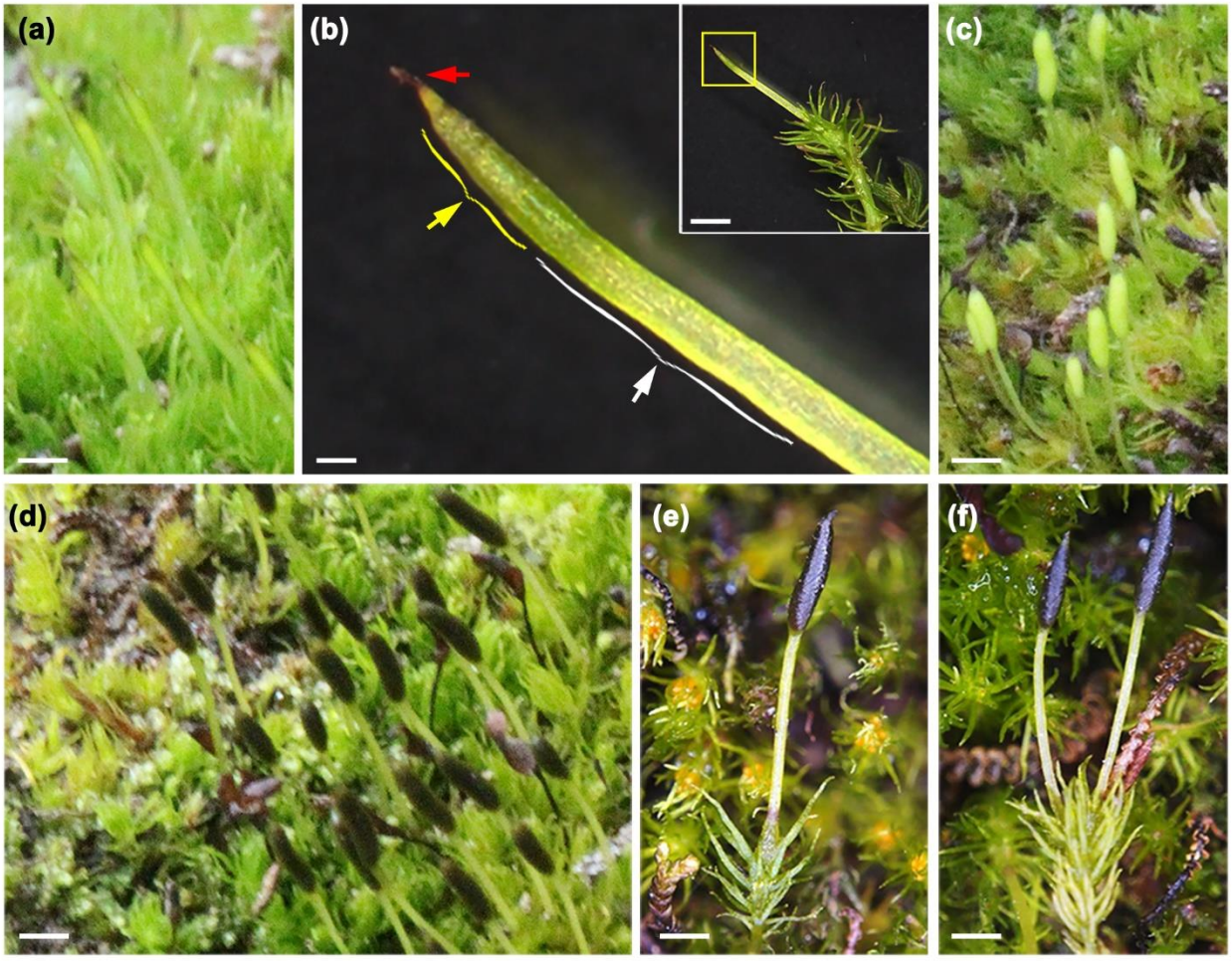
previous year, bar = 200 µm. (**d**) Magnified view of mature antheridia, bar = 200 µm. (**e**) Magnified view of mature and dehisced antheridia. White arrow marks a dehisced antheridium, bar = 200 µm. (a) and (b) were made using EOS-1D X Mark II (Canon) with macro lens (MP-E65 mm, Canon), (c) was made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



196
197

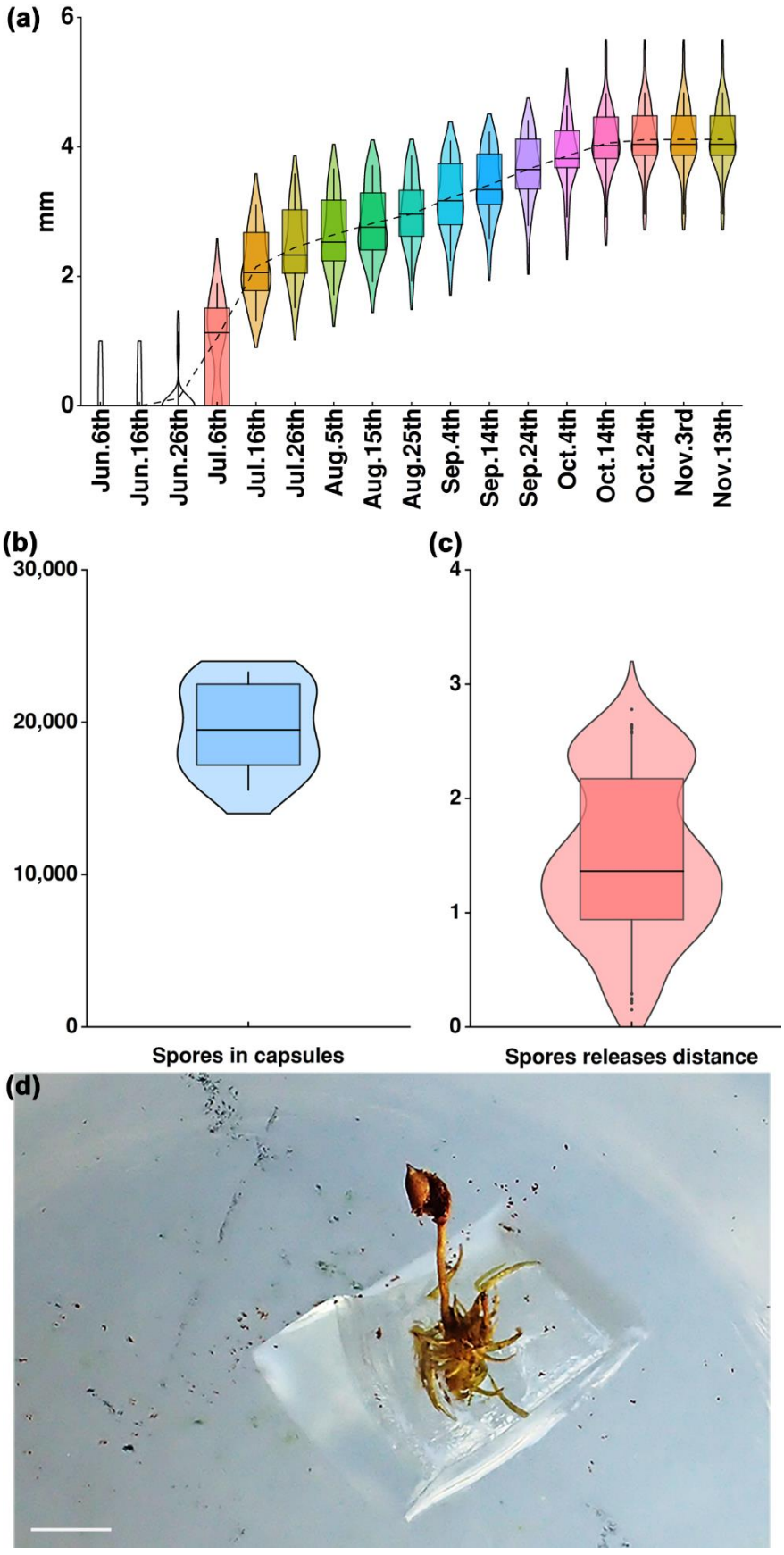
Data S1 Q. Archegonia development in *T. lepidozioides*

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



Data S1 R. Sporophyte development in *T. lepidozioides*

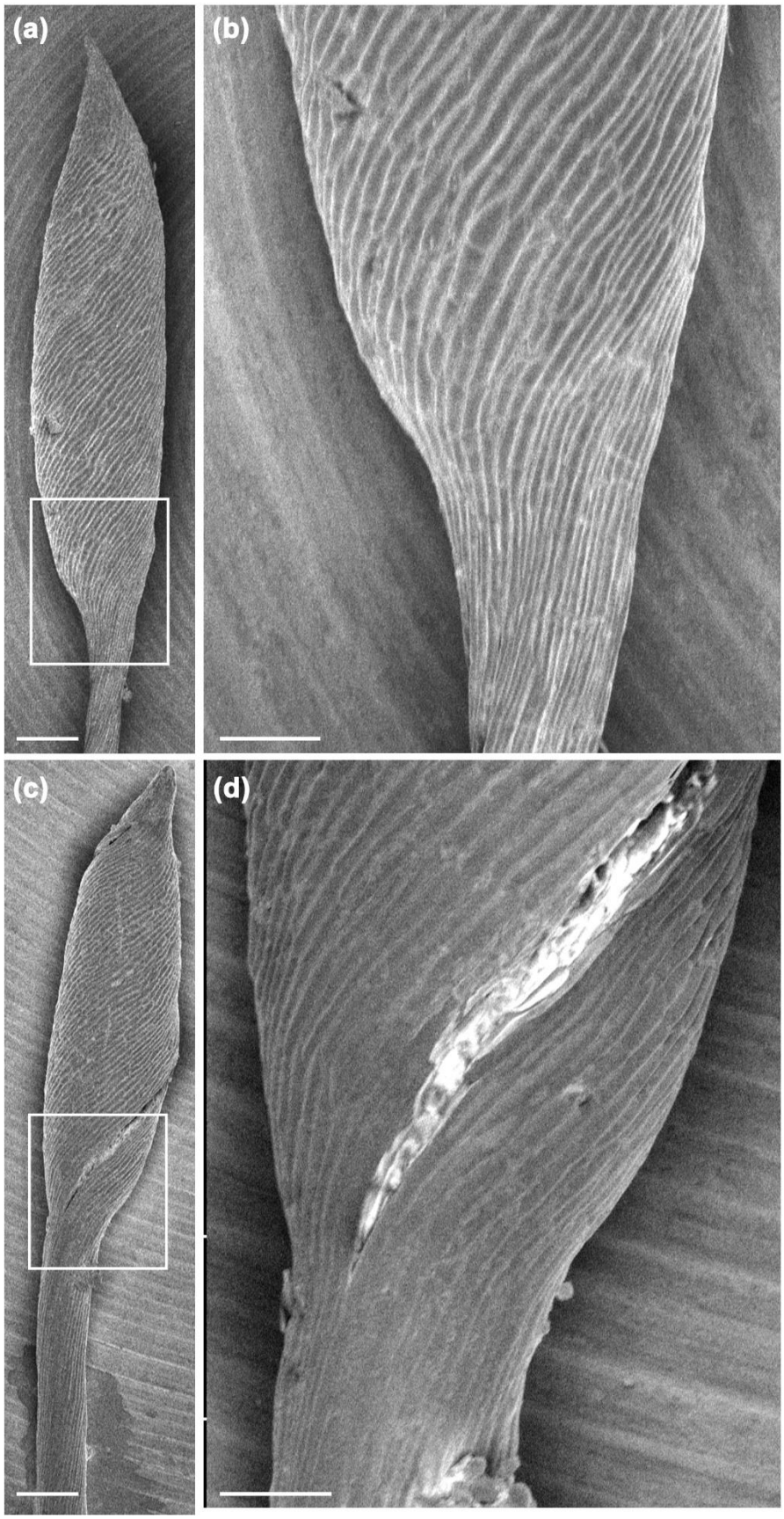
Development of sporophytes in the field. (**a**) Young sporophytes with unexpanded capsules in early July, bar = 1 mm. (**b**) A young sporophyte developing (magnified from top right yellow box, bar = 200 µm) from a fertilized archegonium with the calyptra (red arrow) on the apiculus of the unexpanded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


(a)
mm
0
2
4
6
Jun.6th
Jun.16th
Jun.26th
Jul.6th
Jul.16th
Jul.26th
Aug.5th
Aug.15th
Aug.25th
Sep.4th
Sep.14th
Sep.24th
Oct.4th
Oct.14th
Oct.24th
Nov.3rd
Nov.13th
(b)
30,000
20,000
10,000
0
Spores in capsules
(c)
4
3
2
1
0
Spores releases distance
(d)

219

Data S1 S. Capsule features of *T. lepidozioides*

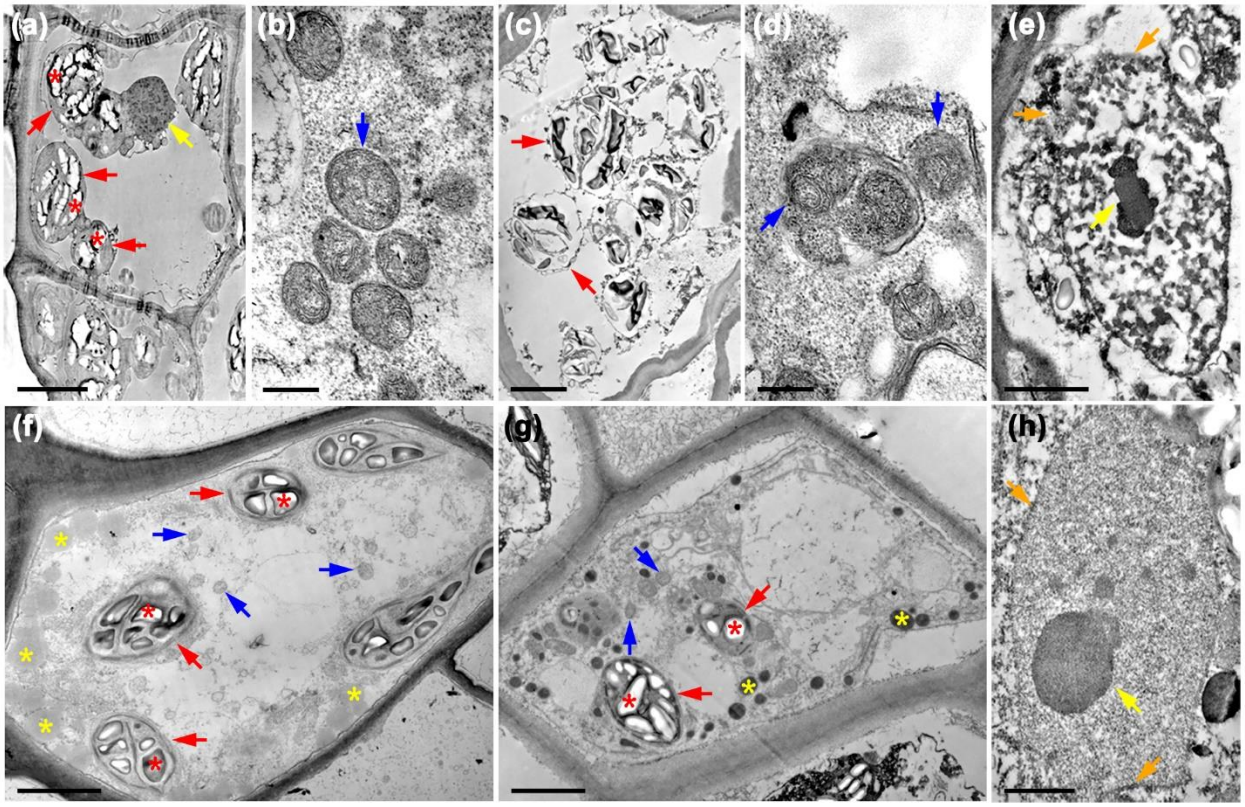
(**a**) Analysis of sporophyte growth and development. Growth was measured with ImageJ from time-lapse photos in every 10 days in 2014 (n = 40). The central mark of each box indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. (**b**) Spore numbers in one capsule (n = 8). (**c**) Spore dispersion distance (in cm) under laboratory conditions in (d), all radius data were from 4 capsules with 30 individual measurements in each. The central mark of each box indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. (**d**) Experimental setting for (c). Erect stem with mature capsule collected from the field and inserted into 1 cm^3 agar in a petri dish. The dish was placed in a chamber with silica gel to mimic dehiscence. The distance from spore to erect bottom were measured with ImageJ, photograph was made using a stereo microscope (M125 C, Leica), bar = 5 mm.



234

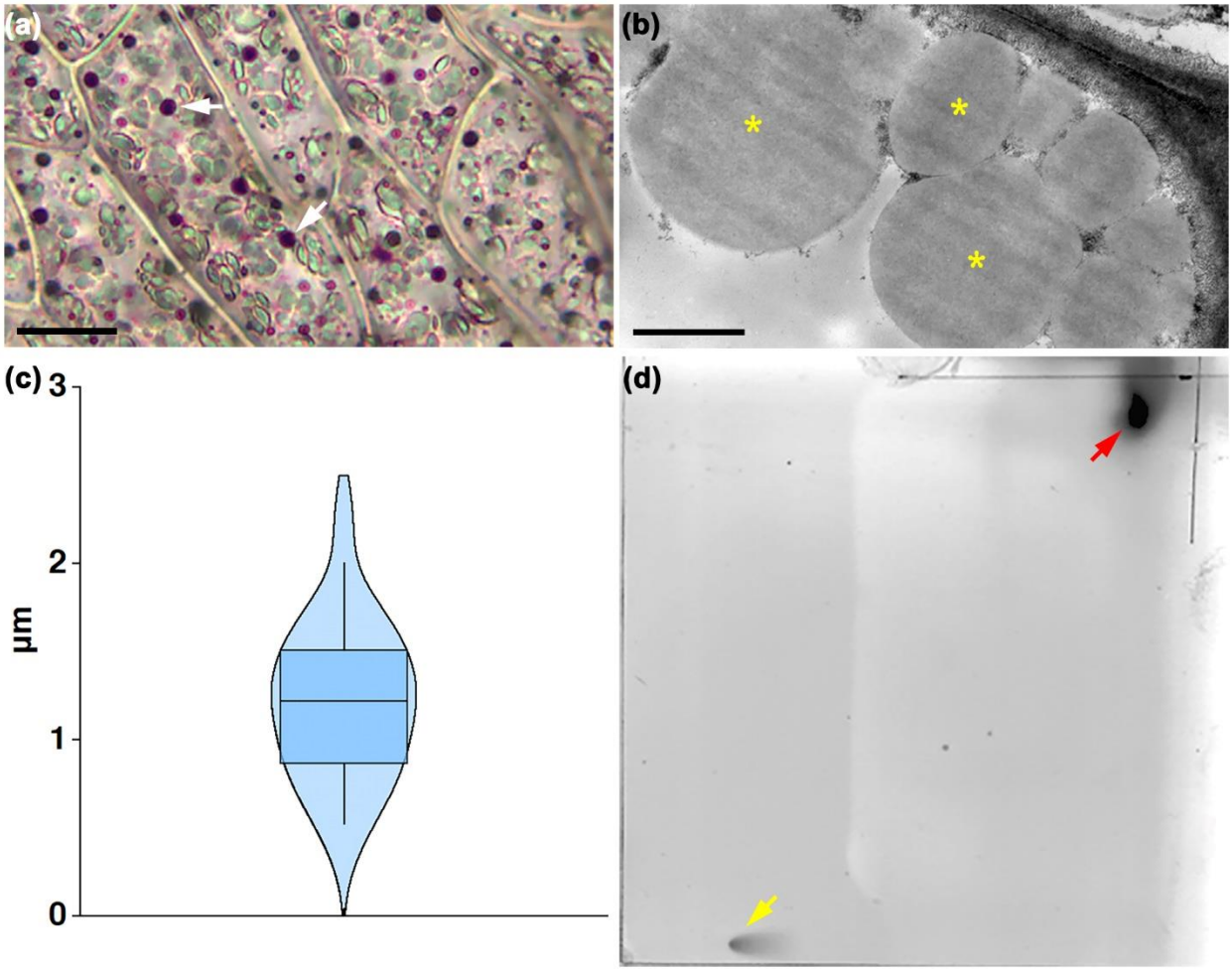
Data S1 T. SEM of *T. lepidozioides* capsules

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



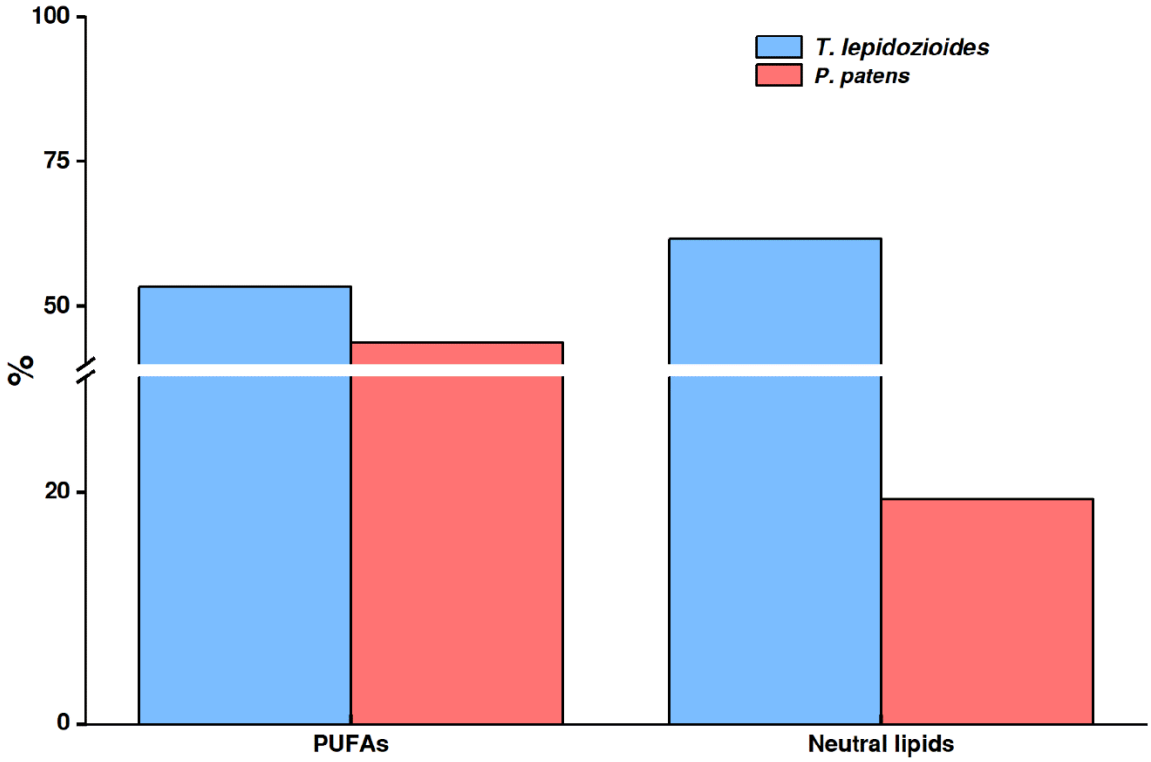
Data S1 U. TEM of cells before and after UV-B radiation in *P. patens* and *T. lepidozioides*

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



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

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



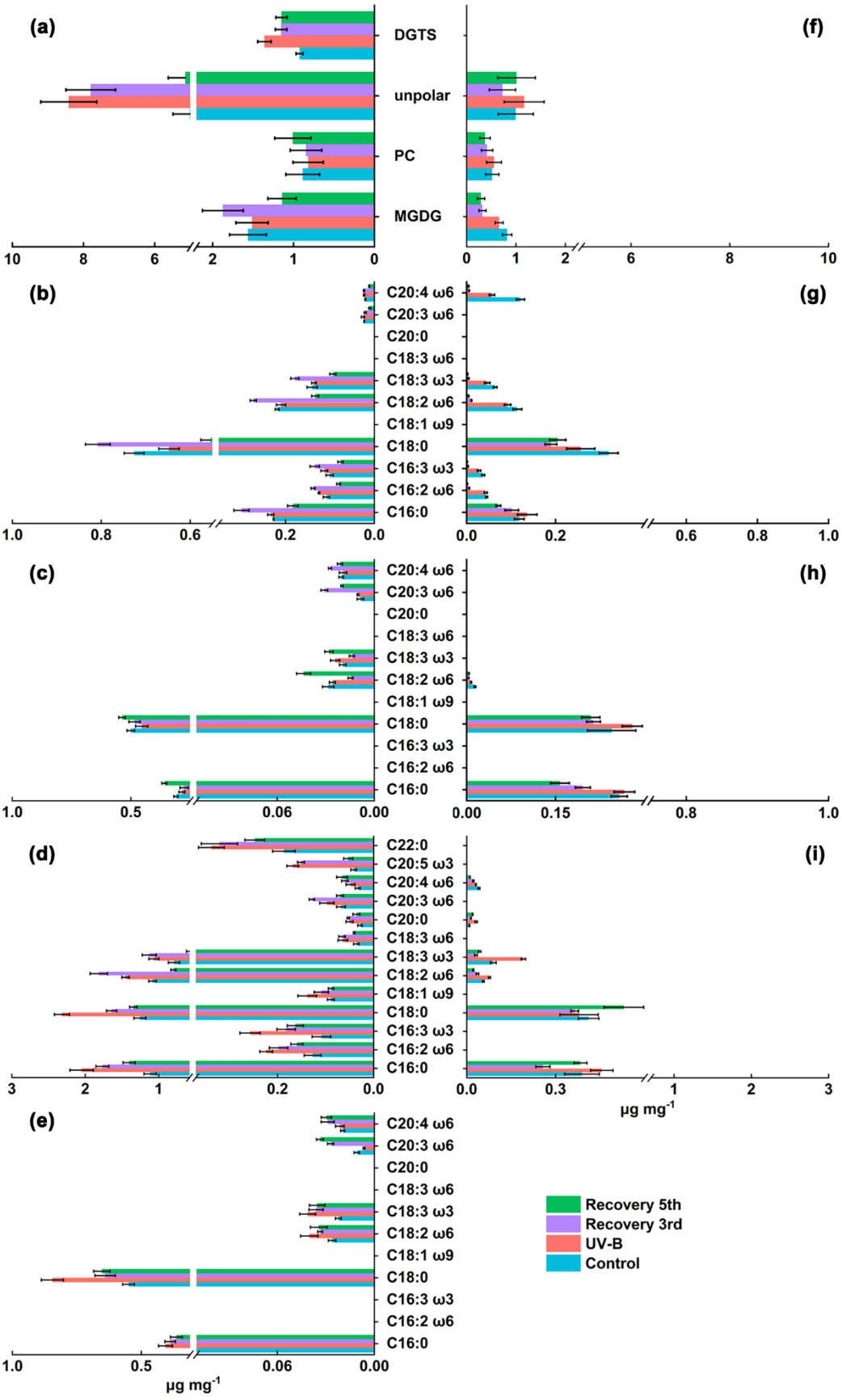
272

273

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

275 *T. lepidozioides* 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.

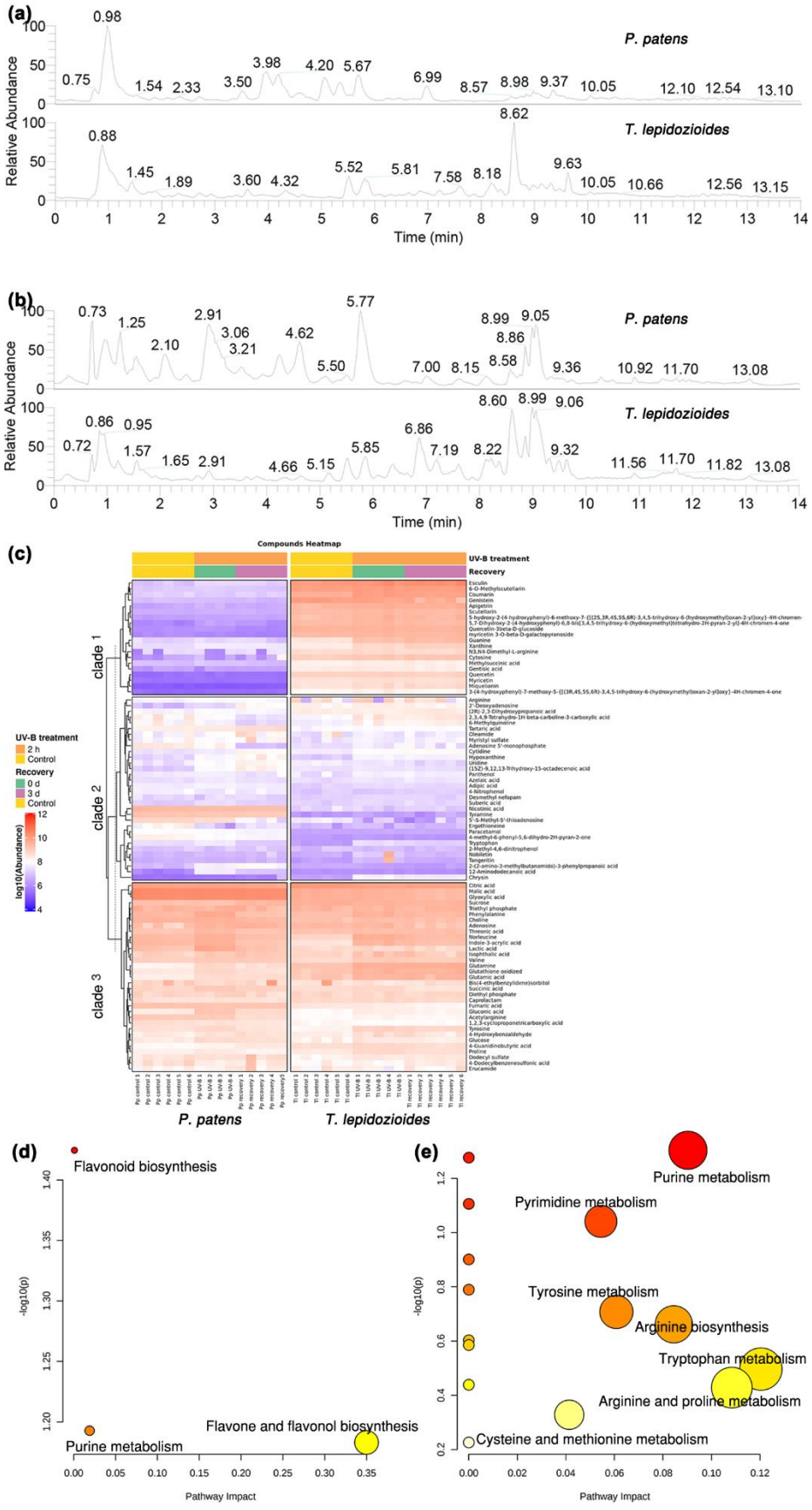
277


(a)
DGTS
unpolar
PC
MGDG
10 8 6 2 1 0
(f)
0 1 2 6 8 10
(b)
C20:4 ω6
C20:3 ω6
C20:0
C18:3 ω6
C18:3 ω3
C18:2 ω6
C18:1 ω9
C18:0
C16:3 ω3
C16:2 ω6
C16:0
1.0 0.8 0.6 0.2 0.0
(g)
0.0 0.2 0.6 0.8 1.0
(c)
1.0 0.5 0.06 0.00
(h)
0.00 0.15 0.8 1.0
(d)
C22:0
C20:5 ω3
3 2 1 0.2 0.0
(i)
0.0 0.3 1 2 3
μg mg-1
(e)
1.0 0.5 0.06 0.00
μg mg-1
Recovery 5th
Recovery 3rd
UV-B
Control

278

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

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


(a)
P. patens
T. lepidozioides
Relative Abundance
Time (min)
(b)
P. patens
T. lepidozioides
Relative Abundance
Time (min)
(c)
Compounds Heatmap
UV-B treatment
Recovery
clade 1
clade 2
clade 3
UV-B treatment
2 h
Control
Recovery
0 d
3 d
Control
log10(Abundance)
P. patens
T. lepidozioides
(d)
Flavonoid biosynthesis
Flavone and flavonol biosynthesis
Purine metabolism
-log10(p)
Pathway Impact
(e)
Purine metabolism
Pyrimidine metabolism
Tyrosine metabolism
Arginine biosynthesis
Tryptophan metabolism
Arginine and proline metabolism
Cysteine and methionine metabolism
-log10(p)
Pathway Impact

292

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

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

SUPPLEMENTAL REFERENCES

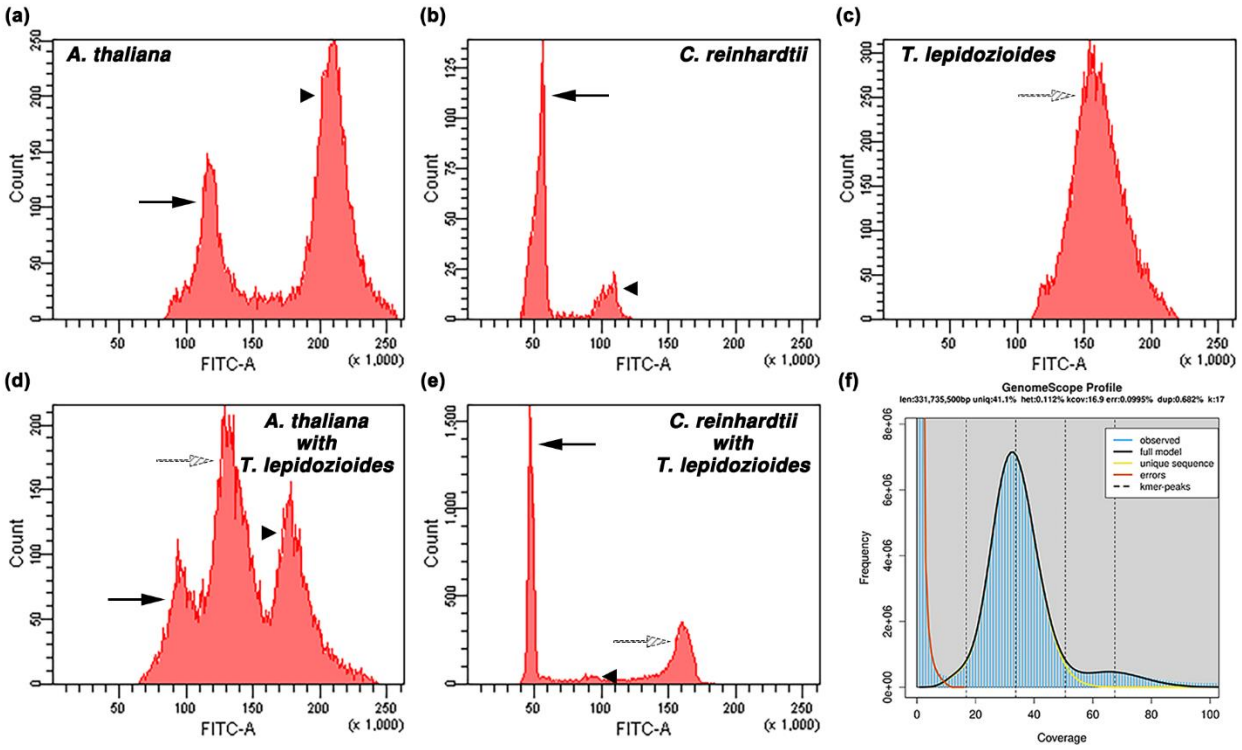
S1. Liu, C.R., Newell, G., and White, M. (2016). On the selection of thresholds for predicting species occurrence with presence-only data. Ecol. Evol. *6*, 337–348. 10.1002/ece3.1878.

S2. Chater, C.C., Caine, R.S., Tomek, M., Wallace, S., Kamisugi, Y., Cuming, A.C., Lang, D., MacAlister, C.A., Casson, S., Bergmann, D.C., et al. (2016). Origin and function of stomata in the moss *Physcomitrella patens*. Nat. Plants *2*, 16179, 16179. 10.1038/nplants.2016.179.

S3. Leventis, P.A., and Grinstein, S. (2010). The distribution and function of phosphatidylserine in cellular membranes. In Annual Review of Biophysics, Vol 39, D.C. Rees, K.A. Dill, and J.R. Williamson, eds. (Annual Reviews), pp. 407–427. 10.1146/annurev.biophys.093008.131234.

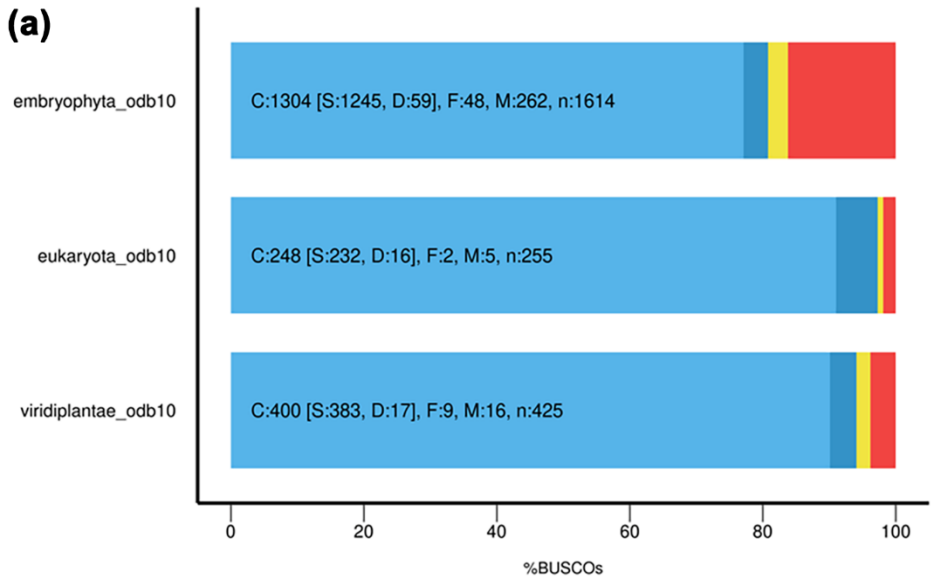
1 **Supplement data 2**

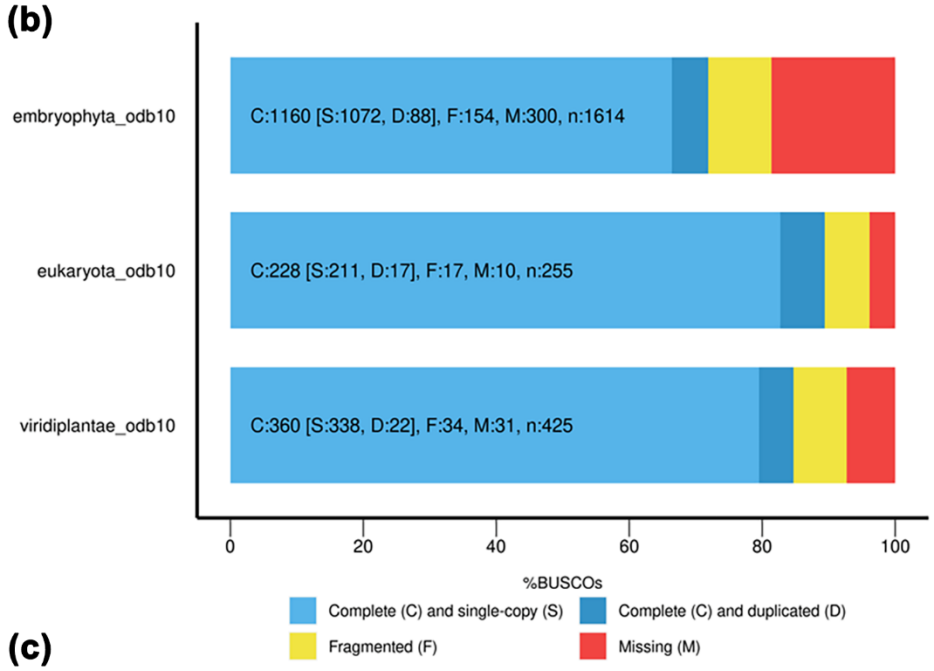
2

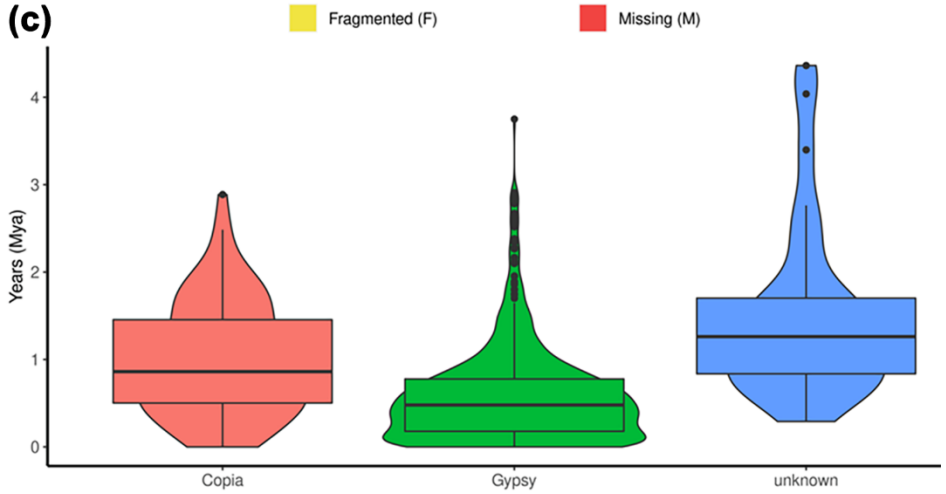


Data S2 A. Genome size estimations of *T. lepidozioides*

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

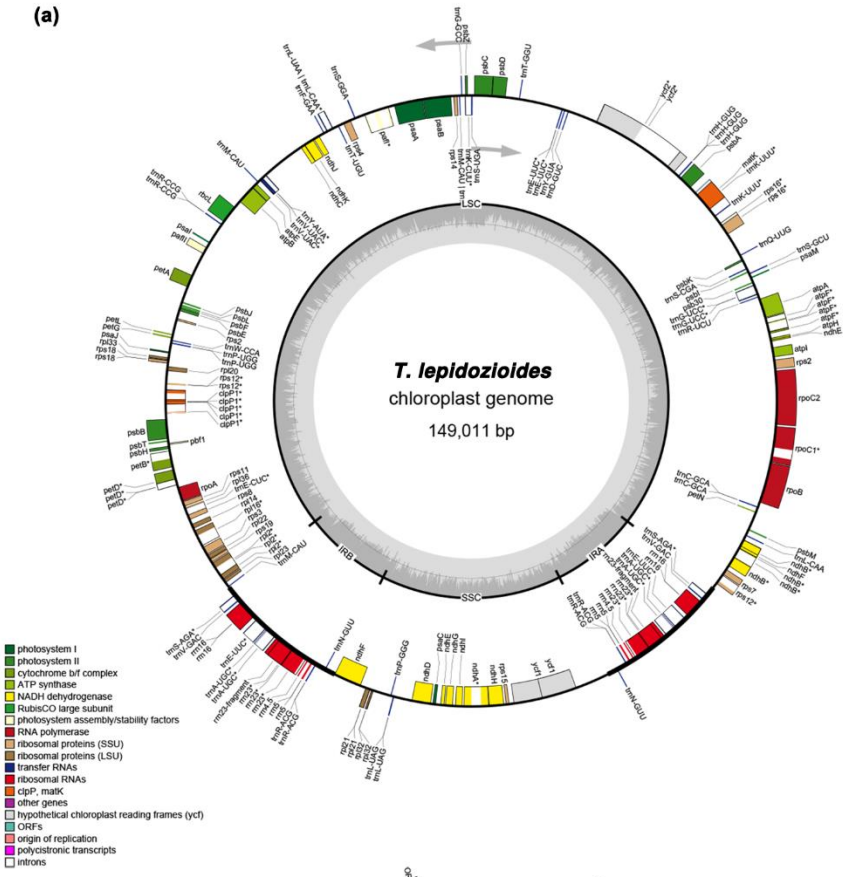

(a)
embryophyta_odb10
C:1304 [S:1245, D:59], F:48, M:262, n:1614
eukaryota_odb10
C:248 [S:232, D:16], F:2, M:5, n:255
viridiplantae_odb10
C:400 [S:383, D:17], F:9, M:16, n:425
0
20
40
60
80
100
%BUSCOs

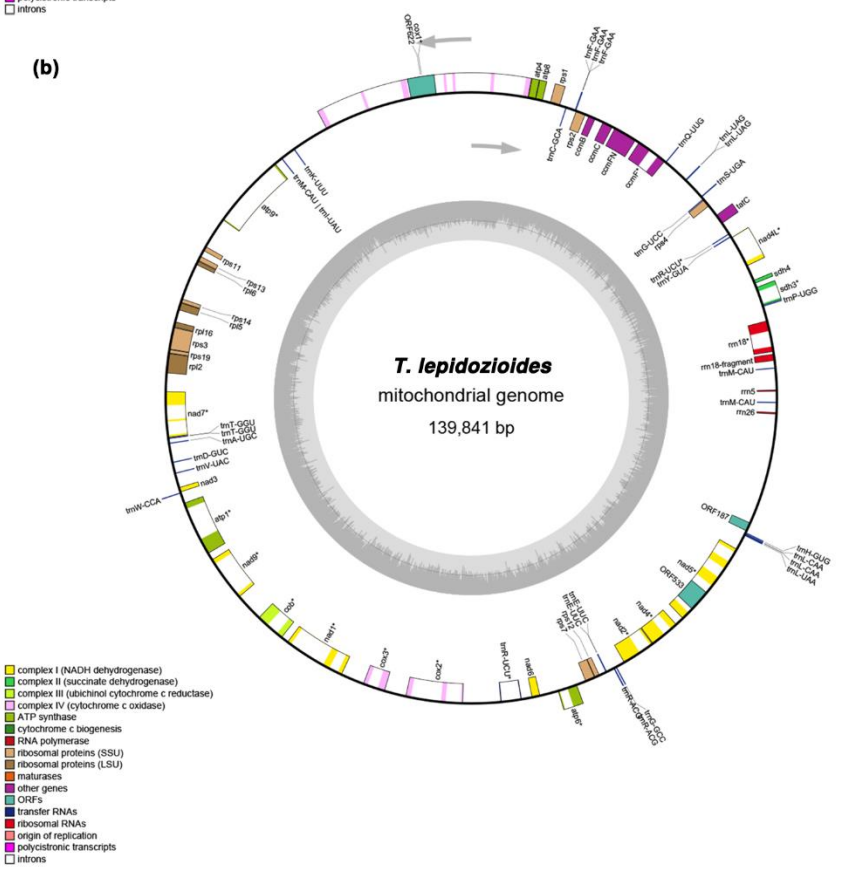

(b)
embryophyta_odb10
C:1160 [S:1072, D:88], F:154, M:300, n:1614
eukaryota_odb10
C:228 [S:211, D:17], F:17, M:10, n:255
viridiplantae_odb10
C:360 [S:338, D:22], F:34, M:31, n:425
0
20
40
60
80
100
%BUSCOs
Complete (C) and single-copy (S)
Complete (C) and duplicated (D)
Fragmented (F)
Missing (M)


(c)
Years (Mya)
0
1
2
3
4
Copia
Gypsy
unknown

Data S2 B. Genome features of *T. lepidozioides*

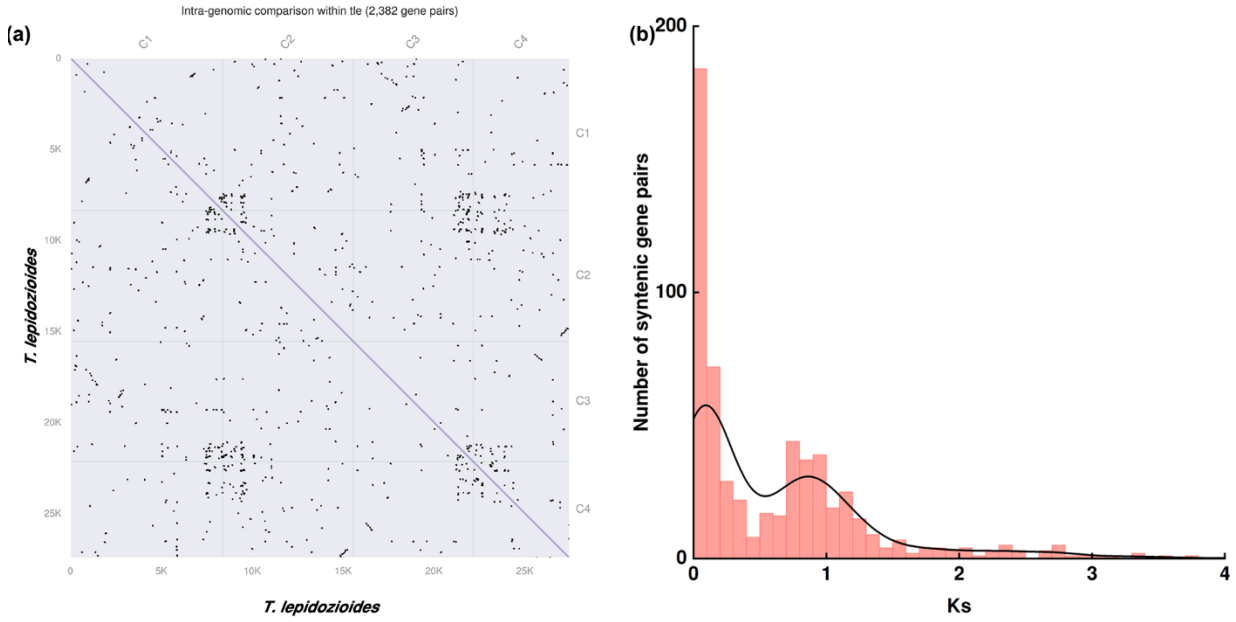
BUSCO assessment of assembly (**a**) and annotation (**b**). The latest all eukaryotic organism gene set eukaryote odb10, all plants gene set viridiplantae odb10 and all land plants gene set embryophyta odb10, were used as the reference, respectively, and Arabidopsis was used in the -sp option, 97.3%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


(a)
T. lepidozioides
chloroplast genome
149,011 bp
LSC
IRB
SSC
IRA
photosystem I
photosystem II
cytochrome b/f complex
ATP synthase
NADH dehydrogenase
RubisCO large subunit
photosystem assembly/stability factors
RNA polymerase
ribosomal proteins (SSU)
ribosomal proteins (LSU)
transfer RNAs
ribosomal RNAs
clpP, matK
other genes
hypothetical chloroplast reading frames (ycf)
ORFs
origin of replication
polycistronic transcripts
introns


(b)
T. lepidozioides
mitochondrial genome
139,841 bp
complex I (NADH dehydrogenase)
complex II (succinate dehydrogenase)
complex III (ubichinol cytochrome c reductase)
complex IV (cytochrome c oxidase)
ATP synthase
cytochrome c biogenesis
RNA polymerase
ribosomal proteins (SSU)
ribosomal proteins (LSU)
maturases
other genes
ORFs
transfer RNAs
ribosomal RNAs
origin of replication
polycistronic transcripts
introns

Data S2 C. Organellar genomes of *T. lepidozioides*

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



47

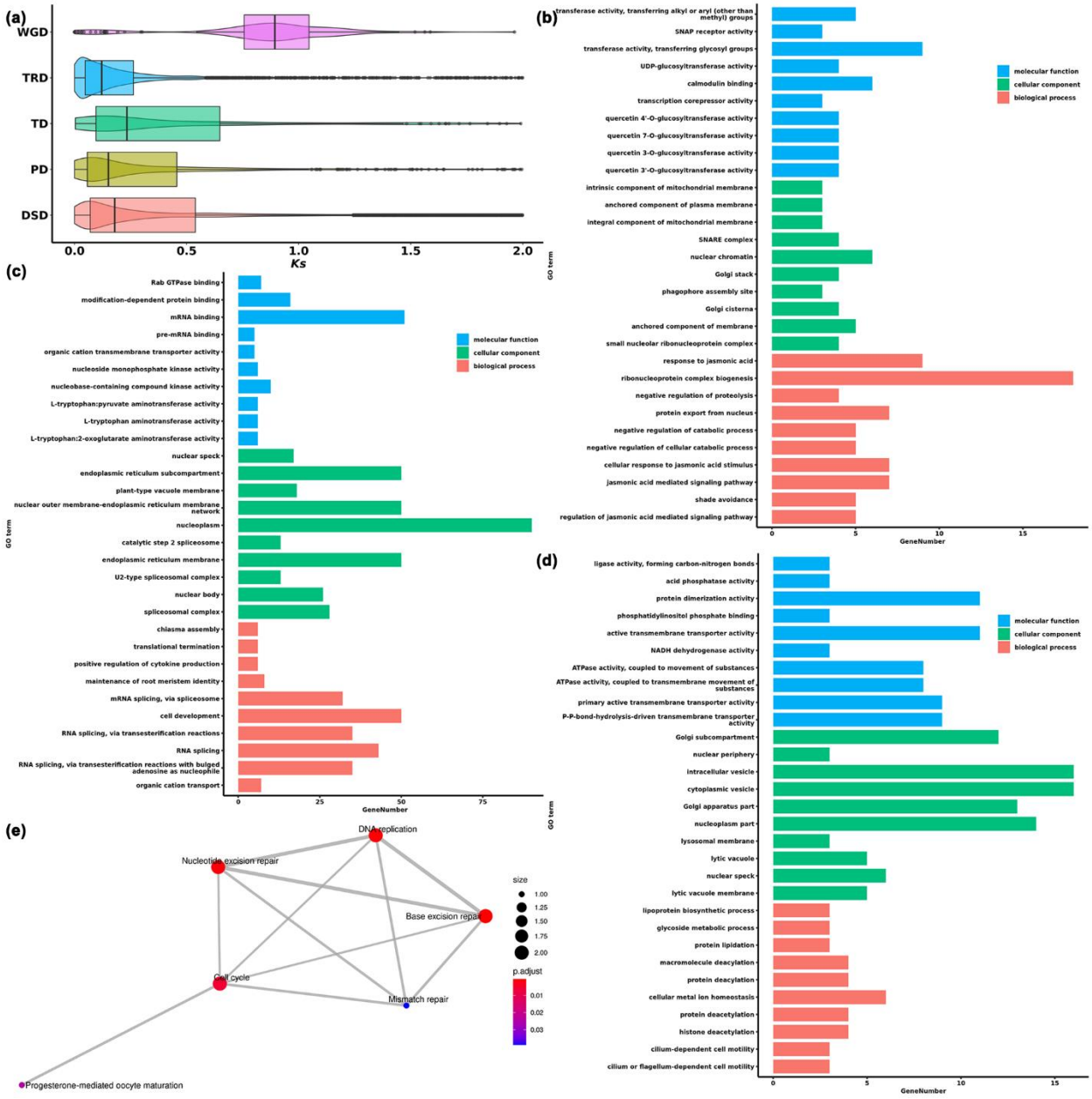
48

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

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

51 in *T. lepidozioides*.

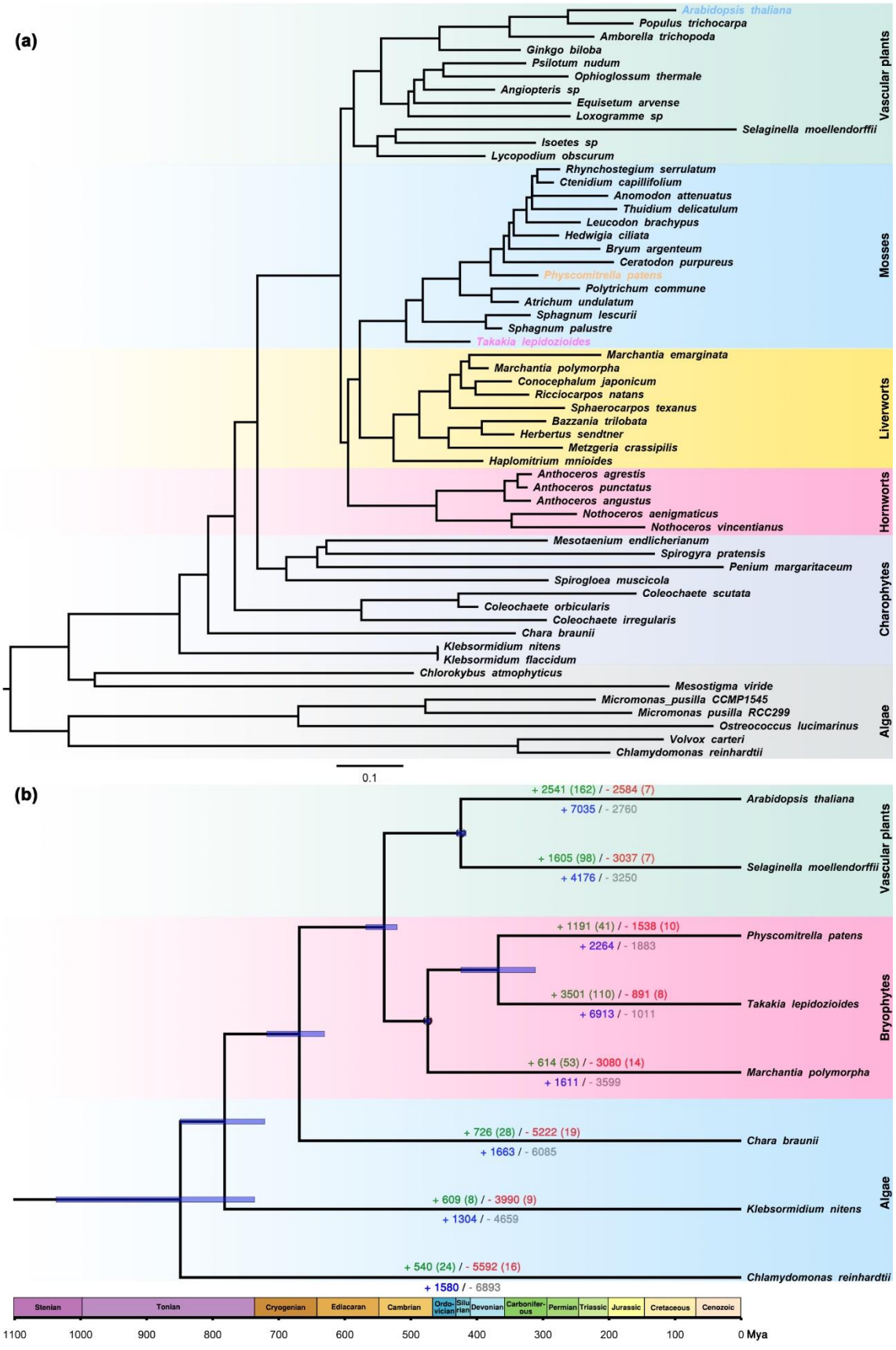
52



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

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

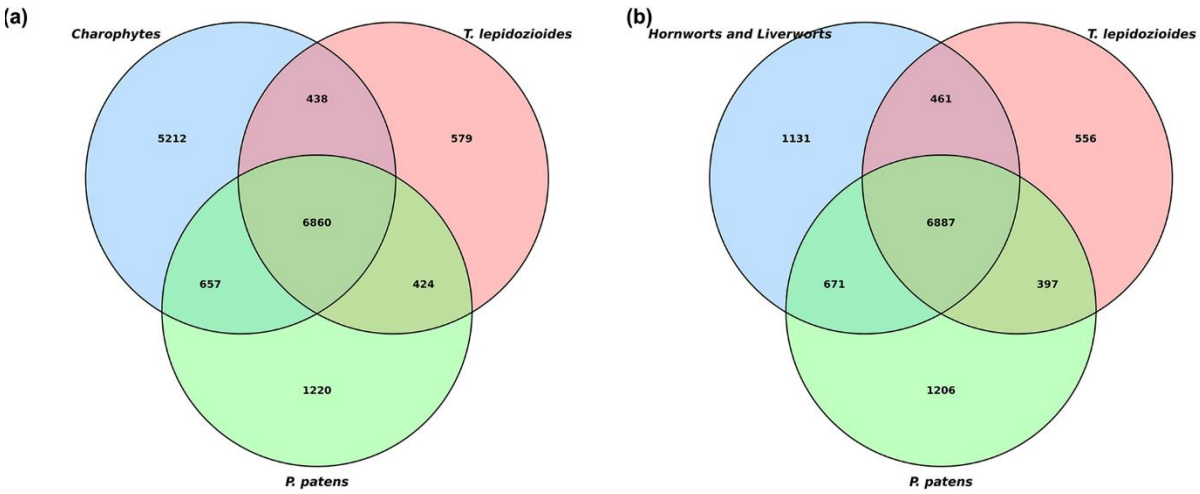
ontology (GO) enrichment of PD in *T. lepidozioides* (**c**) GO enrichment of TRD in *T. lepidozioides*. (**d**) GO enrichment of WGD in *T. lepidozioides*. (**e**) Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment of singleton genes in *T. lepidozioides*.


(a)
Arabidopsis thaliana
Populus trichocarpa
Amborella trichopoda
Ginkgo biloba
Psilotum nudum
Ophioglossum thermale
Angiopteris sp
Equisetum arvense
Loxogramme sp
Selaginella moellendorffii
Isoetes sp
Lycopodium obscurum
Vascular plants
Rhynchostegium serrulatum
Ctenidium capillifolium
Anomodon attenuatus
Thuidium delicatulum
Leucodon brachypus
Hedwigia ciliata
Bryum argenteum
Ceratodon purpureus
Physcomitrella patens
Polytrichum commune
Atrichum undulatum
Sphagnum lescurii
Sphagnum palustre
Takakia lepidozioides
Mosses
Marchantia emarginata
Marchantia polymorpha
Conocephalum japonicum
Ricciocarpos natans
Sphaerocarpos texanus
Bazzania trilobata
Herbertus sendtner
Metzgeria crassipilis
Haplomitrium mnioides
Liverworts
Anthoceros agrestis
Anthoceros punctatus
Anthoceros angustus
Nothoceros aenigmaticus
Nothoceros vincentianus
Hornworts
Mesotaenium endlicherianum
Spirogyra pratensis
Penium margaritaceum
Spirogloea muscicola
Coleochaete scutata
Coleochaete orbicularis
Coleochaete irregularis
Chara braunii
Klebsormidium nitens
Klebsormidum flaccidum
Charophytes
Chlorokybus atmophyticus
Mesostigma viride
Micromonas_pusilla CCMP1545
Micromonas pusilla RCC299
Ostreococcus lucimarinus
Volvox carteri
Chlamydomonas reinhardtii
Algae
0.1
(b)
+ 2541 (162) / - 2584 (7)
+ 7035 / - 2760
Arabidopsis thaliana
+ 1605 (98) / - 3037 (7)
+ 4176 / - 3250
Selaginella moellendorffii
Vascular plants
+ 1191 (41) / - 1538 (10)
+ 2264 / - 1883
Physcomitrella patens
+ 3501 (110) / - 891 (8)
+ 6913 / - 1011
Takakia lepidozioides
+ 614 (53) / - 3080 (14)
+ 1611 / - 3599
Marchantia polymorpha
Bryophytes
+ 726 (28) / - 5222 (19)
+ 1663 / - 6085
Chara braunii
+ 609 (8) / - 3990 (9)
+ 1304 / - 4659
Klebsormidium nitens
+ 540 (24) / - 5592 (16)
+ 1580 / - 6893
Chlamydomonas reinhardtii
Algae
Stenian
Tonian
Cryogenian
Ediacaran
Cambrian
Ordo-vician
Silu-rian
Devonian
Carbonifer-ous
Permian
Triassic
Jurassic
Cretaceous
Cenozoic
1100
1000
900
800
700
600
500
400
300
200
100
0 Mya

66

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

(**a**) Phylogenetic tree of 56 plant species based on 818 low-copy orthologs amino acid sequence. The tree of 56 plant species inferred from five plant clades reveals *T. lepidozioides* as the sister group to all other mosses, rather than to liverworts or to the LCA of mosses and liverworts. (**b**) *T. lepidozioides* time tree with gene family expansion/contraction analysis compared to representatives of seven other plant lineages, coloured numbers above branches show gene family expansions (+) and contractions (-), coloured numbers under branches show gene gains (+) and losses (-), light purple bars show 95% confidence interval lengths, black dots on nodes mark fossil calibration points.



78

79

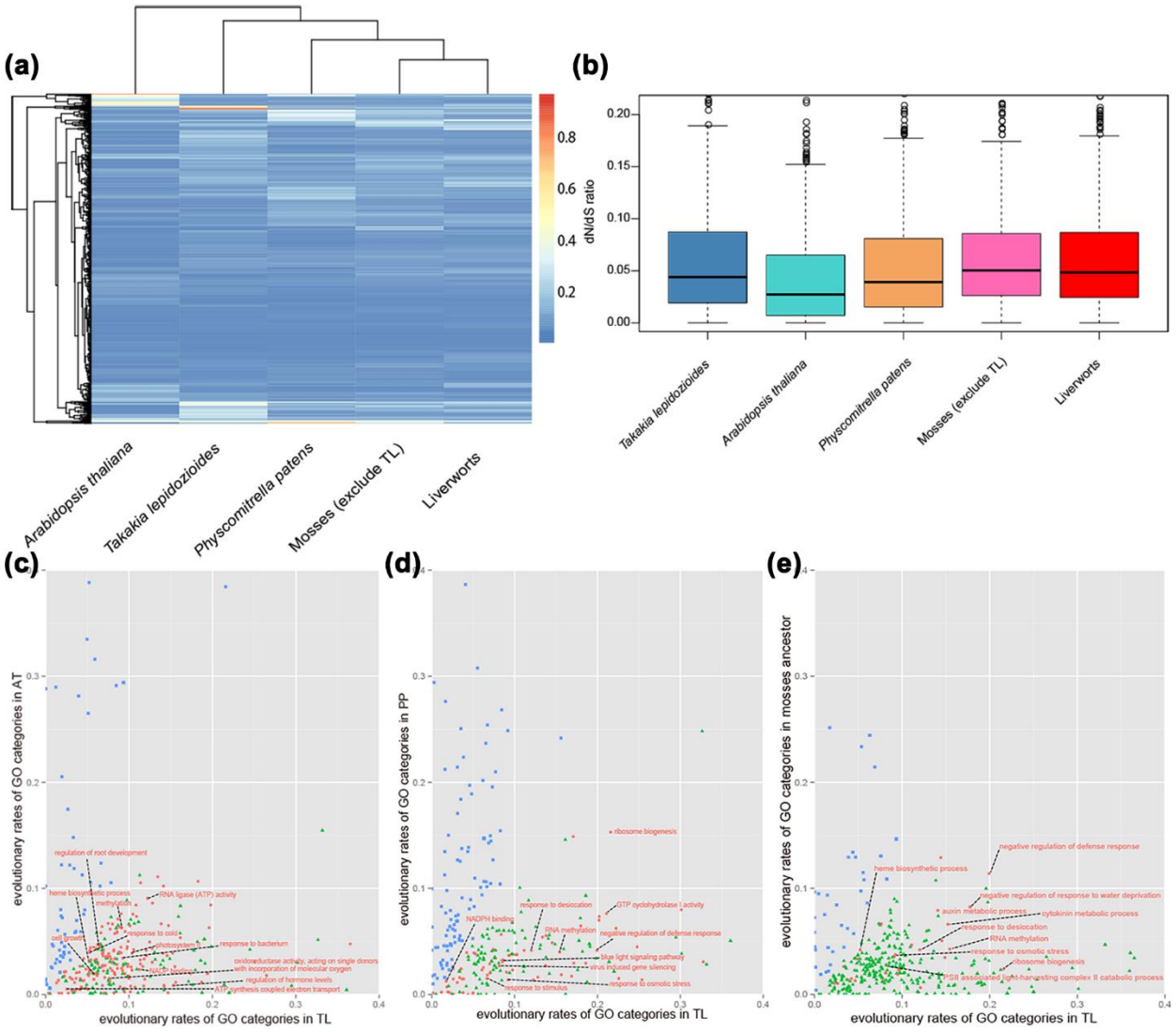
80 **Data S2 G. Gene families characters of *T. lepidozioides***

81 (**a**) Gene families of *T. lepidozioides*, *P. patens* and *Charophytes* (the union of *K. nitens*, *C. braunii*,

82 *M. endlicherianum* and *S. muscicola* sets). (**b**) Gene families of *T. lepidozioides*, *P. patens* and

83 *hornworts and liverworts* (the union of *A. angustus* and *M. polymorpha* sets).

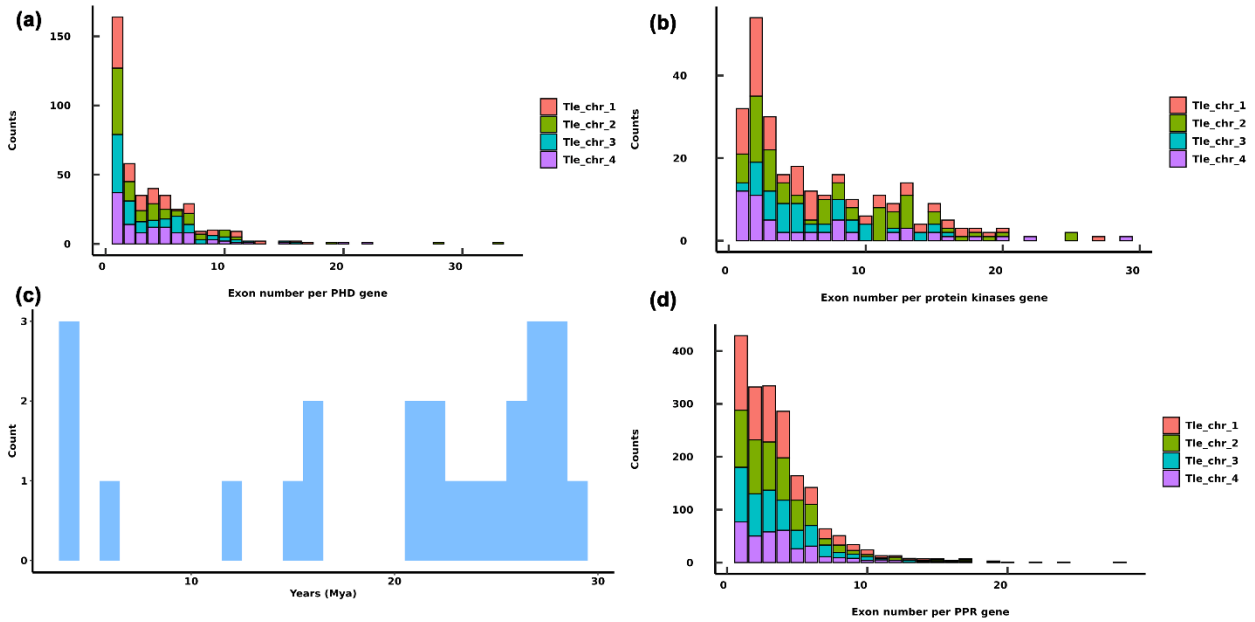
84



Data S2 H. Positive selection and fast evolution

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

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

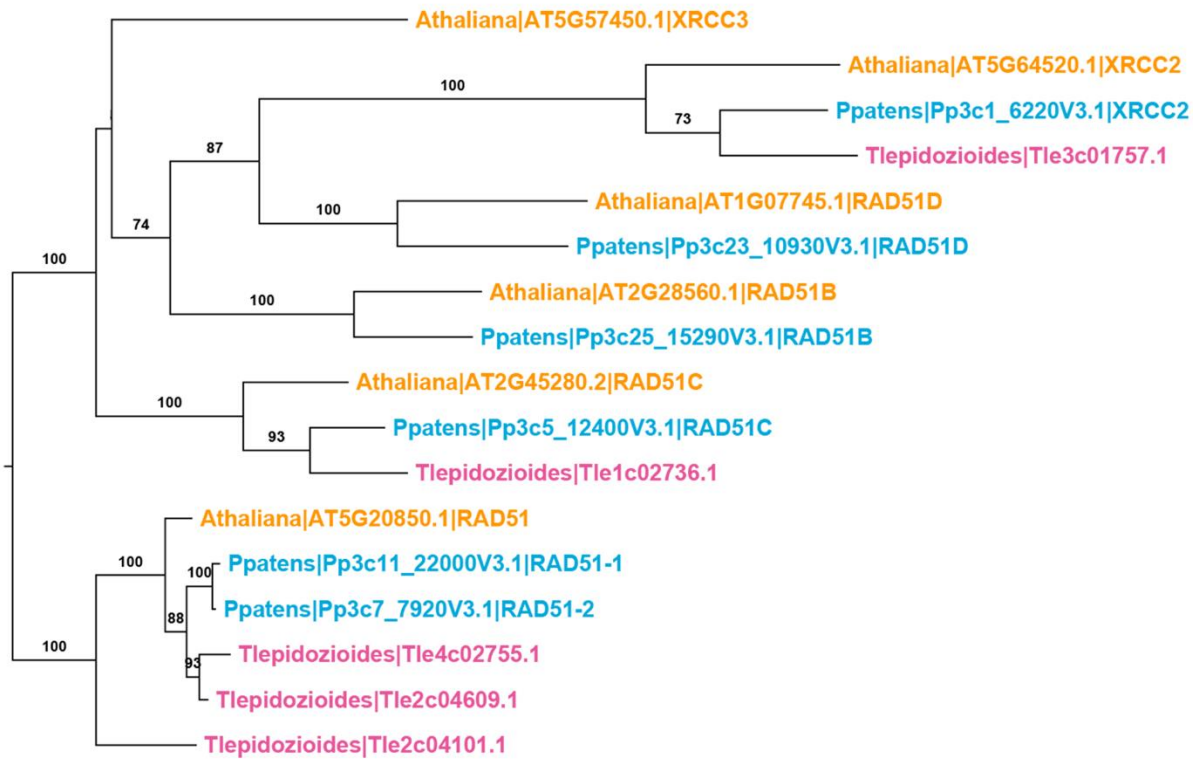


106
107

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

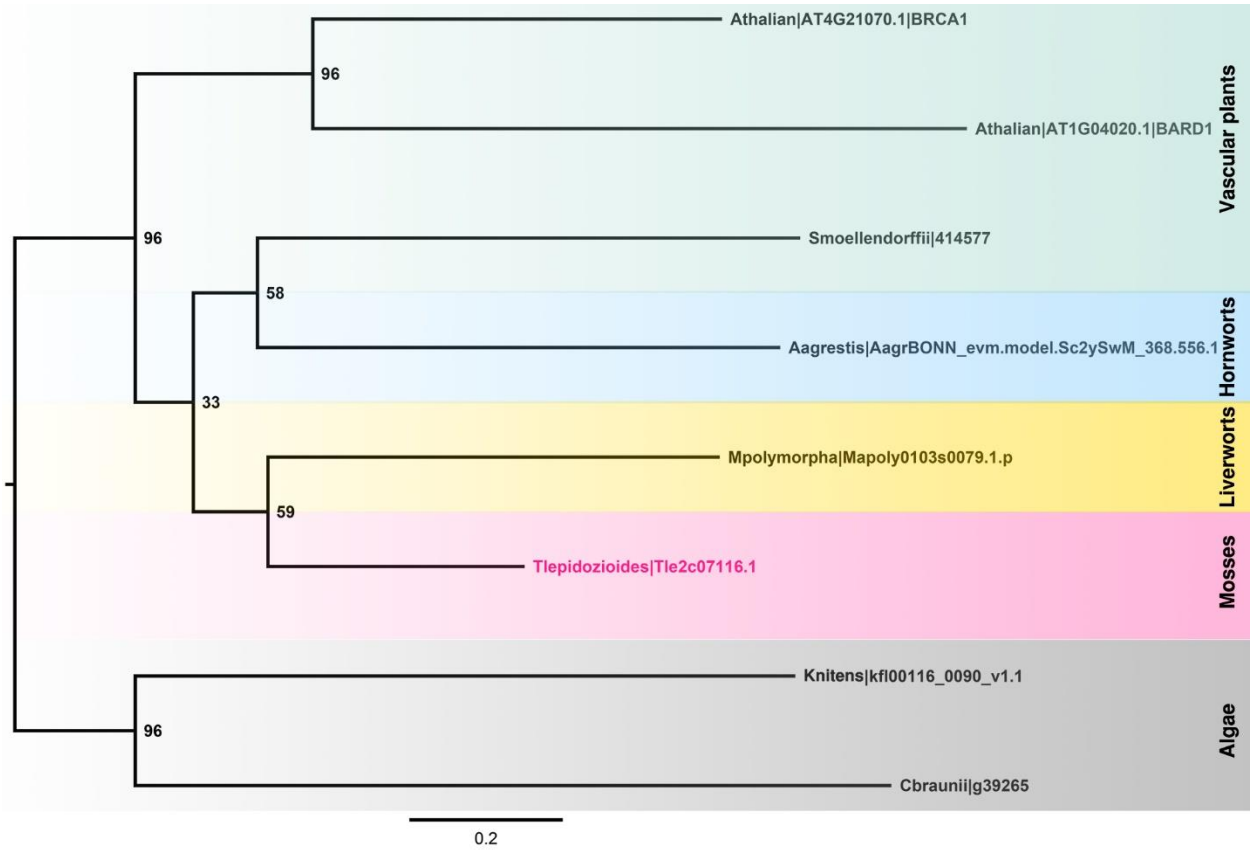
109 (**a**) Statistics of exon numbers in each PHD. (**b**) Statistics of exon numbers in each protein kinase.
110 (**c**) Numbers of protein kinases which are younger than 30 Mya. (**d**) Statistics of exon numbers in
111 each PPR.

112



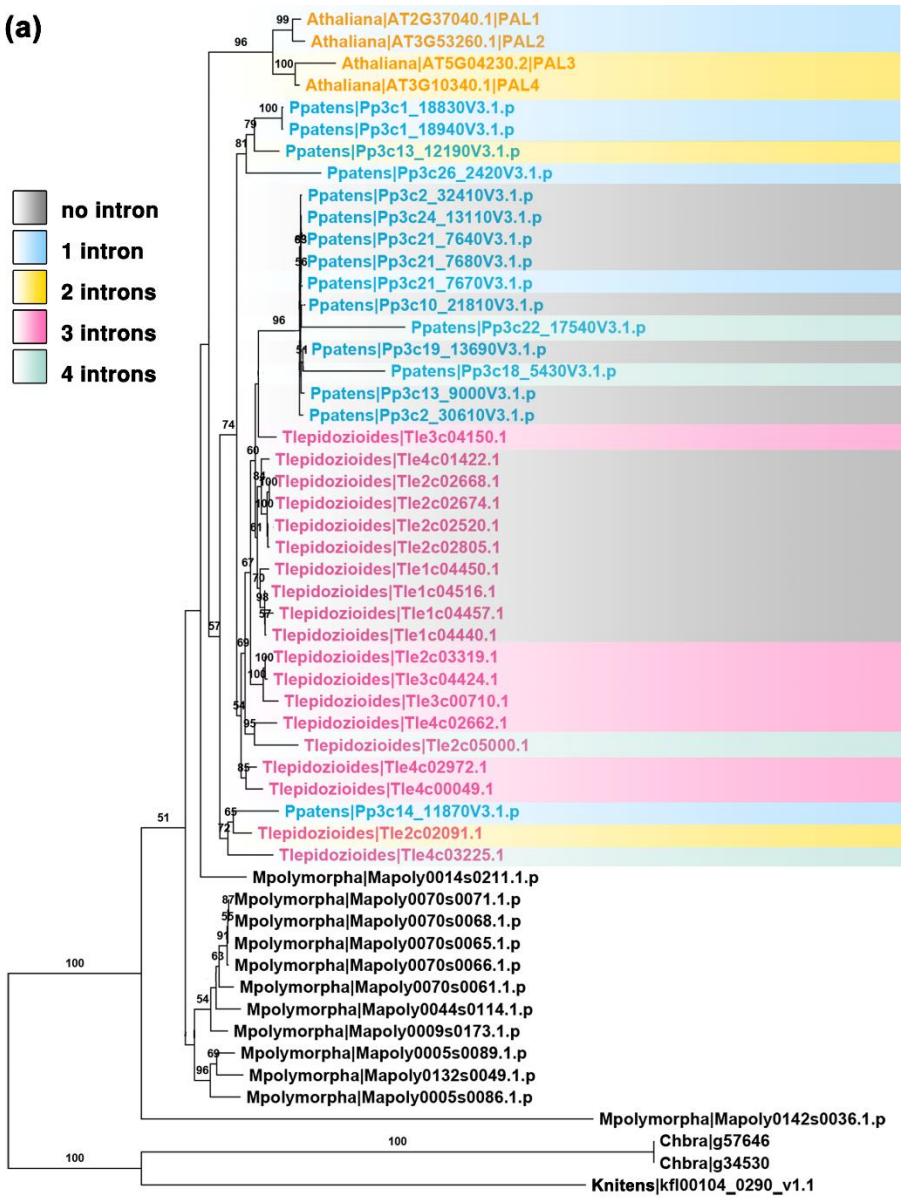
Data S2 J. Phylogenetic trees of RAD51

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



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

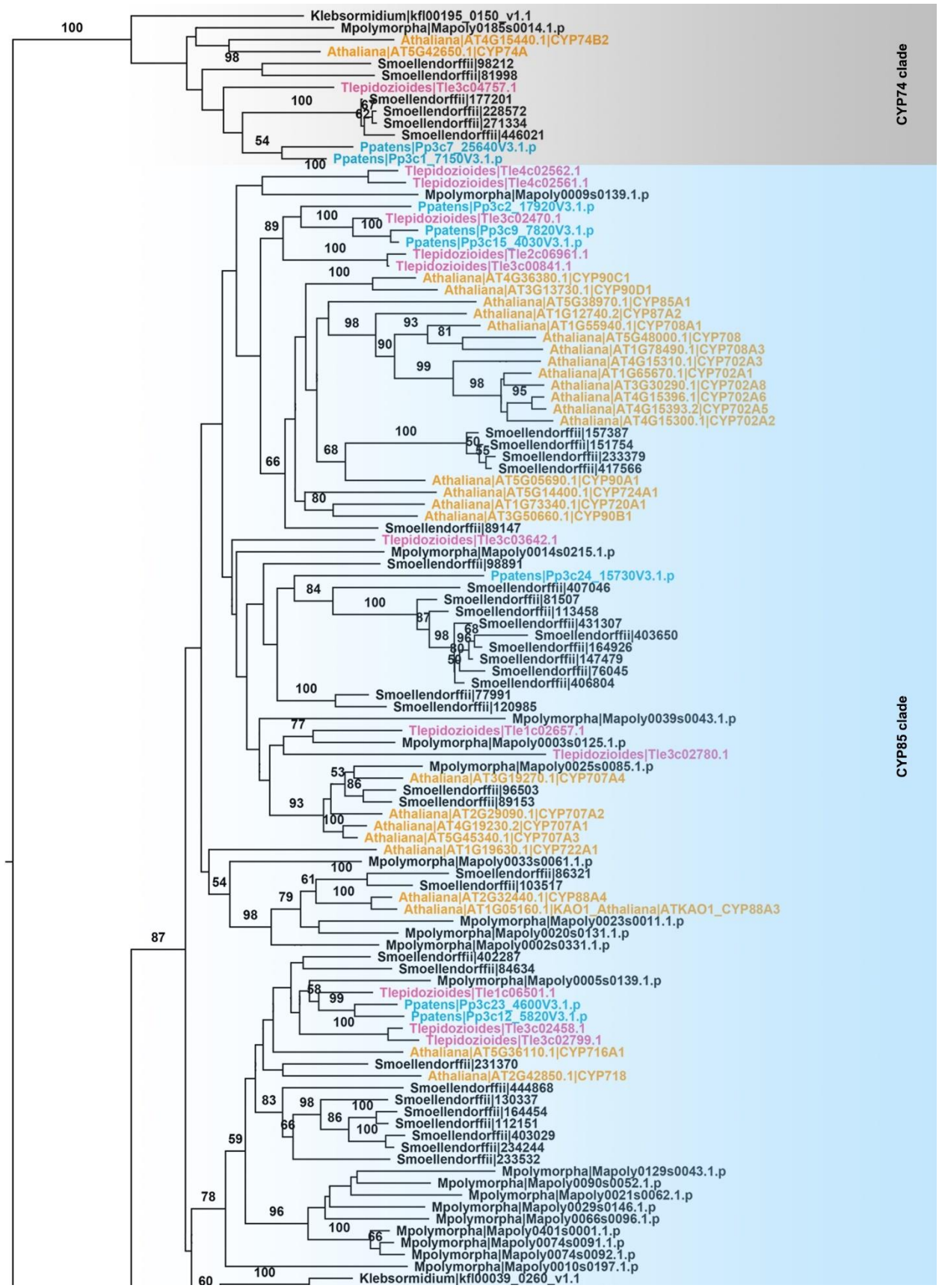
BRCA1 (AT4G21070) were found using blastp with identity cutoff 50%, e-value cutoff 1×10^{-20}. Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates. Numbered nodes show the percentage of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


(a)
no intron
1 intron
2 introns
3 introns
4 introns
Athaliana|AT2G37040.1|PAL1
Athaliana|AT3G53260.1|PAL2
Athaliana|AT5G04230.2|PAL3
Athaliana|AT3G10340.1|PAL4
Ppatens|Pp3c1_18830V3.1.p
Ppatens|Pp3c1_18940V3.1.p
Ppatens|Pp3c13_12190V3.1.p
Ppatens|Pp3c26_2420V3.1.p
Ppatens|Pp3c2_32410V3.1.p
Ppatens|Pp3c24_13110V3.1.p
Ppatens|Pp3c21_7640V3.1.p
Ppatens|Pp3c21_7680V3.1.p
Ppatens|Pp3c21_7670V3.1.p
Ppatens|Pp3c10_21810V3.1.p
Ppatens|Pp3c22_17540V3.1.p
Ppatens|Pp3c19_13690V3.1.p
Ppatens|Pp3c18_5430V3.1.p
Ppatens|Pp3c13_9000V3.1.p
Ppatens|Pp3c2_30610V3.1.p
Tlepidozioides|Tle3c04150.1
Tlepidozioides|Tle4c01422.1
Tlepidozioides|Tle2c02668.1
Tlepidozioides|Tle2c02674.1
Tlepidozioides|Tle2c02520.1
Tlepidozioides|Tle2c02805.1
Tlepidozioides|Tle1c04450.1
Tlepidozioides|Tle1c04516.1
Tlepidozioides|Tle1c04457.1
Tlepidozioides|Tle1c04440.1
Tlepidozioides|Tle2c03319.1
Tlepidozioides|Tle3c04424.1
Tlepidozioides|Tle3c00710.1
Tlepidozioides|Tle4c02662.1
Tlepidozioides|Tle2c05000.1
Tlepidozioides|Tle4c02972.1
Tlepidozioides|Tle4c00049.1
Ppatens|Pp3c14_11870V3.1.p
Tlepidozioides|Tle2c02091.1
Tlepidozioides|Tle4c03225.1
Mpolymorpha|Mapoly0014s0211.1.p
Mpolymorpha|Mapoly0070s0071.1.p
Mpolymorpha|Mapoly0070s0068.1.p
Mpolymorpha|Mapoly0070s0065.1.p
Mpolymorpha|Mapoly0070s0066.1.p
Mpolymorpha|Mapoly0070s0061.1.p
Mpolymorpha|Mapoly0044s0114.1.p
Mpolymorpha|Mapoly0009s0173.1.p
Mpolymorpha|Mapoly0005s0089.1.p
Mpolymorpha|Mapoly0132s0049.1.p
Mpolymorpha|Mapoly0005s0086.1.p
Mpolymorpha|Mapoly0142s0036.1.p
Chbra|g57646
Chbra|g34530
Knitens|kfl00104_0290_v1.1
(b)
PALs Gene Expression Heatmap
UV-B treatment
Recovery
log2(TPM)
10
8
6
4
2
0
UV-B treatment
2 h
Control
Recovery
0 h
12 h
24 h
6 h
Control

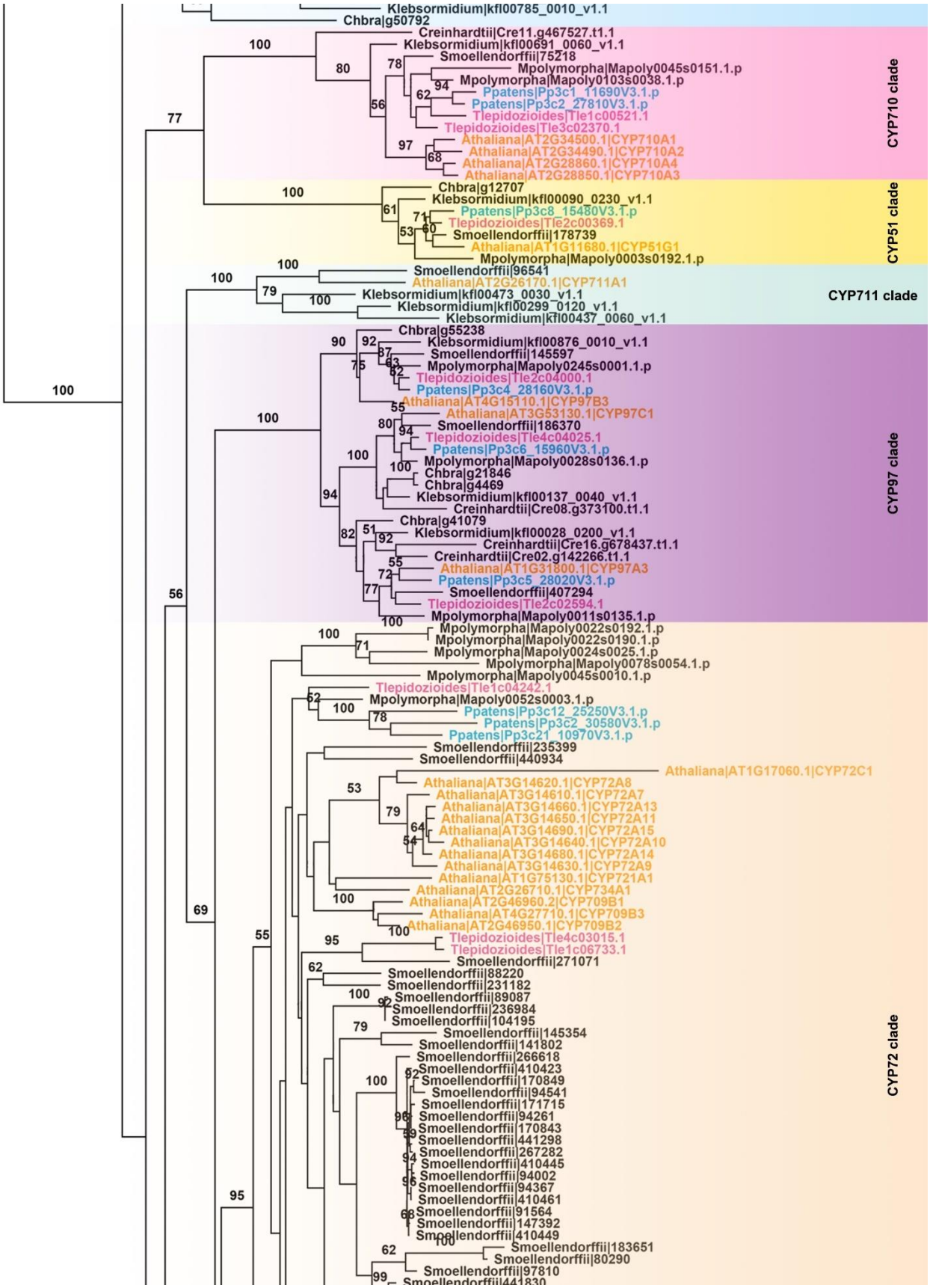
130

Data S2 L. PAL features in *T. lepidozioides*

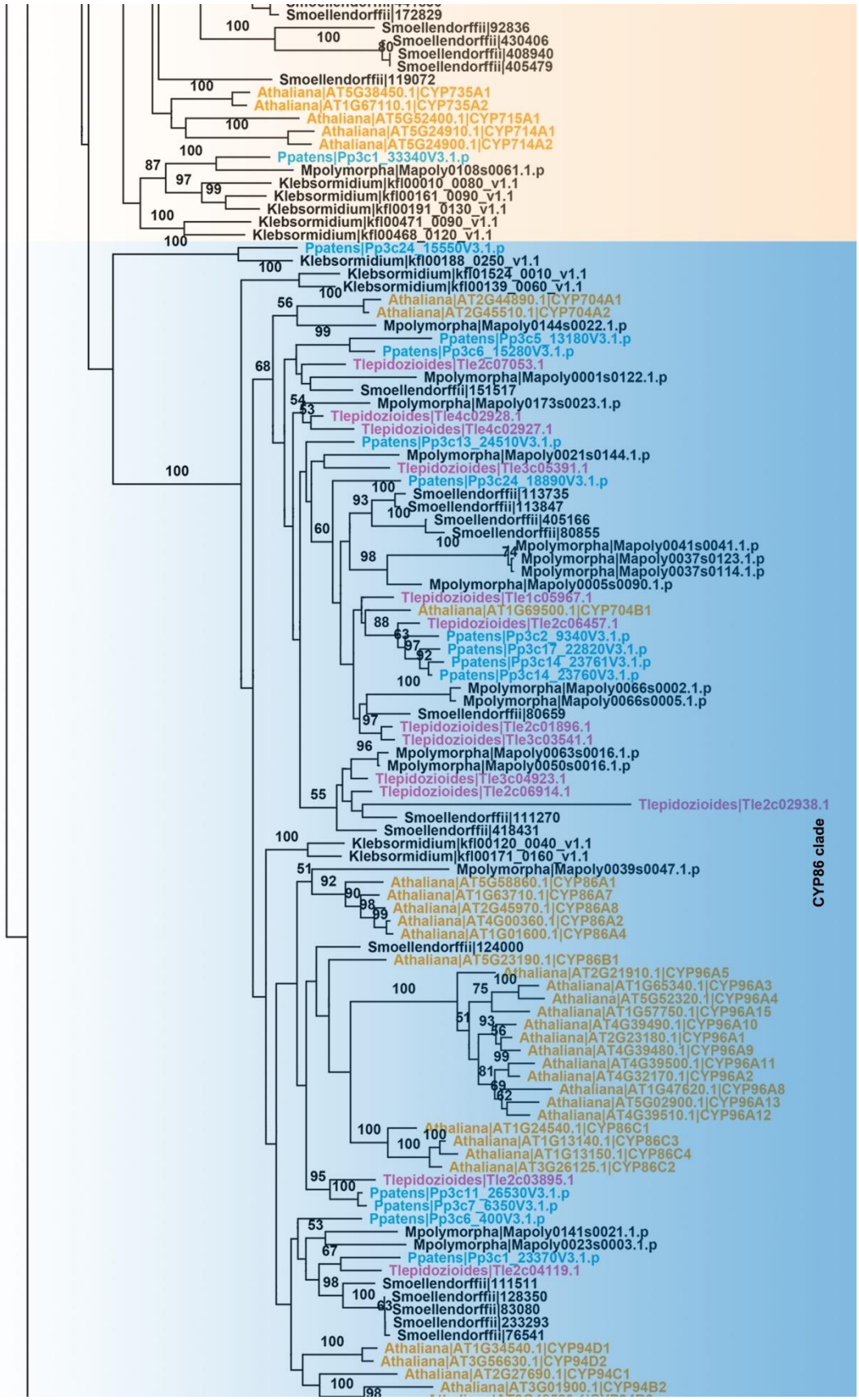
(**a**) Phylogenetic tree of PAL genes. Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates. Numbered nodes show the percentage of bootstraps which support the node, bootstraps with less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


CYP74 clade
CYP85 clade
Klebsormidium|kfl00195_0150_v1.1
Mpolymorpha|Mapoly0185s0014.1.p
Athaliana|AT4G15440.1|CYP74B2
Athaliana|AT5G42650.1|CYP74A
Smoellendorffii|98212
Smoellendorffii|81998
Tlepidozioides|Tle3c04757.1
Smoellendorffii|177201
Smoellendorffii|228572
Smoellendorffii|271334
Smoellendorffii|446021
Ppatens|Pp3c7_25640V3.1.p
Ppatens|Pp3c1_7150V3.1.p
Tlepidozioides|Tle4c02562.1
Tlepidozioides|Tle4c02561.1
Mpolymorpha|Mapoly0009s0139.1.p
Ppatens|Pp3c2_17920V3.1.p
Tlepidozioides|Tle3c02470.1
Ppatens|Pp3c9_7820V3.1.p
Ppatens|Pp3c15_4030V3.1.p
Tlepidozioides|Tle2c06961.1
Tlepidozioides|Tle3c00841.1
Athaliana|AT4G36380.1|CYP90C1
Athaliana|AT3G13730.1|CYP90D1
Athaliana|AT5G38970.1|CYP85A1
Athaliana|AT1G12740.2|CYP87A2
Athaliana|AT1G55940.1|CYP708A1
Athaliana|AT5G48000.1|CYP708
Athaliana|AT1G78490.1|CYP708A3
Athaliana|AT4G15310.1|CYP702A3
Athaliana|AT1G65670.1|CYP702A1
Athaliana|AT3G30290.1|CYP702A8
Athaliana|AT4G15396.1|CYP702A6
Athaliana|AT4G15393.2|CYP702A5
Athaliana|AT4G15300.1|CYP702A2
Smoellendorffii|157387
Smoellendorffii|151754
Smoellendorffii|233379
Smoellendorffii|417566
Athaliana|AT5G05690.1|CYP90A1
Athaliana|AT5G14400.1|CYP724A1
Athaliana|AT1G73340.1|CYP720A1
Athaliana|AT3G50660.1|CYP90B1
Smoellendorffii|89147
Tlepidozioides|Tle3c03642.1
Mpolymorpha|Mapoly0014s0215.1.p
Smoellendorffii|98891
Ppatens|Pp3c24_15730V3.1.p
Smoellendorffii|407046
Smoellendorffii|81507
Smoellendorffii|113458
Smoellendorffii|431307
Smoellendorffii|403650
Smoellendorffii|164926
Smoellendorffii|147479
Smoellendorffii|76045
Smoellendorffii|406804
Smoellendorffii|77991
Smoellendorffii|120985
Mpolymorpha|Mapoly0039s0043.1.p
Tlepidozioides|Tle1c02657.1
Mpolymorpha|Mapoly0003s0125.1.p
Tlepidozioides|Tle3c02780.1
Mpolymorpha|Mapoly0025s0085.1.p
Athaliana|AT3G19270.1|CYP707A4
Smoellendorffii|96503
Smoellendorffii|89153
Athaliana|AT2G29090.1|CYP707A2
Athaliana|AT4G19230.2|CYP707A1
Athaliana|AT5G45340.1|CYP707A3
Athaliana|AT1G19630.1|CYP722A1
Mpolymorpha|Mapoly0033s0061.1.p
Smoellendorffii|86321
Smoellendorffii|103517
Athaliana|AT2G32440.1|CYP88A4
Athaliana|AT1G05160.1|KAO1_Athaliana|ATKAO1_CYP88A3
Mpolymorpha|Mapoly0023s0011.1.p
Mpolymorpha|Mapoly0020s0131.1.p
Mpolymorpha|Mapoly0002s0331.1.p
Smoellendorffii|402287
Smoellendorffii|84634
Mpolymorpha|Mapoly0005s0139.1.p
Tlepidozioides|Tle1c06501.1
Ppatens|Pp3c23_4600V3.1.p
Ppatens|Pp3c12_5820V3.1.p
Tlepidozioides|Tle3c02458.1
Tlepidozioides|Tle3c02799.1
Athaliana|AT5G36110.1|CYP716A1
Smoellendorffii|231370
Athaliana|AT2G42850.1|CYP718
Smoellendorffii|444868
Smoellendorffii|130337
Smoellendorffii|164454
Smoellendorffii|112151
Smoellendorffii|403029
Smoellendorffii|234244
Smoellendorffii|233532
Mpolymorpha|Mapoly0129s0043.1.p
Mpolymorpha|Mapoly0090s0052.1.p
Mpolymorpha|Mapoly0021s0062.1.p
Mpolymorpha|Mapoly0029s0146.1.p
Mpolymorpha|Mapoly0066s0096.1.p
Mpolymorpha|Mapoly0401s0001.1.p
Mpolymorpha|Mapoly0074s0091.1.p
Mpolymorpha|Mapoly0074s0092.1.p
Mpolymorpha|Mapoly0010s0197.1.p
Klebsormidium|kfl00039_0260_v1.1

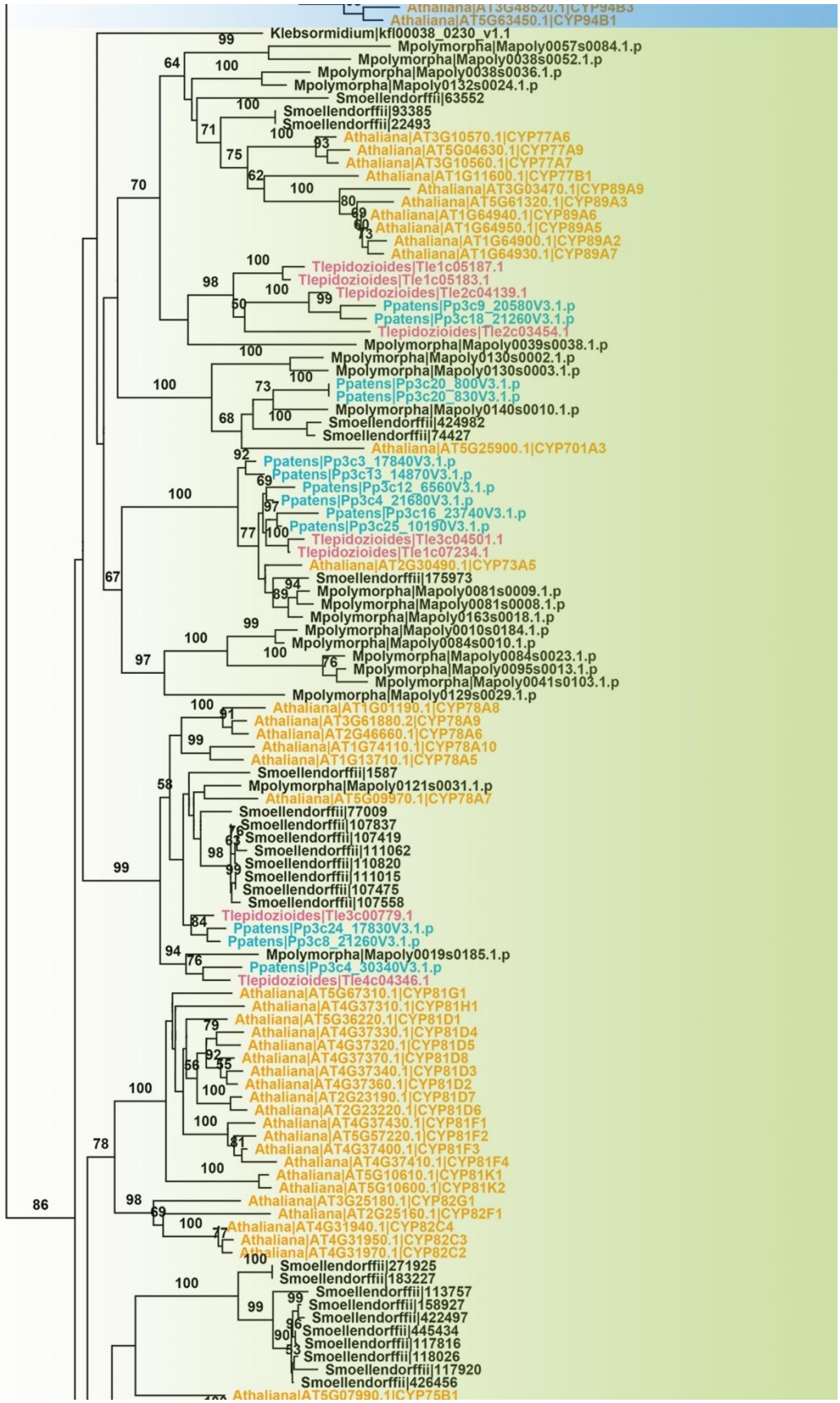
141


Klebsormidium|kfl00785_0010_v1.1
Chbra|g50792
Creinhardtii|Cre11.g467527.t1.1
Klebsormidium|kfl00691_0060_v1.1
Smoellendorffii|75218
Mpolymorpha|Mapoly0045s0151.1.p
Mpolymorpha|Mapoly0103s0038.1.p
Ppatens|Pp3c1_11690V3.1.p
Ppatens|Pp3c2_27810V3.1.p
Tlepidozioides|Tle1c00521.1
Tlepidozioides|Tle3c02370.1
Athaliana|AT2G34500.1|CYP710A1
Athaliana|AT2G34490.1|CYP710A2
Athaliana|AT2G28860.1|CYP710A4
Athaliana|AT2G28850.1|CYP710A3
CYP710 clade
Chbra|g12707
Klebsormidium|kfl00090_0230_v1.1
Ppatens|Pp3c8_15480V3.1.p
Tlepidozioides|Tle2c00369.1
Smoellendorffii|178739
Athaliana|AT1G11680.1|CYP51G1
Mpolymorpha|Mapoly0003s0192.1.p
CYP51 clade
Smoellendorffii|96541
Athaliana|AT2G26170.1|CYP711A1
Klebsormidium|kfl00473_0030_v1.1
Klebsormidium|kfl00299_0120_v1.1
Klebsormidium|kfl00437_0060_v1.1
CYP711 clade
Chbra|g55238
Klebsormidium|kfl00876_0010_v1.1
Smoellendorffii|145597
Mpolymorpha|Mapoly0245s0001.1.p
Tlepidozioides|Tle2c04000.1
Ppatens|Pp3c4_28160V3.1.p
Athaliana|AT4G15110.1|CYP97B3
Athaliana|AT3G53130.1|CYP97C1
Smoellendorffii|186370
Tlepidozioides|Tle4c04025.1
Ppatens|Pp3c6_15960V3.1.p
Mpolymorpha|Mapoly0028s0136.1.p
Chbra|g21846
Chbra|g4469
Klebsormidium|kfl00137_0040_v1.1
Creinhardtii|Cre08.g373100.t1.1
Chbra|g41079
Klebsormidium|kfl00028_0200_v1.1
Creinhardtii|Cre16.g678437.t1.1
Creinhardtii|Cre02.g142266.t1.1
Athaliana|AT1G31800.1|CYP97A3
Ppatens|Pp3c5_28020V3.1.p
Smoellendorffii|407294
Tlepidozioides|Tle2c02594.1
Mpolymorpha|Mapoly0011s0135.1.p
CYP97 clade
Mpolymorpha|Mapoly0022s0192.1.p
Mpolymorpha|Mapoly0022s0190.1.p
Mpolymorpha|Mapoly0024s0025.1.p
Mpolymorpha|Mapoly0078s0054.1.p
Mpolymorpha|Mapoly0045s0010.1.p
Tlepidozioides|Tle1c04242.1
Mpolymorpha|Mapoly0052s0003.1.p
Ppatens|Pp3c12_25250V3.1.p
Ppatens|Pp3c2_30580V3.1.p
Ppatens|Pp3c21_10970V3.1.p
Smoellendorffii|235399
Smoellendorffii|440934
Athaliana|AT1G17060.1|CYP72C1
Athaliana|AT3G14620.1|CYP72A8
Athaliana|AT3G14610.1|CYP72A7
Athaliana|AT3G14660.1|CYP72A13
Athaliana|AT3G14650.1|CYP72A11
Athaliana|AT3G14690.1|CYP72A15
Athaliana|AT3G14640.1|CYP72A10
Athaliana|AT3G14680.1|CYP72A14
Athaliana|AT3G14630.1|CYP72A9
Athaliana|AT1G75130.1|CYP721A1
Athaliana|AT2G26710.1|CYP734A1
Athaliana|AT2G46960.2|CYP709B1
Athaliana|AT4G27710.1|CYP709B3
Athaliana|AT2G46950.1|CYP709B2
Tlepidozioides|Tle4c03015.1
Tlepidozioides|Tle1c06733.1
Smoellendorffii|271071
Smoellendorffii|88220
Smoellendorffii|231182
Smoellendorffii|89087
Smoellendorffii|236984
Smoellendorffii|104195
Smoellendorffii|145354
Smoellendorffii|141802
Smoellendorffii|266618
Smoellendorffii|410423
Smoellendorffii|170849
Smoellendorffii|94541
Smoellendorffii|171715
Smoellendorffii|94261
Smoellendorffii|170843
Smoellendorffii|441298
Smoellendorffii|267282
Smoellendorffii|410445
Smoellendorffii|94002
Smoellendorffii|94367
Smoellendorffii|410461
Smoellendorffii|91564
Smoellendorffii|147392
Smoellendorffii|410449
Smoellendorffii|183651
Smoellendorffii|80290
Smoellendorffii|97810
CYP72 clade

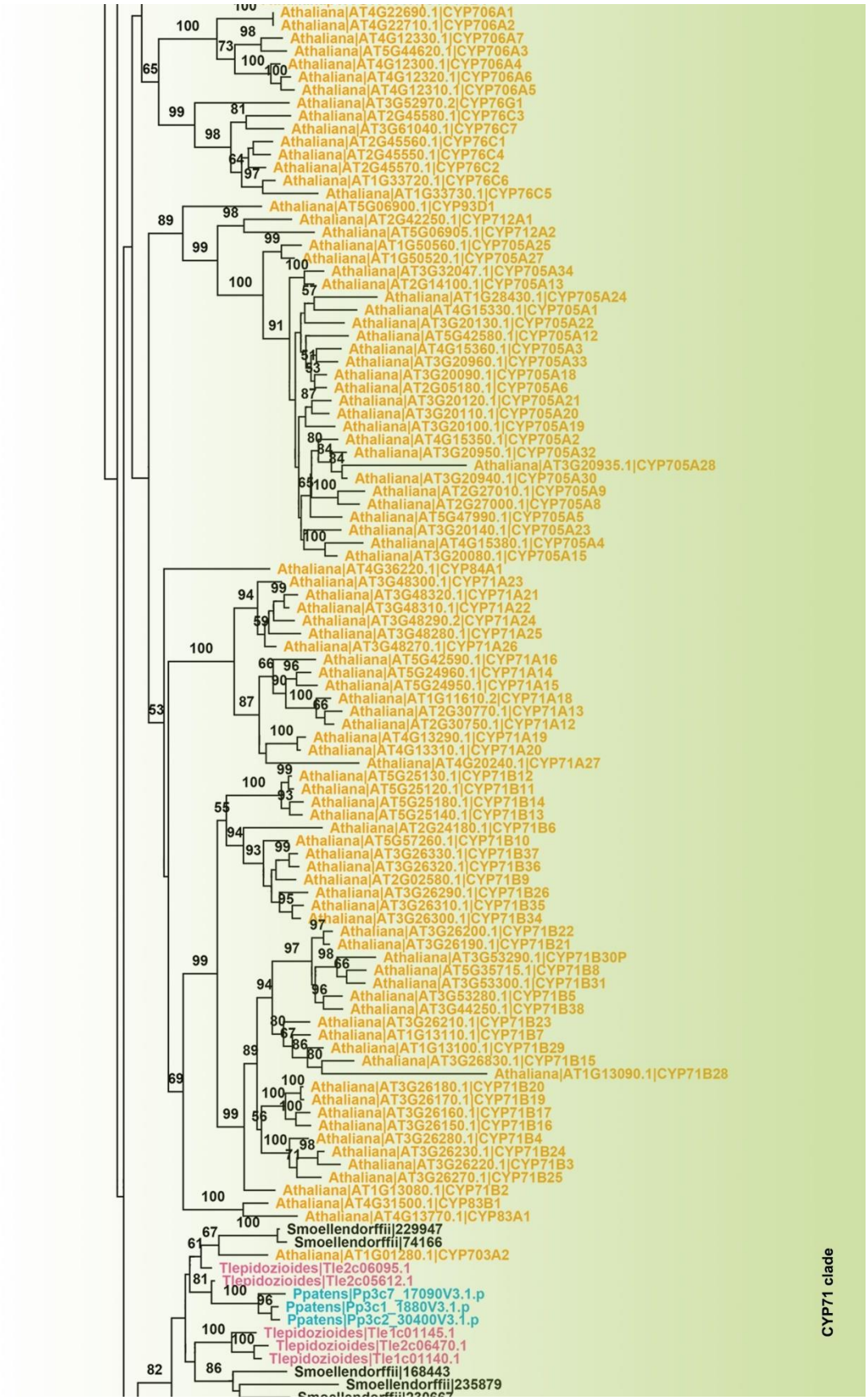
142


Smoellendorffii|172829
Smoellendorffii|92836
Smoellendorffii|430406
Smoellendorffii|408940
Smoellendorffii|405479
Smoellendorffii|119072
Athaliana|AT5G38450.1|CYP735A1
Athaliana|AT1G67110.1|CYP735A2
Athaliana|AT5G52400.1|CYP715A1
Athaliana|AT5G24910.1|CYP714A1
Athaliana|AT5G24900.1|CYP714A2
Ppatens|Pp3c1_33340V3.1.p
Mpolymorpha|Mapoly0108s0061.1.p
Klebsormidium|kfl00010_0080_v1.1
Klebsormidium|kfl00161_0090_v1.1
Klebsormidium|kfl00191_0130_v1.1
Klebsormidium|kfl00471_0090_v1.1
Klebsormidium|kfl00468_0120_v1.1
Ppatens|Pp3c24_15550V3.1.p
Klebsormidium|kfl00188_0250_v1.1
Klebsormidium|kfl01524_0010_v1.1
Klebsormidium|kfl00139_0060_v1.1
Athaliana|AT2G44890.1|CYP704A1
Athaliana|AT2G45510.1|CYP704A2
Mpolymorpha|Mapoly0144s0022.1.p
Ppatens|Pp3c5_13180V3.1.p
Ppatens|Pp3c6_15280V3.1.p
Tlepidozioides|Tle2c07053.1
Mpolymorpha|Mapoly0001s0122.1.p
Smoellendorffii|151517
Mpolymorpha|Mapoly0173s0023.1.p
Tlepidozioides|Tle4c02928.1
Tlepidozioides|Tle4c02927.1
Ppatens|Pp3c13_24510V3.1.p
Mpolymorpha|Mapoly0021s0144.1.p
Tlepidozioides|Tle3c05391.1
Ppatens|Pp3c24_18890V3.1.p
Smoellendorffii|113735
Smoellendorffii|113847
Smoellendorffii|405166
Smoellendorffii|80855
Mpolymorpha|Mapoly0041s0041.1.p
Mpolymorpha|Mapoly0037s0123.1.p
Mpolymorpha|Mapoly0037s0114.1.p
Mpolymorpha|Mapoly0005s0090.1.p
Tlepidozioides|Tle1c05967.1
Athaliana|AT1G69500.1|CYP704B1
Tlepidozioides|Tle2c06457.1
Ppatens|Pp3c2_9340V3.1.p
Ppatens|Pp3c17_22820V3.1.p
Ppatens|Pp3c14_23761V3.1.p
Ppatens|Pp3c14_23760V3.1.p
Mpolymorpha|Mapoly0066s0002.1.p
Mpolymorpha|Mapoly0066s0005.1.p
Smoellendorffii|80659
Tlepidozioides|Tle2c01896.1
Tlepidozioides|Tle3c03541.1
Mpolymorpha|Mapoly0063s0016.1.p
Mpolymorpha|Mapoly0050s0016.1.p
Tlepidozioides|Tle3c04923.1
Tlepidozioides|Tle2c06914.1
Tlepidozioides|Tle2c02938.1
Smoellendorffii|111270
Smoellendorffii|418431
Klebsormidium|kfl00120_0040_v1.1
Klebsormidium|kfl00171_0160_v1.1
Mpolymorpha|Mapoly0039s0047.1.p
Athaliana|AT5G58860.1|CYP86A1
Athaliana|AT1G63710.1|CYP86A7
Athaliana|AT2G45970.1|CYP86A8
Athaliana|AT4G00360.1|CYP86A2
Athaliana|AT1G01600.1|CYP86A4
Smoellendorffii|124000
Athaliana|AT5G23190.1|CYP86B1
Athaliana|AT2G21910.1|CYP96A5
Athaliana|AT1G65340.1|CYP96A3
Athaliana|AT5G52320.1|CYP96A4
Athaliana|AT1G57750.1|CYP96A15
Athaliana|AT4G39490.1|CYP96A10
Athaliana|AT2G23180.1|CYP96A1
Athaliana|AT4G39480.1|CYP96A9
Athaliana|AT4G39500.1|CYP96A11
Athaliana|AT4G32170.1|CYP96A2
Athaliana|AT1G47620.1|CYP96A8
Athaliana|AT5G02900.1|CYP96A13
Athaliana|AT4G39510.1|CYP96A12
Athaliana|AT1G24540.1|CYP86C1
Athaliana|AT1G13140.1|CYP86C3
Athaliana|AT1G13150.1|CYP86C4
Athaliana|AT3G26125.1|CYP86C2
Tlepidozioides|Tle2c03895.1
Ppatens|Pp3c11_26530V3.1.p
Ppatens|Pp3c7_6350V3.1.p
Ppatens|Pp3c6_400V3.1.p
Mpolymorpha|Mapoly0141s0021.1.p
Mpolymorpha|Mapoly0023s0003.1.p
Ppatens|Pp3c1_23370V3.1.p
Tlepidozioides|Tle2c04119.1
Smoellendorffii|111511
Smoellendorffii|128350
Smoellendorffii|83080
Smoellendorffii|233293
Smoellendorffii|76541
Athaliana|AT1G34540.1|CYP94D1
Athaliana|AT3G56630.1|CYP94D2
Athaliana|AT2G27690.1|CYP94C1
Athaliana|AT3G01900.1|CYP94B2
CYP86 clade

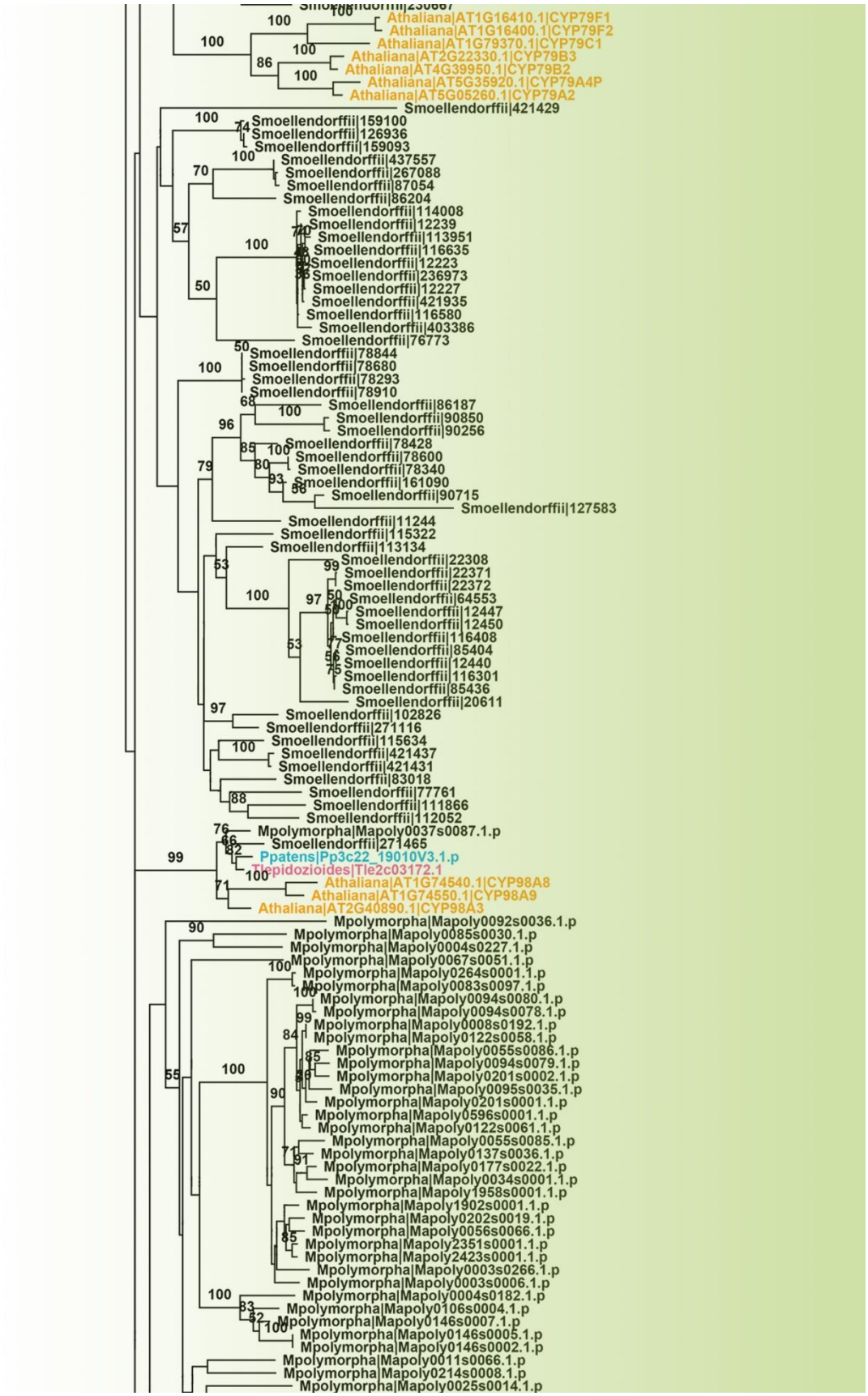
143


Athaliana|AT5G63450.1|CYP94B1
Klebsormidium|kfl00038_0230_v1.1
Mpolymorpha|Mapoly0057s0084.1.p
Mpolymorpha|Mapoly0038s0052.1.p
Mpolymorpha|Mapoly0038s0036.1.p
Mpolymorpha|Mapoly0132s0024.1.p
Smoellendorffii|63552
Smoellendorffii|93385
Smoellendorffii|22493
Athaliana|AT3G10570.1|CYP77A6
Athaliana|AT5G04630.1|CYP77A9
Athaliana|AT3G10560.1|CYP77A7
Athaliana|AT1G11600.1|CYP77B1
Athaliana|AT3G03470.1|CYP89A9
Athaliana|AT5G61320.1|CYP89A3
Athaliana|AT1G64940.1|CYP89A6
Athaliana|AT1G64950.1|CYP89A5
Athaliana|AT1G64900.1|CYP89A2
Athaliana|AT1G64930.1|CYP89A7
Tlepidozioides|Tle1c05187.1
Tlepidozioides|Tle1c05183.1
Tlepidozioides|Tle2c04139.1
Ppatens|Pp3c9_20580V3.1.p
Ppatens|Pp3c18_21260V3.1.p
Tlepidozioides|Tle2c03454.1
Mpolymorpha|Mapoly0039s0038.1.p
Mpolymorpha|Mapoly0130s0002.1.p
Mpolymorpha|Mapoly0130s0003.1.p
Ppatens|Pp3c20_800V3.1.p
Ppatens|Pp3c20_830V3.1.p
Mpolymorpha|Mapoly0140s0010.1.p
Smoellendorffii|424982
Smoellendorffii|74427
Athaliana|AT5G25900.1|CYP701A3
Ppatens|Pp3c3_17840V3.1.p
Ppatens|Pp3c13_14870V3.1.p
Ppatens|Pp3c12_6560V3.1.p
Ppatens|Pp3c4_21680V3.1.p
Ppatens|Pp3c16_23740V3.1.p
Ppatens|Pp3c25_10190V3.1.p
Tlepidozioides|Tle3c04501.1
Tlepidozioides|Tle1c07234.1
Athaliana|AT2G30490.1|CYP73A5
Smoellendorffii|175973
Mpolymorpha|Mapoly0081s0009.1.p
Mpolymorpha|Mapoly0081s0008.1.p
Mpolymorpha|Mapoly0163s0018.1.p
Mpolymorpha|Mapoly0010s0184.1.p
Mpolymorpha|Mapoly0084s0010.1.p
Mpolymorpha|Mapoly0084s0023.1.p
Mpolymorpha|Mapoly0095s0013.1.p
Mpolymorpha|Mapoly0041s0103.1.p
Mpolymorpha|Mapoly0129s0029.1.p
Athaliana|AT1G01190.1|CYP78A8
Athaliana|AT3G61880.2|CYP78A9
Athaliana|AT2G46660.1|CYP78A6
Athaliana|AT1G74110.1|CYP78A10
Athaliana|AT1G13710.1|CYP78A5
Smoellendorffii|1587
Mpolymorpha|Mapoly0121s0031.1.p
Athaliana|AT5G09970.1|CYP78A7
Smoellendorffii|77009
Smoellendorffii|107837
Smoellendorffii|107419
Smoellendorffii|111062
Smoellendorffii|110820
Smoellendorffii|111015
Smoellendorffii|107475
Smoellendorffii|107558
Tlepidozioides|Tle3c00779.1
Ppatens|Pp3c24_17830V3.1.p
Ppatens|Pp3c8_21260V3.1.p
Mpolymorpha|Mapoly0019s0185.1.p
Ppatens|Pp3c4_30340V3.1.p
Tlepidozioides|Tle4c04346.1
Athaliana|AT5G67310.1|CYP81G1
Athaliana|AT4G37310.1|CYP81H1
Athaliana|AT5G36220.1|CYP81D1
Athaliana|AT4G37330.1|CYP81D4
Athaliana|AT4G37320.1|CYP81D5
Athaliana|AT4G37370.1|CYP81D8
Athaliana|AT4G37340.1|CYP81D3
Athaliana|AT4G37360.1|CYP81D2
Athaliana|AT2G23190.1|CYP81D7
Athaliana|AT2G23220.1|CYP81D6
Athaliana|AT4G37430.1|CYP81F1
Athaliana|AT5G57220.1|CYP81F2
Athaliana|AT4G37400.1|CYP81F3
Athaliana|AT4G37410.1|CYP81F4
Athaliana|AT5G10610.1|CYP81K1
Athaliana|AT5G10600.1|CYP81K2
Athaliana|AT3G25180.1|CYP82G1
Athaliana|AT2G25160.1|CYP82F1
Athaliana|AT4G31940.1|CYP82C4
Athaliana|AT4G31950.1|CYP82C3
Athaliana|AT4G31970.1|CYP82C2
Smoellendorffii|271925
Smoellendorffii|183227
Smoellendorffii|113757
Smoellendorffii|158927
Smoellendorffii|422497
Smoellendorffii|445434
Smoellendorffii|117816
Smoellendorffii|118026
Smoellendorffii|117920
Smoellendorffii|426456
Athaliana|AT5G07990.1|CYP75B1

144


Athaliana|AT4G22690.1|CYP706A1
Athaliana|AT4G22710.1|CYP706A2
Athaliana|AT4G12330.1|CYP706A7
Athaliana|AT5G44620.1|CYP706A3
Athaliana|AT4G12300.1|CYP706A4
Athaliana|AT4G12320.1|CYP706A6
Athaliana|AT4G12310.1|CYP706A5
Athaliana|AT3G52970.2|CYP76G1
Athaliana|AT2G45580.1|CYP76C3
Athaliana|AT3G61040.1|CYP76C7
Athaliana|AT2G45560.1|CYP76C1
Athaliana|AT2G45550.1|CYP76C4
Athaliana|AT2G45570.1|CYP76C2
Athaliana|AT1G33720.1|CYP76C6
Athaliana|AT1G33730.1|CYP76C5
Athaliana|AT5G06900.1|CYP93D1
Athaliana|AT2G42250.1|CYP712A1
Athaliana|AT5G06905.1|CYP712A2
Athaliana|AT1G50560.1|CYP705A25
Athaliana|AT1G50520.1|CYP705A27
Athaliana|AT3G32047.1|CYP705A34
Athaliana|AT2G14100.1|CYP705A13
Athaliana|AT1G28430.1|CYP705A24
Athaliana|AT4G15330.1|CYP705A1
Athaliana|AT3G20130.1|CYP705A22
Athaliana|AT5G42580.1|CYP705A12
Athaliana|AT4G15360.1|CYP705A3
Athaliana|AT3G20960.1|CYP705A33
Athaliana|AT3G20090.1|CYP705A18
Athaliana|AT2G05180.1|CYP705A6
Athaliana|AT3G20120.1|CYP705A21
Athaliana|AT3G20110.1|CYP705A20
Athaliana|AT3G20100.1|CYP705A19
Athaliana|AT4G15350.1|CYP705A2
Athaliana|AT3G20950.1|CYP705A32
Athaliana|AT3G20935.1|CYP705A28
Athaliana|AT3G20940.1|CYP705A30
Athaliana|AT2G27010.1|CYP705A9
Athaliana|AT2G27000.1|CYP705A8
Athaliana|AT5G47990.1|CYP705A5
Athaliana|AT3G20140.1|CYP705A23
Athaliana|AT4G15380.1|CYP705A4
Athaliana|AT3G20080.1|CYP705A15
Athaliana|AT4G36220.1|CYP84A1
Athaliana|AT3G48300.1|CYP71A23
Athaliana|AT3G48320.1|CYP71A21
Athaliana|AT3G48310.1|CYP71A22
Athaliana|AT3G48290.2|CYP71A24
Athaliana|AT3G48280.1|CYP71A25
Athaliana|AT3G48270.1|CYP71A26
Athaliana|AT5G42590.1|CYP71A16
Athaliana|AT5G24960.1|CYP71A14
Athaliana|AT5G24950.1|CYP71A15
Athaliana|AT1G11610.2|CYP71A18
Athaliana|AT2G30770.1|CYP71A13
Athaliana|AT2G30750.1|CYP71A12
Athaliana|AT4G13290.1|CYP71A19
Athaliana|AT4G13310.1|CYP71A20
Athaliana|AT4G20240.1|CYP71A27
Athaliana|AT5G25130.1|CYP71B12
Athaliana|AT5G25120.1|CYP71B11
Athaliana|AT5G25180.1|CYP71B14
Athaliana|AT5G25140.1|CYP71B13
Athaliana|AT2G24180.1|CYP71B6
Athaliana|AT5G57260.1|CYP71B10
Athaliana|AT3G26330.1|CYP71B37
Athaliana|AT3G26320.1|CYP71B36
Athaliana|AT2G02580.1|CYP71B9
Athaliana|AT3G26290.1|CYP71B26
Athaliana|AT3G26310.1|CYP71B35
Athaliana|AT3G26300.1|CYP71B34
Athaliana|AT3G26200.1|CYP71B22
Athaliana|AT3G26190.1|CYP71B21
Athaliana|AT3G53290.1|CYP71B30P
Athaliana|AT5G35715.1|CYP71B8
Athaliana|AT3G53300.1|CYP71B31
Athaliana|AT3G53280.1|CYP71B5
Athaliana|AT3G44250.1|CYP71B38
Athaliana|AT3G26210.1|CYP71B23
Athaliana|AT1G13110.1|CYP71B7
Athaliana|AT1G13100.1|CYP71B29
Athaliana|AT3G26830.1|CYP71B15
Athaliana|AT1G13090.1|CYP71B28
Athaliana|AT3G26180.1|CYP71B20
Athaliana|AT3G26170.1|CYP71B19
Athaliana|AT3G26160.1|CYP71B17
Athaliana|AT3G26150.1|CYP71B16
Athaliana|AT3G26280.1|CYP71B4
Athaliana|AT3G26230.1|CYP71B24
Athaliana|AT3G26220.1|CYP71B3
Athaliana|AT3G26270.1|CYP71B25
Athaliana|AT1G13080.1|CYP71B2
Athaliana|AT4G31500.1|CYP83B1
Athaliana|AT4G13770.1|CYP83A1
Smoellendorffii|229947
Smoellendorffii|74166
Athaliana|AT1G01280.1|CYP703A2
Tlepidozioides|Tle2c06095.1
Tlepidozioides|Tle2c05612.1
Ppatens|Pp3c7_17090V3.1.p
Ppatens|Pp3c1_7880V3.1.p
Ppatens|Pp3c2_30400V3.1.p
Tlepidozioides|Tle1c01145.1
Tlepidozioides|Tle2c06470.1
Tlepidozioides|Tle1c01140.1
Smoellendorffii|168443
Smoellendorffii|235879
CYP71 clade

145


Athaliana|AT1G16410.1|CYP79F1
Athaliana|AT1G16400.1|CYP79F2
Athaliana|AT1G79370.1|CYP79C1
Athaliana|AT2G22330.1|CYP79B3
Athaliana|AT4G39950.1|CYP79B2
Athaliana|AT5G35920.1|CYP79A4P
Athaliana|AT5G05260.1|CYP79A2
Smoellendorffii|421429
Smoellendorffii|159100
Smoellendorffii|126936
Smoellendorffii|159093
Smoellendorffii|437557
Smoellendorffii|267088
Smoellendorffii|87054
Smoellendorffii|86204
Smoellendorffii|114008
Smoellendorffii|12239
Smoellendorffii|113951
Smoellendorffii|116635
Smoellendorffii|12223
Smoellendorffii|236973
Smoellendorffii|12227
Smoellendorffii|421935
Smoellendorffii|116580
Smoellendorffii|403386
Smoellendorffii|76773
Smoellendorffii|78844
Smoellendorffii|78680
Smoellendorffii|78293
Smoellendorffii|78910
Smoellendorffii|86187
Smoellendorffii|90850
Smoellendorffii|90256
Smoellendorffii|78428
Smoellendorffii|78600
Smoellendorffii|78340
Smoellendorffii|161090
Smoellendorffii|90715
Smoellendorffii|127583
Smoellendorffii|11244
Smoellendorffii|115322
Smoellendorffii|113134
Smoellendorffii|22308
Smoellendorffii|22371
Smoellendorffii|22372
Smoellendorffii|64553
Smoellendorffii|12447
Smoellendorffii|12450
Smoellendorffii|116408
Smoellendorffii|85404
Smoellendorffii|12440
Smoellendorffii|116301
Smoellendorffii|85436
Smoellendorffii|20611
Smoellendorffii|102826
Smoellendorffii|271116
Smoellendorffii|115634
Smoellendorffii|421437
Smoellendorffii|421431
Smoellendorffii|83018
Smoellendorffii|77761
Smoellendorffii|111866
Smoellendorffii|112052
Mpolymorpha|Mapoly0037s0087.1.p
Smoellendorffii|271465
Ppatens|Pp3c22_19010V3.1.p
Tlepidozioides|Tle2c03172.1
Athaliana|AT1G74540.1|CYP98A8
Athaliana|AT1G74550.1|CYP98A9
Athaliana|AT2G40890.1|CYP98A3
Mpolymorpha|Mapoly0092s0036.1.p
Mpolymorpha|Mapoly0085s0030.1.p
Mpolymorpha|Mapoly0004s0227.1.p
Mpolymorpha|Mapoly0067s0051.1.p
Mpolymorpha|Mapoly0264s0001.1.p
Mpolymorpha|Mapoly0083s0097.1.p
Mpolymorpha|Mapoly0094s0080.1.p
Mpolymorpha|Mapoly0094s0078.1.p
Mpolymorpha|Mapoly0008s0192.1.p
Mpolymorpha|Mapoly0122s0058.1.p
Mpolymorpha|Mapoly0055s0086.1.p
Mpolymorpha|Mapoly0094s0079.1.p
Mpolymorpha|Mapoly0201s0002.1.p
Mpolymorpha|Mapoly0095s0035.1.p
Mpolymorpha|Mapoly0201s0001.1.p
Mpolymorpha|Mapoly0596s0001.1.p
Mpolymorpha|Mapoly0122s0061.1.p
Mpolymorpha|Mapoly0055s0085.1.p
Mpolymorpha|Mapoly0137s0036.1.p
Mpolymorpha|Mapoly0177s0022.1.p
Mpolymorpha|Mapoly0034s0001.1.p
Mpolymorpha|Mapoly1958s0001.1.p
Mpolymorpha|Mapoly1902s0001.1.p
Mpolymorpha|Mapoly0202s0019.1.p
Mpolymorpha|Mapoly0056s0066.1.p
Mpolymorpha|Mapoly2351s0001.1.p
Mpolymorpha|Mapoly2423s0001.1.p
Mpolymorpha|Mapoly0003s0266.1.p
Mpolymorpha|Mapoly0003s0006.1.p
Mpolymorpha|Mapoly0004s0182.1.p
Mpolymorpha|Mapoly0106s0004.1.p
Mpolymorpha|Mapoly0146s0007.1.p
Mpolymorpha|Mapoly0146s0005.1.p
Mpolymorpha|Mapoly0146s0002.1.p
Mpolymorpha|Mapoly0011s0066.1.p
Mpolymorpha|Mapoly0214s0008.1.p
Mpolymorpha|Mapoly0025s0014.1.p

146

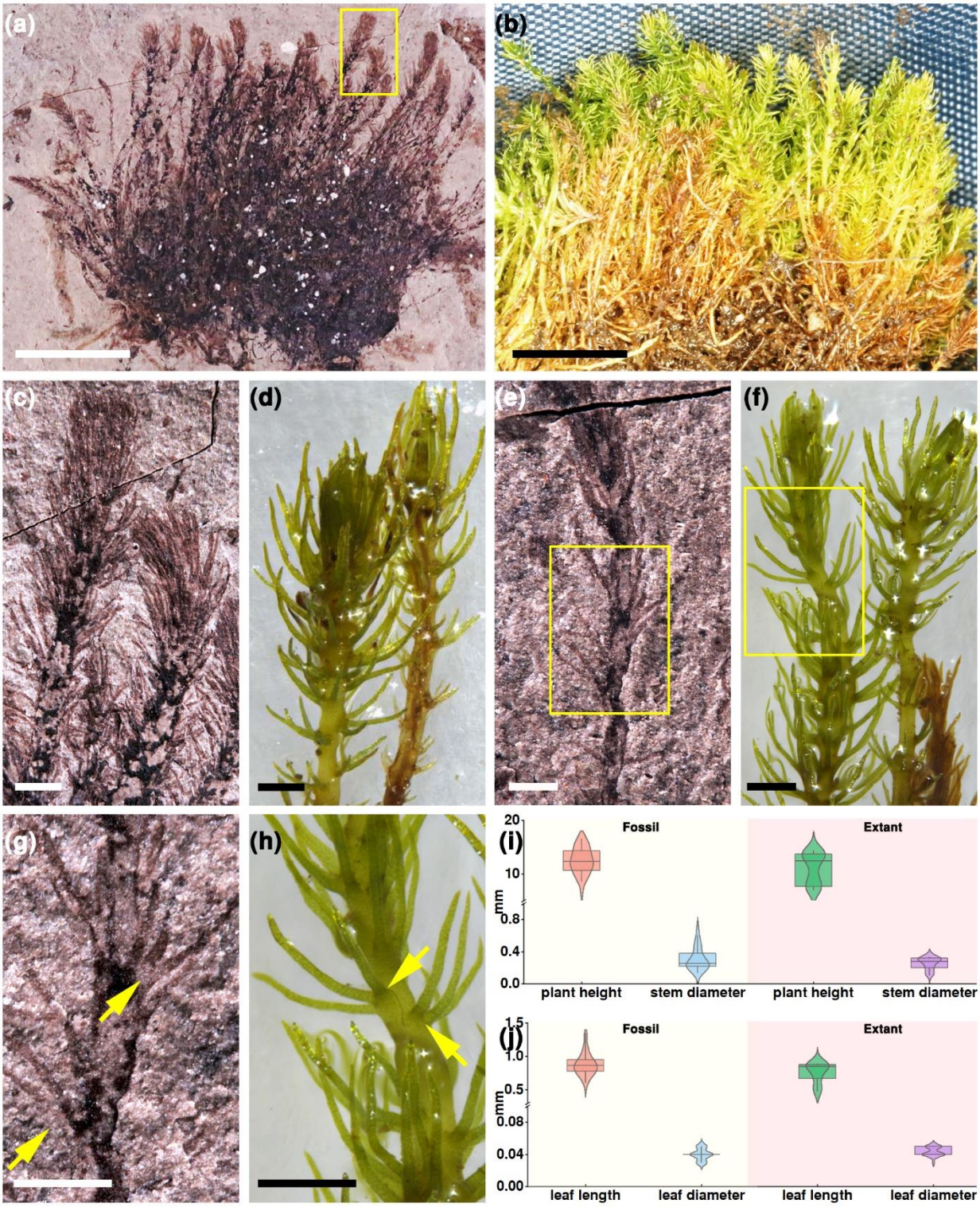
147

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

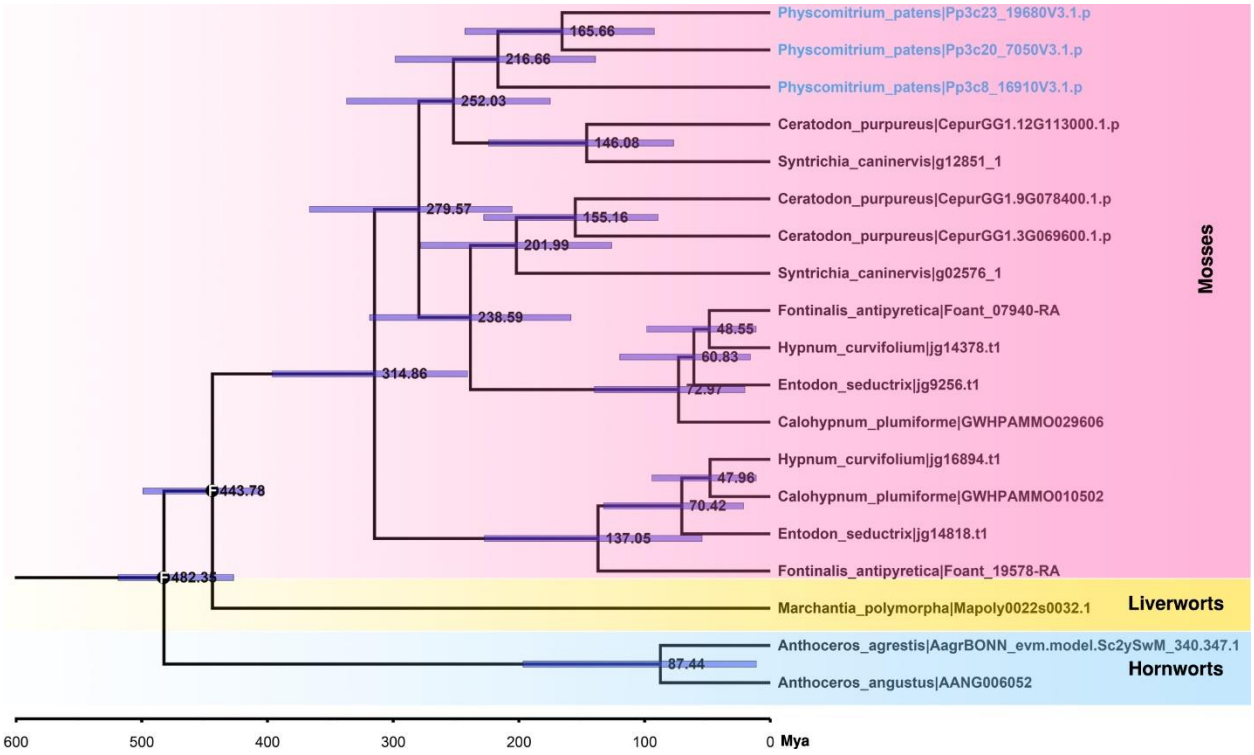
less than 50% support are not shown. Orange font color represents *A. thaliana*, pink font color represents *T. lepidozioides* and blue font color represents *P. patens*. Shade colors as shown on the right.


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
Fossil
Extant
mm
20
10
0.8
0.4
0.0
plant height
stem diameter
plant height
stem diameter
(j)
Fossil
Extant
mm
1.5
1.0
0.5
0.08
0.04
0.00
leaf length
leaf diameter
leaf length
leaf diameter

157
158

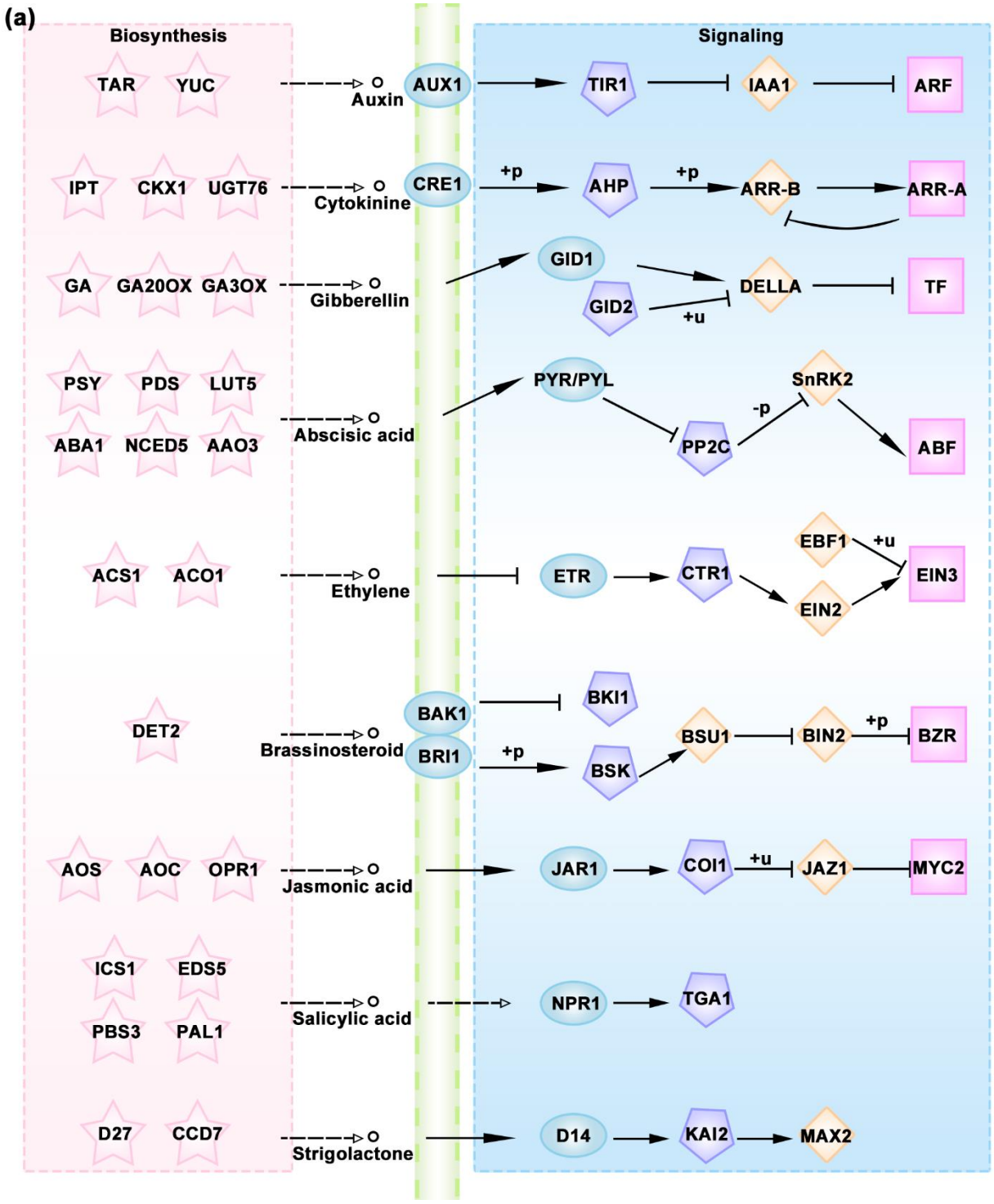
Data S2 N. Plant morphological characteristics of *Takakia* sp. fossil and *T. lepidozioides*

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

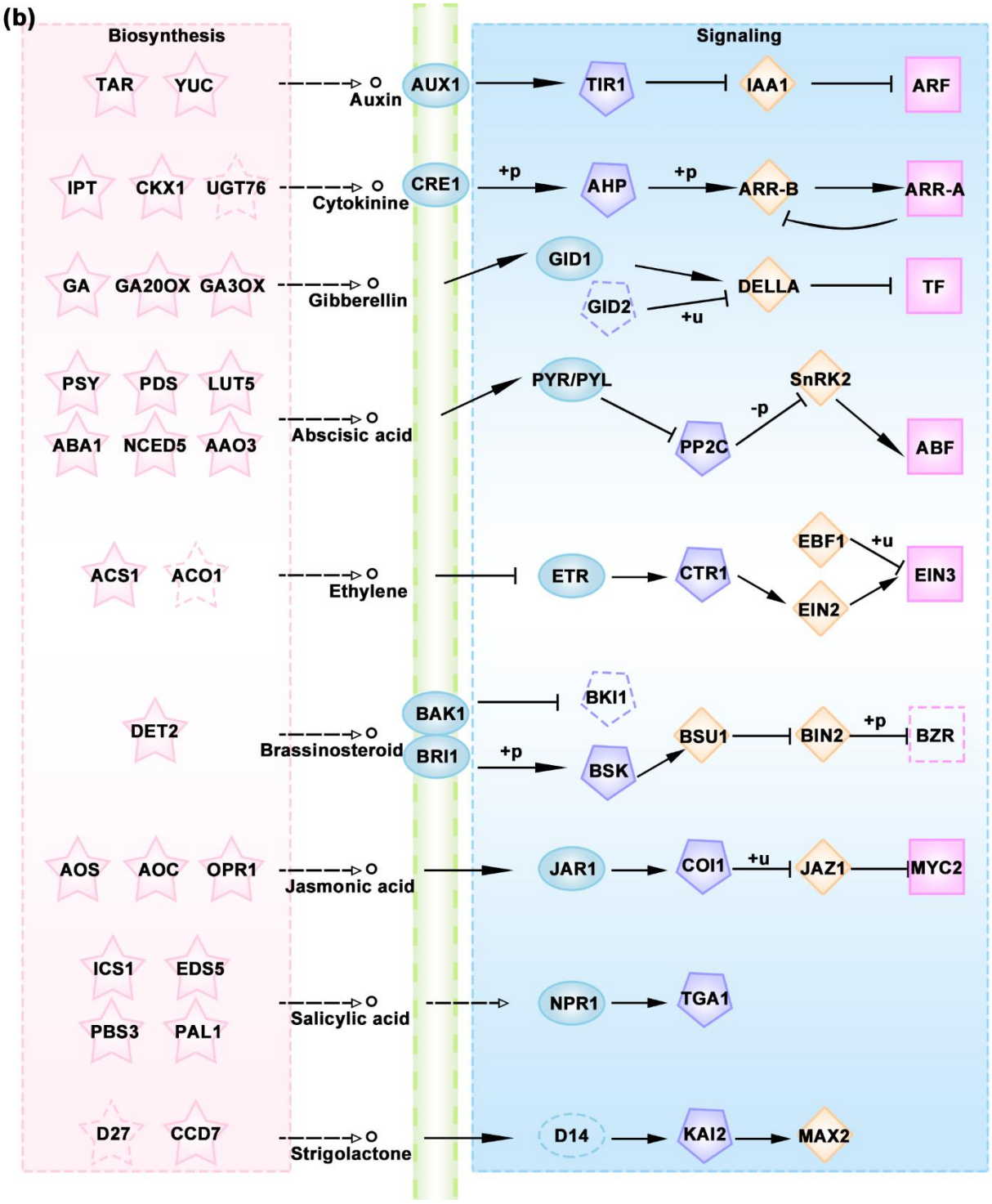


Data S2 O. Phylogenetic timetree of KAN-homologs in bryophytes

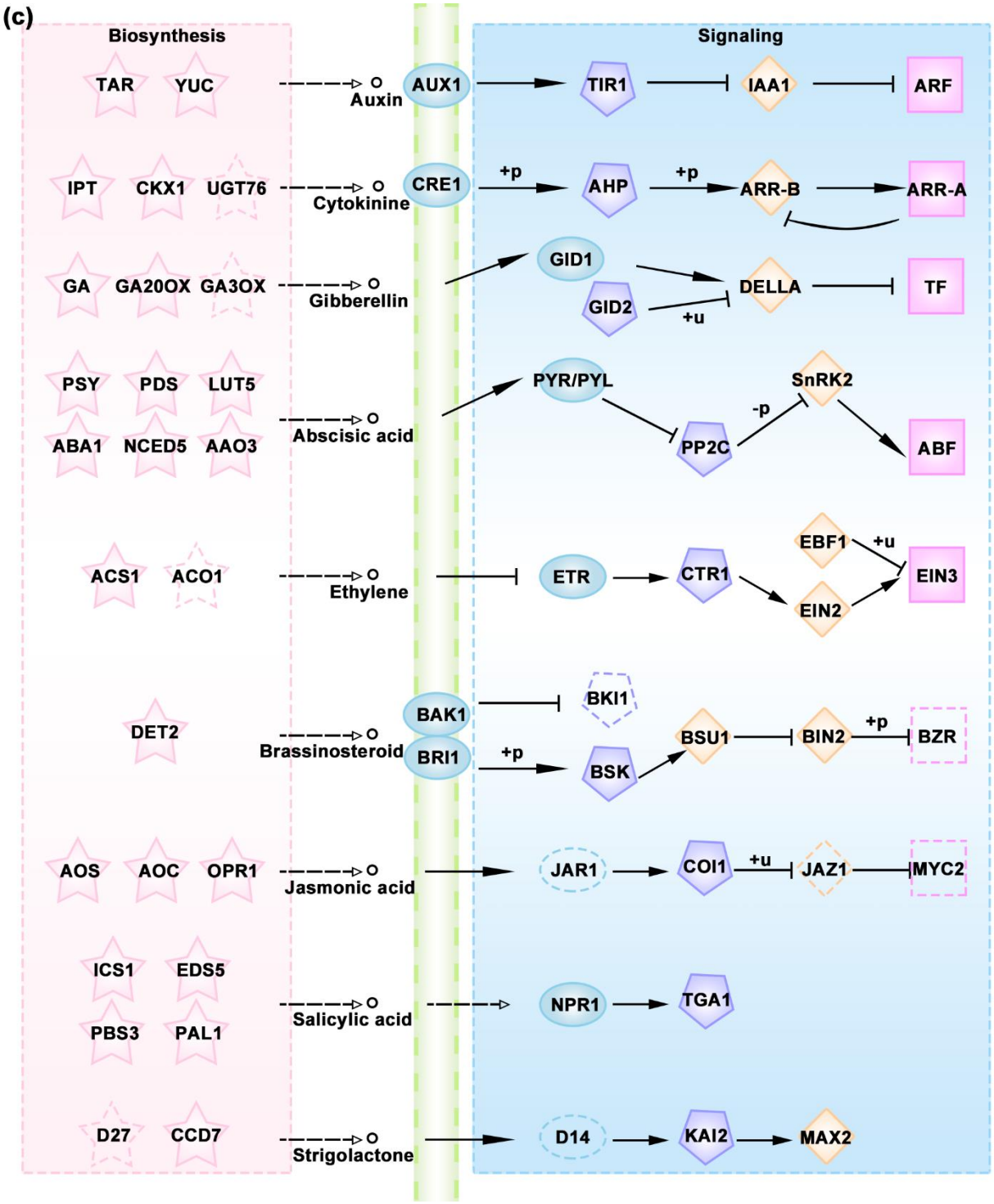
KAN (AT5G16560) were found using blastp with identity cutoff 50%, e-value cutoff 1×10^{-20}. Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates. Numbers on nodes represent divergence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


(a)
Biosynthesis
Signaling
TAR
YUC
Auxin
AUX1
TIR1
IAA1
ARF
IPT
CKX1
UGT76
Cytokinine
CRE1
+p
AHP
+p
ARR-B
ARR-A
GA
GA20OX
GA3OX
Gibberellin
GID1
GID2
+u
DELLA
TF
PSY
PDS
LUT5
ABA1
NCED5
AAO3
Abscisic acid
PYR/PYL
PP2C
-p
SnRK2
ABF
ACS1
ACO1
Ethylene
ETR
CTR1
EBF1
+u
EIN2
EIN3
DET2
Brassinosteroid
BAK1
BRI1
+p
BKI1
BSK
BSU1
BIN2
+p
BZR
AOS
AOC
OPR1
Jasmonic acid
JAR1
COI1
+u
JAZ1
MYC2
ICS1
EDS5
PBS3
PAL1
Salicylic acid
NPR1
TGA1
D27
CCD7
Strigolactone
D14
KAI2
MAX2

182


(b)
Biosynthesis
Signaling
TAR
YUC
Auxin
AUX1
TIR1
IAA1
ARF
IPT
CKX1
UGT76
Cytokinine
CRE1
+p
AHP
+p
ARR-B
ARR-A
GA
GA20OX
GA3OX
Gibberellin
GID1
GID2
+u
DELLA
TF
PSY
PDS
LUT5
ABA1
NCED5
AAO3
Abscisic acid
PYR/PYL
PP2C
-p
SnRK2
ABF
ACS1
ACO1
Ethylene
ETR
CTR1
EBF1
+u
EIN2
EIN3
DET2
Brassinosteroid
BAK1
BRI1
+p
BKI1
BSK
BSU1
BIN2
+p
BZR
AOS
AOC
OPR1
Jasmonic acid
JAR1
COI1
+u
JAZ1
MYC2
ICS1
EDS5
PBS3
PAL1
Salicylic acid
NPR1
TGA1
D27
CCD7
Strigolactone
D14
KAI2
MAX2

183

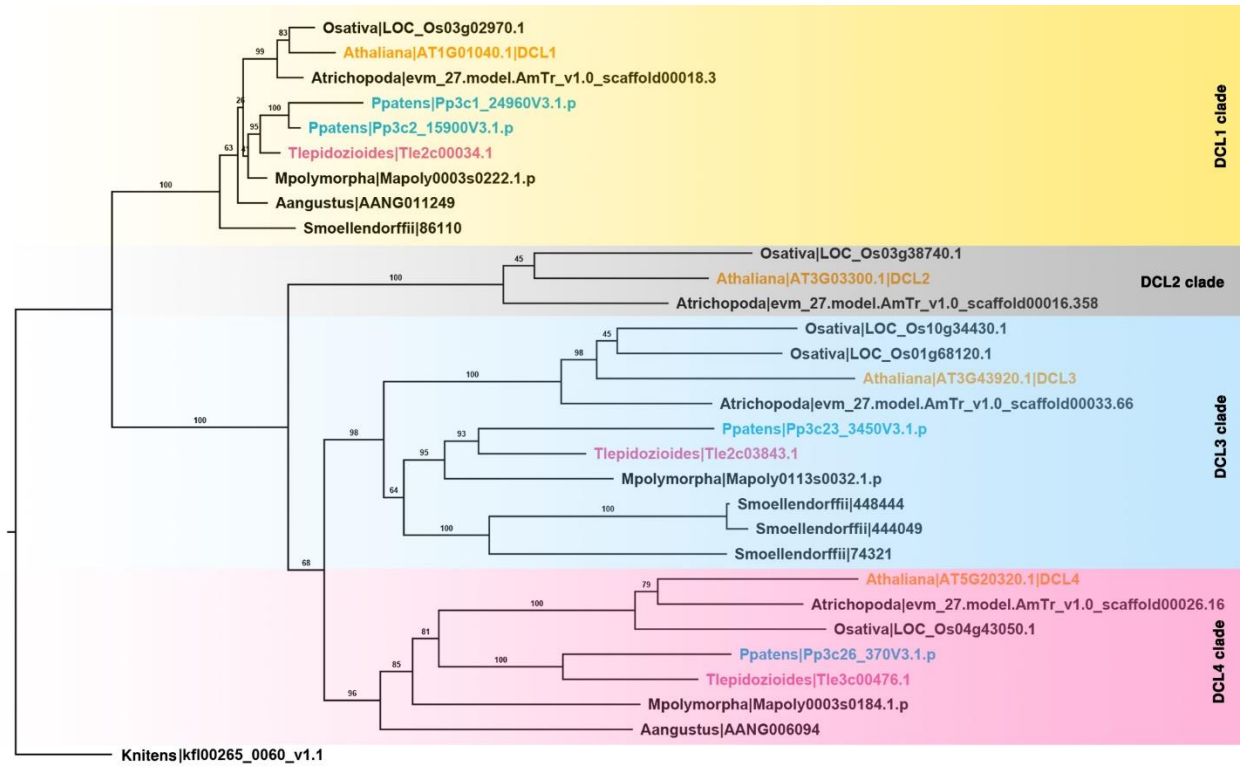

(c)
Biosynthesis
Signaling
TAR
YUC
Auxin
AUX1
TIR1
IAA1
ARF
IPT
CKX1
UGT76
Cytokinine
CRE1
+p
AHP
+p
ARR-B
ARR-A
GA
GA20OX
GA3OX
Gibberellin
GID1
GID2
+u
DELLA
TF
PSY
PDS
LUT5
ABA1
NCED5
AAO3
Abscisic acid
PYR/PYL
PP2C
-p
SnRK2
ABF
ACS1
ACO1
Ethylene
ETR
CTR1
EBF1
+u
EIN2
EIN3
DET2
Brassinosteroid
BAK1
BRI1
+p
BKI1
BSK
BSU1
BIN2
+p
BZR
AOS
AOC
OPR1
Jasmonic acid
JAR1
COI1
+u
JAZ1
MYC2
ICS1
EDS5
PBS3
PAL1
Salicylic acid
NPR1
TGA1
D27
CCD7
Strigolactone
D14
KAI2
MAX2

184

185

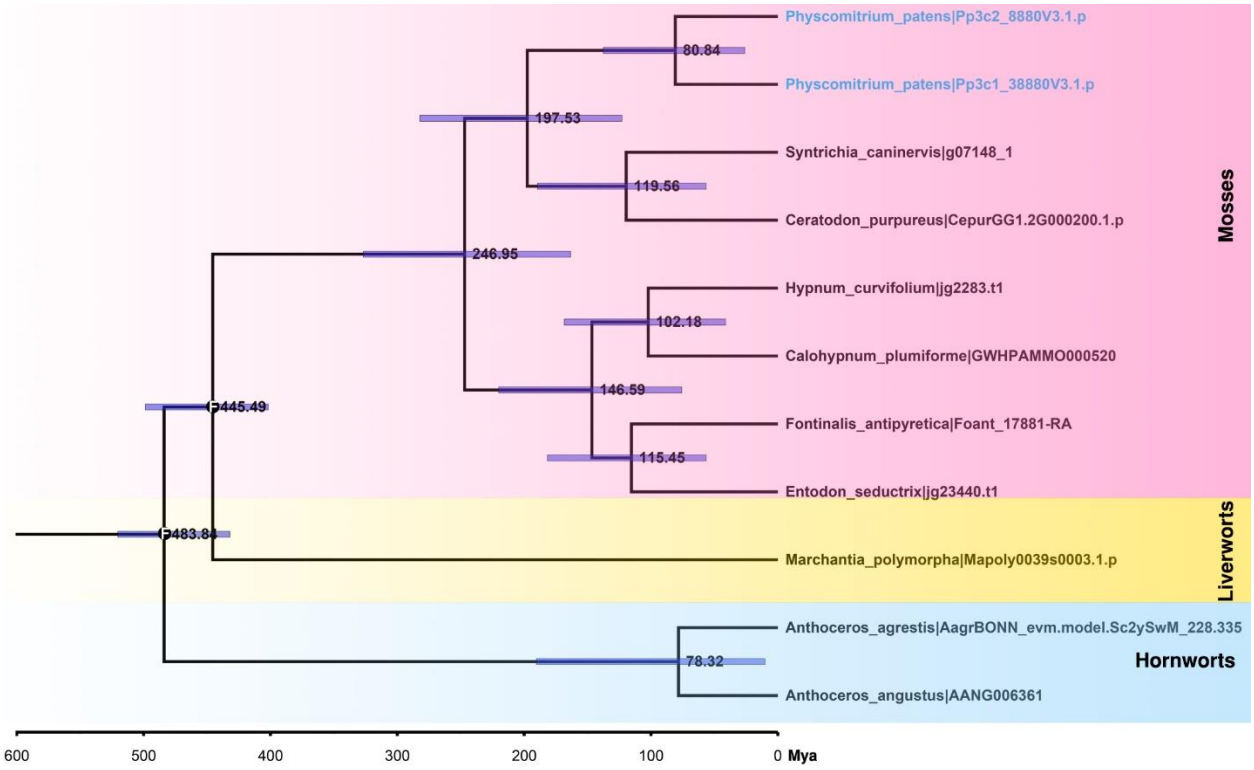
Data S2 P. Comparison of hormone pathway genes in *A. thaliana*, *P. patens* and *T. lepidozioides*

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



Data S2 Q. Phylogenetic trees of Dicer proteins

BLASTP was performed using Dicer of *A. thaliana* sequences to find candidate Dicer proteins. Putative Dicer proteins were selected with "E < 1×10^{-20}, identity > 30%, coverage > 70%". Maximum likelihood trees using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments of proteins are the consensus of 1,000 replicates. Numbered nodes show the percentage of bootstraps which support the node, bootstraps with less than 50% support are not shown. The four clades of DICER-like (DCL) plant proteins are color-coded. Orange font color represents *A. thaliana*, pink font color represents *T. lepidozioides* and blue font color represents *P. patens*. Shade colors as shown on the right.



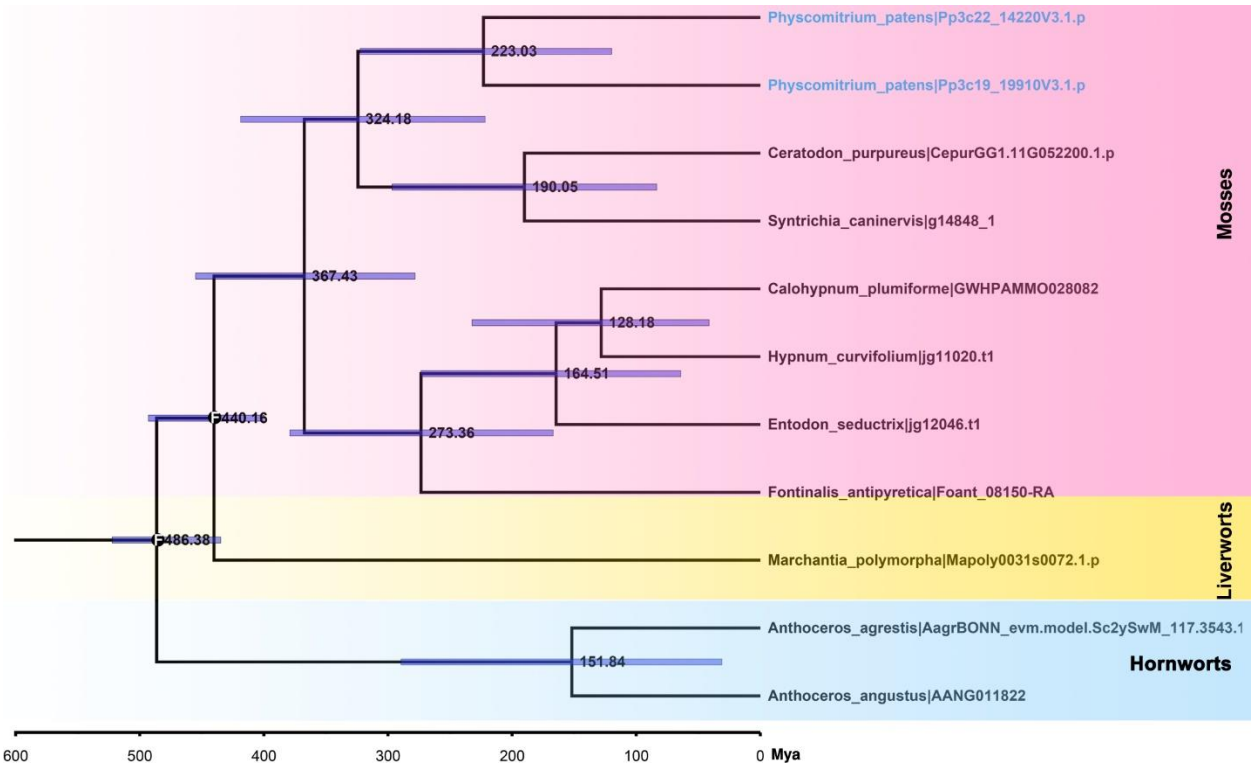
205

206

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

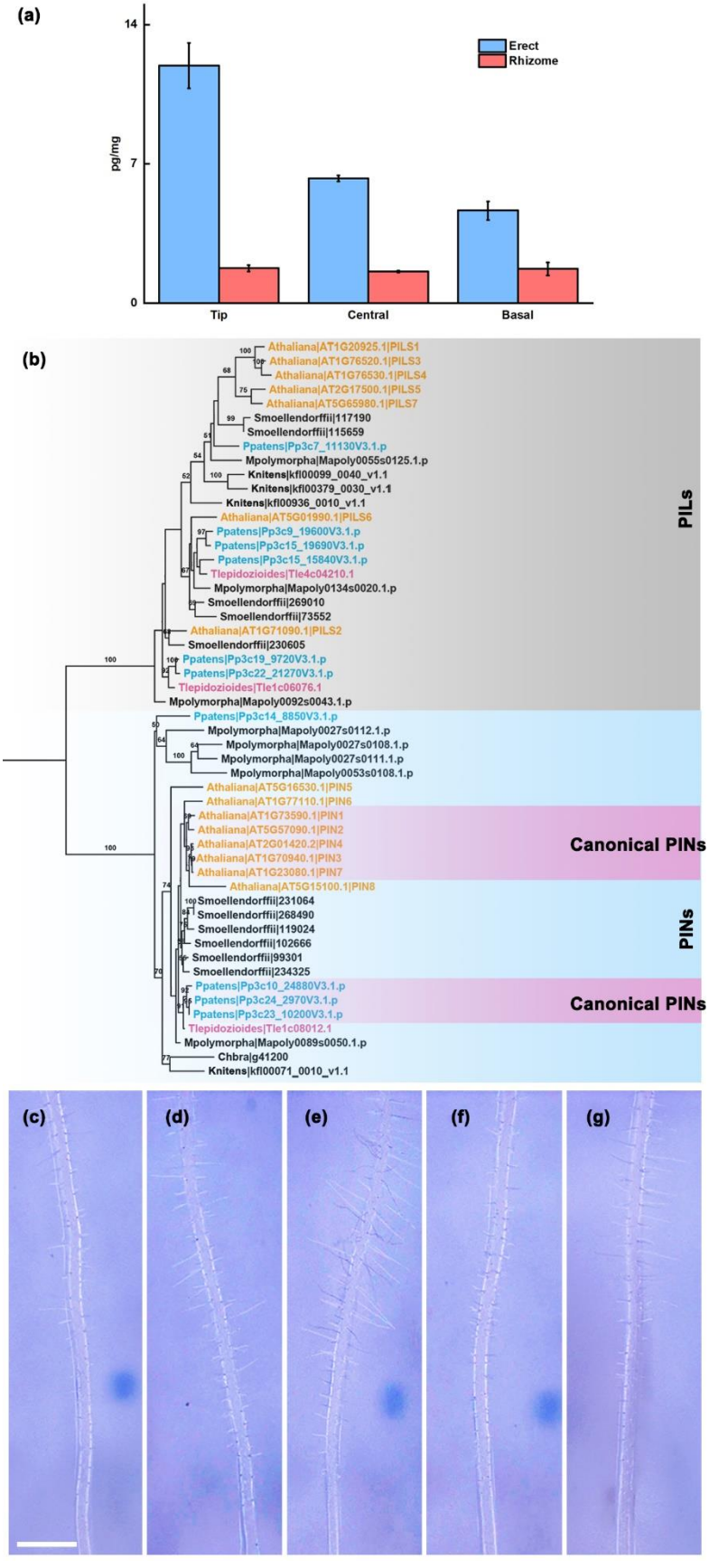
208 RHD6 (AT1G66470) homologs were found using blastp with identity cutoff 50%, e-value cutoff
209 1×10^{-20}. Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon
210 MAFFT alignments of proteins and trimmed by trimAl. Tree is the consensus of 1,000 replicates.
211 Numbers on nodes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

213



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

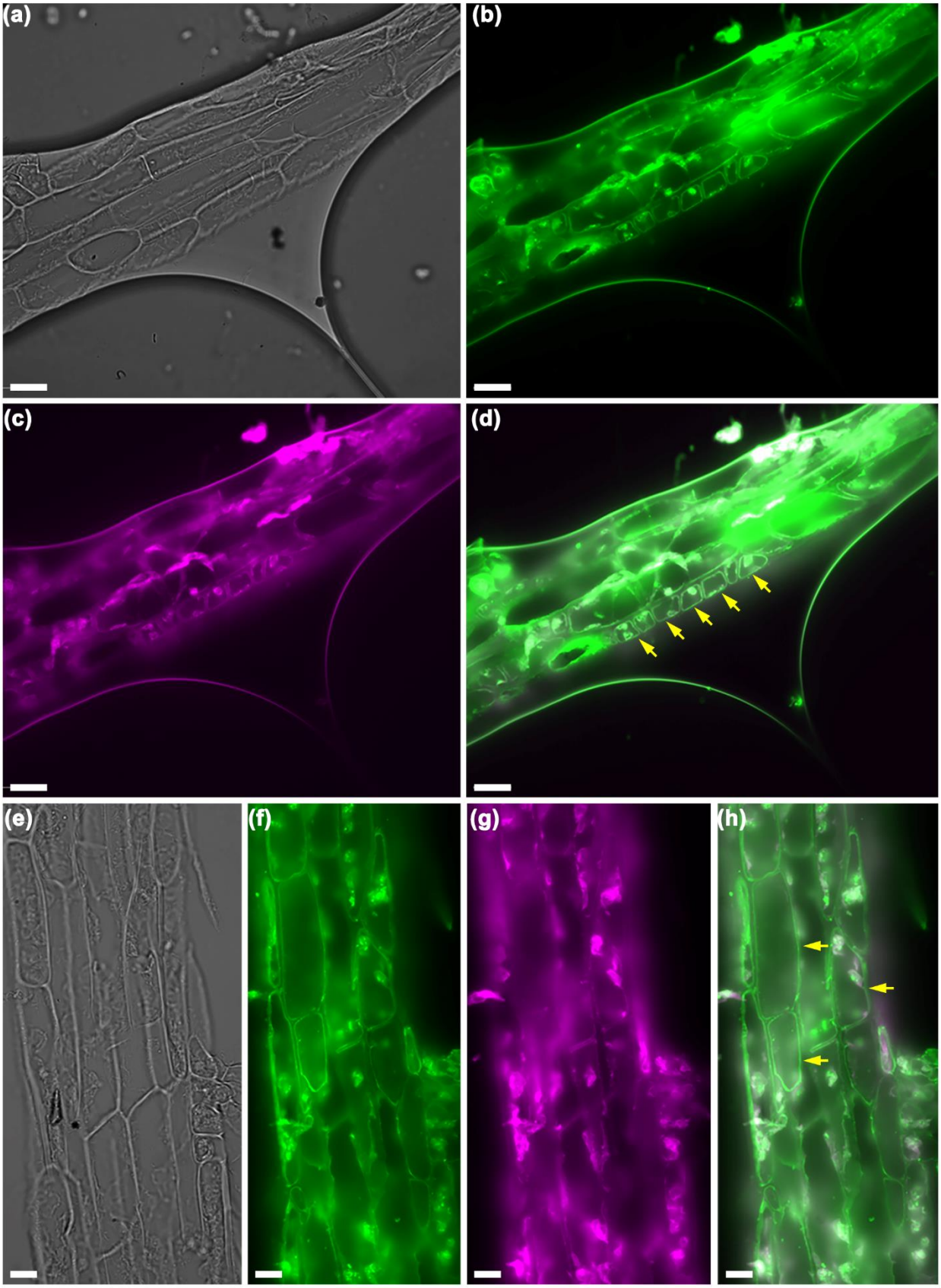
FAMA (AT3G24140) homologs were found using blastp with identity cutoff 30%, e-value cutoff 1×10^{-20}. Then HMMER used to check bHLH domain in FAMA-homolog candidates. Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments of proteins and trimmed by trimAl. Tree is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


(a)
14
7
0
pg/mg
Erect
Rhizome
Tip
Central
Basal
(b)
Athaliana|AT1G20925.1|PILS1
Athaliana|AT1G76520.1|PILS3
Athaliana|AT1G76530.1|PILS4
Athaliana|AT2G17500.1|PILS5
Athaliana|AT5G65980.1|PILS7
Smoellendorffii|117190
Smoellendorffii|115659
Ppatens|Pp3c7_11130V3.1.p
Mpolymorpha|Mapoly0055s0125.1.p
Knitens|kfl00099_0040_v1.1
Knitens|kfl00379_0030_v1.1
Knitens|kfl00936_0010_v1.1
Athaliana|AT5G01990.1|PILS6
Ppatens|Pp3c9_19600V3.1.p
Ppatens|Pp3c15_19690V3.1.p
Ppatens|Pp3c15_15840V3.1.p
Tlepidozioides|Tle4c04210.1
Mpolymorpha|Mapoly0134s0020.1.p
Smoellendorffii|269010
Smoellendorffii|73552
Athaliana|AT1G71090.1|PILS2
Smoellendorffii|230605
Ppatens|Pp3c19_9720V3.1.p
Ppatens|Pp3c22_21270V3.1.p
Tlepidozioides|Tle1c06076.1
Mpolymorpha|Mapoly0092s0043.1.p
PILs
Ppatens|Pp3c14_8850V3.1.p
Mpolymorpha|Mapoly0027s0112.1.p
Mpolymorpha|Mapoly0027s0108.1.p
Mpolymorpha|Mapoly0027s0111.1.p
Mpolymorpha|Mapoly0053s0108.1.p
Athaliana|AT5G16530.1|PIN5
Athaliana|AT1G77110.1|PIN6
Athaliana|AT1G73590.1|PIN1
Athaliana|AT5G57090.1|PIN2
Athaliana|AT2G01420.2|PIN4
Athaliana|AT1G70940.1|PIN3
Athaliana|AT1G23080.1|PIN7
Canonical PINs
Athaliana|AT5G15100.1|PIN8
Smoellendorffii|231064
Smoellendorffii|268490
Smoellendorffii|119024
Smoellendorffii|102666
Smoellendorffii|99301
Smoellendorffii|234325
PINs
Ppatens|Pp3c10_24880V3.1.p
Ppatens|Pp3c24_2970V3.1.p
Ppatens|Pp3c23_10200V3.1.p
Canonical PINs
Tlepidozioides|Tle1c08012.1
Mpolymorpha|Mapoly0089s0050.1.p
Chbra|g41200
Knitens|kfl00071_0010_v1.1
(c)
(d)
(e)
(f)
(g)

224

Data S2 T. IAA content and PINs

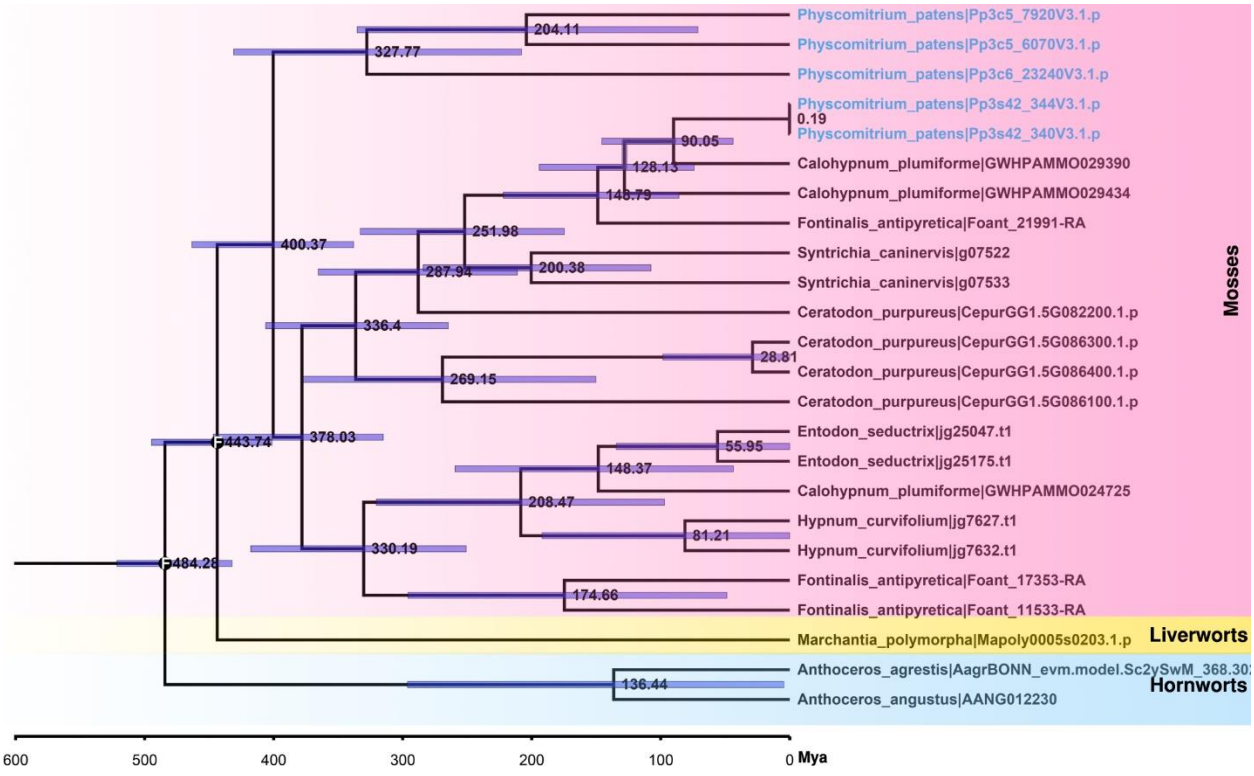
(**a**) Gradient of IAA content from top to base in *T. lepidozioides* stems (mean ± s.d.; n = 5). Auxin concentrations were highest at the top and lowest at the stem base, and there was no detectable gradient in rhizomes. (**b**) Phylogenetic tree of PINs. Maximum likelihood tree using RAxML with the PROTGAMMAJTT model based upon MAFFT alignments of proteins are the consensus of 1,000 replicates. Numbered nodes show the percentage of bootstraps which support the node. Orange font color represents *A. thaliana*, pink font color represents *T. lepidozioides* and blue font color represents *P. patens*. Shade colors as shown on the right. (**c – g**) The root phenotypes of *TlPIN1-3* transgenic *A. thaliana* plants, bar = 200 µm. (**c**) TlPIN1, (**d**) TlPIN2, (**e**) TlPIN3, (**f**) PpPINB, (**g**) AtPIN2.



237

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

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



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

SUPPLEMENTAL REFERENCES

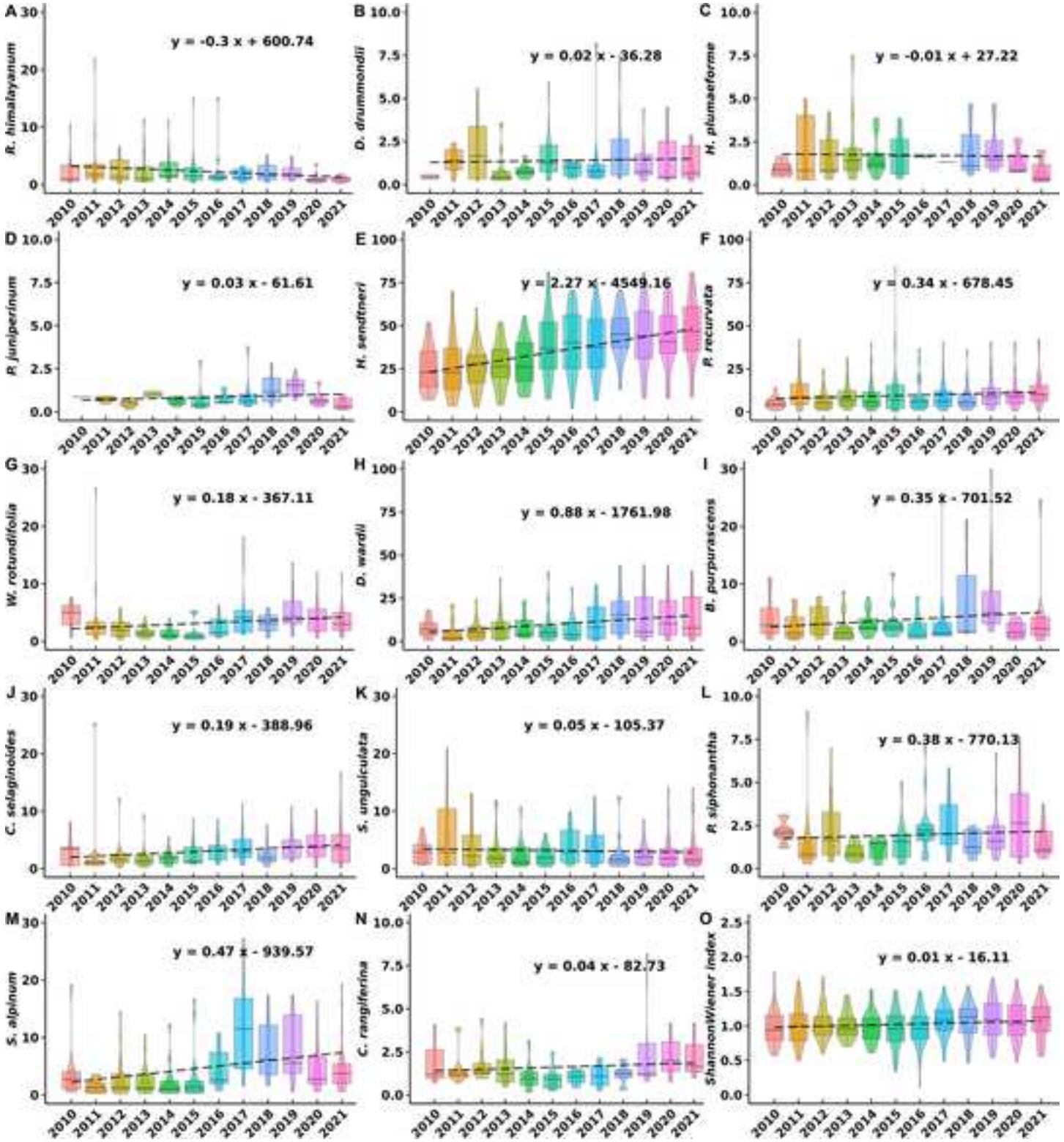
S1. Reski, R., Faust, M., Wang, X.H., Wehe, M., and Abel, W.O. (1994). Genome analysis of the moss *Physcomitrella patens* (Hedw.) B.S.G. Mol. Gen. Genet. *244*, 352-359.

S2. Marcais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. Bioinformatics *27*, 764–770. 10.1093/bioinformatics/btr011.

S3. Vurture, G.W., Sedlazeck, F.J., Nattestad, M., Underwood, C.J., Fang, H., Gurtowski, J., and Schatz, M.C. (2017). GenomeScope: fast reference-free genome profiling from short reads. Bioinformatics *33*, 2202–2204. 10.1093/bioinformatics/btx153.

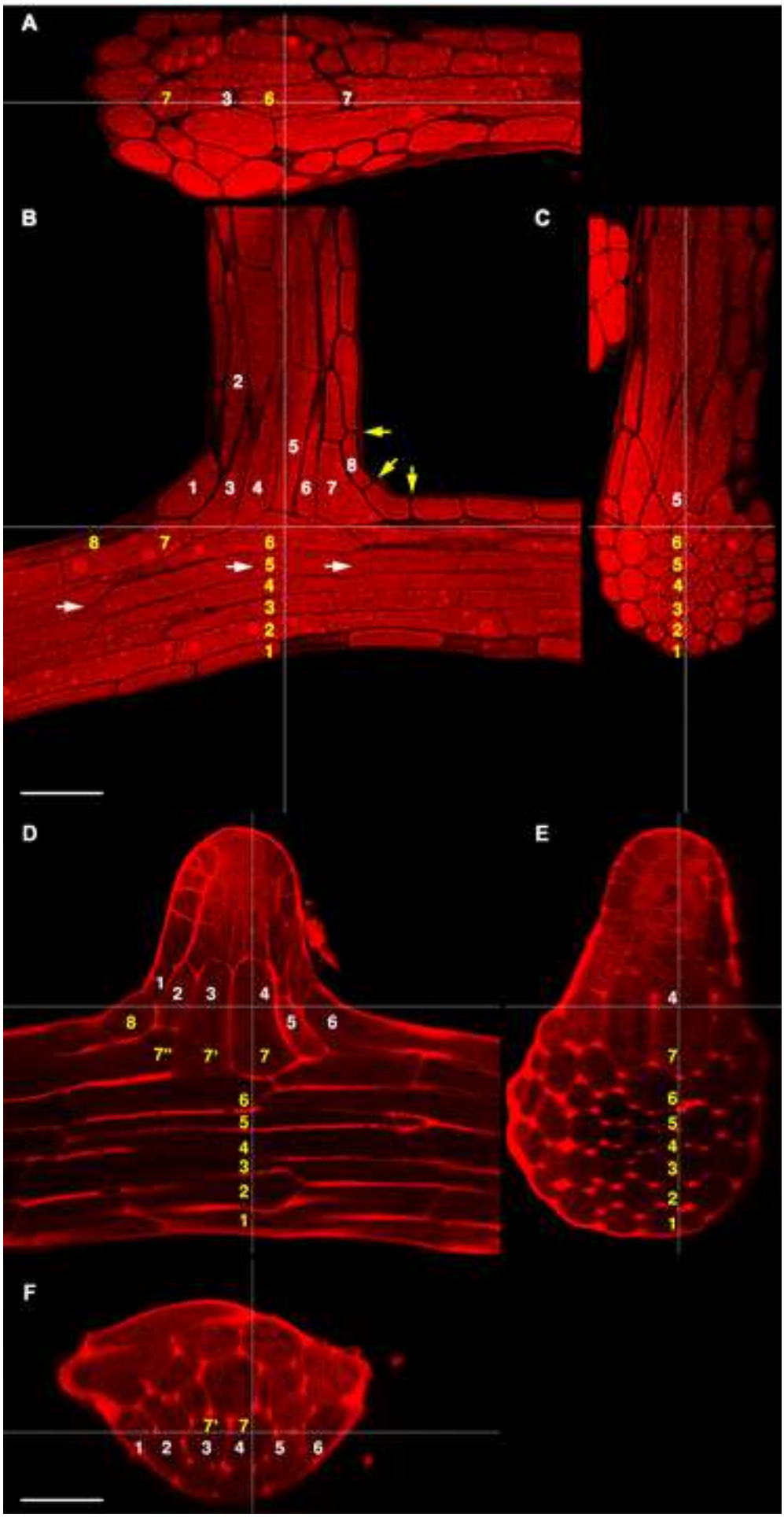
Supplemental Figure 1

Click here to access/download;Supplemental Figure;submit-FigureS1-17.4cm.tif



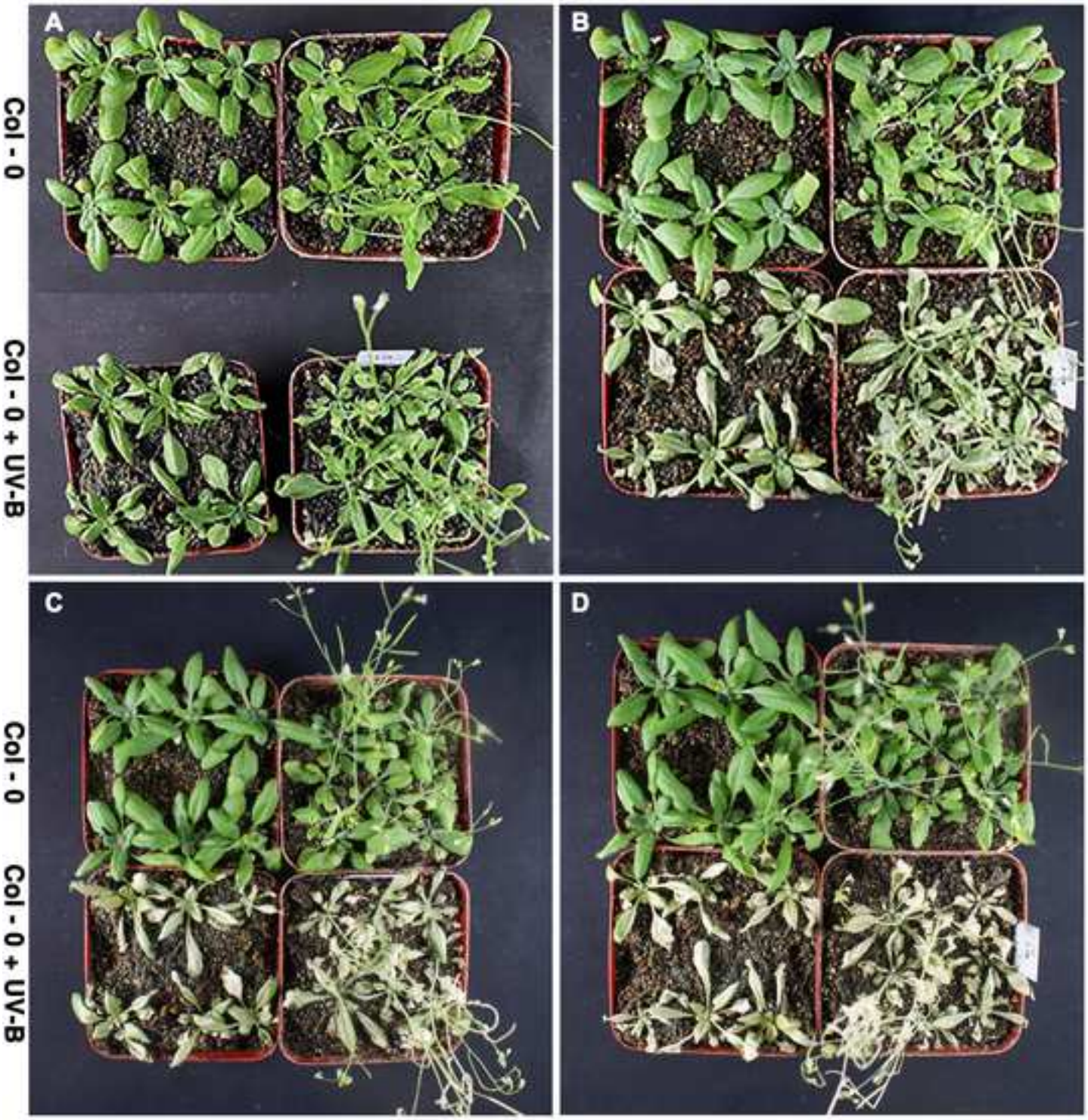
Supplemental Figure 2

Click here to access/download;Supplemental Figure;submit-FigureS2-24.7cm.tif



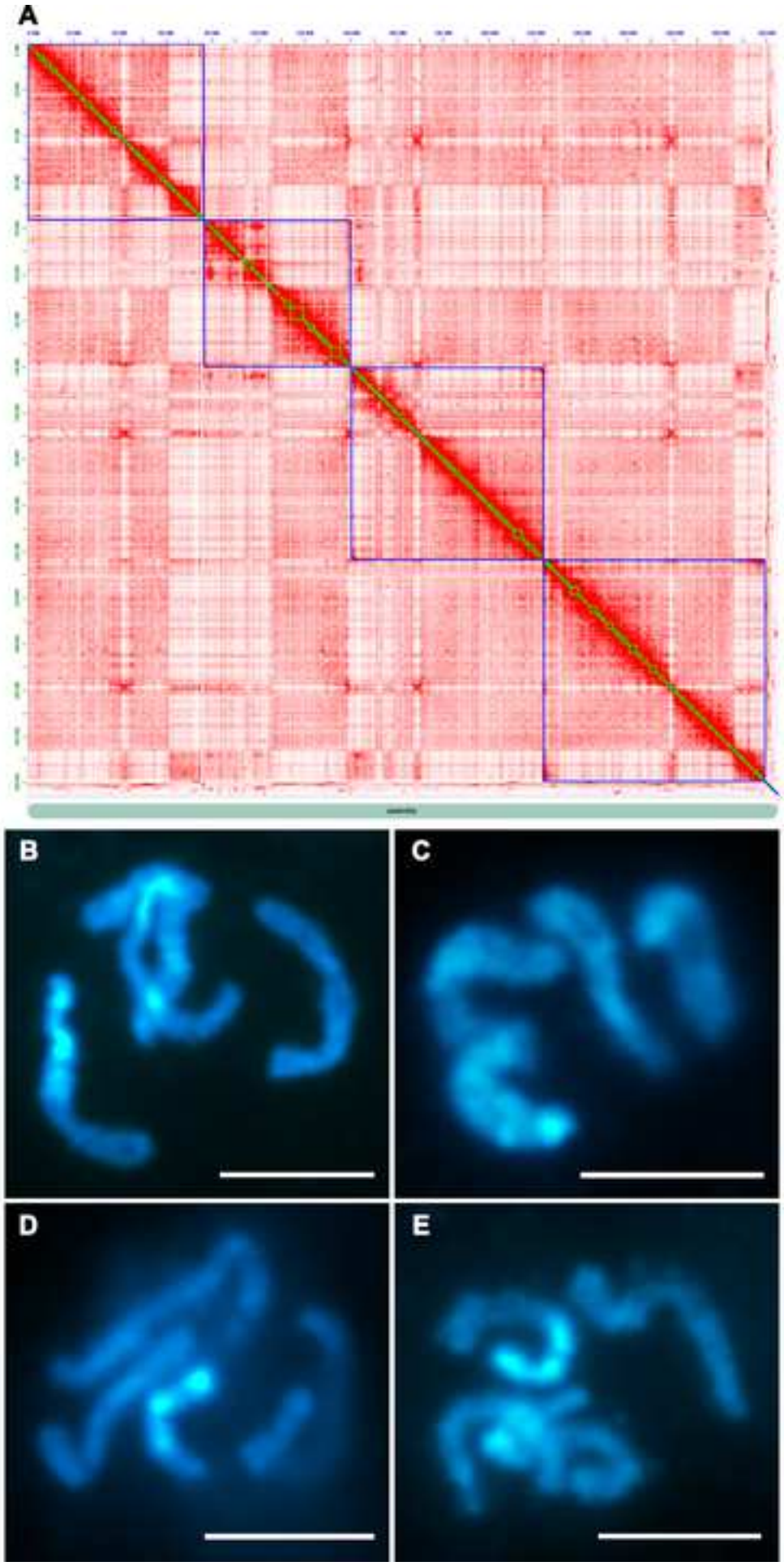
Supplemental Figure 3

Click here to access/download;Supplemental Figure;submit-FigureS3-17.4cm.tif



Supplemental Figure 4

Click here to access/download;Supplemental Figure;submit-FigureS4-24.7cm.tif



Supplemental Figure 5

Click here to access/download;Supplemental Figure;submit-FigureS5-17.4cm.tif