



**HAL**  
open science

## Stomatal regulators are co-opted for seta development in the astomatous liverwort *Marchantia polymorpha*

Kenta Moriya, Makoto Shirakawa, Jeanne Loue-Manifel, Yoriko Matsuda, Yen-Ting Lu, Kentaro Tamura, Yoshito Oka, Tomonao Matsushita, Ikuko Hara-Nishimura, Gwyneth Ingram, et al.

### ► To cite this version:

Kenta Moriya, Makoto Shirakawa, Jeanne Loue-Manifel, Yoriko Matsuda, Yen-Ting Lu, et al.. Stomatal regulators are co-opted for seta development in the astomatous liverwort *Marchantia polymorpha*. 2023. hal-03812124

**HAL Id: hal-03812124**

**<https://cnrs.hal.science/hal-03812124v1>**

Preprint submitted on 27 Oct 2023

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

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

1 **Stomatal regulators are co-opted for the development of setae in the**  
2 **astomatous liverwort *Marchantia polymorpha***

3  
4 **Authors**

5 Kenta C. Moriya<sup>1</sup>, Makoto Shirakawa<sup>2</sup>, Jeanne Loue-Manifel<sup>3,4</sup>, Yoriko Matsuda<sup>5</sup>, Yen-  
6 Ting Lu<sup>2,4</sup>, Kentaro Tamura<sup>6</sup>, Yoshito Oka<sup>1</sup>, Tomonao Matsushita<sup>1</sup>, Ikuko Hara-  
7 Nishimura<sup>7</sup>, Gwyneth Ingram<sup>3</sup>, Ryuichi Nishihama<sup>5,8</sup>, Justin Goodrich<sup>4</sup>, Takayuki  
8 Kohchi<sup>5</sup>, Tomoo Shimada<sup>1,\*</sup>

9  
10 **Affiliations**

11 <sup>1</sup>Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

12 <sup>2</sup>Graduate School of Biological Sciences, Nara Institute of Science and Technology  
13 (NAIST), Ikoma 630-0192, Japan

14 <sup>3</sup>Laboratoire Reproduction et Développement des Plantes, ENS de Lyon, CNRS,  
15 INRAE, UCB Lyon 1, Lyon 69342, France

16  
17 <sup>4</sup>Institute of Molecular Plant Sciences, University of Edinburgh, Daniel Rutherford  
18 Building, Max Born Crescent, Edinburgh EH9 3BF, UK

19 <sup>5</sup>Graduate School of Biostudies, Kyoto University, Kyoto 606-8502, Japan

20 <sup>6</sup>School of Food and Nutritional Sciences, University of Shizuoka, Shizuoka 422-8526,  
21 Japan

22 <sup>7</sup>Faculty of Science and Engineering, Konan University, Kobe 658-8501, Japan

23 <sup>8</sup>Department of Applied Biological Science, Faculty of Science and Technology, Tokyo  
24 University of Science, Noda 278-8510, Japan

25  
26 \*Correspondence: tshimada@gr.bot.kyoto-u.ac.jp

27  
28 **Abstract**

29 The evolution of special types of cells requires the acquisition of new gene regulatory  
30 networks controlled by transcription factors (TFs). In stomatous plants, stomatal  
31 formation is regulated by a TF module formed by subfamilies Ia and IIIb basic helix-  
32 loop-helix TFs (Ia-IIIb bHLH); however, how this module evolved during land plant  
33 diversification remains obscure. Here, we show that, in the astomatous liverwort  
34 *Marchantia polymorpha*, a Ia-IIIb bHLH module regulates the development of a unique  
35 sporophyte tissue, the seta, which is found in mosses and liverworts. The sole Ia bHLH  
36 gene, MpSETA, and a IIIb bHLH gene, MpICE2, regulate cell division and/or  
37 differentiation of seta lineage cells. MpSETA can partially replace the stomatal function  
38 of Ia bHLH TFs in *Arabidopsis thaliana*, suggesting that a common regulatory  
39 mechanism underlies setal and stomatal formation. Our findings reveal the co-option of  
40 a Ia-IIIb bHLH TF module for regulating cell fate determination and/or cell division of  
41 distinct types of cells during the evolution of land plant.

43 Land plants developed unique types of cells and tissues to adapt to the terrestrial  
44 environment during evolution<sup>1,2</sup>. The acquisition of novel cells or tissues leading to  
45 complex body plans is related to the diversification of transcription factors (TFs)<sup>3,4</sup>. Basic  
46 helix-loop-helix (bHLH) TFs represent a TF superfamily that plays key roles in cell fate  
47 determination and cell division during the development of eukaryotes. In land plants, the  
48 number of genes encoding bHLH TFs has increased compared with those of chlorophyte  
49 and charophyte algae, suggesting that bHLH TFs are involved in terrestrial adaptation<sup>5</sup>.  
50 For example, the acquisition of stomata, a special tissue for gas exchange on the epidermis,  
51 is an important adaptation of plants to terrestrial environments; previous studies have  
52 revealed that stomatal formation is regulated by bHLH TFs as master TFs<sup>6-9</sup>. In  
53 *Arabidopsis thaliana*, three TFs belonging to subfamily Ia (SPEECHLESS [SPCH],  
54 MUTE, and FAMA)<sup>6-8</sup> form heterodimers with subfamily IIIb TFs (ICE1, also known as  
55 SCREAM [SCRM], and ICE2, also known as SCRM2) to promote stomatal formation by  
56 regulating downstream gene expression<sup>9</sup>. The molecular mechanism of stomatal  
57 formation by Ia and IIIb bHLH TFs is conserved in the moss *Physcomitrium patens*, in  
58 which stomata play an important role in spore dispersal by promoting dehydration and  
59 dehiscence of the sporangium<sup>10,11</sup>.

60 Recent studies revealed that a gene encoding a Ia bHLH is present in the genome  
61 of the astomatous liverwort *Marchantia polymorpha*<sup>4,12-14</sup>. Despite the importance of Ia  
62 bHLH in stomatal development, its function in plants lacking stomata has not yet been  
63 explored. Here, we report that the Ia bHLH protein, designated as MpSETA, is a master  
64 regulator of the formation of the seta, which is a diploid tissue involved in long-distance  
65 spore dispersal in *M. polymorpha*. Furthermore, we show that Ia and IIIb bHLH positively  
66 regulate setal formation by heterodimerization, similar to their role in stomatal formation  
67 in other land plants. This outcome not only advances our understanding of the  
68 mechanisms of the evolution of plant tissue formation but also provides new insights into  
69 the co-option of gene expression regulatory networks (GRNs).

## 71 Results

### 72 MpSETA is the sole Ia bHLH protein in *M. polymorpha*.

73 To identify Ia bHLH coding genes in *M. polymorpha*, we constructed a phylogenetic tree  
74 of Ia bHLH proteins from various plant species using the bHLH domain and a C-terminal  
75 conserved domain called SMF (also known as the ACT-like domain)<sup>15,16</sup> (Fig. 1a and  
76 Extended Data Fig. 1). Our phylogenetic analysis suggested that MpBHLH35  
77 (Mp2g04160) is the sole Ia bHLH in *M. polymorpha* (Fig. 1a). We named this gene  
78 MpSETA based on its specific expression in the seta tissue of the sporophyte (see below).  
79 Multiple alignments of bHLH proteins revealed that the amino acid residues predicted to  
80 be important for E-box (CANNTG) binding and bHLH dimerization are highly conserved  
81 in MpSETA, although the amino acid sequence of the bHLH domain of MpSETA is  
82 relatively divergent compared to other Ia bHLH (Extended Data Fig. 1b). Additionally,  
83 we found partial sequences of two MpSETA related genes (LcSETA1 and LcSETA2) in the  
84 genome of the Marchantiopsida liverwort *Lunularia cruciata*<sup>17</sup>, although there is no  
85 evidence that these putative MpSETA-like genes are expressed and are functional in this  
86 species (Fig. 1a and Extended Data Fig. 1b-c). The amino acid sequences of the bHLH  
87 and SMF domains are well conserved between MpSETA and LcSETA1. Thus, we can  
88 conclude that Ia bHLH genes are conserved in the genome of Marchantiales liverworts.

89 Since the amino acid sequence of MpSETA is divergent, it is unclear if MpSETA

90 has similar properties as the other Ia bHLH proteins. Therefore, we investigated whether  
91 MpSETA can act as a stomatal regulator replacing AtSPCH, AtMUTE, or AtFAMA. In  
92 this context, MpSETA was expressed under the native promoters of AtSPCH, AtMUTE,  
93 and AtFAMA in *spch-3*, *mute-2*, and *fama-1* backgrounds, respectively (Fig. 1b and  
94 Extended Data Fig. 2a,b). Even though *mute-2* results in arrested meristemoids (self-  
95 renewing stomatal precursors), a few stomata were formed in *mute-2* expressing MpSETA  
96 (Fig. 1b) ( $2.67 \pm 1.56$  and  $2.33 \pm 1.72$  per abaxial side of the cotyledon in the lines #8-4  
97 and #10-11, respectively [mean  $\pm$  s.d.;  $n = 12$ ]). Notably, hydathode pores (a modified  
98 form of stomatal pores) were often found in these lines. This might be due to the high  
99 expression activity of the AtMUTE promoter in the hydathode of the cotyledons<sup>18</sup>.  
100 MpSETA also exhibited the potential to rescue *fama-1*. *A. thaliana fama* mutant displays  
101 caterpillar-like stomatal-lineage cells called “*fama* tumors,” where the terminal  
102 symmetric division occurs more than once<sup>8</sup>. In *fama-1* expressing MpSETA, excess cell  
103 divisions in stomatal lineage were suppressed, although no mature stomata were formed  
104 (Extended Data Fig. 2b,c). Neither stomata nor stomatal-lineage cells were found in *spch-3*  
105 expressing MpSETA (Extended Data Fig. 2a). These results showing that MpSETA is  
106 partially functional in stomatal-cell division and differentiation in *A. thaliana* suggest that  
107 it can interact with AtICE1 and AtSCRM2. Therefore, we tested this ability in yeast two-  
108 hybrid assays and bimolecular fluorescent complementation (BiFC) assays. We found that  
109 MpSETA physically interacted with AtICE1 and AtSCRM2 (Extended Data Fig. 2d,e).  
110 Thus, our findings indicate that MpSETA from the astomatous liverwort is a bona fide Ia  
111 bHLH TF, albeit its amino acid sequence is divergent.

112

### 113 **MpSETA is preferentially expressed in developing sporophyte.**

114 To investigate the expression pattern of MpSETA in *M. polymorpha*, we reanalyzed the  
115 public RNA-seq dataset from several organs<sup>4,19-22</sup> and found that MpSETA was  
116 preferentially expressed in the diploid sporophyte, whereas its expression level was low  
117 in the haploid gametophyte (Fig. 2a).

118 We examined in detail the expression pattern of MpSETA by characterizing the  
119 different stages of development in the *M. polymorpha* sporophyte (Fig. 2b-c); in wild type,  
120 sporophytes are divided into three tissues: foot, seta, and sporangium (Fig. 2b). The foot  
121 plays a key role in nutrient transport between the gametophyte and the sporophyte, while  
122 the seta, which consists of files of elongated cells that form a stalk suspending the  
123 sporangium, plays a pivotal role in spore dispersal<sup>23-25</sup>. After late spore maturation in  
124 sporophyte development, the seta extends through elongation of its cells and thrusts the  
125 sporangium outwards, causing it to break through the surrounding calyptra and  
126 pseudoperianth, which are tissues derived from the archegonium and the archegoniophore,  
127 respectively, to protect the sporophyte during development<sup>23-25</sup>. As a result, spores are  
128 dispersed after sporangium dehiscence by desiccation. Based on the unique  
129 developmental events as shown in previous studies<sup>22,23,26</sup>, we divided the development of  
130 the sporophyte into 10 stages, as follows (Fig. 2c,d): (I) 2-cell stage with the epibasal cell  
131 (the upper cell that forms the foot and the seta) and the hypobasal cell (the lower cell that  
132 forms the sporangium), (II) 4 or 8-cell stage, (III) early-globular stage with a distinct  
133 amphithecium (the outermost tissue that forms the capsule wall) and the endothecium (the  
134 inner archesporial tissue); (IV) late-globular stage; (V) archesporial-tissue stage with the  
135 visible differentiated foot, seta, and archesporial tissues; (VI) sporogenous-cell stage with  
136 elaterocytes and sporogenous cells, which are the precursors of elaters and sporocytes

137 (spore mother cells); (VII) sporocyte stage; (VIII) spore-tetrad stage (after meiosis); (IX)  
138 matured stage; (X) seta-elongated stage. Note that seta and foot are established between  
139 stages V and IX. In stage VIII, we anatomically observed proliferative (symmetric) cell  
140 divisions in the seta region (Fig. 2e). Moreover, cell division activity was detected not in  
141 the foot but rather in the seta of sporophytes between stages VII and VIII using a G2-M  
142 phase reporter line, *proMpCYCB;1:Dbox-GUS*<sup>27</sup> (Fig. 2f). Thus, we concluded that the cell  
143 files in the seta are established by a few proliferative cell divisions of the putative “seta  
144 mother cells” in the later developmental stage.

145 We detected the promoter activity of *MpSETA* in the sporophyte by generating  
146 transformants expressing the  $\beta$ -glucuronidase gene (*GUS*) under the control of the  
147 *MpSETA* promoter (*proMpSETA:GUS*). *GUS* activity was found in developing  
148 sporophytes between stages IV and VII, especially in seta, whereas no *GUS* activity was  
149 found in stages VIII and IX (Fig. 2g). In gametophytic tissues, *GUS* signals were detected  
150 only in young antheridia (Extended Data Fig. 3). These findings suggest that *MpSETA* is  
151 expressed early in seta development and may regulate seta cell division and/or  
152 differentiation rather than the elongation of seta cells.

#### 153 ***Mpseta*<sup>ko</sup> mutants show defects in setal formation in the sporophyte.**

154 We generated loss-of-function mutants of *MpSETA* by homologous recombination-  
155 mediated gene targeting to reveal the function of *MpSETA* *in vivo*<sup>28</sup>. Therefore, we  
156 obtained two independent *MpSETA* knock-out lines (*Mpseta-1*<sup>ko</sup> and *Mpseta-2*<sup>ko</sup>) in  
157 which the genomic regions encoding the bHLH domain were replaced with a  
158 hygromycin-resistance gene cassette (Extended Data Fig. 4a-c). We confirmed the loss of  
159 the full-length transcripts of the *MpSETA* gene by reverse-transcription polymerase chain  
160 reaction (RT-PCR) analysis of homozygous mutant sporophytes produced from crosses  
161 between *Mpseta-1*<sup>ko</sup> or *Mpseta-2*<sup>ko</sup> males and females (Extended Data Fig. 4d). Although  
162 *MpSETA* reporter gene expression was found in the developing antheridia (Extended Data  
163 Fig. 3), no obvious phenotype was observed during sperm formation in *Mpseta*<sup>ko</sup> lines  
164 (Extended Data Fig. 4e), and mutant males were fertile.

165 We crossed males and females of *Mpseta*<sup>ko</sup> and compared the resulting  
166 sporophytes with those of the wild type to investigate phenotypes of *Mpseta*<sup>ko</sup> mutants in  
167 the diploid generation. In longitudinal sections of mature sporophytes of *Mpseta*<sup>ko</sup>, we  
168 found anatomical defects in the development of setal cells (Fig. 3a,b). We did not observe  
169 any elongated seta cells or cell files of seta cells in *Mpseta*<sup>ko</sup> mutants. The SF/SP ratio (a  
170 ratio of the length from the foot to the proximal side of sporangium to the total sporophyte  
171 length) and the number of cells around the seta region were significantly reduced in  
172 *Mpseta*<sup>ko</sup> compared with wild type (Fig. 3c,d). Detailed analysis of earlier stages of  
173 sporophyte development revealed that defects in *Mpseta-1*<sup>ko</sup> setae could be found even at  
174 stage VI, the earliest stage at which putative seta mother cells are unequivocally  
175 recognized (Fig. 3e). Despite the obvious loss of setae, spores and other sporophytic  
176 tissues were normally formed in *Mpseta-1*<sup>ko</sup> (Fig. 3e,f). We concluded that *Mpseta*<sup>ko</sup>  
177 mutants have defects in the differentiation from putative seta precursor cells to seta  
178 mother cells, which prevents the induction of the subsequent proliferative cell divisions.  
179 At one-month postfertilization, the wild-type sporangia were pushed out of the calyptras.  
180 By contrast, *Mpseta*<sup>ko</sup> sporangia remained buried inside the calyptras, presumably due to  
181 the defects in seta development, and hence were not exposed to the outside (Fig. 3g).  
182 Since sporangia are completely wrapped by the calyptras and pseudoperianths, spore  
183

184 dispersal does not occur unless the sporangia are pushed out of the calyptras by elongation  
185 of seta cells<sup>29</sup>. In genetic complementation experiments, we generated a transgenic line  
186 having the genomic region of MpSETA introduced into Mpseta-1<sup>ko</sup> (gMpSETA Mpseta-  
187 1<sup>ko</sup>). The sporophytes obtained by crossing these complemented lines had normal setae,  
188 and their sporangia were pushed out of the calyptra as in the wild type (Fig. 3).  
189 Consequently, these results suggest that MpSETA is essential for setal formation and spore  
190 dispersal in *M. polymorpha*.

191

### 192 **MpICE2 physically interacts with MpSETA and regulates setal development.**

193 Since the interaction between Ia and IIIb bHLHs is evolutionarily conserved<sup>9,10,30</sup>, we  
194 examined whether Ia bHLH and IIIb bHLH physically interact with each other in *M.*  
195 *polymorpha*. We performed phylogenetic analyses to find IIIb bHLH genes in *M.*  
196 *polymorpha* and identified two genes encoding IIIb bHLH proteins in the genome of *M.*  
197 *polymorpha*, both of which are orthologous to AtICE1 and AtSCRM2 (Extended Data Fig.  
198 5). We named these genes MpICE1 (Mp4g04910) and MpICE2 (Mp4g04920). The amino  
199 acid sequences of the bHLH and ACT-like domains of IIIb bHLHs were highly conserved  
200 among land plants (Extended Data Fig. 6). In addition, two putative IIIb bHLH-encoding  
201 genes, LcICE1 and LcICE2, were found in the genome of *L. cruciata* (Extended Data  
202 Figs. 5 and 6).

203 Reanalyzing public RNA-seq data revealed high expression of MpICE2 at 13  
204 days postfertilization in young sporophytes (from stage III to stage IV)<sup>19</sup>, in which  
205 putative seta mother cells were dividing and differentiating, while the expression level of  
206 MpICE1 was almost constant in all tissues (Fig. 4a). Thus, we assumed that MpICE2 may  
207 predominantly function in cooperation with MpSETA during seta cell formation rather  
208 than MpICE1. We expressed *Citrine* (a yellow fluorescent protein; YFP), *GUS*, and  
209 nuclear localization-signal (*NLS*) fusion gene (*Citrine-GUS-NLS*) under the control of the  
210 MpICE2 promoter to confirm the tissue/cell-level expression pattern of MpICE2 in the  
211 sporophyte. In this line, GUS signals were detected overall in stages IV and V sporophytes,  
212 specifically in the seta and foot in stages VI-VIII sporophytes, and only in the foot in  
213 mature sporophytes together with Citrine and GUS signals in gametophytic tissues  
214 (Extended Data Fig. 7).

215 We performed yeast two-hybrid assays using full-length MpSETA in pairwise  
216 combinations with MpICE2 to test the interaction between MpSETA and MpICE2; in this  
217 context, a physical interaction was observed (Fig. 4b). Additionally, BiFC assay was used  
218 to test the interaction between these bHLH TFs. YFP signals were detected in the nuclei  
219 of *Nicotiana benthamiana* leaves coexpressing MpSETA-nYFP and MpICE2-cYFP (Fig.  
220 4c). These results suggest that MpSETA has the potential to interact with MpICE2 in *M.*  
221 *polymorpha*.

222 We generated two independent genome-edited lines using the CRISPR/Cas9  
223 system (Mpice2-2<sup>ge</sup> and Mpice2-6<sup>ge</sup>) to assess the function of MpICE2<sup>31,32</sup>. The first one  
224 retained two amino acid substitutions (L503H and M504L) that were predicted to be  
225 important for the DNA-binding activity of the bHLH domain<sup>33</sup>, while the second had a  
226 frame-shift mutation that caused deletion of the C-terminal half of the bHLH domain and  
227 the whole ACT-like domain (Extended Data Fig. 8). In these Mpice2 mutants, the setae  
228 of sporophytes were not formed (Fig. 4d,e) and the number of cells between foot and  
229 sporangium was significantly reduced (Fig. 4f,g). Additionally, we found that the number  
230 of cells in the seta region of Mpice2-2<sup>ge</sup> was higher than that of Mpice2-6<sup>ge</sup> (Fig. 4g). This

231 is probably because the predicted translational product of *Mpice2-2<sup>se</sup>* has a two-amino  
232 acid substitution in the bHLH domain, and might be partially functional in comparison  
233 with the null allele *Mpice2-6<sup>se</sup>*. In both *Mpice2<sup>se</sup>* mutants, the sporophytes did not emerge  
234 outside of the protective organs derived from the archegonia, similar to the *Mpseta<sup>ko</sup>*  
235 mutants (Fig. 4h). Since the single mutants of *Mpice2<sup>se</sup>* showed almost the same  
236 phenotype as that of the *Mpseta<sup>ko</sup>* mutants, *MpICE1* may not be functionally redundant  
237 with *MpICE2* at least in the setal formation. The *Mpice2<sup>se</sup>* phenotype in the setal region  
238 was completely suppressed by introducing the genomic region of *MpICE2* into *Mpice2-*  
239 *6<sup>se</sup>* (Fig. 4d-h). Therefore, we can conclude that the MpSETA-MpICE2 heterodimer plays  
240 an important role in the setal development of *M. polymorpha*.

241 Next, we tested if *MpICE2* can enhance the rescue of the stomatal phenotype of  
242 *mute-2* by *proAtMUTE:MpSETA*. Overexpression of *MpICE2* in *proAtMUTE:MpSETA*  
243 *mute-2* did not enhance stomatal formation (Extended Data Fig. 9a) ( $3.63 \pm 3.16$  and  
244  $2.64 \pm 1.41$  per abaxial side of the cotyledon in the lines #8-4 and #10-11, respectively  
245 [mean  $\pm$  s.d.;  $n = 11$ ]). Therefore, the MpSETA-MpICE2 heterodimer does not appear to  
246 regulate the expression of stomatal genes in *A. thaliana*. In addition, the expression of  
247 *MpICE1* or *MpICE2* under the control of the *AtICE1* promoter in *ice1-2 scrm2-1* mutants  
248 failed to cause stomatal-lineage cell formation (Extended Data Fig. 9b). Therefore,  
249 *MpICE1* and *MpICE2* cannot act with *AtSPCH* to regulate stomatal formation in *A.*  
250 *thaliana*, despite the similarity of their amino acid sequences with *AtICE1* (Extended  
251 Data Fig. 6).

## 252 Discussion

253 In this article, we showed that two transcription factors, MpSETA (Ia bHLH) and MpICE2  
254 (IIIb bHLH), play a pivotal role in controlling the formation of the diploid tissue seta in  
255 the sporophyte of *M. polymorpha*, which is an astomatous liverwort (Figs. 3 and 4).  
256 Similarly, in other non-liverwort land plants, a module formed by the subfamilies Ia and  
257 IIIb bHLH TFs regulates GRNs in stomatal development<sup>2</sup>. *MpSETA* could partially  
258 complement the defects of *A. thaliana mute* and *fama*, suggesting similar properties of Ia  
259 bHLH TFs from *M. polymorpha* and *A. thaliana*. However, *MpSETA* was unable to  
260 complement the *spch* mutant (Fig. 1b and Extended Data Fig. 2). These results are  
261 consistent with the previous hypothesis stating that the ancestral Ia bHLH proteins had a  
262 MUTE- and FAMA-like function<sup>34,35</sup>. Although the nature of *MpSETA* expressing cells is  
263 still unknown, *MpSETA* may function as a regulator of cell differentiation and asymmetric  
264 cell division during setal formation, similar to its role in stomatal formation (Fig. 5a).

265 Although stomata and setae are completely different morphologically, they both  
266 play a common role in promoting sporangial dehiscence and spore dispersal. In mosses  
267 and hornworts, stomata are present in the epidermis of the sporangia and function in the  
268 desiccation of the sporangia by gas exchange to promote its dehiscence and spore  
269 dispersal<sup>25,36</sup>. In *A. thaliana*, *AtICE1* controls stomatal development on anther epidermis  
270 and can regulate dehydration and dehiscence of the anther<sup>37</sup>. However, the seta is the  
271 tissue that supports the sporangia and has a role through cell elongation in thrusting the  
272 sporangia outside the surrounding maternal tissues to permit long-distance dispersal of  
273 spores<sup>24,25</sup>. In this study, we found that the lack of setae in *Mpseta<sup>ko</sup>* and *Mpice2<sup>se</sup>* mutants  
274 prevented sporangia dehiscence and spore dispersal due to the inability of the sporangia  
275 to break through the protective organs around it (Figs. 3 and 4). Thus, the Ia-IIIb bHLH  
276 module has a common role in the development of the tissues involved in spore or pollen  
277

278 dispersal.

279 Previous research suggested that land plant evolution occurred through the reuse  
280 and/or modification of preexisting GRNs<sup>20,38-43</sup>. Here, we hypothesize that the Ia-IIIb  
281 bHLH module was primarily used in stomatal formation and secondarily co-opted to setal  
282 formation in the common ancestor of “Setaphyta” (Fig. 5b). Since only mosses and  
283 liverworts have setae, a mosses-liverworts clade is called Setaphyta<sup>1,44,45</sup>. However, the  
284 process of setal development differs between mosses and liverworts; whereas a transient  
285 intercalary meristem, called the seta meristem, produces seta in mosses, the body plan,  
286 including seta, foot, and sporangium, is established by formative cell division at an early  
287 stage in liverworts<sup>25</sup>. Therefore, whether the Ia-IIIb bHLH module is involved in the  
288 formation of seta in mosses, such as *P. patens*<sup>46,47</sup>, should be tested.

289 In *A. thaliana*, the AtFAMA-AtICE1 heterodimer regulates not only the  
290 development of stomata but also of myrosin idioblasts, which are adjacent to vascular  
291 tissues and contribute to defense against herbivores<sup>41,48,49</sup>. Since myrosin cells are present  
292 only in the Brassicales, the AtFAMA-AtICE1 module for stomatal formation was co-  
293 opted for myrosin cell development during the evolution of Brassicales (Fig. 5b). The  
294 AtFAMA-AtICE1 module is thought to regulate the expression of different genes in  
295 stomatal and myrosin cell lineages<sup>41,48,49</sup>; however, the detailed mechanism is not yet  
296 known. Liverworts have already lost several stomatal-related genes<sup>4,50</sup>, such as the  
297 leucine-rich-repeat receptor-like gene *TOO MANY MOUTHS (TMM)* and the secreted  
298 peptide gene *EPIDERMAL PATTERNING FACTOR 1/2 (EPF1/2)*. The mechanisms  
299 underlying the development of setal-cell lineage might be different from the ones related  
300 to stomatal-cell lineages. AtMUTE directly regulates cell-cycle related genes (cyclin and  
301 cyclin-dependent kinase genes) and several stomatal-related TF genes<sup>51</sup>, and AtSPCH  
302 directly regulates the stomatal-related genes and brassinosteroid pathway genes<sup>52</sup>. In *M.*  
303 *polymorpha*, orthologues of many of the Ia-IIIb bHLH target genes, including *CYCB*,  
304 *CYCD*, *CDKB*, *ERECTA*, and BZR/BES family TF genes, are conserved<sup>4</sup>. RNA-seq and  
305 ChIP-seq analyses could identify genes that function downstream of the MpSETA-  
306 MpICE2 module, and help to clarify whether other stomatal formation-related genes are  
307 involved in setal formation.

308 Are *SMF* genes and/or MpSETA orthologs conserved in the Jungermanniopsida  
309 or Haplomitriopsida liverworts? A BLAST search of 1,000 plant transcriptomes  
310 (OneKP)<sup>53</sup>, using AtFAMA as query, revealed the absence of Ia bHLH TFs in liverworts  
311 except for MpSETA. Since the liverwort transcriptome samples used in OneKP often do  
312 not contain sporophytes, it is difficult to detect genes that are specifically or transiently  
313 expressed in developing sporophytes, such as Ia bHLH. Since the Ricciaceae species have  
314 lost seta, it will be important to investigate whether these species have Ia bHLH and  
315 whether the Ia bHLHs are functional to understand the evolution of setae. Genome  
316 analyses of various plant species in the future could be useful to understand the  
317 relationship between Ia bHLH diversification and stomata/seta formation.

318

## 319 **Methods**

### 320 **Phylogenetic analysis**

321 The classifications of bHLH are according to Pires and Dolan<sup>5</sup>. Amino acid sequence  
322 information was retrieved from MarpolBase (<https://marchantia.info>), Phytozome v.13  
323 (<https://phytozome-next.jgi.doe.gov/>), OneKP (<https://db.cngb.org/onekp/>), TAIR  
324 (<http://www.arabidopsis.org/>), and NCBI

325 (<https://www.ncbi.nlm.nih.gov/genome/?term=PRJNA701193>). The bHLH domain and  
326 C-terminal ACT-like domain were aligned using MAFFT<sup>54</sup> v.6.864  
327 (<https://www.genome.jp/tools-bin/mafft>) with the default parameters, and the alignment  
328 gaps were removed manually. Amino acid sequences were visualized with Jalview<sup>55</sup>  
329 v.2.11.2.1. BLAST searches of NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) were used  
330 to predict the amino acids important for nucleotide binding and dimer formation in the  
331 bHLH domains. Phylogenetic tree constructions were performed using the maximum-  
332 likelihood algorithm on MEGA 7<sup>56</sup> with the JTT+G+I substitution model for Ia and  
333 LG+G+I substitution model for IIIb bHLH. Bootstrap analyses with 1000 replicates were  
334 performed in each analysis to assess the statistical support for the topology. Subfamilies  
335 III(a+c) and III(d+e) bHLHs were chosen as the outgroup of the phylogeny for Ia and IIIb  
336 bHLH, respectively.

337

### 338 **Plant materials and growth condition**

339 Male and female accessions of *M. polymorpha* L., Takaragaike-1 (Tak-1), and Tak-2,  
340 respectively, were maintained asexually. A female progeny backcrossed to Tak-1 for three  
341 backcross generations (BC3-38) and male Tak-1 were used as wild type. Gemmae and  
342 thalli were cultured on half-strength Gamborg's B5 media containing 1% (w/v) agar and  
343 1% (w/v) sucrose under 50-60  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  continuous white light at 22°C. To  
344 induce gametangiophore development, 10 to 12-day-old thalli were transferred to 16-h-  
345 light/8-h-dark conditions with 50-60  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  white light and 50-60  $\mu\text{mol}$   
346  $\text{photons m}^{-2} \text{s}^{-1}$  far-red light emitted from diodes (IR LED STICK 18W, Namoto) at 18°C  
347 and incubated for 1 month.

348 The *Arabidopsis thaliana* Columbia-0 (Col-0) accession was used as wild type  
349 except for *mute-2* where Wassilewskija-4 (Ws-4) was used. Seeds were surface sterilized  
350 with 70% ethanol and then sown onto half-strength MS media containing 0.5% (w/v)  
351 gellan gum and 1% (w/v) sucrose. The seeds were incubated at 22°C under 50-60  $\mu\text{mol}$   
352  $\text{photons m}^{-2} \text{s}^{-1}$  continuous white light.

353

### 354 **Complementation tests of *A. thaliana* stomatal defective mutants**

355 T-DNA insertion mutants *spch-3* (SAIL\_36\_B06) and *fama-1* (SALK\_100073) were  
356 obtained from the Arabidopsis Biological Resource Center (ABRC); *mute-2*  
357 (FLAG\_225D03) from the French National Institute for Agricultural Research (INRA).  
358 *ice1-2 scrm2-1*<sup>9</sup> was provided by K.U. Torii. To construct *proAtSPCH:MpSETA*, firstly the  
359 *AtSPCH* promoter (2,572 bp upstream of the translational initiation site) was amplified  
360 from Col-0 gDNA, and the *MpSETA* CDS was amplified from cDNA derived from Tak-  
361 1 antheridiophores. *proAtSPCH* and *MpSETA* CDS fragments were fused by PCR using  
362 PrimeSTAR GXL polymerase (Takara Bio), and the resultant PCR fragment was  
363 subcloned into pENTR1A entry clones (Invitrogen) at the SalI and EcoRV restriction sites  
364 using In-Fusion HD Cloning Kit (Takara Bio) and transferred into the destination vector  
365 pGWB501<sup>57</sup> using Gateway LR Clonase II Enzyme mix (Thermo Fisher Scientific).  
366 *proAtMUTE:MpSETA* and *proAtFAMA:MpSETA* constructs were generated by LR  
367 recombination of the R4pGWB501<sup>58</sup> or R4pGWB601<sup>58</sup> with a pENTR1A containing the  
368 *MpSETA* CDS at the EcoRI sites and either pENTR5'/TOPO\_ *proAtMUTE* harboring  
369 *proAtMUTE*, 1,930 bp upstream of the translational initiation site, or  
370 pENTR5'/TOPO\_ *proAtFAMA* harboring *proAtFAMA*, 3,105 bp upstream of the  
371 translational initiation site (R4pGWB501\_ *proAtMUTE:MpSETA* and

372 R4pGWB601\_proAtFAMA:MpSETA). To construct *proAtICE1:MpICE1* or  
373 *proAtICE1:MpICE2*, firstly the *AtICE1* promoter (2,578 bp upstream of the translational  
374 initiation site) was amplified from Col-0 gDNA, and *MpICE1* or *MpICE2* CDS was  
375 amplified from cDNA derived from Tak-1 thalli. *proAtICE1* and *MpICE1* or *MpICE2* CDS  
376 fragments were fused by PCR using PrimeSTAR GXL polymerase, and the resultant PCR  
377 fragment was subcloned into pENTR1A entry clones at the Sall and EcoRI restriction  
378 sites using an In-Fusion HD Cloning Kit and transferred into the destination vector  
379 pGWB501<sup>57</sup> using Gateway LR Clonase II Enzyme mix. The resultant plasmids were  
380 introduced into *spch-3/+*, *mute-2/+*, *fama-1/+*, or *ice1-2/+ scrm2-1* heterozygous plants  
381 by the previously described method<sup>59</sup> using *Agrobacterium tumefaciens* strain GV3101.  
382 We confirmed that all transformants had a single insertion event by segregation analyses.  
383 We used T<sub>3</sub> or T<sub>4</sub> homozygous plants. For *MpICE2* overexpression analyses, the *MpICE2*  
384 CDS was transferred into the destination vector pFAST-R02<sup>60</sup> using Gateway Clonase II  
385 Enzyme mix, and the resultant plasmid was introduced into Ws-4, *mute-2/+*, and  
386 *proAtMUTE:MpSETA mute-2/+* (#8-4 and #10-11). T<sub>1</sub> seeds expressing TagRFP were  
387 selected, and T<sub>1</sub> plants were used to the analyses. We stained the cotyledons by FM4-64  
388 and observed them using an LSM780 laser scanning microscope (Carl Zeiss). Images  
389 were processed with Fiji (NIH). The sequences of primers used in this study are shown  
390 in Supplementary Tables 1 and 2.

391

### 392 **Gene expression analysis**

393 Publicly available transcriptome data were downloaded from the Sequence Read Archive  
394 (SRA) repository. Accession numbers include: sporelings (SRR4450254, SRR4450255,  
395 SRR4450256)<sup>4</sup>, male thalli (DRR118949, DRR118950, DRR118951)<sup>20</sup>, female thalli  
396 (DRR118943, DRR118944, DRR118945)<sup>20</sup>, antheridiophores (DRR050346,  
397 DRR050347, DRR050348)<sup>21</sup>, archegoniophores (DRR050351, DRR050352,  
398 DRR050353)<sup>21</sup>, antheridia (DRR050349, DRR050350)<sup>21</sup>, archegonia (DRR209029,  
399 DRR209030, DRR209031, DRR209031)<sup>22</sup>, 13 DPF sporophytes (SRR1553297,  
400 SRR1553298, SRR1553299)<sup>19</sup>, mature sporophytes (SRR896223)<sup>4</sup>. RNA-seq data were  
401 preprocessed to filter out low-quality sequences using fastp<sup>61</sup> v.0.20.0. The sequence  
402 reads were mapped to the *M. polymorpha* genome v.6.1 (<https://marchantia.info>) by  
403 STAR<sup>62</sup> v.2.7.8a with default parameters. The post-processing of SAM/BAM files was  
404 performed using SAMtools<sup>63</sup> v.1.11. The read counts for each gene were used to calculate  
405 transcript per million (TPM) by using RSEM<sup>64</sup> v.1.3.0 with default parameters. Plots were  
406 created using Rstudio v.1.4.1106 (<https://www.rstudio.com/>).

407

### 408 **Histology**

409 Sporophytes (stage I and II) were stained with 4',6-diamidino-2-phenylindol (DAPI) as  
410 described previously<sup>22</sup> and were observed using an LSM780 laser scanning microscope.

411 Plant samples were fixed with 2% (w/v) paraformaldehyde and 2% (v/v)  
412 glutaraldehyde in 0.05 M cacodylate buffer (pH 7.4) for 2 hours at room temperature,  
413 post-fixed with 2% (v/v) osmium tetroxide in 0.1 M cacodylate buffer for 2 hours at room  
414 temperature, dehydrated in an ethanol series, substituted with acetone, and then embedded  
415 in Spurr's resin (Polysciences). The Spurr's blocks were cut into semi-thin sections (0.75-  
416 2 μm) with glass knives on an ultramicrotome Leica Ultracut UCT (Leica Microsystems)  
417 and stained with a solution containing 1% (w/v) sodium tetra-borate and 1% (w/v)  
418 toluidine blue O. Sections were mounted on MAS-coated glass slides (Matsunami Glass).

419 Images were obtained using a VB-7010 (KEYENCE)/AxioCam HRc (Zeiss) and were  
420 processed with Fiji or Adobe Photoshop Elements 9 (Adobe Systems).

421

### 422 **Histochemical GUS staining**

423 To construct *proMpSETA:GUS*, we amplified the genomic fragment of the 4,194 bp  
424 upstream region of the translational initiation site from Tak-1 gDNA using PrimeSTAR  
425 Max DNA polymerase (Takara Bio), subcloned it into pENTR1A at the EcoRI restriction  
426 sites using an In-Fusion HD Cloning Kit and then transferred it into the destination vector  
427 pMpGWB104<sup>65</sup> using Gateway LR Clonase II Enzyme mix (Thermo Fisher Scientific).  
428 To construct *proMpICE2:Citrine-GUS-NLS*, firstly we amplified the genomic fragment  
429 containing the 3,060 bp upstream region of the translational initiation site from Tak-1  
430 gDNA, *Citrine* ORF from pMpGWB107<sup>65</sup>, and *GUS-NLS* ORF from pPZP211\_35S-NG-  
431 GUS-NLS-nosT<sup>66</sup>. These fragments were fused by PCR, subcloned into pENTR1A at the  
432 Sall and EcoRV sites using an In-Fusion HD Cloning Kit, and transferred into  
433 pMpGWB101<sup>65</sup>. The resultant plasmids were introduced into Tak-1 accession by the  
434 previously described method<sup>67</sup> using *A. tumefaciens* strain GV2260.

435 The tissues of *proMpSETA:GUS* or *proMpICE2:Citrine-GUS-NLS* plants were  
436 vacuum-infiltrated and incubated at 37°C overnight in GUS staining solution containing  
437 10 mM EDTA (pH 8.0), 100 mM NaH<sub>2</sub>PO<sub>4</sub> (pH 7.0), 0.1% (v/v) Triton X-100, 0.5 g  
438 L<sup>-1</sup> 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-Gluc), 5 mM potassium-  
439 ferrocyanide, and 5 mM potassium-ferricyanide. Samples were washed in 70% (v/v)  
440 ethanol and cleared with chloral-hydrate / glycerol solution.

441

### 442 **Generation of *Mpseta*<sup>ko</sup> mutants**

443 Tak-1 genomic sequences of 3,125 bp upstream and 3,101 bp downstream of MpSETA  
444 bHLH domain coding region were amplified by PCR with a PrimeSTAR Max DNA  
445 polymerase and inserted into PacI and AseI sites of pJHY-TMp1<sup>28</sup>, respectively. The  
446 vector was introduced into germinating F<sub>1</sub> spores from Tak-1 and Tak-2 cross via *A.*  
447 *tumefaciens* strain GV2260 as previously described<sup>68</sup>. The transformed T<sub>1</sub> plants carrying  
448 the targeted insertion were selected by PCR using GoTaq DNA polymerase (Promega).  
449 As T<sub>1</sub> plants of *Mpseta-1*<sup>ko</sup> and *Mpseta-2*<sup>ko</sup> were both females, male mutants were  
450 obtained from F<sub>1</sub> sporelings by crossing female mutants with Tak-1.

451

### 452 **Reverse transcription PCR**

453 For gene expression analysis, 21 DPF sporophytes were collected from wild type,  
454 *Mpseta-1*<sup>ko</sup>, and *Mpseta-2*<sup>ko</sup>, and total RNA was extracted by RNeasy Plant Mini Kit  
455 (Qiagen) according to the manufacturer's protocol. The quality and quantity of total RNA  
456 were evaluated with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific).  
457 First-strand cDNA was synthesized using ReverTra Ace qPCR RT Master Mix with  
458 gDNA Remover (Toyobo), and semiquantitative RT-PCR was undertaken using *MpEF1α*  
459 as a loading control<sup>69</sup>.

460

### 461 **Yeast two-hybrid assay**

462 The coding sequences of MpSETA and MpICE2 were amplified from cDNA derived  
463 from mRNA of Tak-1 thalli by PCR using PrimeSTAR Max DNA polymerase or  
464 PrimeSTAR GXL polymerase (Takara Bio). The coding sequences of AtICE1 and  
465 AtSCRM2 were amplified from cDNA derived from Col-0 leaves using PrimeSTAR

466 GXL polymerase. The resultant PCR fragments were subcloned into pENTR1A at the  
467 EcoRI sites or Sall and EcoRV sites using an In-Fusion HD Cloning Kit. To generate a  
468 bait destination vector, pDEST-GBKT7-Amp<sup>r</sup>, *Amp<sup>r</sup>* was amplified from pDEST-  
469 GADT7<sup>70</sup> by PCR using PrimeSTAR Max DNA polymerase, and the fragment was  
470 cloned into the SfoI site of pDEST-GBKT7<sup>70</sup>. The inserted fragments, MpSETA,  
471 MpICE2, AtICE1, and AtSCRM2 were transferred into pDEST-GADT7 and/or pDEST-  
472 GBKT7-Amp<sup>r</sup> using Gateway LR Clonase II Enzyme mix. Bait and prey constructs were  
473 co-transformed into the yeast strain Y2HGGold (Clontech) using the Frozen-EZ Yeast  
474 Transformation II Kit (Zymo Research) and the transformants were grown on solid SD  
475 media lacking Leu and Trp (SD-LW). To examine the interaction between the bait and  
476 prey proteins, transformants were grown on solid SD media lacking Leu, Trp, His, and  
477 adenine with 40 mg L<sup>-1</sup> X- $\alpha$ -gal and 200  $\mu$ g L<sup>-1</sup> Aureobasidin A (SD-LWHA/X/AbA) at  
478 30°C. pDEST-GBKT7-Amp<sup>r</sup> and pDEST-GADT7 were used as negative controls  
479 (empty).

480

#### 481 **BiFC**

482 The coding sequences of MpSETA, MpICE2, AtICE1 and AtSCRM2 subcloned into  
483 pENTR1A described above were transferred into the pB4GWnY and/or  
484 pB4GWcY/pB4cYGW binary vector<sup>71</sup> using LR reaction to be fused with N-terminal or  
485 C-terminal half of YFP (MpSETA-*nYFP*, MpICE2-*cYFP*, *cYFP*-AtICE1, and *cYFP*-  
486 AtSCRM2 driven by the *CaMV35S* promoter). Transformed *A. tumefaciens* strain  
487 GV3101 cells harboring expression vectors were cultured and resuspended in distilled  
488 water to a final optical density of OD<sub>600</sub> = 1.0. Mixed *Agrobacterium* cultures were  
489 infiltrated into 4-week-old *Nicotiana benthamiana* leaves. Nuclei were stained with DAPI  
490 at the 1 mg L<sup>-1</sup> concentration for 30 min. Samples were observed 1.5 days post-  
491 inoculation (DPI) by an LSM780 laser scanning microscope. pB4GWnY and  
492 pB4GWcY/pB4cYGW were used as a negative control (empty).

493

#### 494 **Generation of *Mpice2*<sup>ge</sup> mutants**

495 To generate *Mpice2*<sup>ge</sup> mutants, the MpICE2 locus encoding the bHLH domain was edited  
496 using CRISPR/Cas9 based genome-editing system as previously described<sup>31</sup>. Two sgRNA  
497 were designed to generate *Mpice2*<sup>ge</sup> mutants. The oligonucleotides encoding sgRNA were  
498 cloned into pMpGE\_En03<sup>31</sup> between BsaI sites and then introduced into pMpGE010<sup>31</sup>.  
499 Male and female mutants that do not harbor T-DNA containing *Cas9* were obtained from  
500 F<sub>1</sub> sporelings crossed with wild type and *Mpice2-2*<sup>ge</sup> or *Mpice2-6*<sup>ge</sup>.

501

#### 502 **Complementation tests of *Mpseta*<sup>ko</sup> and *Mpice2*<sup>ge</sup>**

503 To construct gMpSETA for *Mpseta-1*<sup>ko</sup> complementation, the genomic region containing  
504 the 4,194 bp upstream region and coding sequences was amplified from Tak-1 genomic  
505 DNA. The fragment was cloned into pENTR1A between EcoRI sites and then introduced  
506 into pMpGWB301<sup>65</sup>. The resultant plasmids were introduced into female *Mpseta-1*<sup>ko</sup>.  
507 Male gMpSETA *Mpseta-1*<sup>ko</sup> was obtained from F<sub>1</sub> sporelings produced from crosses  
508 between female gMpSETA *Mpseta-1*<sup>ko</sup> and male *Mpseta-1*<sup>ko</sup>. To construct gMpICE2 for  
509 *Mpice2-6*<sup>ge</sup>, the genomic region containing 3,060 bp upstream regions and coding  
510 sequences was amplified from Tak-1 genomic DNA. The fragment was cloned into  
511 pENTR1A between Sall and EcoRV sites and then introduced into pMpGWB101<sup>65</sup>. The  
512 resultant plasmids were introduced into female *Mpice2-6*<sup>ge</sup>. A male gMpICE2 *Mpice2-*

513 *6<sup>ge</sup>* was obtained from a F<sub>1</sub> sporeling derived from a cross between a female gMpICE2  
514 Mpice2-6<sup>ge</sup> and a male Mpice2-6<sup>ge</sup>.

515

516

## 517 **References**

- 518 1. Ligrone, R., Duckett, J. G. & Renzaglia, K. S. Major transitions in the evolution of  
519 early land plants: A bryological perspective. *Ann. Bot.* **109**, 851–871 (2012).
- 520 2. Romani, F. & Moreno, J. E. Molecular mechanisms involved in functional  
521 macroevolution of plant transcription factors. *New Phytol.* **230**, 1345–1353 (2021).
- 522 3. Catarino, B., Hetherington, A. J., Emms, D. M., Kelly, S. & Dolan, L. The stepwise  
523 increase in the number of transcription factor families in the precambrian predated  
524 the diversification of plants on land. *Mol. Biol. Evol.* **33**, 2815–2819 (2016).
- 525 4. Bowman, J. L. *et al.* Insights into land plant evolution garnered from the  
526 *Marchantia polymorpha* genome. *Cell* **171**, 287-304.e15 (2017).
- 527 5. Pires, N. & Dolan, L. Origin and Diversification of Basic-Helix-Loop-Helix  
528 Proteins in Plants. *Mol. Biol. Evol.* **5**, 911–912 (2010).
- 529 6. MacAlister, C. A., Ohashi-Ito, K. & Bergmann, D. C. Transcription factor control  
530 of asymmetric cell divisions that establish the stomatal lineage. *Nature* **445**, 537–  
531 540 (2007).
- 532 7. Pillitteri, L. J., Sloan, D. B., Bogenschutz, N. L. & Torii, K. U. Termination of  
533 asymmetric cell division and differentiation of stomata. *Nature* **445**, 501–505  
534 (2007).
- 535 8. Ohashi-Ito, K. & Bergmann, D. C. *Arabidopsis* FAMA controls the final  
536 proliferation/differentiation switch during stomatal development. *Plant Cell* **18**,  
537 2493–2505 (2006).
- 538 9. Kanaoka, M. M. *et al.* SCREAM/ICE1 and SCREAM2 specify three cell-state  
539 transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* **20**,  
540 1775–1785 (2008).
- 541 10. Chater, C. C. *et al.* Origin and function of stomata in the moss *Physcomitrella*  
542 *patens*. *Nat. Plants* **2**, 16179 (2016).
- 543 11. Caine, R. S. *et al.* Stomata and sporophytes of the model moss *Physcomitrium*  
544 *patens*. *Front. Plant Sci.* **11**, 643 (2020).
- 545 12. Chater, C. C. C., Caine, R. S., Fleming, A. J. & Gray, J. E. Origins and evolution  
546 of stomatal development. *Plant Physiol.* **174**, 624–638 (2017).
- 547 13. Harris, B. J. *et al.* Phylogenomic evidence for the monophyly of bryophytes and  
548 the reductive evolution of stomata. *Curr. Biol.* **30**, 2001-2012.e2 (2020).
- 549 14. Flores-Sandoval, E., Romani, F. & Bowman, J. L. Co-expression and  
550 transcriptome analysis of *Marchantia polymorpha* transcription factors supports  
551 class C ARFs as independent actors of an ancient auxin regulatory module. *Front.*  
552 *Plant Sci.* **9**, 1345 (2018).
- 553 15. Feller, A., Hernandez, J. M. & Grotewold, E. An ACT-like domain participates in  
554 the dimerization of several plant basic-helix-loop-helix transcription factors. *J.*  
555 *Biol. Chem.* **281**, 28964–28974 (2006).
- 556 16. Seo, H. *et al.* Intragenic suppressors unravel the role of the SCREAM ACT-like  
557 domain for bHLH partner selectivity in stomatal development. *Proc. Natl. Acad.*  
558 *Sci. U. S. A.* **119**, e2117774119 (2022).
- 559 17. Linde, A. M., Eklund, D. M., Cronberg, N., Bowman, J. L. & Lagercrantz, U. Rates

- 560 and patterns of molecular evolution in bryophyte genomes, with focus on complex  
 561 thalloid liverworts, Marchantiopsida. *Mol. Phylogenet. Evol.* **165**, 107295 (2021).
- 562 18. Pillitteri, L. J., Bogenschutz, N. L. & Torii, K. U. The bHLH protein, MUTE,  
 563 controls differentiation of stomata and the hydathode pore in *Arabidopsis*. *Plant*  
 564 *Cell Physiol.* **49**, 934–943 (2008).
- 565 19. Frank, M. H. & Scanlon, M. J. Transcriptomic evidence for the evolution of shoot  
 566 meristem function in sporophyte-dominant land plants through concerted selection  
 567 of ancestral gametophytic and sporophytic genetic programs. *Mol. Biol. Evol.* **32**,  
 568 355–367 (2015).
- 569 20. Yamaoka, S. *et al.* Generative cell specification requires transcription factors  
 570 evolutionarily conserved in land plants. *Curr. Biol.* **28**, 479–486.e5 (2018).
- 571 21. Higo, A. *et al.* Transcriptional framework of male gametogenesis in the liverwort  
 572 *Marchantia polymorpha* L. *Plant Cell Physiol.* **57**, 325–338 (2016).
- 573 22. Hisanaga, T. *et al.* Deep evolutionary origin of gamete-directed zygote activation  
 574 by KNOX/BELL transcription factors in green plants. *Elife* **10**, e57090 (2021).
- 575 23. Shimamura, M. *Marchantia polymorpha*: Taxonomy, phylogeny and morphology  
 576 of a model system. *Plant Cell Physiol.* **57**, 230–256 (2016).
- 577 24. Haig, D. Filial mistletoes: The functional morphology of moss sporophytes. *Ann.*  
 578 *Bot.* **111**, 337–345 (2013).
- 579 25. Ligrone, R., Duckett, J. G. & Renzaglia, K. S. The origin of the sporophyte shoot  
 580 in land plants: A bryological perspective. *Ann. Bot.* **110**, 935–941 (2012).
- 581 26. Durand, E. J. The development of the sexual organs and sporogonium of  
 582 *Marchantia polymorpha*. *Bull. Torrey Bot. Club* **7**, 321–335 (1908).
- 583 27. Hernández-García, J. *et al.* Coordination between growth and stress responses by  
 584 DELLA in the liverwort *Marchantia polymorpha*. *Curr. Biol.* **31**, 3678–3686.e11  
 585 (2021).
- 586 28. Ishizaki, K., Johzuka-Hisatomi, Y., Ishida, S., Iida, S. & Kohchi, T. Homologous  
 587 recombination-mediated gene targeting in the liverwort *Marchantia polymorpha*  
 588 L. *Sci. Rep.* **3**, 1532 (2013).
- 589 29. Shaw, J. & Renzaglia, K. Phylogeny and diversification of bryophytes. *Am. J. Bot.*  
 590 **91**, 1557–1581 (2004).
- 591 30. Wu, Z. *et al.* Multiple transcriptional factors control stomata development in rice.  
 592 *New Phytol.* **223**, 220–232 (2019).
- 593 31. Sugano, S. S. *et al.* Efficient CRISPR/Cas9-based genome editing and its  
 594 application to conditional genetic analysis in *Marchantia polymorpha*. *PLoS One*  
 595 **13**, e0205117 (2018).
- 596 32. Sugano, S. S. *et al.* CRISPR/Cas9-mediated targeted mutagenesis in the liverwort  
 597 *Marchantia polymorpha* L. *Plant Cell Physiol.* **55**, 475–481 (2014).
- 598 33. Chinnusamy, V. *et al.* ICE1: A regulator of cold-induced transcriptome and  
 599 freezing tolerance in *Arabidopsis*. *Genes Dev.* **17**, 1043–1054 (2003).
- 600 34. Davies, K. A. & Bergmann, D. C. Functional specialization of stomatal bHLHs  
 601 through modification of DNA-binding and phosphoregulation potential. *Proc. Natl.*  
 602 *Acad. Sci. U. S. A.* **111**, 15585–15590 (2014).
- 603 35. MacAlister, C. A. & Bergmann, D. C. Sequence and function of basic helix-loop-  
 604 helix proteins required for stomatal development in *Arabidopsis* are deeply  
 605 conserved in land plants. *Evol. Dev.* **13**, 182–192 (2011).
- 606 36. Merced, A. & Renzaglia, K. S. Structure, function and evolution of stomata from

- 607 a bryological perspective. *Bryophyt. Divers. Evol.* **39**, 7–20 (2017).
- 608 37. Wei, D. *et al.* INDUCER OF CBF EXPRESSION 1 is a male fertility regulator  
609 impacting anther dehydration in *Arabidopsis*. *PLoS Genet.* **14**, e1007695 (2018).
- 610 38. Higo, A. *et al.* Transcription factor DUO1 generated by neo-functionalization is  
611 associated with evolution of sperm differentiation in plants. *Nat. Commun.* **9**, 5283  
612 (2018).
- 613 39. Menand, B. *et al.* An ancient mechanism controls the development of cells with a  
614 rooting function in land plants. *Science* **316**, 1477–1480 (2007).
- 615 40. Koshimizu, S. *et al.* *Physcomitrella* MADS-box genes regulate water supply and  
616 sperm movement for fertilization. *Nat. Plants* **4**, 36–45 (2018).
- 617 41. Shirakawa, M. *et al.* FAMA is an essential component for the differentiation of  
618 two distinct cell types, myrosin cells and guard cells, in *Arabidopsis*. *Plant Cell* **26**,  
619 4039–4052 (2014).
- 620 42. Yasui, Y. *et al.* GEMMA CUP-ASSOCIATED MYB1, an ortholog of axillary  
621 meristem regulators, is essential in vegetative reproduction in *Marchantia*  
622 *polymorpha*. *Curr. Biol.* **29**, 3987–3995.e5 (2019).
- 623 43. Proust, H. *et al.* RSL class I genes controlled the development of epidermal  
624 structures in the common ancestor of land plants. *Curr. Biol.* **26**, 93–99 (2016).
- 625 44. Puttick, M. N. *et al.* The interrelationships of land plants and the nature of the  
626 ancestral embryophyte. *Curr. Biol.* **28**, 733–745.e2 (2018).
- 627 45. Renzaglia, K. & Garbary, D. J. Motile gametes of land plants: Diversity,  
628 development, and evolution. *CRC. Crit. Rev. Plant Sci.* **20**, 107–213 (2001).
- 629 46. Sakakibara, K., Nishiyama, T., Deguchi, H. & Hasebe, M. Class 1 KNOX genes  
630 are not involved in shoot development in the moss *Physcomitrella patens* but do  
631 function in sporophyte development. *Evol. Dev.* **10**, 555–566 (2008).
- 632 47. Coudert, Y., Novák, O. & Harrison, C. J. A KNOX-cytokinin regulatory module  
633 predates the origin of indeterminate vascular plants. *Curr. Biol.* **29**, 2743–2750.e5  
634 (2019).
- 635 48. Li, M. & Sack, F. D. Myrosin idioblast cell fate and development are regulated by  
636 the *Arabidopsis* transcription factor FAMA, the auxin pathway, and vesicular  
637 trafficking. *Plant Cell* **26**, 4053–4066 (2014).
- 638 49. Shirakawa, M., Tanida, M. & Ito, T. The cell differentiation of idioblast myrosin  
639 cells: Similarities with vascular and guard cells. *Front. Plant Sci.* **12**, 829541  
640 (2021).
- 641 50. Li, F. W. *et al.* *Anthoceros* genomes illuminate the origin of land plants and the  
642 unique biology of hornworts. *Nat. Plants* **6**, 259–272 (2020).
- 643 51. Han, S. *et al.* MUTE directly orchestrates cell-state switch and the single  
644 symmetric division to create stomata. *Dev. Cell* **45**, 303–315.e5 (2018).
- 645 52. Lau, O. S. *et al.* Direct roles of SPEECHLESS in the specification of stomatal self-  
646 renewing cells. *Science* **345**, 1605–1609 (2014).
- 647 53. Leebens-Mack, J. H. *et al.* One thousand plant transcriptomes and the  
648 phylogenomics of green plants. *Nature* **574**, 679–685 (2019).
- 649 54. Katoh, K. & Toh, H. Recent developments in the MAFFT multiple sequence  
650 alignment program. *Brief. Bioinform.* **9**, 286–298 (2008).
- 651 55. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J.  
652 Jalview Version 2-A multiple sequence alignment editor and analysis workbench.  
653 *Bioinformatics* **25**, 1189–1191 (2009).

- 654 56. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics  
655 Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
- 656 57. Nakagawa, T. *et al.* Improved gateway binary vectors: High-performance vectors  
657 for creation of fusion constructs in transgenic analysis of plants. *Biosci. Biotechnol.*  
658 *Biochem.* **71**, 2095–2100 (2007).
- 659 58. Nakamura, S. *et al.* Gateway binary vectors with the bialaphos resistance gene, bar,  
660 as a selection marker for plant transformation. *Biosci. Biotechnol. Biochem.* **74**,  
661 1315–1319 (2010).
- 662 59. Clough, S. J. & Bent, A. F. Floral dip: A simplified method for *Agrobacterium*-  
663 mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
- 664 60. Shimada, T. L., Shimada, T. & Hara-Nishimura, I. A rapid and non-destructive  
665 screenable marker, FAST, for identifying transformed seeds of *Arabidopsis*  
666 *thaliana*. *Plant J.* **61**, 519–528 (2010).
- 667 61. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: An ultra-fast all-in-one FASTQ  
668 preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
- 669 62. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**,  
670 15–21 (2013).
- 671 63. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics*  
672 **25**, 2078–2079 (2009).
- 673 64. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq  
674 data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
- 675 65. Ishizaki, K. *et al.* Development of gateway binary vector series with four different  
676 selection markers for the liverwort *Marchantia polymorpha*. *PLoS One* **10**,  
677 e0138876 (2015).
- 678 66. Matsushita, T., Mochizuki, N. & Nagatani, A. Dimers of the N-terminal domain of  
679 phytochrome B are functional in the nucleus. *Nature* **424**, 571–574 (2003).
- 680 67. Tsuboyama, S., Nona, S., Ezura, H. & Kodama, Y. Improved G-AgarTrap : A  
681 highly efficient transformation method for intact gemmalings of the liverwort  
682 *Marchantia polymorpha*. *Sci. Rep.* **8**, 10800 (2018).
- 683 68. Ishizaki, K., Chiyoda, S., Yamato, K. T. & Kohchi, T. *Agrobacterium*-mediated  
684 transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging  
685 model for plant biology. *Plant Cell Physiol.* **49**, 1084–1091 (2008).
- 686 69. Rövekamp, M., Bowman, J. L. & Grossniklaus, U. *Marchantia* MpRKD regulates  
687 the gametophyte-sporophyte transition by keeping egg cells quiescent in the  
688 absence of fertilization. *Curr. Biol.* **26**, 1782–1789 (2016).
- 689 70. Rossignol, P., Collier, S., Bush, M., Shaw, P. & Doonan, J. H. *Arabidopsis* POT1A  
690 interacts with TERT-V(18), an N-terminal splicing variant of telomerase. *J. Cell*  
691 *Sci.* **120**, 3678–3687 (2007).
- 692 71. Hino, T. *et al.* Two Sec13p homologs, AtSec13A and AtSec13B, redundantly  
693 contribute to the formation of COPII transport vesicles in *Arabidopsis thaliana*.  
694 *Biosci. Biotechnol. Biochem.* **75**, 1848–1852 (2011).

695

## 696 Acknowledgements

697 We thank Tsuyoshi Nakagawa (Shimane University, Japan), Shoji Mano (National  
698 Institute for Basic Biology, Japan), Shigeo S. Sugano (National Institute of Advanced  
699 Industrial Science and Technology, Japan), and Keiko U. Torii (The University of Texas  
700 at Austin, USA) for sharing the materials. We also thank Keiji Nakajima (Nara Institute

701 of Science and Technology, Japan) for sharing the figures of plants in Fig. 5b. We are  
702 grateful to James Raymond for critical readings of this manuscript. This work was  
703 supported by MEXT/JSPS KAKENHI grants to M.S. (JP19K06722 and JP20H05416),  
704 to K.T. (JP26711017 and JP18K06283), to Y.O. (JP18K19964), to T.M. (JP20H05905  
705 and JP20H05906), to I.H.-N. (JP15H05776), to R.N. (JP20H04884) and to T.S.  
706 (JP18K06284); Grants-in-Aid JSPS Fellows to K.C.M. (JP21J14990) and to M.S.  
707 (JP12J05453) and; the Takeda Science Foundation, the Kato Memorial Bioscience  
708 Foundation, and the Ohsumi Frontier Science Foundation to M.S. J.L.-M. and Y-T Lu  
709 were supported by Ph.D. studentships from the Darwin Trust of Edinburgh.

710

#### 711 **Author contributions**

712 K.C.M. and T.S. conceived and designed the research in general; K.C.M. performed most  
713 of the experiments and analyzed the data; M.S. and Y.M. performed the experiments on  
714 *MpSETA*; J.L.-M., Y.-T. L., G.I., and J.G. performed the experiments on *MpICE2*; K.T.,  
715 Y.O., T.M., I.H.-N., R.N., J.G., and T.K. supervised the experiments; K.C.M. and T.S.  
716 wrote the manuscript; All authors read, edited, and approved the manuscript.

717

#### 718 **Competing interests**

719 The authors declare no competing interests.

720

721 **Figure legends**

722 **Fig. 1 | MpSETA is the only bHLH transcription factor that belongs to subfamily Ia**  
723 **in *M. polymorpha*.** **a**, A maximum-likelihood bHLH phylogenetic tree of subfamilies Ia,  
724 Ib(1) (purple), II (gray), and III(a+c) (outgroup) is shown. Numbers at branches indicate  
725 bootstrap values calculated from 1,000 replicates. Ia bHLHs are divided into four groups:  
726 TEC1/6 clade (light blue), MYC70 clade (cream-yellow), SMF/FAMA clade (magenta),  
727 and SPCH/MUTE clade (green). Species are abbreviated as follows: Mp, *M. polymorpha*  
728 (liverwort); Lc, *L. cruciata* (liverwort); Pp, *P. patens* (moss); Cepur, *Ceratodon purpureus*  
729 (moss); Aagr, *Anthoceros agrestis* (hornwort); Sm, *Selaginella moellendorffii*  
730 (lycophyte); AmTr, *Amborella trichopoda* (basal angiosperm); Os, *Oryza sativa*  
731 (monocot); At, *A. thaliana* (dicot). An arrow indicates MpSETA/MpBHLH35  
732 (Mp2g04160), and arrowheads indicate bHLH TFs mentioned as FAMA-like bHLH TFs  
733 in a previous study<sup>14</sup>. Amino acid sequences from only *A. thaliana* and *M. polymorpha*  
734 were used for subfamilies Ib(1), II, and III(a+c). **b**, Confocal images of *A. thaliana* abaxial  
735 cotyledons of wild type (Ws-4), *mute-2*, and *proAtMUTE:MpSETA mute-2* at 9 days after  
736 stratification (DAS). The upper panels and lower panels show the middle area and tip area  
737 of the cotyledons, respectively (left image). Arrowheads and asterisks indicate stomata  
738 and hydathode pores, respectively. Bars, 100  $\mu$ m (**b**, confocal images), and 1 mm (**b**, left).

739  
740 **Fig. 2 | MpSETA is preferentially expressed in developing sporophyte.** **a**, Box plot  
741 showing the expression profiles of MpSETA across nine *M. polymorpha* tissues. Y-axis  
742 shows transcripts per million (TPM). Sporeling and thalli are vegetative gametophytic (*n*)  
743 organs. Antheridiophores and archegoniophores are haploid male and female  
744 reproductive receptacles, respectively. Antheridia and archegonia are the organs that  
745 produce sperms and egg cells, respectively. Sporophytes are diploid (*2n*) organs  
746 developed after fertilization. **b**, Tissue section of mature *M. polymorpha* sporophyte  
747 (stage IX). **c**, Developmental stages of *M. polymorpha* sporophytes after 4–40 days'  
748 postfertilization. (I) sporophyte differentiating an epibasal cell and a hypobasal cell, (II)  
749 4-cell or 8-cell sporophyte, (III) sporophyte differentiating amphithecium and  
750 endothecium, (IV) later globular sporophyte, (V) sporophyte differentiating foot, seta,  
751 and archesporial tissues, (VI) sporophyte differentiating sporogenous cells and  
752 elaterocytes (or elater mother cell), (VII) sporophyte differentiating sporocytes (or spore  
753 mother cells), (VIII) sporophyte differentiating spore tetrads after meiosis, (IX) mature  
754 sporophyte before the seta elongation, (X) mature sporophyte after seta elongation.  
755 Arrows indicate the endothecium. Arrowheads indicate the cell wall of the first cell  
756 division. **d**, Magnified images showing the foot and seta for each stage of wild-type  
757 sporophytes. Same images as used in (c). **e**, The formation of the cell files of seta in the  
758 wild type stage VIII sporophyte. Enlarged images of the square areas are shown in (d).  
759 Arrowheads indicate the plane of proliferative cell divisions. **f**, Histochemical detection  
760 of  $\beta$ -glucuronidase (GUS) activity in the sporophytes crossed with male wild type and  
761 female *proMpCYCB;1:Dbox-GUS*. **g**, Histochemical detection of  $\beta$ -glucuronidase (GUS)  
762 activity driven by MpSETA promoter in the developing seta region from stage IV to stage  
763 IX. f, foot; s, seta; at, archesporial tissue; sp, sporangium; ca, calyptra; p, pseudoperianth.  
764 Arrowheads indicate the cell wall of the first cell division. Bars, 50  $\mu$ m (**c**, top), and 100  
765  $\mu$ m (**b**, **c**, bottom, **d**, **f** and **g**).  
766

767 **Fig. 3 | *Mpseta*<sup>ko</sup> mutants show developmental defects in the seta. a,b**, Longitudinal  
768 sections of sporophytes of indicated genotypes. Enlarged images of the square areas are  
769 shown in (b). Dashed and solid arrows indicate the length of the sporophyte (SP) and the  
770 length from the boundary between gametophyte and sporophyte to the proximal side of  
771 sporangium (SF), respectively. **c,d**, Quantitative data showing the SF/SP length ratio and  
772 the number of cells in the seta region. Since the boundary between seta and foot is unclear,  
773 cells excluding transfer cells (the cells at the boundary between gametophyte and  
774 sporophyte) were used to count the number of cells. The different letters indicate  
775 significantly different mean values at  $p < 0.01$  (Tukey's HSD test). ( $n = 6$  independent  
776 lines). **e**, Comparative analyses of the tissue development between wild type and *Mpseta*-  
777 *I*<sup>ko</sup>. Enlarged images of longitudinal sections of stage VI, stage VII, and stage VIII  
778 sporophytes. **f**, Sporogenesis process in wild type and *Mpseta-I*<sup>ko</sup>. Sporocytes are the cells  
779 before the pre-meiotic stage, and spore tetrads are the cells immediately after meiosis. **g**,  
780 One month-post-fertilization archegoniophores that produce mature sporophytes.  
781 Arrowheads indicate sporangia exposed to the outside of calyptra due to elongation of  
782 seta cells. Bars, 50  $\mu\text{m}$  (f), 100  $\mu\text{m}$  (a, b, e), and 5 mm (e).

783  
784 **Fig. 4 | *MpICE2* regulates setal development in cooperation with *MpSETA*. a**. Box  
785 plot showing the expression profiles of *MpICE1* and *MpICE2* across nine *M. polymorpha*  
786 tissues. Y-axis shows transcripts per million (TPM). **b**, Y2H assays showing the  
787 interaction between *MpSETA* and *MpICE2*. *MpSETA* fused with GAL4 DNA-binding  
788 domain (DBD) was used as bait, while *MpICE2* fused with GAL4 activation domain  
789 (AD) was used as prey. DBD alone and AD alone were used as the negative control. **c**,  
790 BiFC assays showing the interaction between *MpSETA* and *MpICE2* in *N. benthamiana*  
791 leaf epidermal cells. *MpSETA* was fused to the N-terminal fragment of EYFP (nYFP),  
792 while *MpICE2* was fused to the C-terminal fragment of EYFP (cYFP). nYFP alone and  
793 cYFP alone were used as the negative control. Nuclei were stained by 4',6-diamidino-2-  
794 phenylindole (DAPI). **d,e**, Longitudinal sections of sporophytes of indicated genotypes.  
795 Enlarged images of the square areas are shown in (e). Dashed and solid arrows indicate  
796 the length of the sporophyte (SP) and the length from the boundary between gametophyte  
797 and sporophyte to the proximal side of sporangium (SF), respectively. **f,g**, Quantitative  
798 data showing the SF/SP length ratio and the number of cells in the seta region. Since the  
799 boundary between seta and foot is unclear, cells excluding transfer cells (the cells at the  
800 boundary between gametophyte and sporophyte) were used to count the number of cells.  
801 The different letters indicate significantly different mean values at  $p < 0.01$  (Tukey's HSD  
802 test) ( $n = 6$  or 7 independent lines). **h**, One month-post-fertilization archegoniophores that  
803 produce mature sporophytes. Arrowheads indicate sporangia exposed to the outside of  
804 calyptra due to elongation of seta cells. Bars, 10  $\mu\text{m}$  (c, YFP, DAPI, and Merge), 100  $\mu\text{m}$   
805 (c, left, d, and e), and 5 mm (f).

806  
807 **Fig. 5 | Function and co-option of the Ia-IIIb bHLH module during the evolution of**  
808 **land plants. a**, Schematic model comparing the molecular functions of the Ia-IIIb bHLH  
809 TF modules in *A. thaliana* and *M. polymorpha* during cell fate determination. In *A.*  
810 *thaliana*, the heterodimer of Ia bHLHs (SPCH, MUTE, and FAMA) and IIIb bHLHs  
811 (ICE1 and SCRM2) control the development of stomata. In *M. polymorpha*, the  
812 heterodimer of Ia bHLH (*MpSETA*) and IIIb bHLH (*MpICE2*) regulates cell  
813 differentiation and cell division in the seta precursor transition and might directly or

814 indirectly be involved in the symmetric division of a putative seta mother cell. **b**, An  
815 evolutionary model for the Ia-IIIb bHLH TF module. Co-option of the Ia-IIIb bHLH  
816 module might have occurred multiple times independently during the evolution of land  
817 plants. First, stomata and a transcriptional module consisting of Ia-IIIb bHLHs evolved  
818 in the common ancestor of land plants. In the ancestral plant, the Ia bHLHs may have had  
819 MUTE- and FAMA-like functions. Second, the Ia-IIIb bHLH TF module might have been  
820 co-opted to regulate setal development in the ancestor of the Setophyta. Third, after  
821 mosses and liverworts diverged, the common ancestor of liverworts lost its stomata. A co-  
822 option of the Ia-IIIb bHLH module occurred in the Brassicales plants for regulating  
823 myrosin idioblast development.

824

#### 825 **Extended Data**

#### 826 **Extended Data Fig. 1 | Comparison of the domain architecture of Ia bHLHs in land**

827 **plants. a**, A diagram of the domain architecture of MpSETA (*M. polymorpha*), PpSMF1,  
828 PpSMF2 (*P. patens*), AtSPCH, AtMUTE, and AtFAMA (*A. thaliana*). While no PEST  
829 domain was identified, MpSETA has a bHLH domain and SMF domain conserved at the  
830 C-terminus like other Ia bHLH proteins. SMF domain is structurally considered to be the  
831 ACT-like domain, which is a putative domain for protein-protein dimerization. **b**,  
832 Sequence alignment of the bHLH domain of Ia bHLH proteins. Ia bHLHs are surrounded  
833 by a black box, and others are Ib(1) bHLHs. Asterisks indicate amino acids that are  
834 assumed to be important for binding to the E-box (CANNTG), and the triangles indicate  
835 amino acids that are assumed to be important for dimerization of the bHLH domain. The  
836 yellow box indicates the LxCxE motif, which is a binding motif with Retinoblastoma-  
837 related (RBR). **c**, Sequence alignment of the C-terminal SMF domain of Ia bHLH proteins.

838

#### 839 **Extended Data Fig. 2 | Function of MpSETA in *A. thaliana* Ia bHLH mutants. a**,

840 Confocal images of *A. thaliana* abaxial cotyledons of wild type (Col-0), *spch-3*, and  
841 *proAtSPCH:MpSETA spch-3* at 9 days after stratification (DAS). **b**, Confocal images of *A.*  
842 *thaliana* abaxial cotyledons of wild type (Col-0), *fama-1*, and *proAtFAMA:MpSETA fama-*  
843 *1* at 9 DAS. Brackets and arrows indicate *fama* tumors and stomatal-lineage cells,  
844 respectively. **c**, Quantitative data of the distribution of the number of cell divisions that  
845 occurred in the stomatal lineage in each genotype. ( $n > 320$  cells per genotype, 9 DAS  
846 cotyledons). **d**, Y2H assays in which the MpSETA fused with GAL4 DNA-binding  
847 domain (DBD) was used as bait, and the AtICE1 and AtSCRM2 fused with GAL4  
848 activation domain (AD) were used as prey. DBD alone and AD alone were used as the  
849 negative controls. **e**, BiFC assays showing the interaction between MpSETA and AtICE1  
850 or AtSCRM2 in *N. benthamiana* leaf epidermal cells. MpSETA was fused to the N-  
851 terminal fragment of EYFP (nYFP), while AtICE1 or AtSCRM2 was fused to the C-  
852 terminal fragment of EYFP (cYFP). nYFP alone and cYFP alone were used as the  
853 negative control. Nuclei were stained by DAPI. Bars, 10  $\mu\text{m}$  (**e**), and 100  $\mu\text{m}$  (**a,b**).

854

#### 855 **Extended Data Fig. 3 | Expression analysis of MpSETA in the gametophytic tissues.**

856 Histochemical detection of  $\beta$ -glucuronidase (GUS) activity driven by MpSETA promoter  
857 in the developing antheridia. Bars, 1 mm.

858

#### 859 **Extended Data Fig. 4 | Generation and phenotypes of MpSETA knock-out lines. a**,

860 Structure of the MpSETA locus disrupted by homologous recombination. Knock-out lines

861 have a deletion in the bHLH domain coding region. White boxes indicate the exons of the  
862 MpSETA coding sequence. DT-A, diphtheria toxin A fragment gene; *Hgr<sup>R</sup>*, hygromycin  
863 resistant gene. **b**, Genotyping of the *Mpseta<sup>ko</sup>* lines used in this study to distinguish the  
864 sex. *rbm27*, a male-specific marker; *rhf73*, a female-specific marker. **c**, Genotyping of the  
865 *Mpseta<sup>ko</sup>* lines. The position of primers used for PCR is shown in **(a)**. M, Male; F, Female.  
866 **d**, RT-PCR to confirm the loss of the full-length MpSETA transcript in *Mpseta<sup>ko</sup>* lines in  
867 21 DPF sporophytes. MpEF1 $\alpha$  was used as an internal control. **e**, Spermatogenesis  
868 process in WT and *Mpseta<sup>ko</sup>* lines. All images are at the same scale. Bars, 10  $\mu$ m (**e**).

869  
870 **Extended Data Fig. 5 | Phylogenetic tree of IIIb bHLH TFs.** A maximum-likelihood  
871 bHLH phylogenetic tree of subfamilies IIIb, III (a+c) (light blue), and III(d+e) (outgroup)  
872 is shown. Numbers at branches indicate bootstrap values calculated from 1,000 replicates.  
873 IIIb bHLHs are divided into 2 groups: ICE/SCRM clade (orange) and NFL clade  
874 (magenta). Species are abbreviated as follows: Mp, *M. polymorpha* (liverwort); Lc, *L.*  
875 *cruciata* (liverwort); Pp, *P. patens* (moss); Cepur, *Ceratodon purpureus* (moss); Agr,  
876 *Anthoceros agrestis* (hornwort); Sm, *Selaginella moellendorffii* (lycophyte); AmTr,  
877 *Amborella trichopoda* (basal angiosperm); Os, *Oryza sativa* (monocot); At, *A. thaliana*  
878 (dicot). Arrows indicate MpICE1 (Mp4g04910) and MpICE2 (Mp4g04920). For the  
879 phylogenetic construction of subfamilies III(a+c) and III(d+e), we used the amino acid  
880 sequences from only *A. thaliana* and *M. polymorpha*.

881  
882 **Extended Data Fig. 6 | Comparison of the domain architecture of IIIb bHLHs in**  
883 **land plants.** **a**, A diagram of the domain architecture of MpICE1, MpICE2 (*M.*  
884 *polymorpha*), PpSCRM1 (*P. patens*), AtICE1, and AtSCRM2 (*A. thaliana*). MpICE1 and  
885 MpICE2 have a bHLH domain and ACT-like domain conserved at the C-terminus like  
886 other IIIb bHLH proteins. **b**, Sequence alignment of the bHLH domain of IIIb bHLH  
887 proteins. IIIb bHLHs are surrounded by a black box, and others are outgroup. Asterisks  
888 indicate amino acids that are assumed to be important for binding to the E-box  
889 (CANNTG). **c**, Sequence alignment of the C-terminal ACT-like domain of IIIb bHLH  
890 proteins.

891  
892 **Extended Data Fig. 7 | The expression analysis of MpICE2.** **a**, Histochemical detection  
893 of  $\beta$ -glucuronidase (GUS) activity driven by MpICE2 promoter in the vegetative thallus.  
894 **b**, Confocal images of the dorsal epidermis of *proMpICE2:Citrine-GUS-NLS* line. Upper  
895 and lower panels indicate the epidermis around the apical notch and the epidermis around  
896 the midrib, respectively. Arrows indicate the air pores. **c,d**, Histochemical detection of  
897 GUS activity driven by MpICE2 promoter in the gametophytic reproductive organs. An  
898 antheridiophore (**c**) and an archegoniophore (**d**) are shown. **e**, Expression pattern of  
899 MpICE2 in the developing sporophytes. f, foot; s, seta; at, archesporial tissue; sp,  
900 sporangium; ca, calyptra; p, pseudoperianth (*n*). Arrowheads indicate the cell wall of the  
901 first cell division. Bars, 5 mm (**c** and **d**), 100  $\mu$ m (**b** and **e**).

902  
903 **Extended Data Fig. 8 | Generation of Mpice2 mutants by CRISPR/Cas9.** **a**, Schematic  
904 representation of the MpICE2 gene and the resulting mutations in the obtained  
905 CRISPR/Cas9-generated alleles. Gray, white, and blue boxes indicate the coding  
906 sequences (CDS), the untranslated regions (UTR), and the bHLH domain coding region,  
907 respectively. **b**, Sequence alignment of putative translational products of wild type and

908 *Mpice2<sup>ge</sup>* mutants. Asterisks indicate the amino acids that are assumed to be important for  
909 binding to E-box.

910

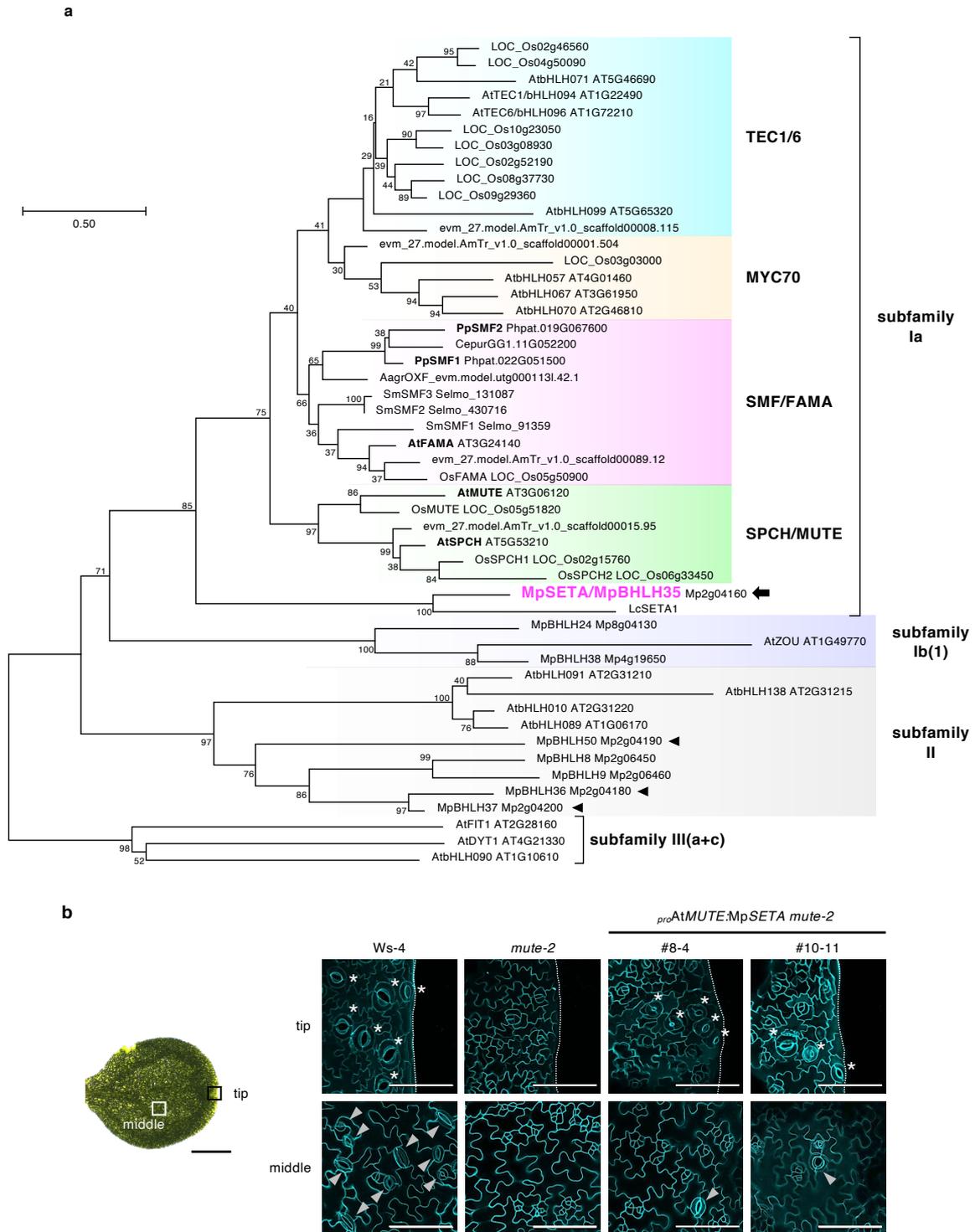
911 **Extended Data Fig. 9 | Functional analysis of MpICE1 and MpICE2 in *A. thaliana***  
912 **mutants. a**, Confocal images of *A. thaliana* abaxial cotyledons of wild type (Ws-4), *ice1-*  
913 *2 scrm2-1*, and *proAtMUTE:MpSETA mute-2* expressing MpICE2 at 9 DAS. Arrowheads  
914 and asterisks indicate stomata and hydathode pores, respectively. **b**, Confocal images of  
915 *A. thaliana* abaxial leaves of wild type (Col-0), *ice1-2 scrm2-1*, *proAtICE1:MpICE1 ice1-*  
916 *2 scrm2-1*, and *proAtICE1:MpICE2 ice1-2 scrm2-1* at 13 DAS. Bars, 100  $\mu$ m.

917

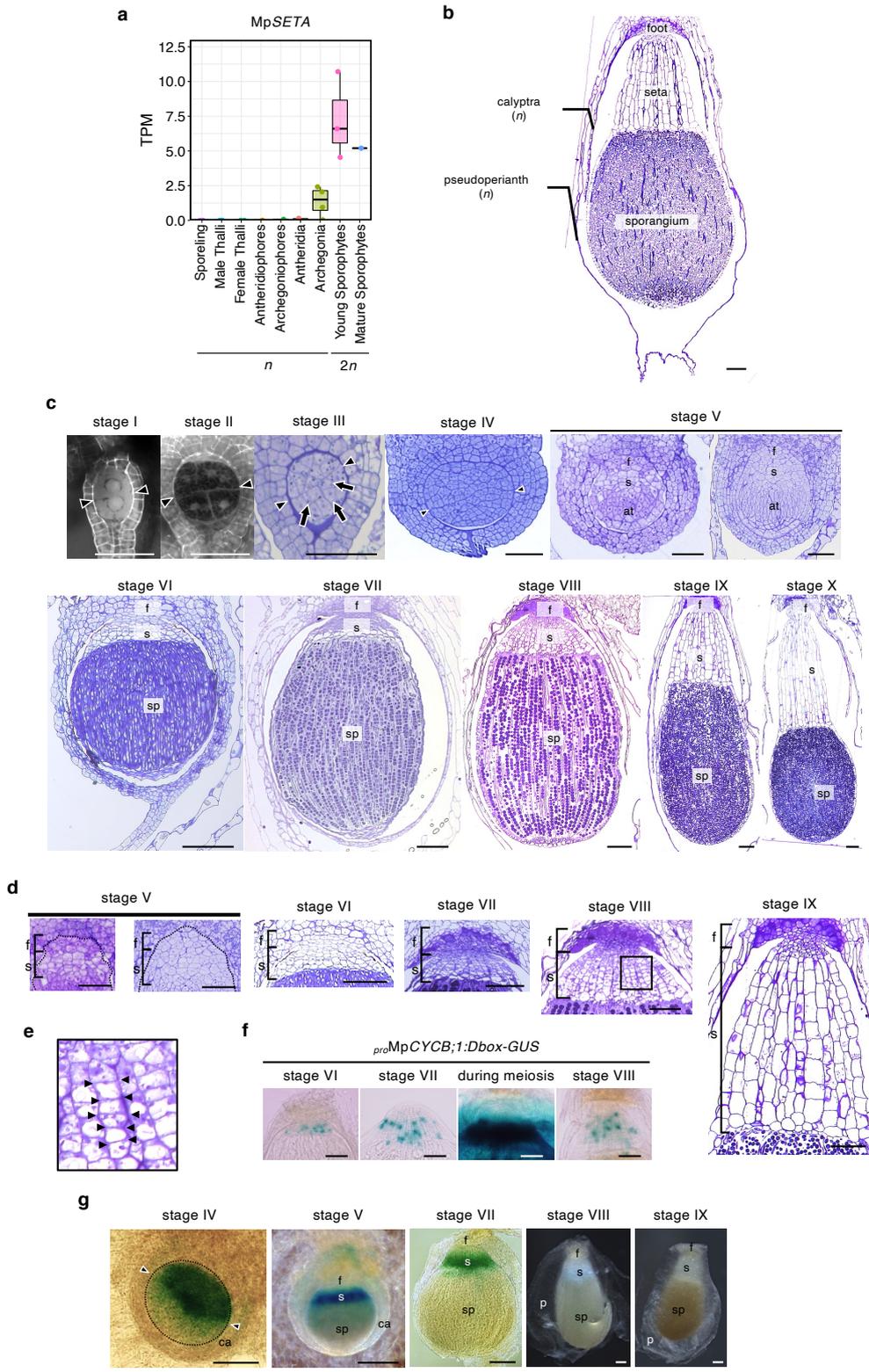
918

919 **Supplementary Information**

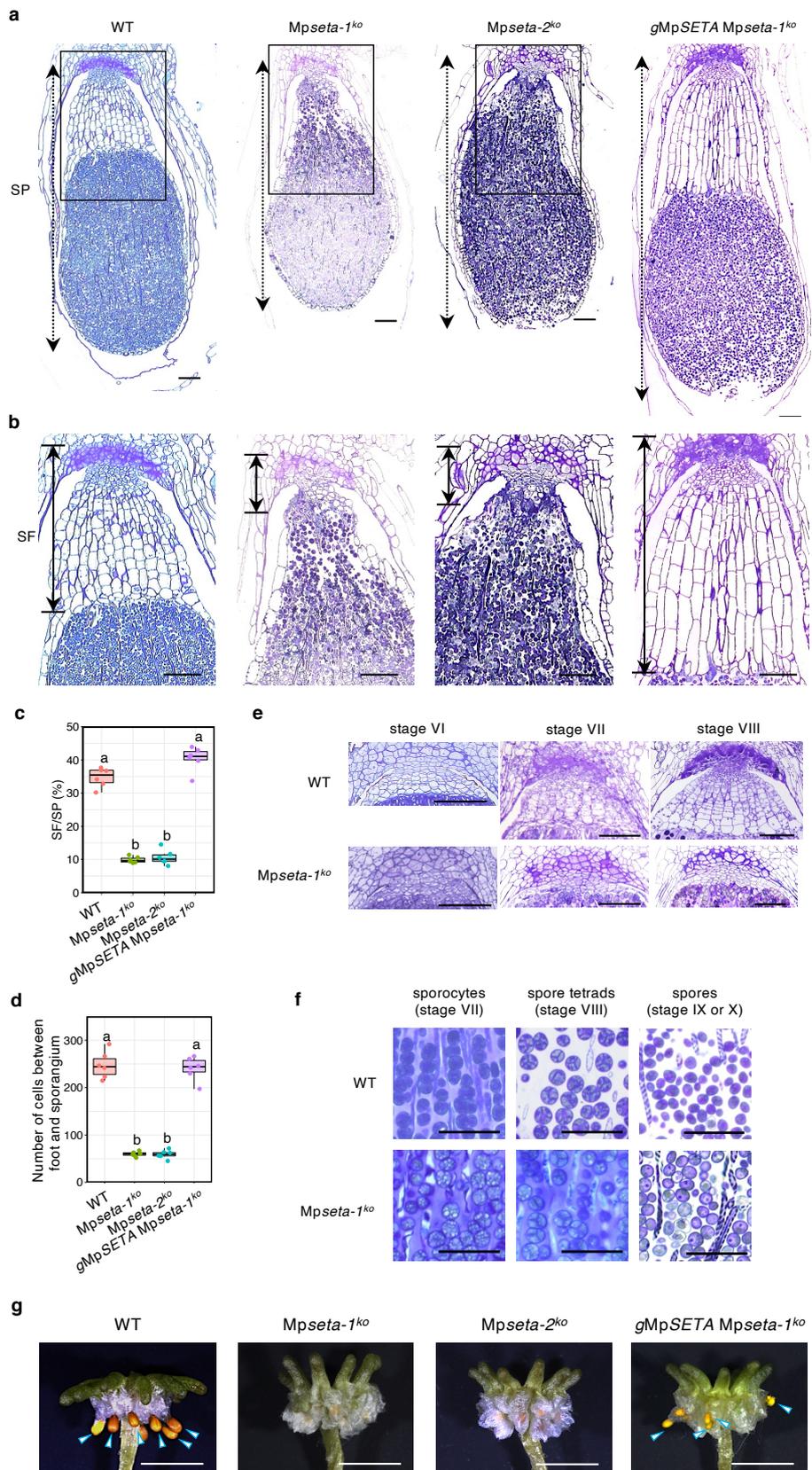
920 Supplementary Tables 1 and 2



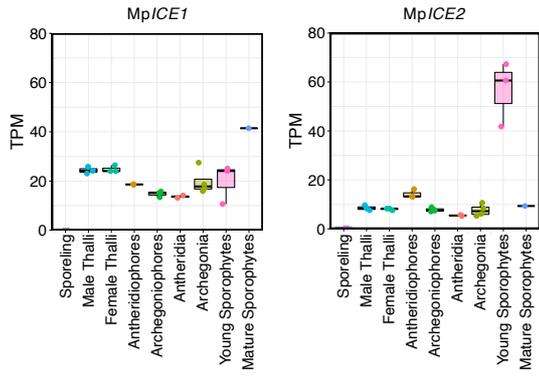
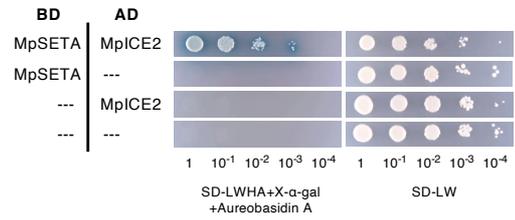
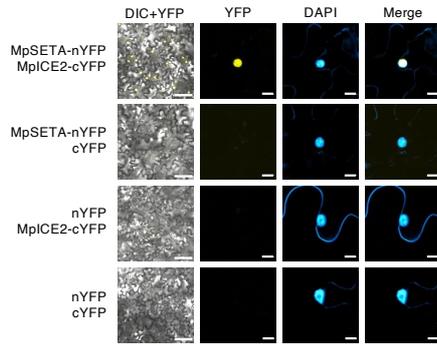
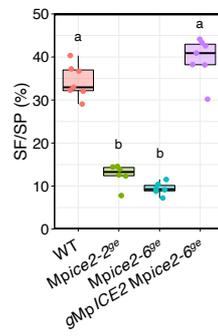
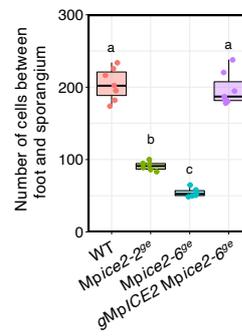
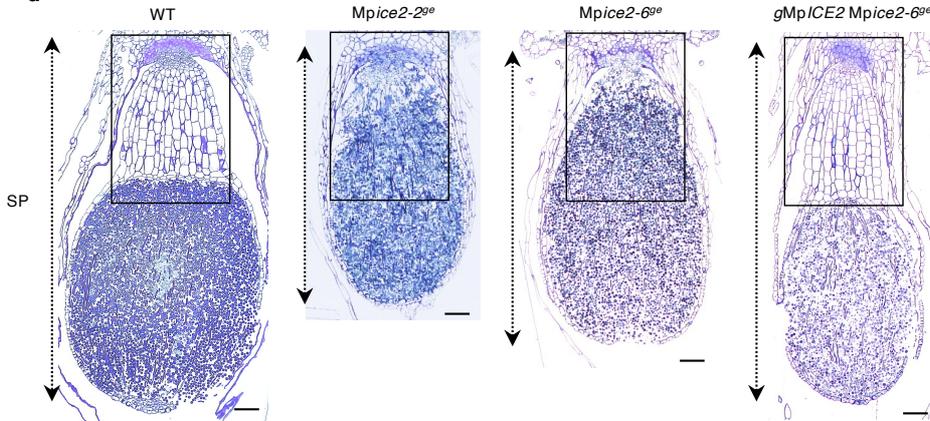
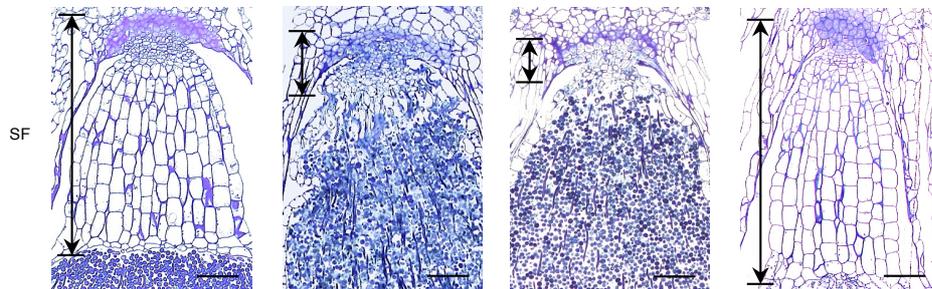
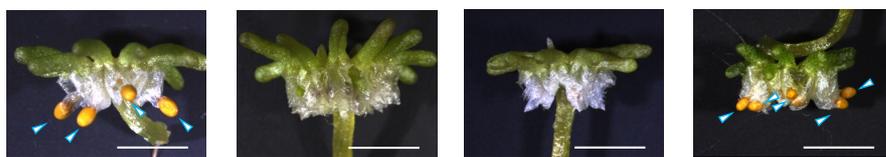
**Fig. 1 | MpSETA is the only bHLH transcription factor that belongs to subfamily Ia in *M. polymorpha*.** **a**, A maximum-likelihood bHLH phylogenetic tree of subfamilies Ia, Ib(1) (purple), II (gray), and III(a+c) (outgroup) is shown. Numbers at branches indicate bootstrap values calculated from 1,000 replicates. Ia bHLHs are divided into four groups: TEC1/6 clade (light blue), MYC70 clade (cream-yellow), SMF/FAMA clade (magenta), and SPCH/MUTE clade (green). Species are abbreviated as follows: Mp, *M. polymorpha* (liverwort); Lc, *L. cruciata* (liverwort); Pp, *P. patens* (moss); Cepur, *Ceratodon purpureus* (moss); Aagr, *Anthoceros agrestis* (hornwort); Sm, *Selaginella moellendorffii* (lycophyte); AmTr, *Amborella trichopoda* (basal angiosperm); Os, *Oryza sativa* (monocot); At, *A. thaliana* (dicot). An arrow indicates MpSETA/MpBHLH35 (Mp2g04160), and arrowheads indicate bHLH TFs mentioned as FAMA-like bHLH TFs in a previous study<sup>14</sup>. Amino acid sequences from only *A. thaliana* and *M. polymorpha* were used for subfamilies Ib(1), II, and III(a+c). **b**, Confocal images of *A. thaliana* abaxial cotyledons of wild type (Ws-4), *mute-2*, and *proAtMUTE:MpSETA mute-2* at 9 days after stratification (DAS). The upper panels and lower panels show the middle area and tip area of the cotyledons, respectively (left image). Arrowheads and asterisks indicate stomata and hydathode pores, respectively. Bars, 100  $\mu$ m (b, confocal images), and 1 mm (b, left).



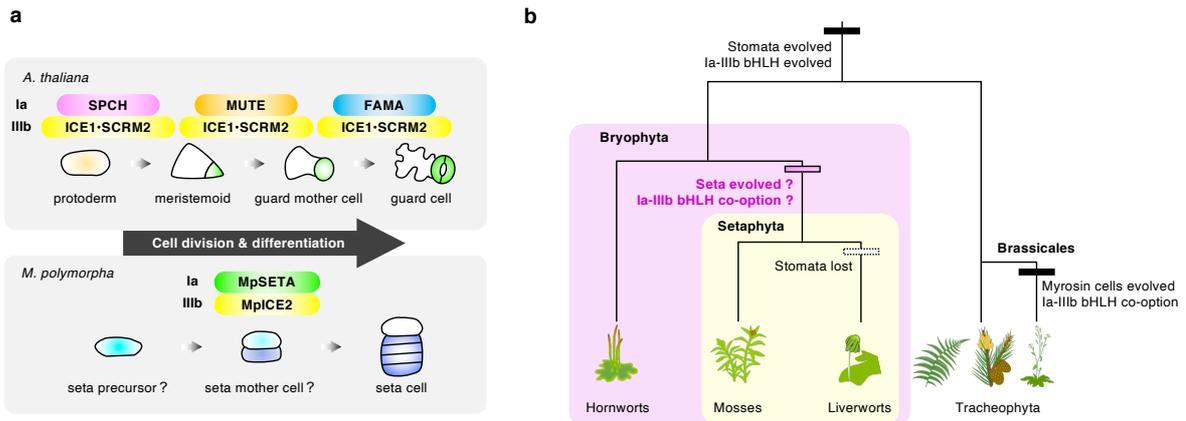
**Fig. 2 | MpSETA is preferentially expressed in developing sporophyte.** **a**, Box plot showing the expression profiles of MpSETA across nine *M. polymorpha* tissues. Y-axis shows transcripts per million (TPM). Sporeling and thalli are vegetative gametophytic ( $n$ ) organs. Antheridiophores and archegoniophores are haploid male and female reproductive receptacles, respectively. Antheridia and archegonia are the organs that produce sperms and egg cells, respectively. Sporophytes are diploid ( $2n$ ) organs developed after fertilization. **b**, Tissue section of mature *M. polymorpha* sporophyte (stage IX). **c**, Developmental stages of *M. polymorpha* sporophytes after 4–40 days' postfertilization. (I) sporophyte differentiating an epibasal cell and a hypobasal cell, (II) 4-cell or 8-cell sporophyte, (III) sporophyte differentiating amphithecium and endothecium, (IV) later globular sporophyte, (V) sporophyte differentiating foot, seta, and archesporial tissues, (VI) sporophyte differentiating sporogenous cells and elaterocytes (or elater mother cell), (VII) sporophyte differentiating sporocytes (or spore mother cells), (VIII) sporophyte differentiating spore tetrads after meiosis, (IX) mature sporophyte before the seta elongation, (X) mature sporophyte after seta elongation. Arrows indicate the endothecium. Arrowheads indicate the cell wall of the first cell division. **d**, Magnified images showing the foot and seta for each stage of wild-type sporophytes. Same images as used in (c). **e**, The formation of the cell files of seta in the wild type stage VIII sporophyte. Enlarged images of the square areas are shown in (d). Arrowheads indicate the plane of proliferative cell divisions. **f**, Histochemical detection of  $\beta$ -glucuronidase (GUS) activity in the sporophytes crossed with male wild type and female *proMpCYCB;1:Dbox-GUS*. **g**, Histochemical detection of  $\beta$ -glucuronidase (GUS) activity driven by MpSETA promoter in the developing seta region from stage IV to stage IX. f, foot; s, seta; at, archesporial tissue; sp, sporangium; ca, calyptra; p, pseudoperianth. Arrowheads indicate the cell wall of the first cell division. Bars, 50  $\mu$ m (c, top), 100  $\mu$ m (d and f), and 200  $\mu$ m (b, c, bottom, and g).



**Fig. 3 | *Mpseta*<sup>ko</sup> mutants show developmental defects in the seta.** **a,b**, Longitudinal sections of sporophytes of indicated genotypes. Enlarged images of the square areas are shown in **(b)**. Dashed and solid arrows indicate the length of the sporophyte (SP) and the length from the boundary between gametophyte and sporophyte to the proximal side of sporangium (SF), respectively. **c,d**, Quantitative data showing the SF/SP length ratio and the number of cells in the seta region. Since the boundary between seta and foot is unclear, cells excluding transfer cells (the cells at the boundary between gametophyte and sporophyte) were used to count the number of cells. The different letters indicate significantly different mean values at  $p < 0.01$  (Tukey's HSD test). ( $n = 6$  independent lines). **e**, Comparative analyses of the tissue development between wild type and *Mpseta-1*<sup>ko</sup>. Enlarged images of longitudinal sections of stage VI, stage VII and stage VIII sporophytes. **f**, Sporogenesis process in wild type and *Mpseta-1*<sup>ko</sup>. Sporocytes are the cells before the pre-meiotic stage, and spore tetrads are the cells immediately after meiosis. **g**, One month-post-fertilization archegoniophores that produce mature sporophytes. Arrowheads indicate sporangia exposed to the outside of calyptra due to elongation of seta cells. Bars, 50  $\mu\text{m}$  (**f**), 100  $\mu\text{m}$  (**e**), 200  $\mu\text{m}$  (**a** and **b**), and 5 mm (**e**).

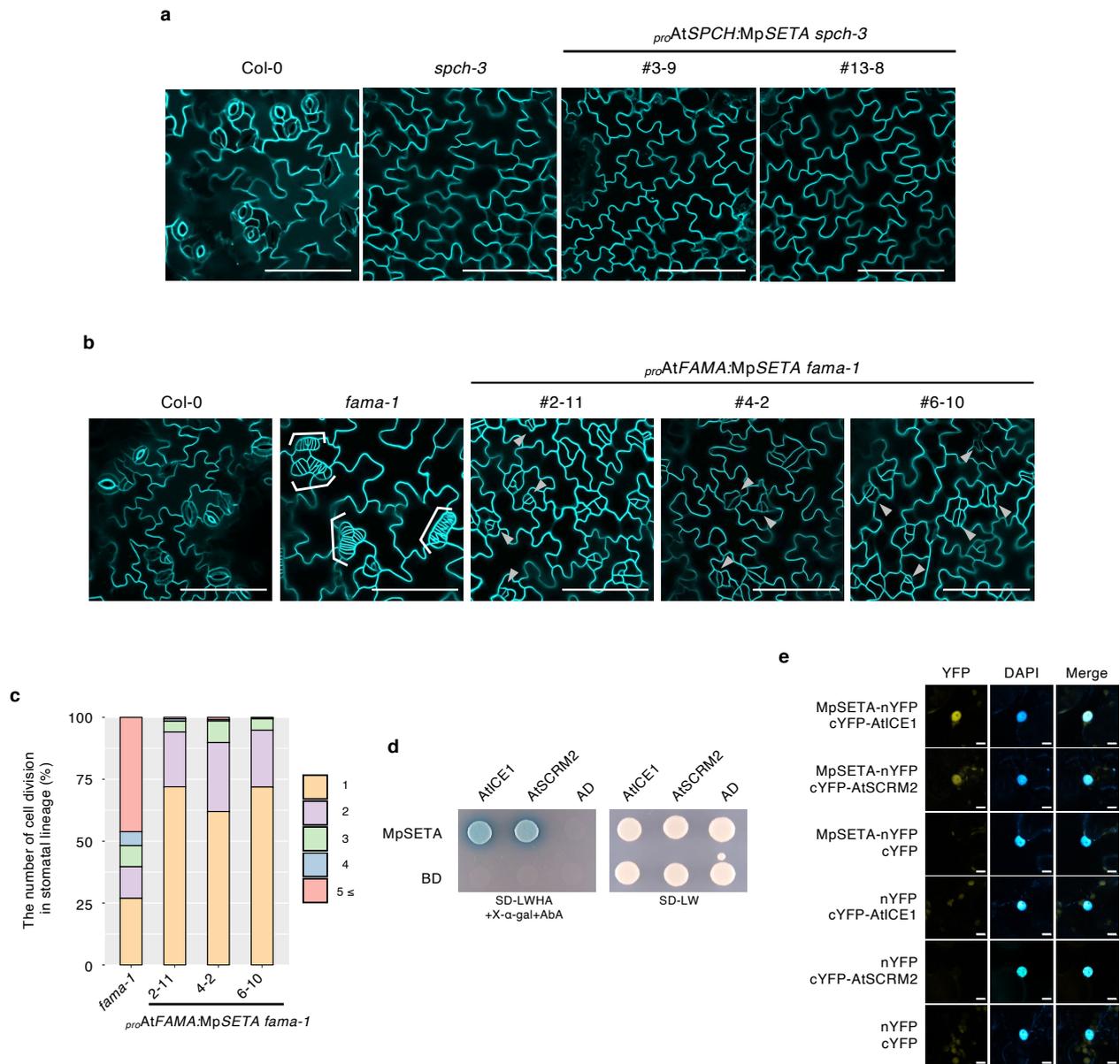
**a****b****c****f****g****d****e****h**

**Fig. 4 | MpICE2 regulates setal development in cooperation with MpSETA.** **a.** Box plot showing the expression profiles of MpICE1 and MpICE2 across nine *M. polymorpha* tissues. Y-axis shows transcripts per million (TPM). **b.** Y2H assays showing the interaction between MpSETA and MpICE2. MpSETA fused with GAL4 DNA-binding domain (DBD) was used as bait, while MpICE2 fused with GAL4 activation domain (AD) was used as prey. DBD alone and AD alone were used as the negative control. **c.** BiFC assays showing the interaction between MpSETA and MpICE2 in *N. benthamiana* leaf epidermal cells. MpSETA was fused to the N-terminal fragment of EYFP (nYFP), while MpICE2 was fused to the C-terminal fragment of EYFP (cYFP). nYFP alone and cYFP alone were used as the negative control. Nuclei were stained by 4',6-diamidino-2-phenylindole (DAPI). **d,e.** Longitudinal sections of sporophytes of indicated genotypes. Enlarged images of the square areas are shown in **(e)**. Dashed and solid arrows indicate the length of the sporophyte (SP) and the length from the boundary between gametophyte and sporophyte to the proximal side of sporangium (SF), respectively. **f,g.** Quantitative data showing the SF/SP length ratio and the number of cells in the seta region. Since the boundary between seta and foot is unclear, cells excluding transfer cells (the cells at the boundary between gametophyte and sporophyte) were used to count the number of cells. The different letters indicate significantly different mean values at  $p < 0.01$  (Tukey's HSD test) ( $n = 6$  or  $7$  independent lines). **h.** One month-post-fertilization archegoniophores that produce mature sporophytes. Arrowheads indicate sporangia exposed to the outside of calyptra due to elongation of seta cells. Bars, 10  $\mu\text{m}$  (**c**, YFP, DAPI, and Merge), 100  $\mu\text{m}$  (**c**, left, **d**, and **e**), and 5 mm (**f**).

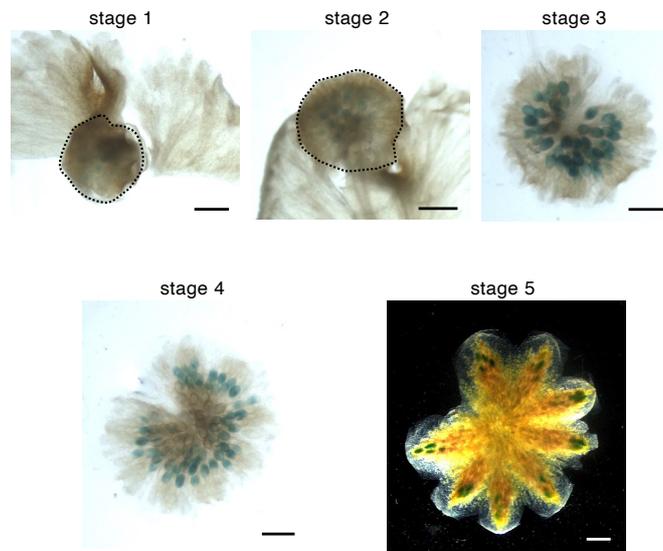


**Fig. 5 | Function and co-option of the la-IIIb bHLH module during the evolution of land plants. a**, Schematic model comparing the molecular functions of the la-IIIb bHLH TF modules in *A. thaliana* and *M. polymorpha* during cell fate determination. In *A. thaliana*, the heterodimer of la bHLHs (SPCH, MUTE, and FAMA) and IIIb bHLHs (ICE1 and SCRM2) control the development of stomata. In *M. polymorpha*, the heterodimer of la bHLH (MpSETA) and IIIb bHLH (MpICE2) regulates cell differentiation and cell division in the seta precursor transition and might directly or indirectly be involved in the symmetric division of a putative seta mother cell. **b**, An evolutionary model for the la-IIIb bHLH TF module. Co-option of the la-IIIb bHLH module might have occurred multiple times independently during the evolution of land plants. First, stomata and a transcriptional module consisting of la-IIIb bHLHs evolved in the common ancestor of land plants. In the ancestral plant, the la bHLHs may have had MUTE- and FAMA-like functions. Second, the la-IIIb bHLH TF module might have been co-opted to regulate setal development in the ancestor of the Setophyta. Third, after mosses and liverworts diverged, the common ancestor of liverworts lost its stomata. A co-option of the la-IIIb bHLH module occurred in the Brassicales plants for regulating myrosin idioblast development.

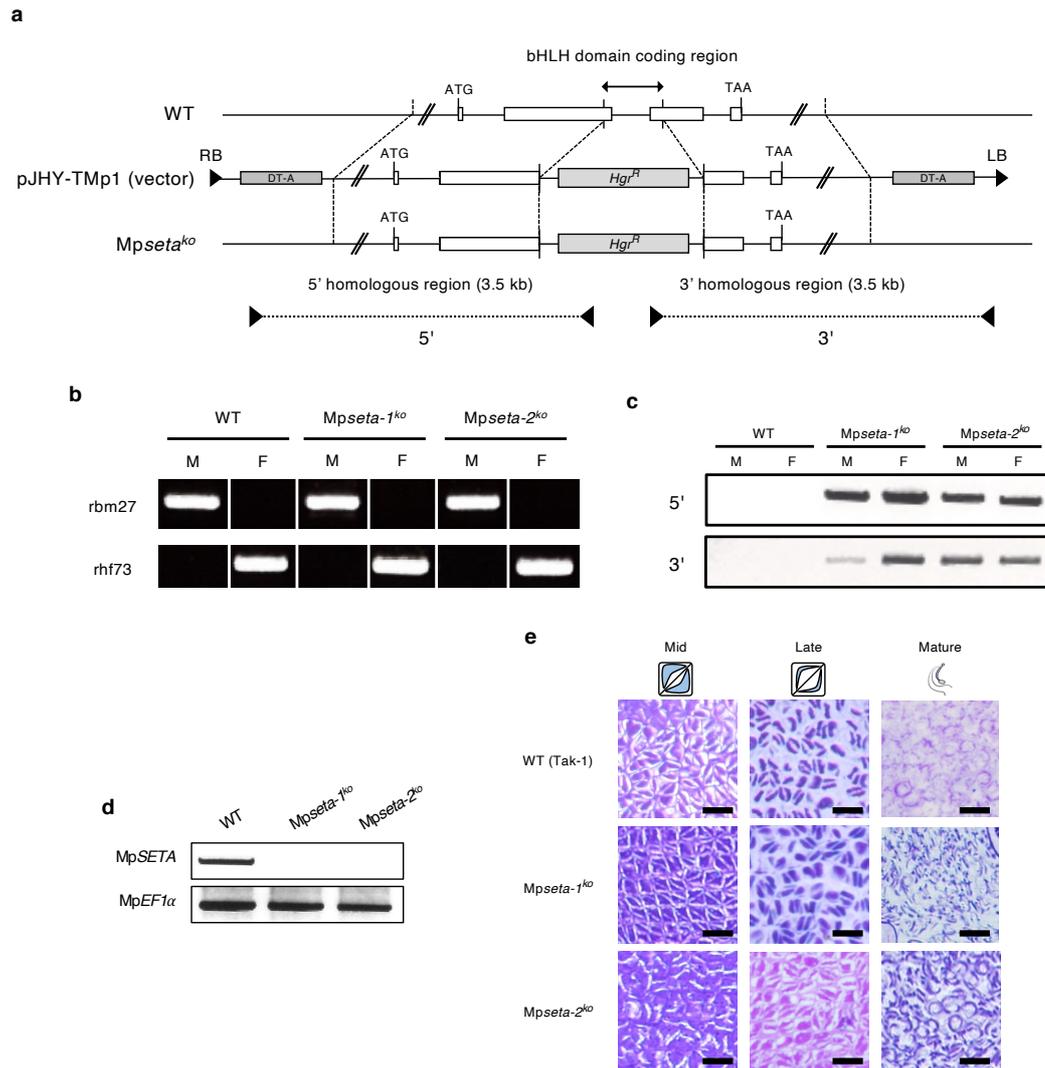




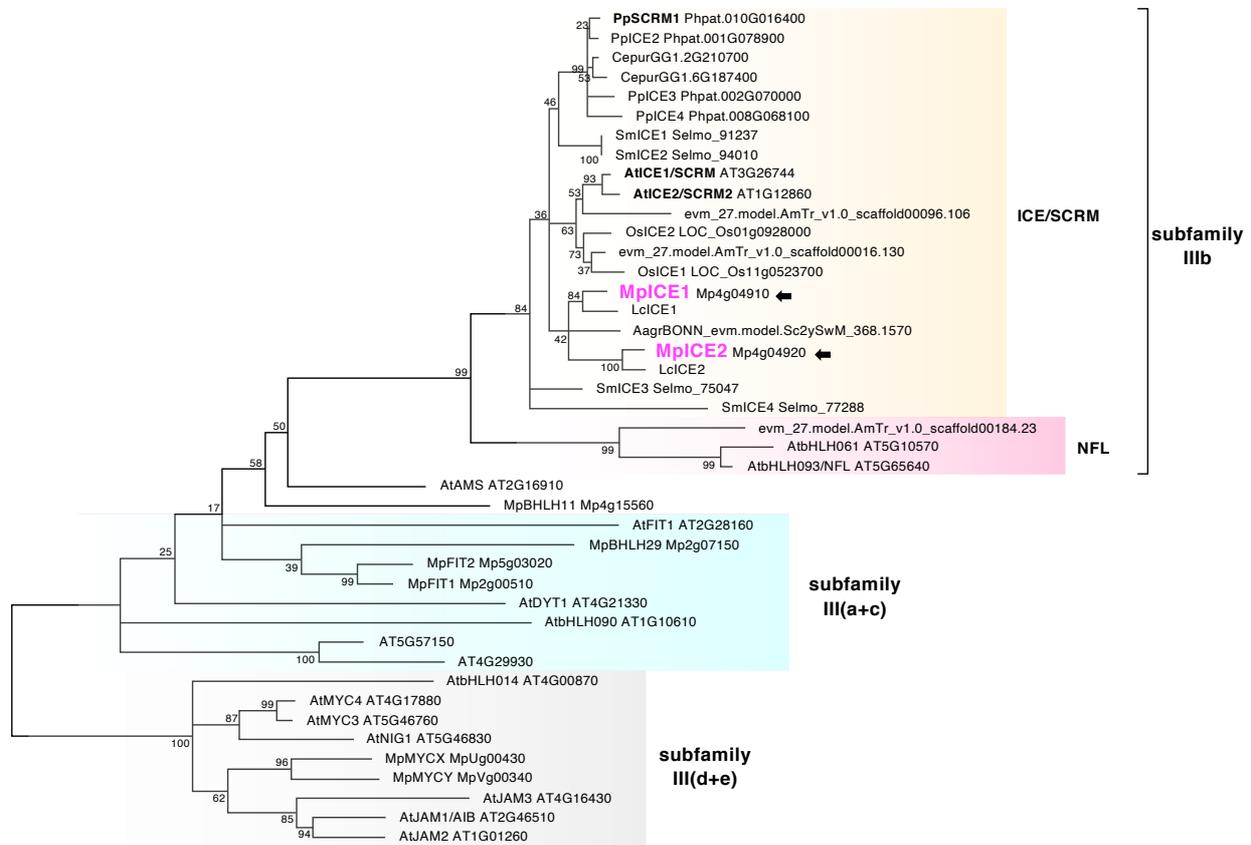
**Extended Data Fig. 2 | Function of MpSETA in *A. thaliana* la bHLH mutants.** **a**, Confocal images of *A. thaliana* abaxial cotyledons of wild type (Col-0), *spch-3*, and *proAtSPCH:MpSETA spch-3* at 9 days after stratification (DAS). **b**, Confocal images of *A. thaliana* abaxial cotyledons of wild type (Col-0), *fama-1*, and *proAtFAMA:MpSETA fama-1* at 9 DAS. Brackets and arrows indicate *fama* tumors and stomatal-lineage cells, respectively. **c**, Quantitative data of the distribution of the number of cell divisions that occurred in the stomatal lineage in each genotype. ( $n > 320$  cells per genotype, 9 DAS cotyledons). **d**, Y2H assays in which the MpSETA fused with GAL4 DNA-binding domain (DBD) was used as bait, and the AtICE1 and AtSCRM2 fused with GAL4 activation domain (AD) were used as prey. DBD alone and AD alone were used as the negative controls. **e**, BiFC assays showing the interaction between MpSETA and AtICE1 or AtSCRM2 in *N. benthamiana* leaf epidermal cells. MpSETA was fused to the N-terminal fragment of EYFP (nYFP), while AtICE1 or AtSCRM2 was fused to the C-terminal fragment of EYFP (cYFP). nYFP alone and cYFP alone were used as the negative control. Nuclei were stained by DAPI. Bars, 10  $\mu\text{m}$  (**e**), and 100  $\mu\text{m}$  (**a,b**).



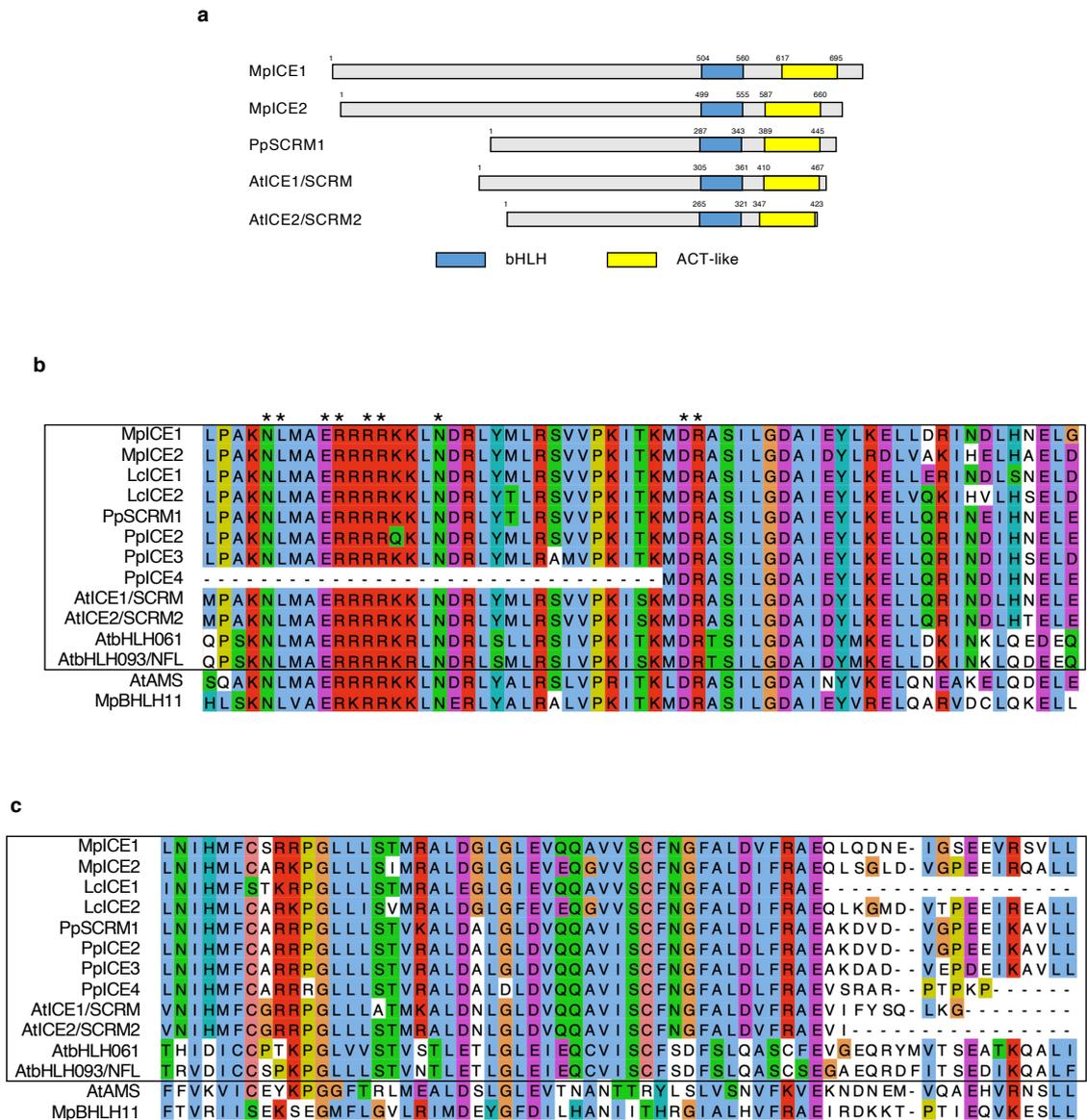
**Extended Data Fig. 3 | Expression analysis of *MpSETA* in the gametophytic tissues.** Histochemical detection of  $\beta$ -glucuronidase (GUS) activity driven by *MpSETA* promoter in the developing antheridia. Bars, 1 mm.



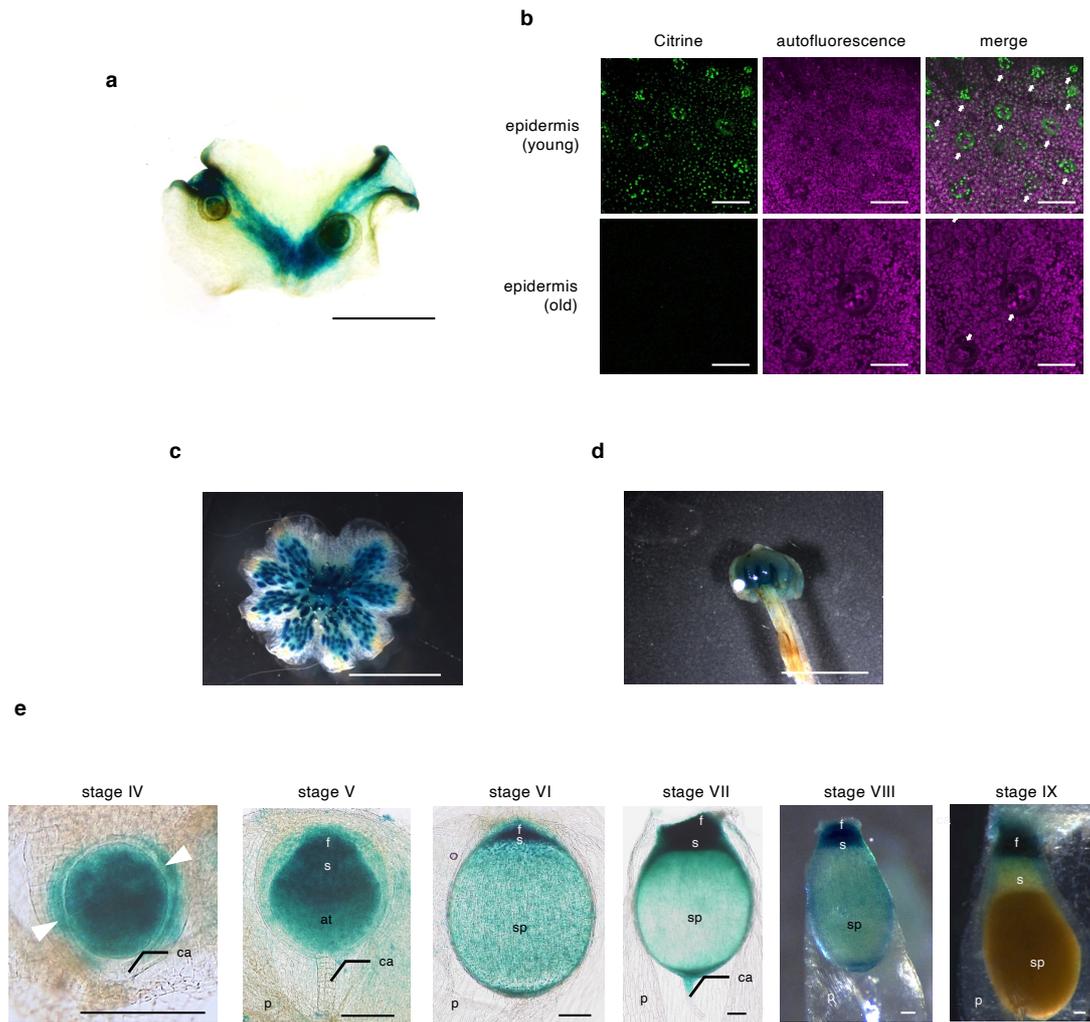
**Extended Data Fig. 4 | Generation and phenotypes of MpSETA knock-out lines.** **a**, Structure of the MpSETA locus disrupted by homologous recombination. Knock-out lines have a deletion in the bHLH domain coding region. White boxes indicate the exons of the MpSETA coding sequence. DT-A, diphtheria toxin A fragment gene; *Hgr<sup>R</sup>*, hygromycin resistant gene. **b**, Genotyping of the Mpseta<sup>ko</sup> lines used in this study to distinguish the sex. rbm27, a male-specific marker; rhf73, a female-specific marker. **c**, Genotyping of the Mpseta<sup>ko</sup> lines. The position of primers used for PCR is shown in (a). M, Male; F, Female. **d**, RT-PCR to confirm the loss of the full-length MpSETA transcript in Mpseta<sup>ko</sup> lines in 21 DPF sporophytes. MpEF1 $\alpha$  was used as an internal control. **e**, Spermatogenesis process in WT and Mpseta<sup>ko</sup> lines. All images are at the same scale. Bars, 10  $\mu$ m (e).



**Extended Data Fig. 5 | Phylogenetic tree of IIIb bHLH TFs.** A maximum-likelihood bHLH phylogenetic tree of subfamilies IIIb, III (a+c) (light blue), and III(d+e) (outgroup) is shown. Numbers at branches indicate bootstrap values calculated from 1,000 replicates. IIIb bHLHs are divided into 2 groups: ICE/SCRM clade (orange) and NFL clade (magenta). Species are abbreviated as follows: Mp, *M. polymorpha* (liverwort); Lc, *L. cruciata* (liverwort); Pp, *P. patens* (moss); Cepur, *Ceratodon purpureus* (moss); Aagr, *Anthoceros agrestis* (hornwort); Sm, *Selaginella moellendorffii* (lycophyte); AmTr, *Amborella trichopoda* (basal angiosperm); Os, *Oryza sativa* (monocot); At, *A. thaliana* (dicot). Arrows indicate MpICE1 (Mp4g04910) and MpICE2 (Mp4g04920). For the phylogenetic construction of subfamilies III(a+c) and III(d+e), we used the amino acid sequences from only *A. thaliana* and *M. polymorpha*.

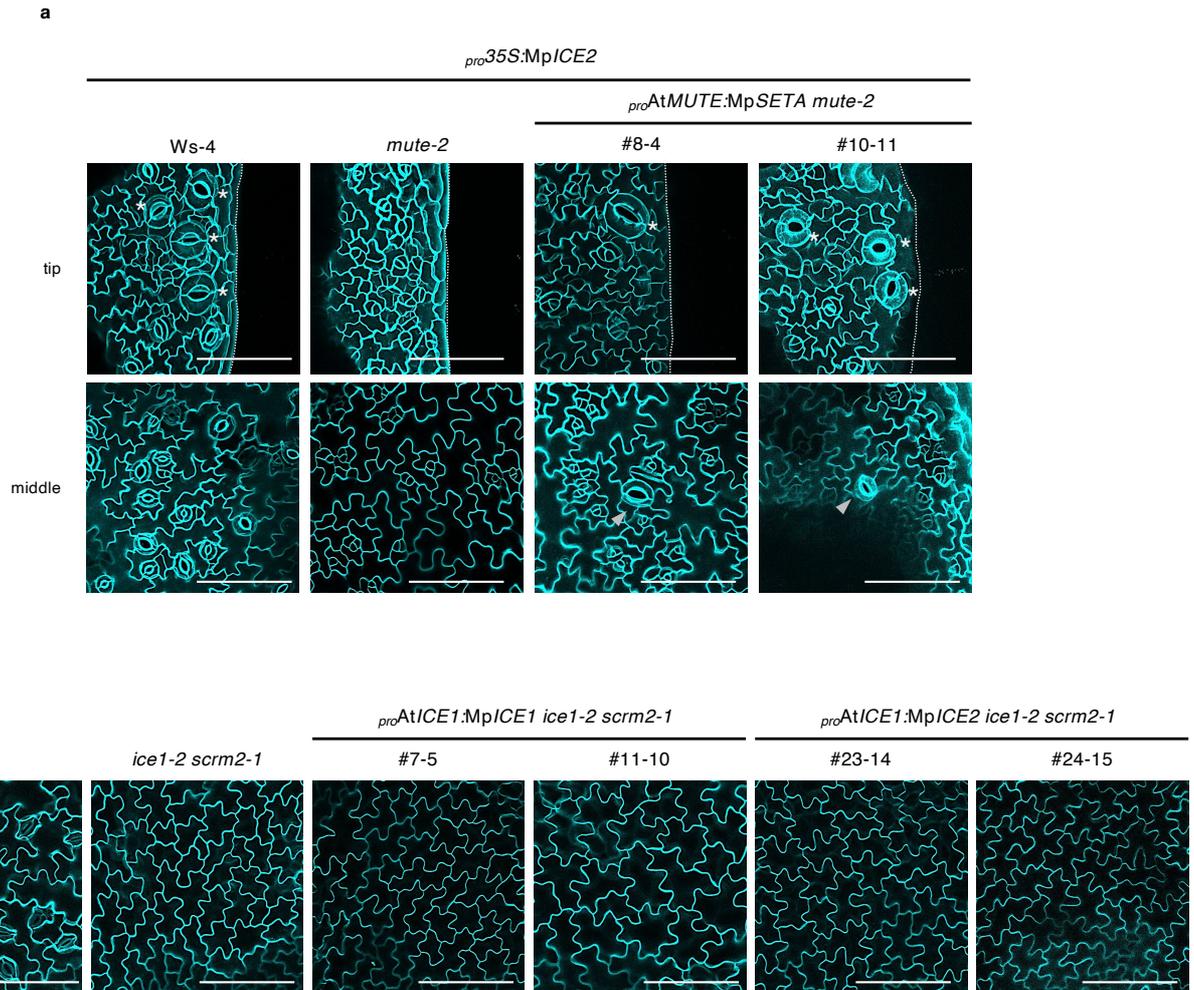


**Extended Data Fig. 6 | Comparison of the domain architecture of IIIb bHLHs in land plants. a**, A diagram of the domain architecture of MpiCE1, MpiCE2 (*M. polymorpha*), PpSCRM1 (*P. patens*), AtICE1, and AtSCRM2 (*A. thaliana*). MpiCE1 and MpiCE2 have a bHLH domain and ACT-like domain conserved at the C-terminus like other IIIb bHLH proteins. **b**, Sequence alignment of the bHLH domain of IIIb bHLH proteins. IIIb bHLHs are surrounded by a black box, and others are outgroup. Asterisks indicate amino acids that are assumed to be important for binding to the E-box (CANNTG). **c**, Sequence alignment of the C-terminal ACT-like domain of IIIb bHLH proteins.



**Extended Data Fig. 7 | The expression analysis of *MpICE2*.** **a**, Histochemical detection of  $\beta$ -glucuronidase (GUS) activity driven by *MpICE2* promoter in the vegetative thallus. **b**, Confocal images of the dorsal epidermis of *proMpICE2:Citrine-GUS-NLS* line. Upper and lower panels indicate the epidermis around the apical notch and the epidermis around the midrib, respectively. Arrows indicate the air pores. **c,d**, Histochemical detection of GUS activity driven by *MpICE2* promoter in the gametophytic reproductive organs. An antheridiophore (**c**) and an archegoniophore (**d**) are shown. **e**, Expression pattern of *MpICE2* in the developing sporophytes. f, foot; s, seta; at, archesporial tissue; sp, sporangium; ca, calyptra; p, pseudoperianth (*n*). Arrowheads indicate the cell wall of the first cell division. Bars, 5 mm (**c** and **d**), 100  $\mu$ m (**b** and **e**).





**Extended Data Fig. 9 | Functional analysis of MpICE1 and MpICE2 in *A. thaliana* mutants. a**, Confocal images of *A. thaliana* abaxial cotyledons of wild type (*Ws-4*), *ice1-2 scrm2-1*, and *proAtMUTE:MpSETA mute-2* expressing MpICE2 at 9 DAS. Arrowheads and asterisks indicate stomata and hydathode pores, respectively. **b**, Confocal images of *A. thaliana* abaxial leaves of wild type (Col-0), *ice1-2 scrm2-1*, *proAtICE1:MpICE1 ice1-2 scrm2-1*, and *proAtICE1:MpICE2 ice1-2 scrm2-1* at 13 DAS. Bars, 100  $\mu$ m.