Supplementary Information

Origin and function of stomata in the moss *Physcomitrella patens*

Caspar C. Chater[1‡], Robert S. Caine[2‡], Marta Tomek[3], Simon Wallace[4], Yasuko Kamisugi[5], Andrew C. Cuming[5], Daniel Lang[3], Cora A. MacAlister[6], Stuart Casson[7], Dominique C. Bergmann[8], Eva L. Decker[3], Wolfgang Frank[9], Julie E. Gray[7], Andrew Fleming[2], Ralf Reski[3,10*] & David J. Beerling[2*]

[1]Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de Mexico, Cuernavaca, Mexico.

[2]Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

[3]Plant Biotechnology, Faculty of Biology, University of Freiburg, Schaenzlestr. 1, 79104 Freiburg, Germany

[4]Royal College of Veterinary Surgeons, Belgravia House, 62-64 Horseferry Rd, London SW1P 2AF, UK

[5]Centre for Plant Sciences, University of Leeds, Leeds, LS2 9JT, UK

[6]Department of Molecular Cellular and Developmental Biology, University of Michigan, Ann Arbor, Michigan, 48109-1048, USA

[7]Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, UK

[8]HHMI and Department of Biology, Stanford University, Stanford, CA 94305-5020, USA

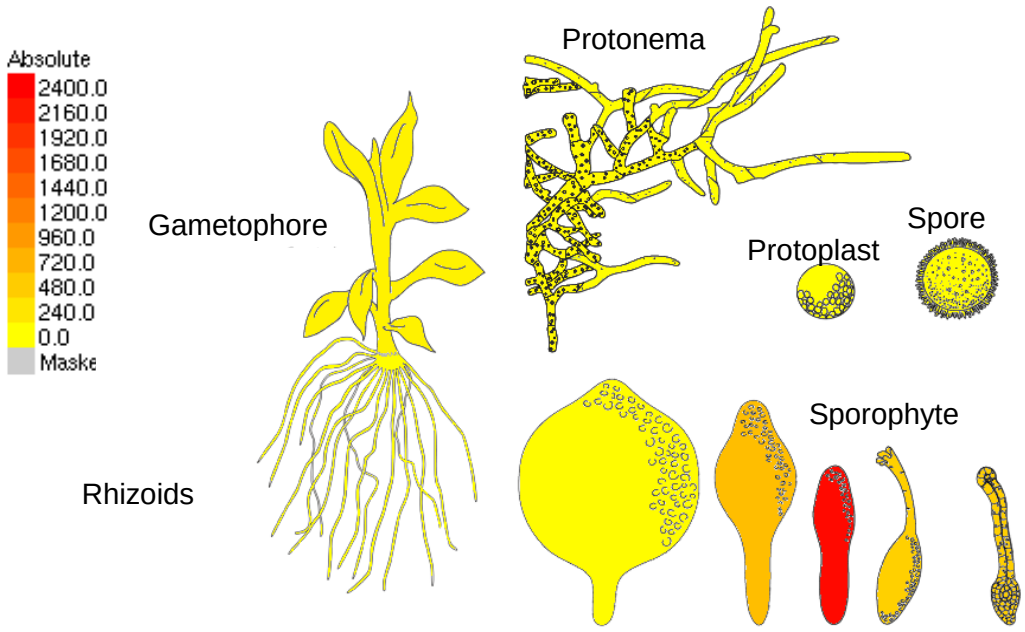
[9]Plant Molecular Cell Biology, Faculty of Biology, Ludwig-Maximilians-Universität München, LMU Biocenter, Großhaderner Straße 2, 82152 Planegg-Martinsried, Germany

[10]BIOSS – Centre for Biological Signalling Studies, 79104 Freiburg, Germany

*Corresponding authors. Emails: d.j.beerling@sheffield.ac.uk, ralf.reski@biologie.uni-freiburg.de

Supplementary Information contains:

(a) *PpSMF1*



(b) *PpSMF2*

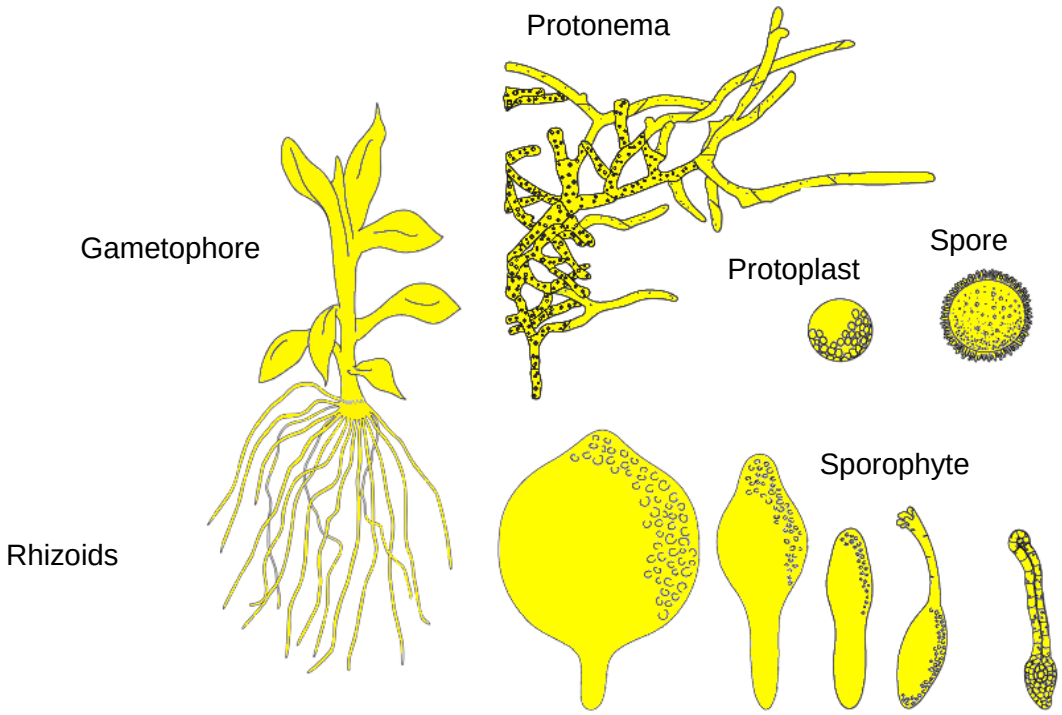
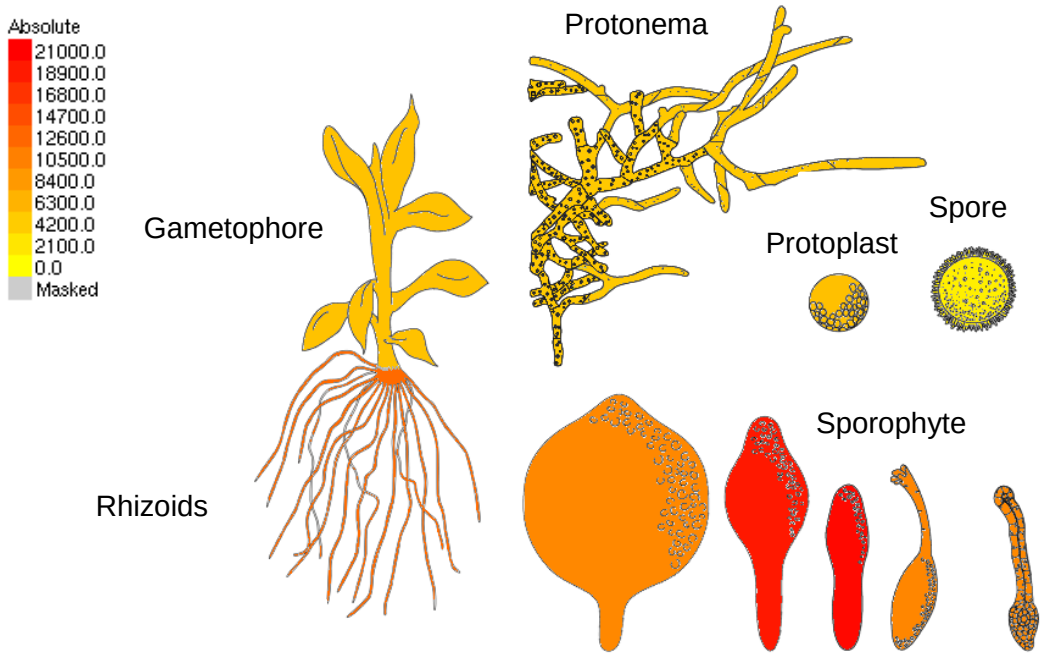
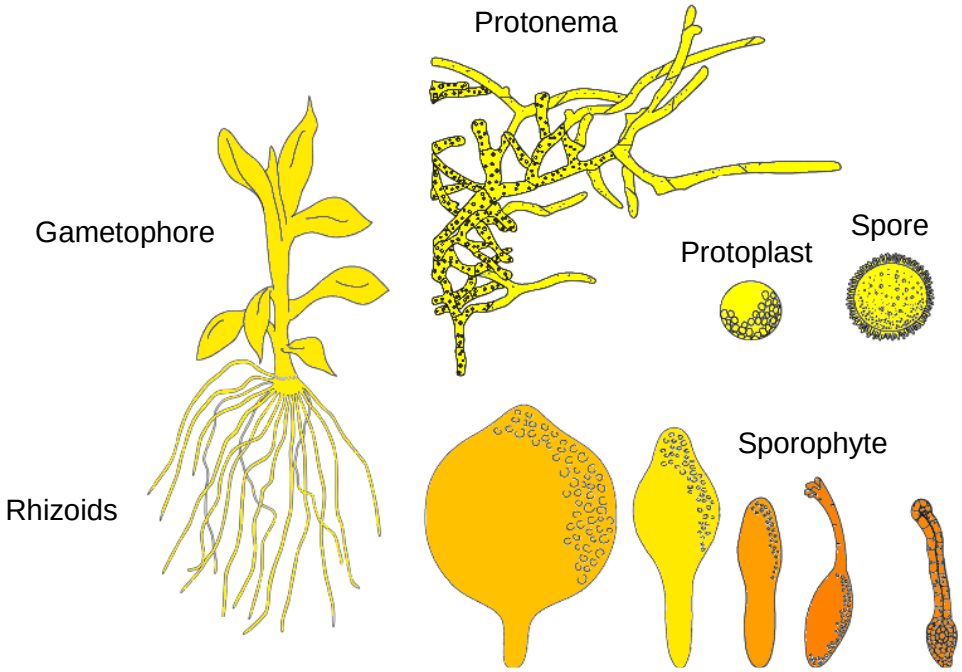


Figure S1. Expression profiles of *PpSMF1* and *PpSMF2* adapted from *Physcomitrella patens* eFP browser at bar.utoronto.ca[1]. For (**a**) *PpSMF1* (Pp1s71_321V6.1) maximal expression observed in the developing sporophyte and (**b**) *PpSMF2* (Pp1s519_13V6.1) maximal expression observed in rhizoids. Expression levels for both schematics were set to a signal threshold of 2400 to illustrate the differences in expression profiles between *PpSMF1* and *PpSMF2*. Bright red colouring represents high expression, orange represent moderate expression and bright yellow represents little to no expression. Under our conditions we found *PpSMF2* to be upregulated in the sporophyte relative to the protonema.

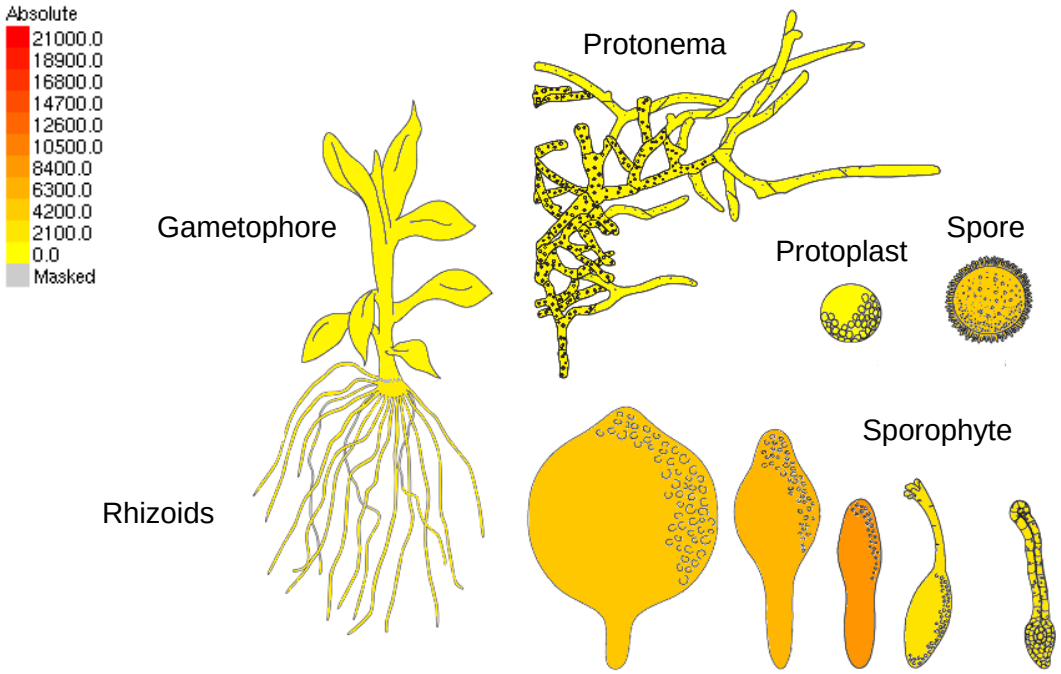
(a) *PpSCRM1* (Pp3c10_4280V3.9)



(b) Pp3c2_16410V3.3 (Pp1s30_360V6.1)



(c) Pp3c_20960V3.1 (Pp1s257_22V6.1)



(d) Pp3c8_18070V3.1 (Pp1s212_62V6.1)

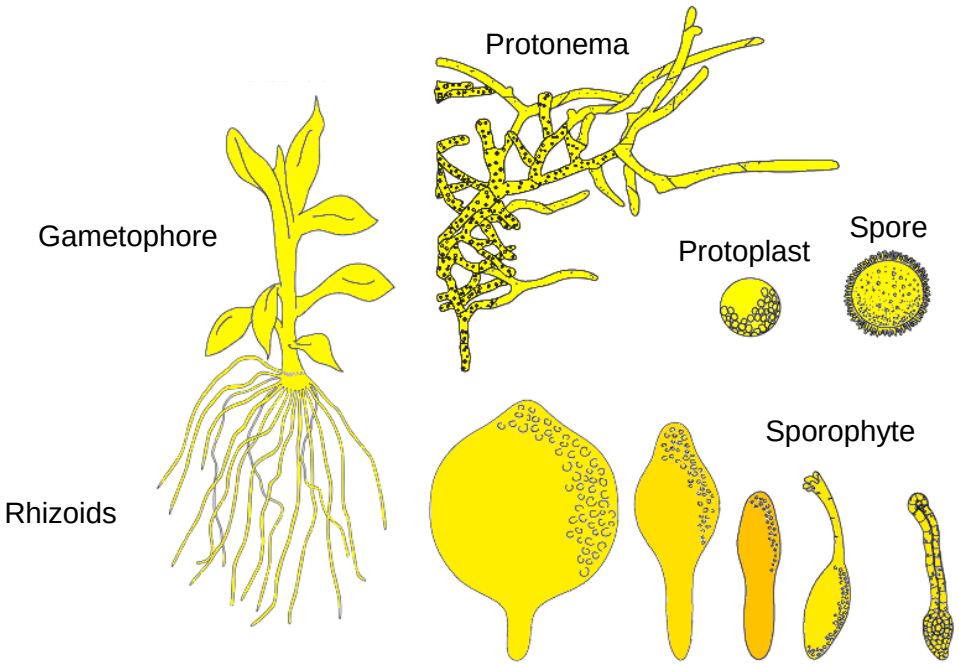


Figure S2. Expression profiles of *PpSCRM1*, and three PpSCRM inparalogues according to the *Physcomitrella patens* eFP browser at bar.utoronto.ca[1]. (a) *For PpSCRM1* (Pp1s51_178V6.1/Pp3c10_4260V3.9) maximum expression is observed in the developing sporophyte. The remaining three inparalogous genes (b) Pp3c2_16410V3.3 (Pp1s30_360V6.1), (c) Pp3c_20960V3.1 (Pp1s257_22V6.1) and (d) Pp3c8_18070V3.1 (Pp1s212_62V6.1) also showed maximum expression in the developing sporophyte, but comparatively to *PpSCRM1* observed expression was considerably less. Expression levels for all four schematics were set to a common signal threshold of 21000 to illustrate the differences in expression between *PpSCRM1* and the three other identified inparalogues. Bright red colouring represents high expression, orange represent moderate expression and bright yellow represents little to no expression.

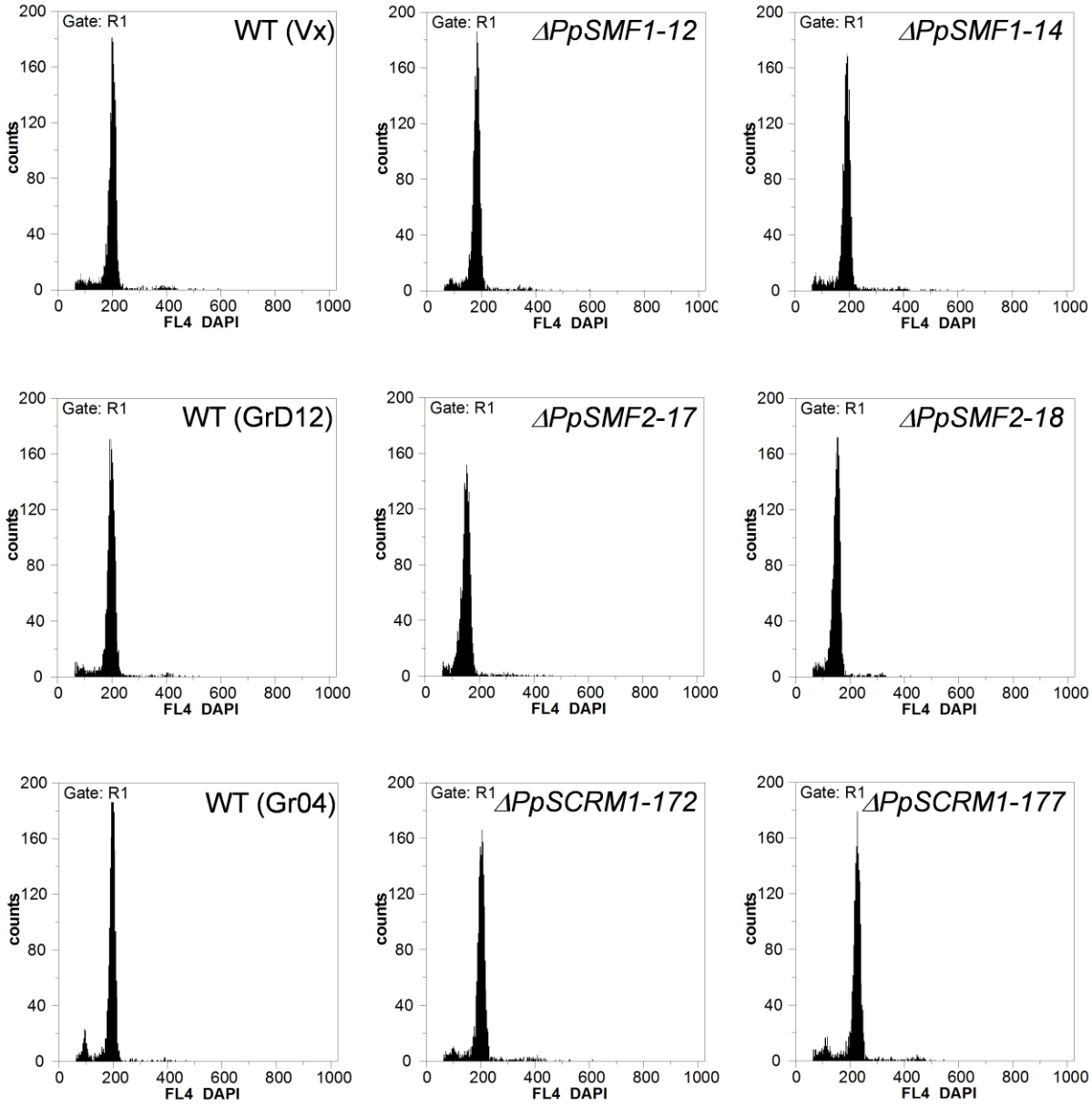


Figure S3. Flow cytometry (FCM) analysis of the different Δ*PpSMF1*, Δ*PpSMF2*, and Δ*PpSCRM1* knockout lines as well as the WT strains which were the genotypes used for gene targeting[2]. The histograms show a dominant peak around 200 for all lines, indicating they maintained the regular haploid stage.

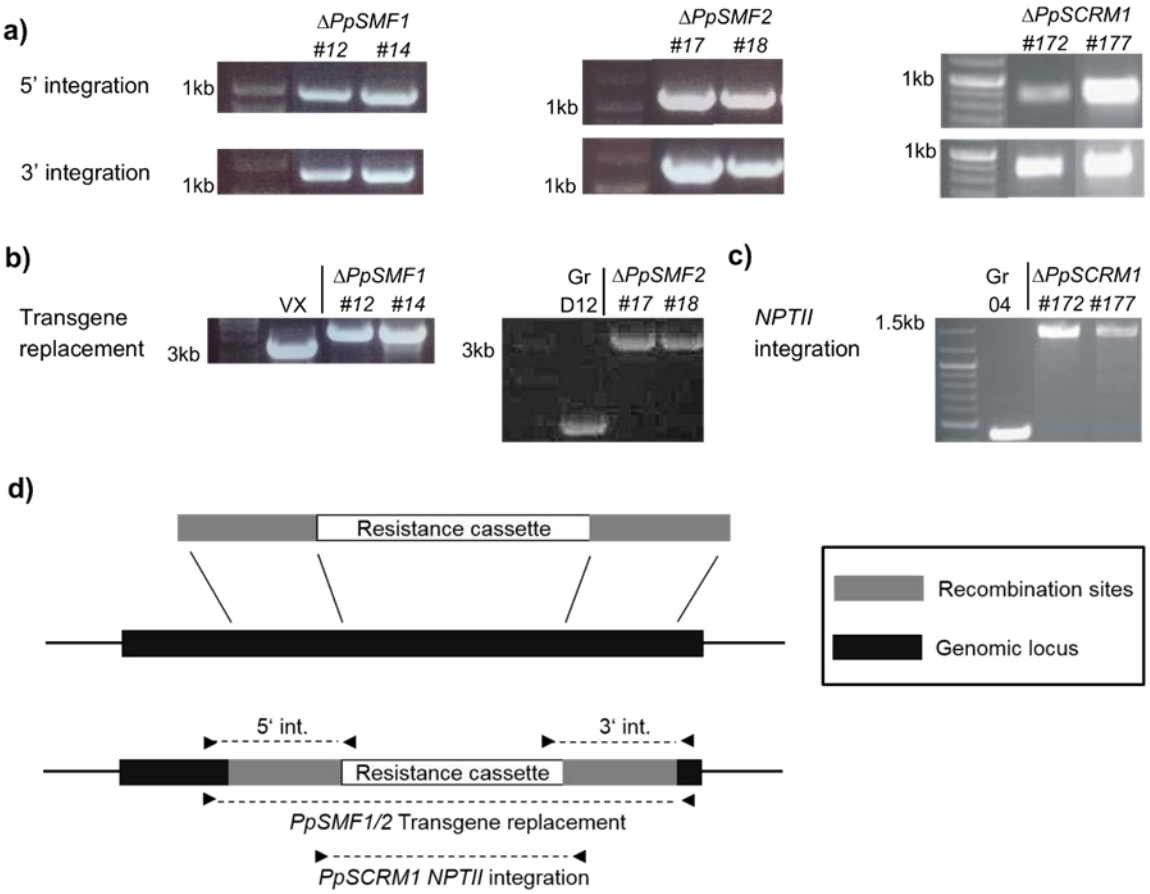
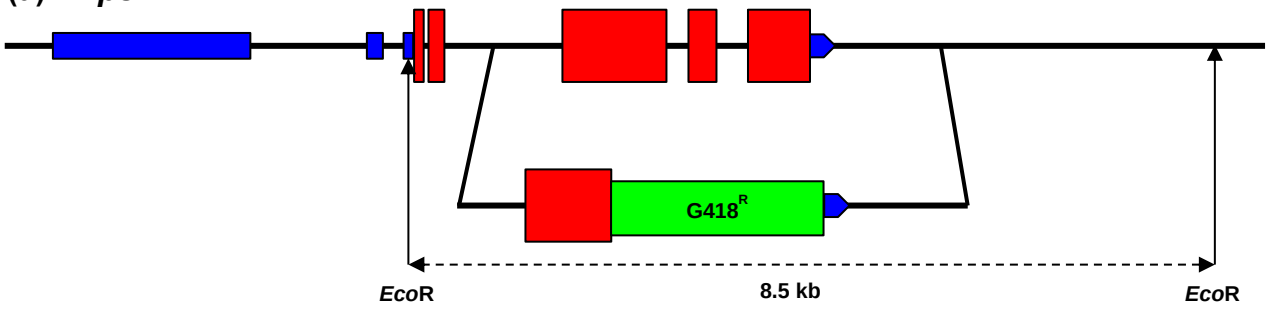
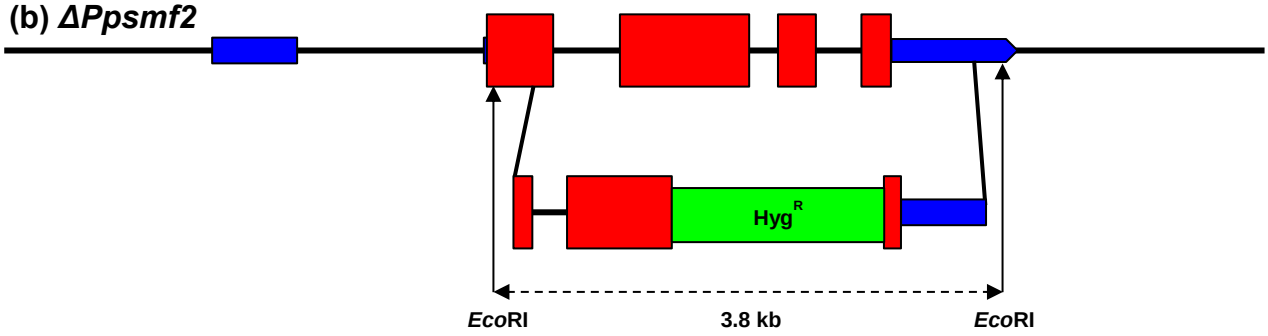


Figure S4. Confirmation of transgene targeting to the loci of *PpSMF1, PpSMF2* and *PpSCRM1* in the corresponding mutants. (a) PCR amplification of genomic regions at the 5' and 3' of the respective target genes showing insertion of the transgene resistance cassette in *ΔPpsmf1*, *ΔPpsmf2* and *ΔPpscrm1* lines. (b) PCR amplification showing a single transgene replacement of the genomic regions of *PpSMF1* and *PpSMF2* in *ΔPpsmf1* and *ΔPpsmf2* lines. (c) PCR amplification illustrating the integration of the *NPTII* selection cassette into the PpSCRM1 gene locus of the *ΔPpscrm1* lines. (d) Schematic highlighting the PCR amplicons presented in (a) to (c). Arrows represent approximate positions targeted by the forward and reverse primers to determine correct transgene targeting.

(a) *ΔPpsmf1*



(b) *ΔPpsmf2*



(c) Southern blot

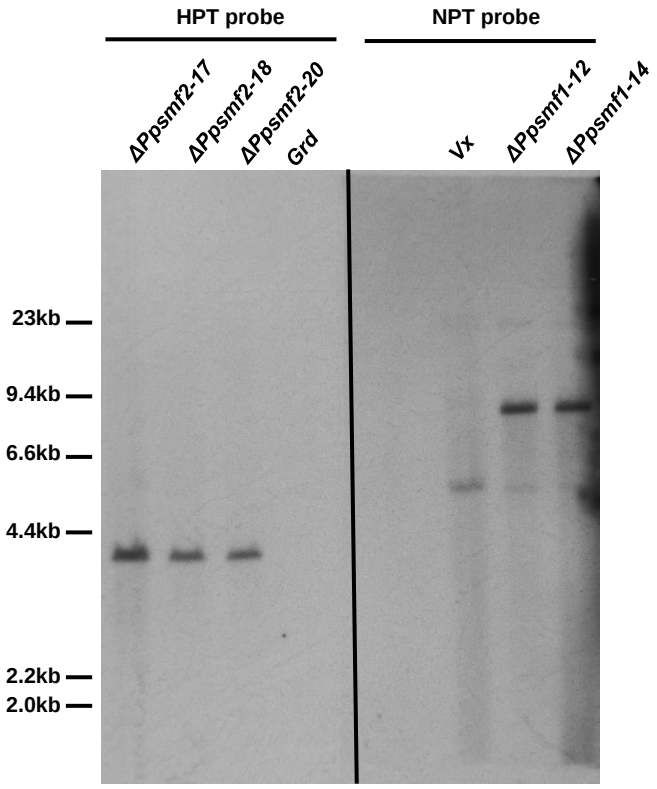


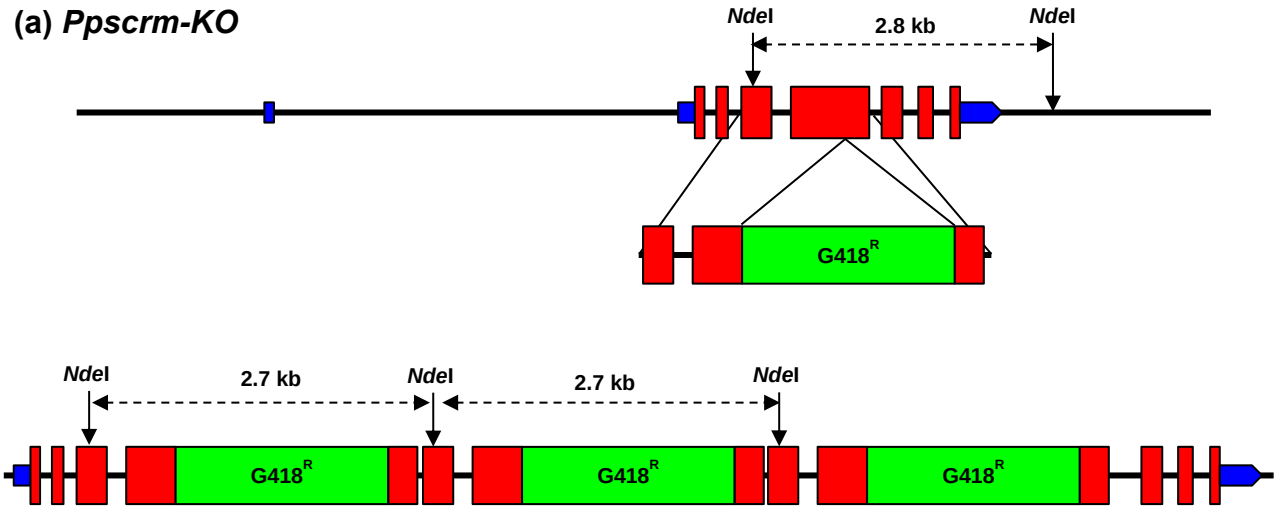
Figure S5. Southern blot analysis of mutant smf lines of *Physcomitrella*.

(a) Schematic of the Wild-type *PpSMF1* locus and the targeting construct indicating the recombination positions resulting in allele replacement in *ΔPpsmf1*.

(b) Schematic of the Wild-type *PpSMF2* locus and the targeting construct indicating the recombination positions resulting in allele replacement in *ΔPpsmf2*.

(c) Southern blot analysis of independent lines for *ΔPpsmf1* and *ΔPpsmf2* all displaying single copy insertions of the relevant cassette. Genomic DNA was digested with EcoRI and probed with an HPT probe for detection of the hygromycin selection cassette in the *ΔPpsmf2* lines and with an NPT probe to detect the kanamycin selection cassette in *ΔPpsmf1* lines. In each case a single hybridising fragment of the expected size was detected, indicating a single-copy gene replacement in each line. *Grd* and *Vx* correspond to Wild-type DNA of the Gransden and Villersexel genotypes, respectively. The additional faint hybridisation signals seen in the *ΔPpsmf1* lines correspond to cross-hybridisation with endogenous WT DNA.

(a) *Ppscrm-KO*



(b) Multicopy tandem integration at the *PpSCRM* locus

(c) Southern blot analysis of *ΔPpscrm-KO* lines.

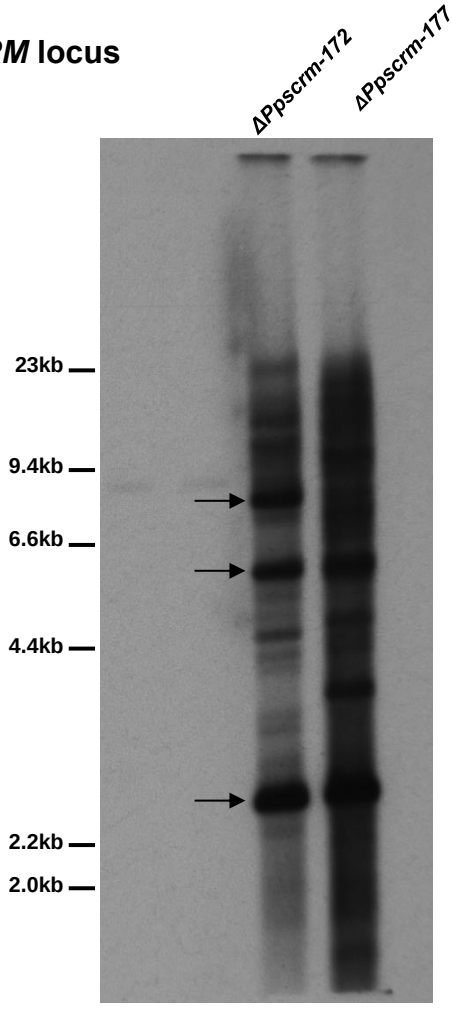


Figure S6. Southern blot analysis of mutant *scrm* lines of *Physcomitrella*.

(a) Schematic of Wild-type *PpSCRM* locus and the targeting construct indicating the recombination positions resulting in allele replacement in *ΔPpscrm*.

(b) Schematic of targeted replacement of the *PpSCRM* gene by a multiply tandem repeated copy of the targeting construct in *ΔPpscrm.*

(c) Two independent multiple disruption *ΔPpscrm* lines. Genomic DNA was digested with *Nde*I and probed with an NPT probe for detection of the kanamycin selection cassette. Multiple hybridising fragments indicate (i) that the targeted gene replacement comprised integration of tandem multiple copies, indicated by the strongly hybridising fragments at 2.7kb, 5.4 kb and 8.1 kb (arrowed) that would correspond to a NPTII monomer, dimer and trimer, respectively. The additional fragments likely result from adventitious integration at off-target sites in the genome, of vectors that may additionally have undergone fragmentation and/or rearrangement. (Note in the two tracks immediately adjacent to the *ΔPpscrm* DNA can be faintly seen the single-copy *ΔPpsmf1* fragments revealed in this very much shorter exposure of the filter in Figure S5)

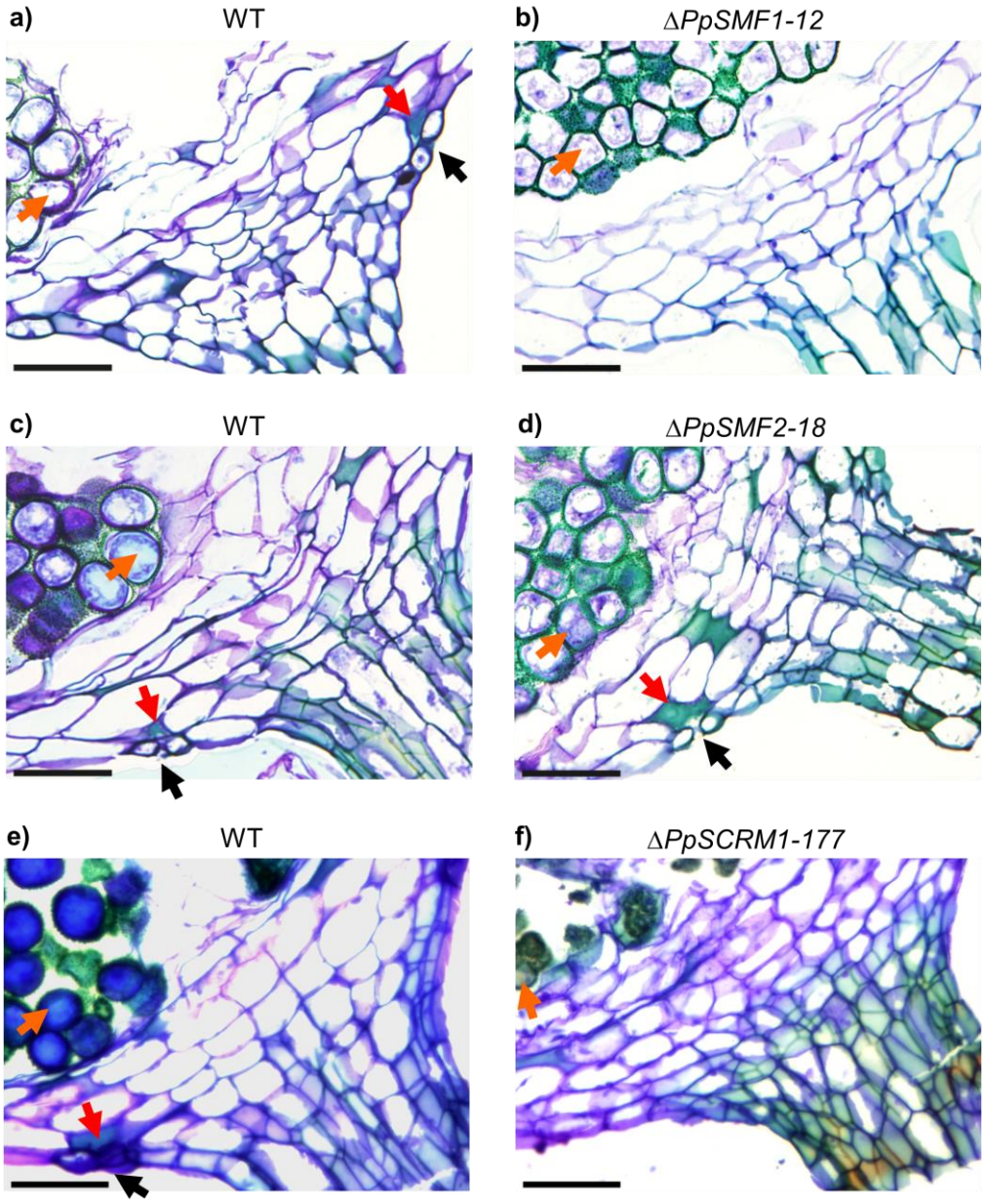


Figure S7. Sub-stomatal cavities are absent from mature Δ*PpSMF1* and Δ*PpSCRM1* mutant sporophytes but present in both WT and Δ*PpSMF2* sporophytes. (a) Cross-section of the basal region of a WT (Villersexel K3) sporophyte with a representative stoma (black arrow) and sub-stomatal cavity (red arrow). A representative mature spore is indicated (orange arrow) in the above spore sac. (b) Equivalent Δ*PpSMF1* sporophyte capsule showing an absence of stomata or sub-stomatal cavities. Mature spores can be seen in the spore sac. (c) Cross-section of the basal region of a WT (Gransden D12) sporophyte with representative stoma, underlying sub-stomatal cavity and spore (arrow coding as before). (d) A mutant Δ*PpSMF2* sporophyte capsule which, like the WT backgrounds, displays stomata with sub-

stomatal cavities and mature spores. (e) Cross-section of the basal region of a WT (Gransden 2004) sporophyte with representative stoma, underlying sub-stomatal cavity and spore sac (arrow coding as before). (f) Equivalent Δ*PpSCRM1* sporophyte capsule showing an absence of stomata and associated sub-stomatal cavities with mature spores highlighted in the above. The scale bars equal 50µm in all images.

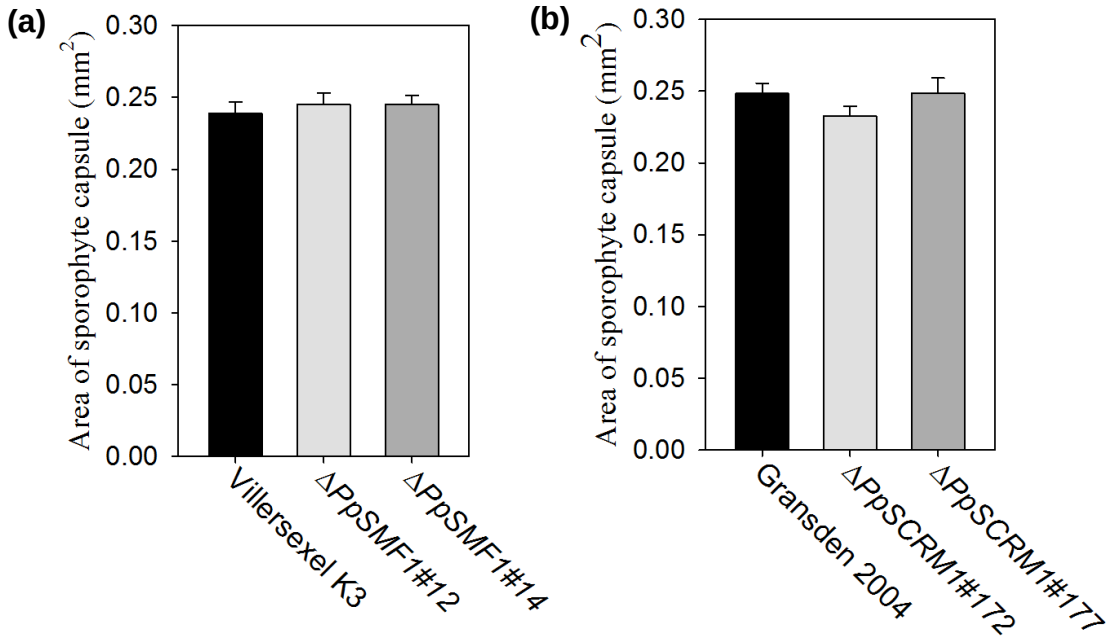


Figure S8. Mature spore capsule size (area) does not differ between (a) Δ*PpSMF1* mutants and corresponding wild-type or between (b) Δ*PpSCRM1* mutants and corresponding wild-type. For each wild-type and mutant line 50 mature browning sporophyte capsules randomly collected from 2 peat pellets were analyzed (25 capsules from each pellet). In the case of Δ*PpSCRM1#177*, 25 sporophytes were analyzed. Capsules were collected 5 ½ weeks after the application of water. Sample diameter was measured using the oval function of Image J to measure the main body of the sporophyte capsule minus the protruding area (the undifferentiated operculum) where the calyptra sits during sporophyte expansion. No significant differences were found when One-way ANOVA indicated no significant differences mutant lines and corresponding wild-types ($P > 0.05$).

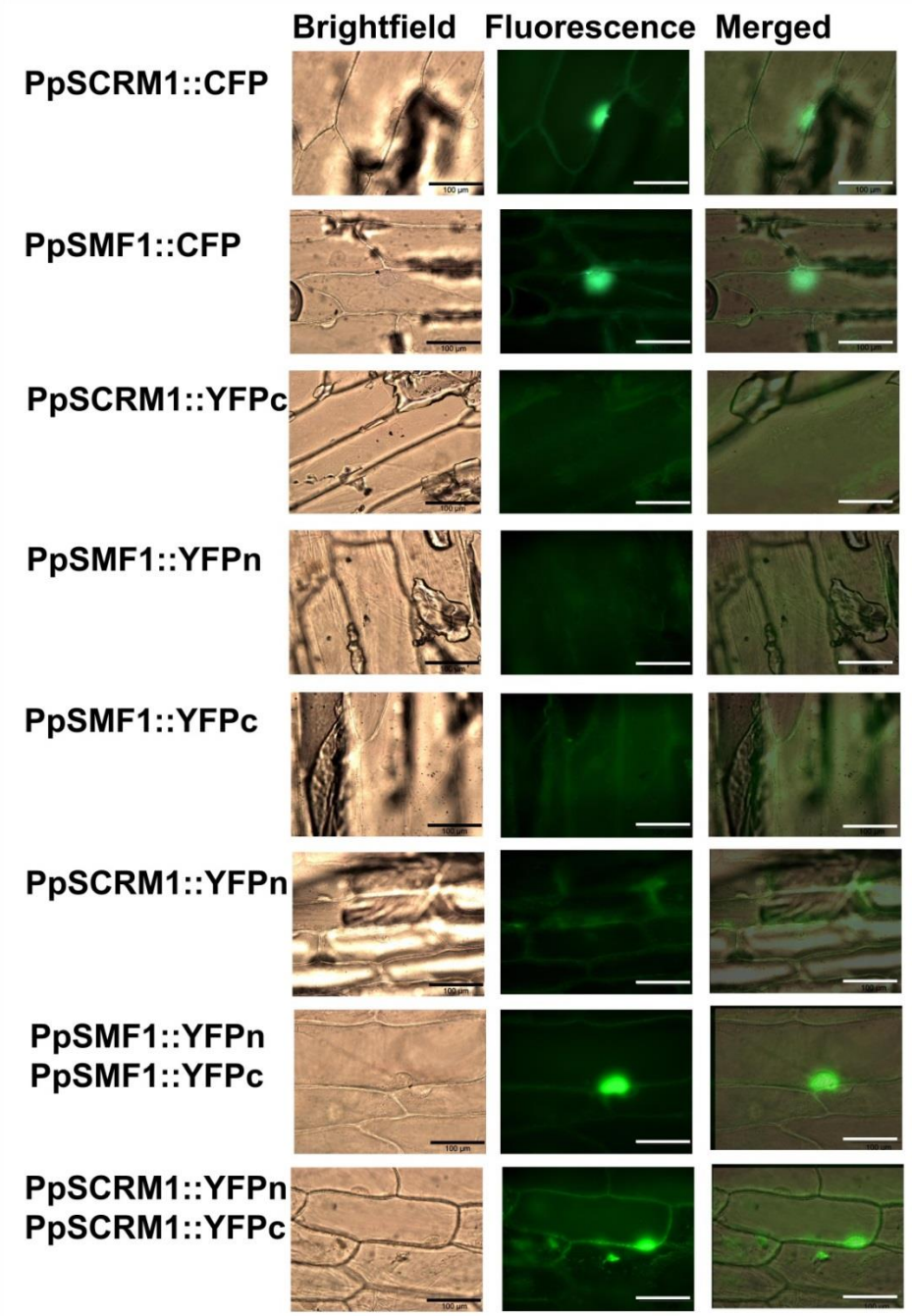


Figure S9. Bimolecular fluorescence complementation assays demonstrating PpSMF1 and PpSCRM1 protein-protein interactions. Representative bright-field microscopy, fluorescence and overlay/merged images of BiFC analysis showing pairwise combinations of bHLH constructs, each fused with full-length CFP molecule or a complementary, half-YFP molecule (YFPn fusion and YFPc fusions). In the intact *Allium cepa* PpSMF1::CFP and PpSCRM1::CFP controls showed strong nuclear localisation. Individual half-YFP controls gave no signal, thus confirming the specific interaction. PpSMF1 formed homodimers in a similar manner to AtMUTE; a property which has been lost in AtSPCH and AtFAMA[3]. Similarly, PpSCRM1 was able to homodimerise, unlike AtSCRMs. Scale bars indicate 100 µm.

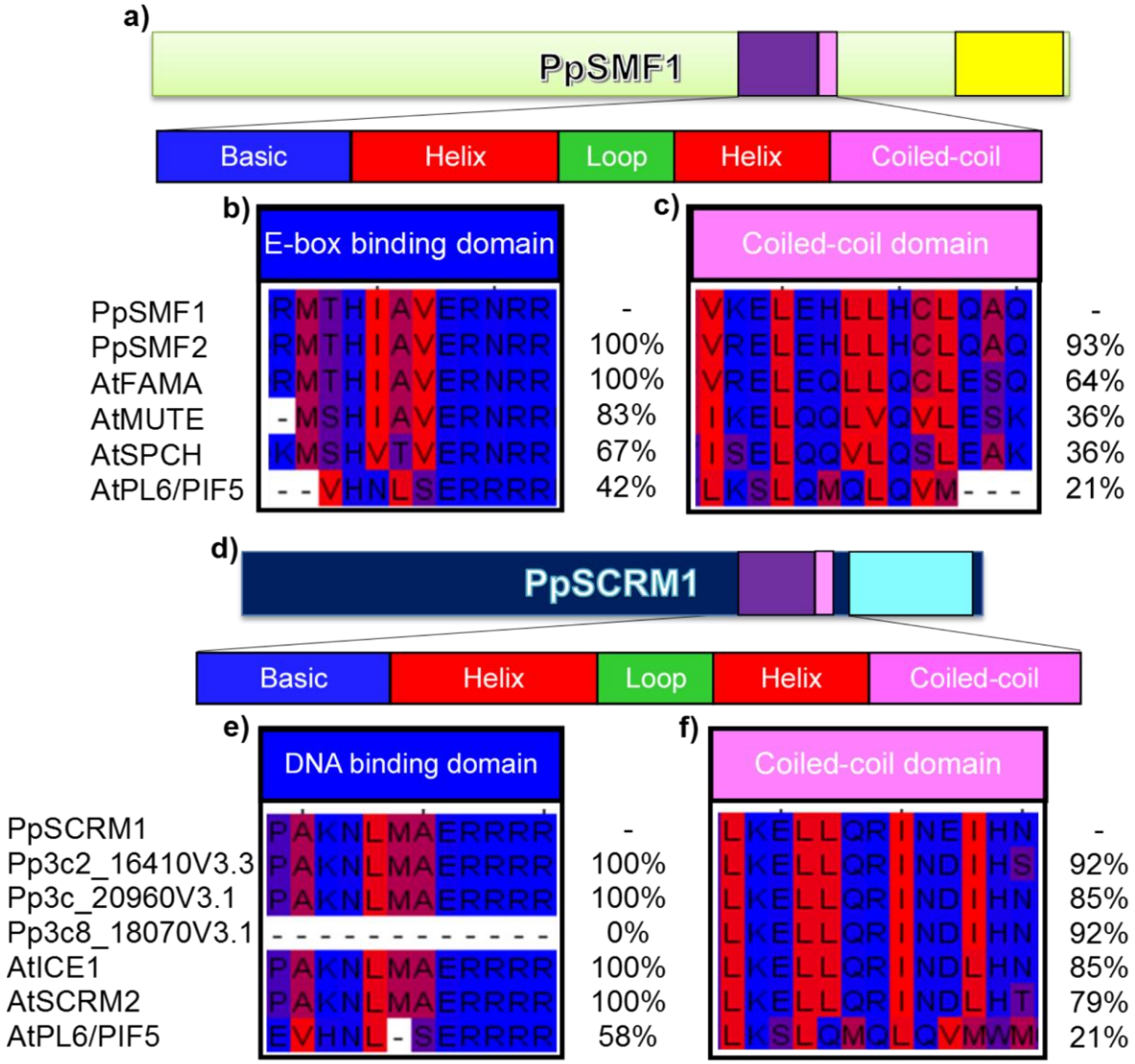


Figure S10. Conservation of amino acid motifs in stomatal development genes. (a) *P. patens* SMF1 protein with bHLH (purple) and SMF (yellow) domains highlighted.(b) Amino acid sequence alignment of Ebox binding domains of *P. patens* and *A. thaliana* proteins regulating stoma development. (c) Amino acid sequence alignment of coiled-coil domains of *P. patens* and *A. thaliana* proteins regulating stoma development (d) *P. patens* SCRM1 protein with bHLH (purple) and SCRM (light blue) domains highlighted. (e) Amino acid sequence alignment of the DNA binding domains of *P. patens* and *A. thaliana* SCRM homologues. The non-stomatal-associated bHLH transcription factor AtPL6/PIF5 was used as an outgroup in panels b, c, e and f to show the high level of conservation between related SMFs and SCRMs and their corresponding orthologues. Percentages relate to amino acid identity when aligned with *P. patens* SMF1 (b and c) or SCRM1 (e and f). All sequence alignments were performed on Jalview[4] using the MUSCLE alignment algorithm[5]. Predicted coiled-coil domains were identified using JNet prediction software in JalView[4]. Amino acids residues marked red have hydrophobic properties, as the colour becomes bluer the residues concerned have more hydrophilic properties.

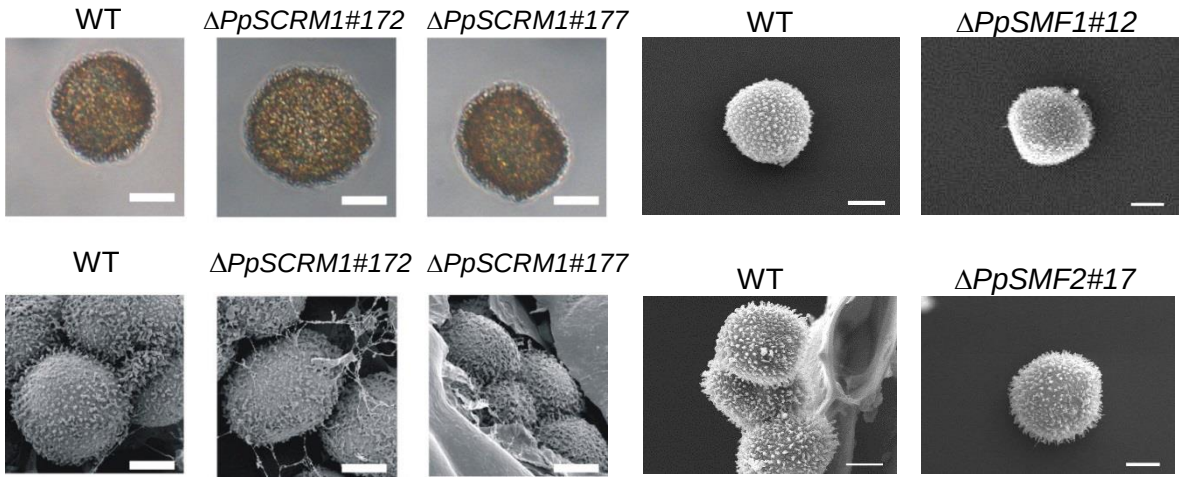


Figure S11. Spore morphology of wild-types and mutant lines. Left upper panels, light microscopy images of wild-type, Δ*PpSCRM1#172* and Δ*PpSCRM1#177* spores. Left lower panels, SEM images of wild-type, Δ*PpSCRM1#172* and Δ*PpSCRM1#177* spores. Right upper panels, SEM images of wild-type and Δ*PpSMF1#12* spores. Right lower panel, SEM images of WT and Δ*PpSMF2#17* spores. Scale bars indicate 10 µm. Light microscopy (Δ*PpSCRM1* lines only) and electron scanning microscopy analyses showed no difference in size, shape or surface structure between wild types and corresponding mutant backgrounds. Absence of stomata did not influence the spore morphology in either *ΔPpSMF1*, *ΔPpSMF2*, or *ΔPpSCRM1*.

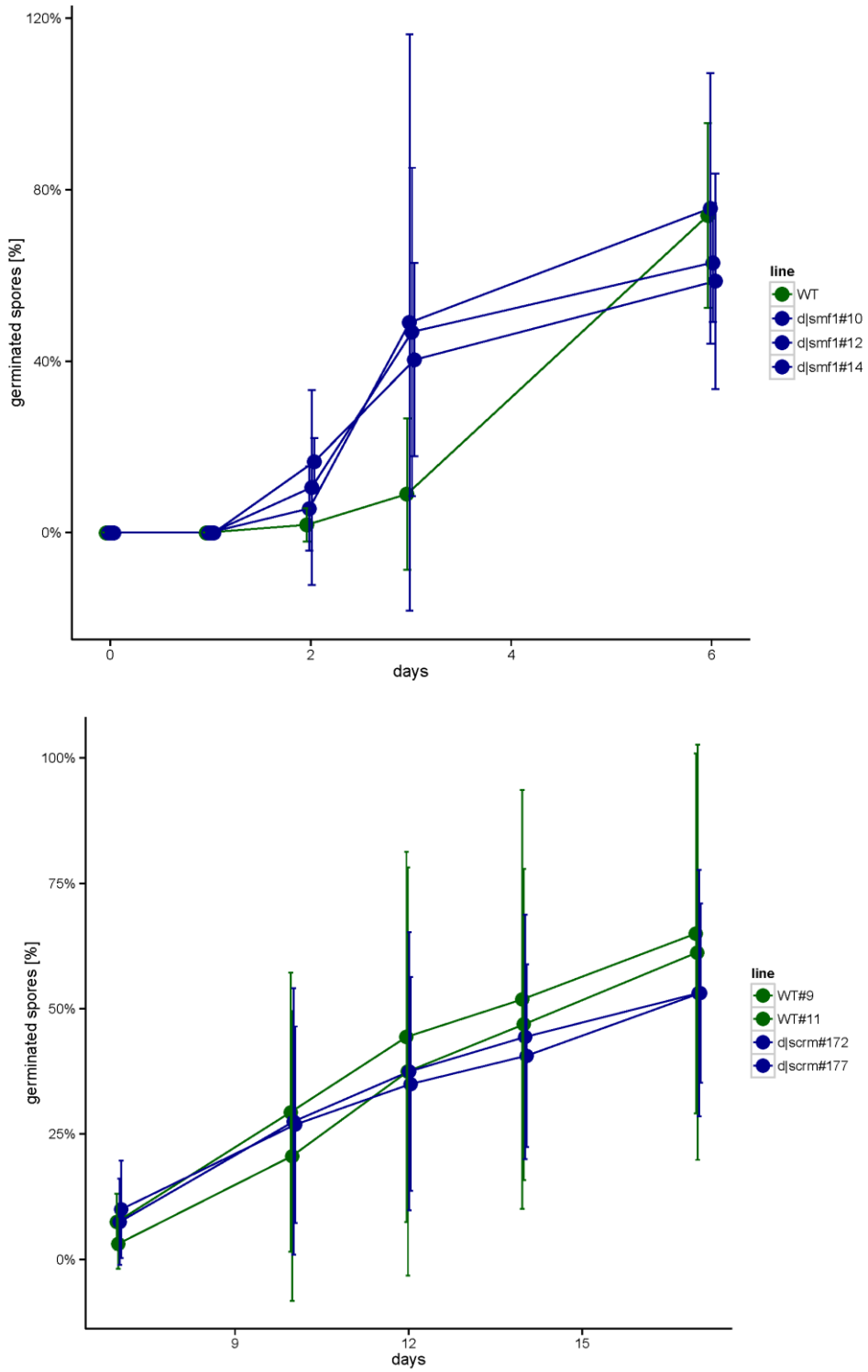
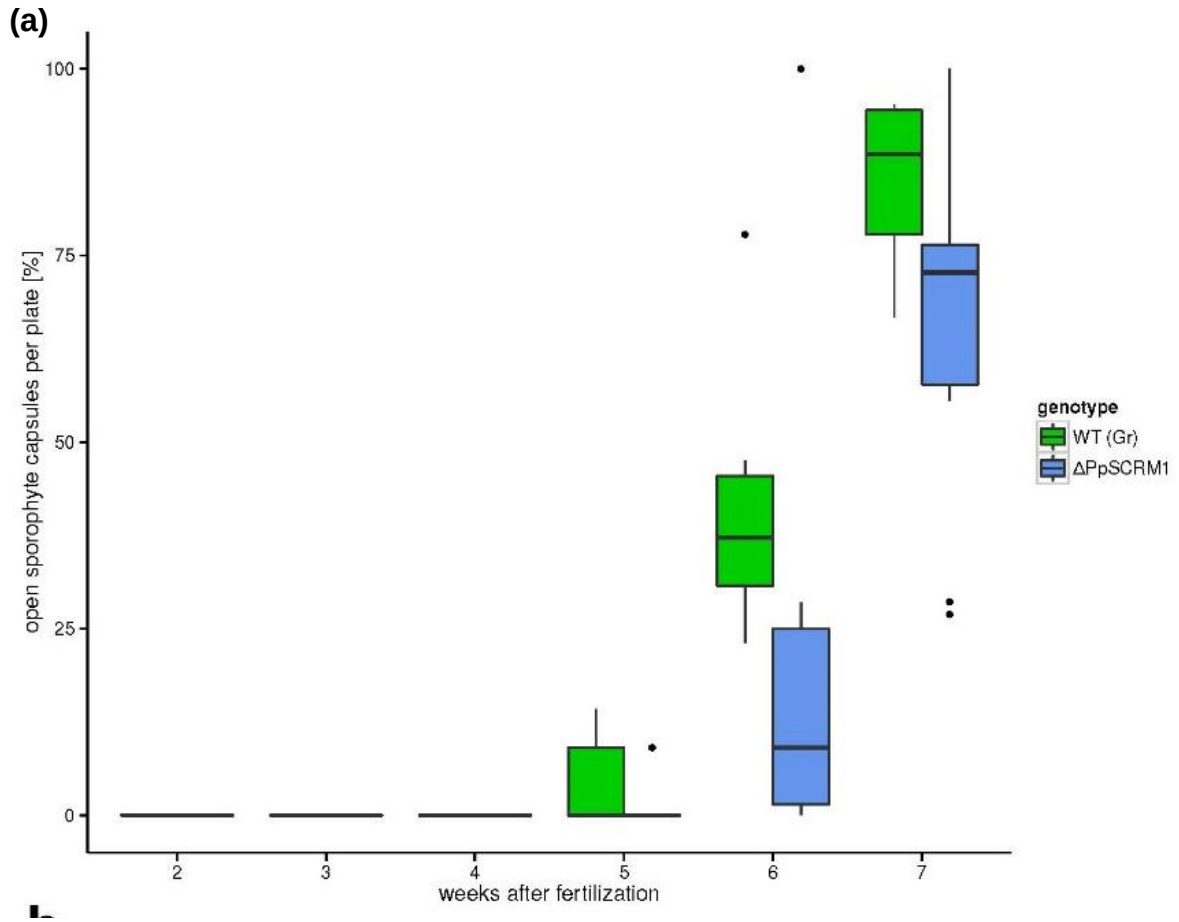


Figure S12. Loss of stomata in *Physcomitrella* has no effect on spore germination success. Generalized linear modelling indicates no significant differences for comparisons of WT versus *ΔPpSMF1* (P = 0.442) or WT versus *ΔPpSCRM1* (P = 0.748). Bars show means ± 95% confidence limits based on spore germination assessments on at least 40 spores per line.


(a)
100
75
50
25
0
open sporophyte capsules per plate [%]
2
3
4
5
6
7
weeks after fertilization
genotype
WT (Gr)
ΔPpSCRM1

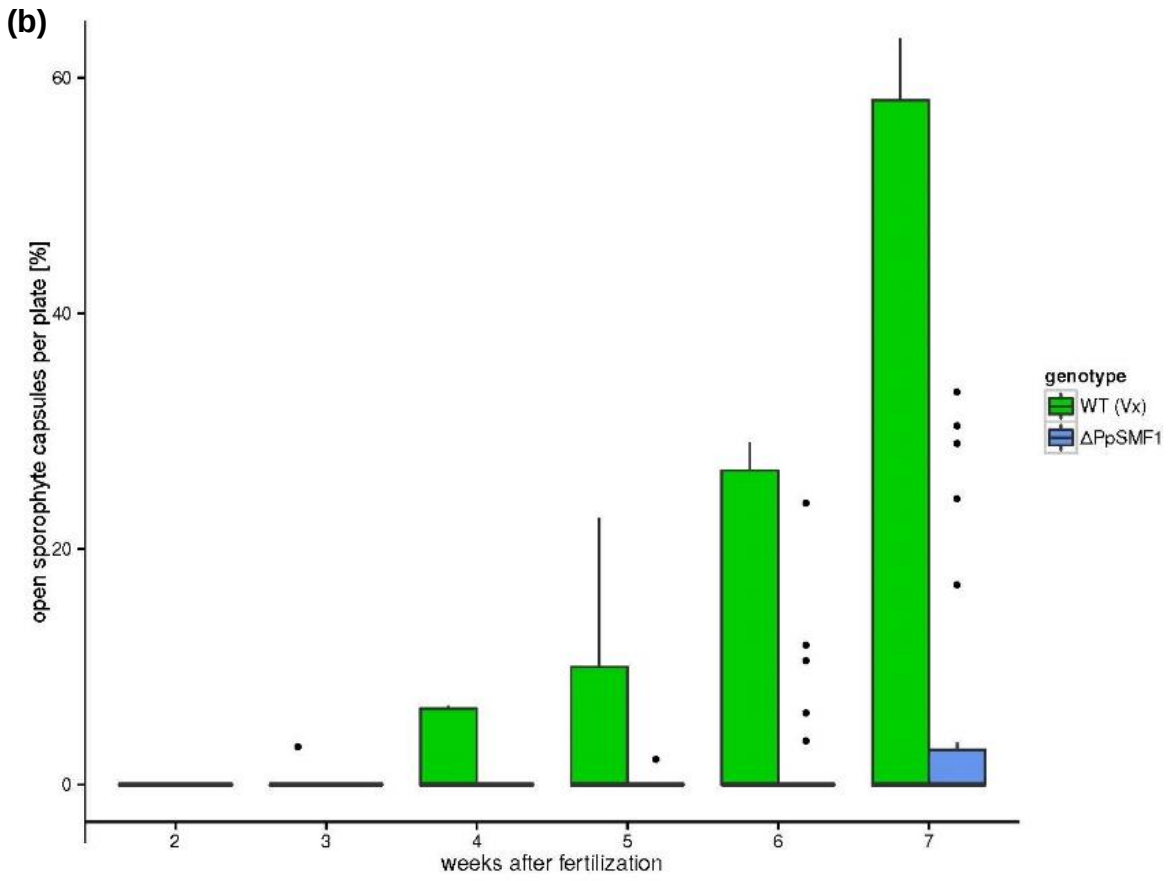

(b)
60
40
20
0
open sporophyte capsules per plate [%]
2
3
4
5
6
7
weeks after fertilization
genotype
WT (Vx)
ΔPpSMF1

Figure S13. Loss of *PpSMF1* and the *PpSCRM1* gene functions delays dehiscence of spore capsules. Box-whisker plots of the percentages of ruptured sporophyte capsules in the WT, Δ*PpSMF1* and Δ*PpSCRM1* lines over a developmental time series experiment ranging from second and seventh week after induction of fertilization. The plots depict the same data as Figure 4, splitting up the different wild-type controls. Vertical lines within boxes mark the median. The boxes indicate the upper (75 %) and lower (25 %) quartiles. The whiskers indicate the ranges of the minimal and maximal values.

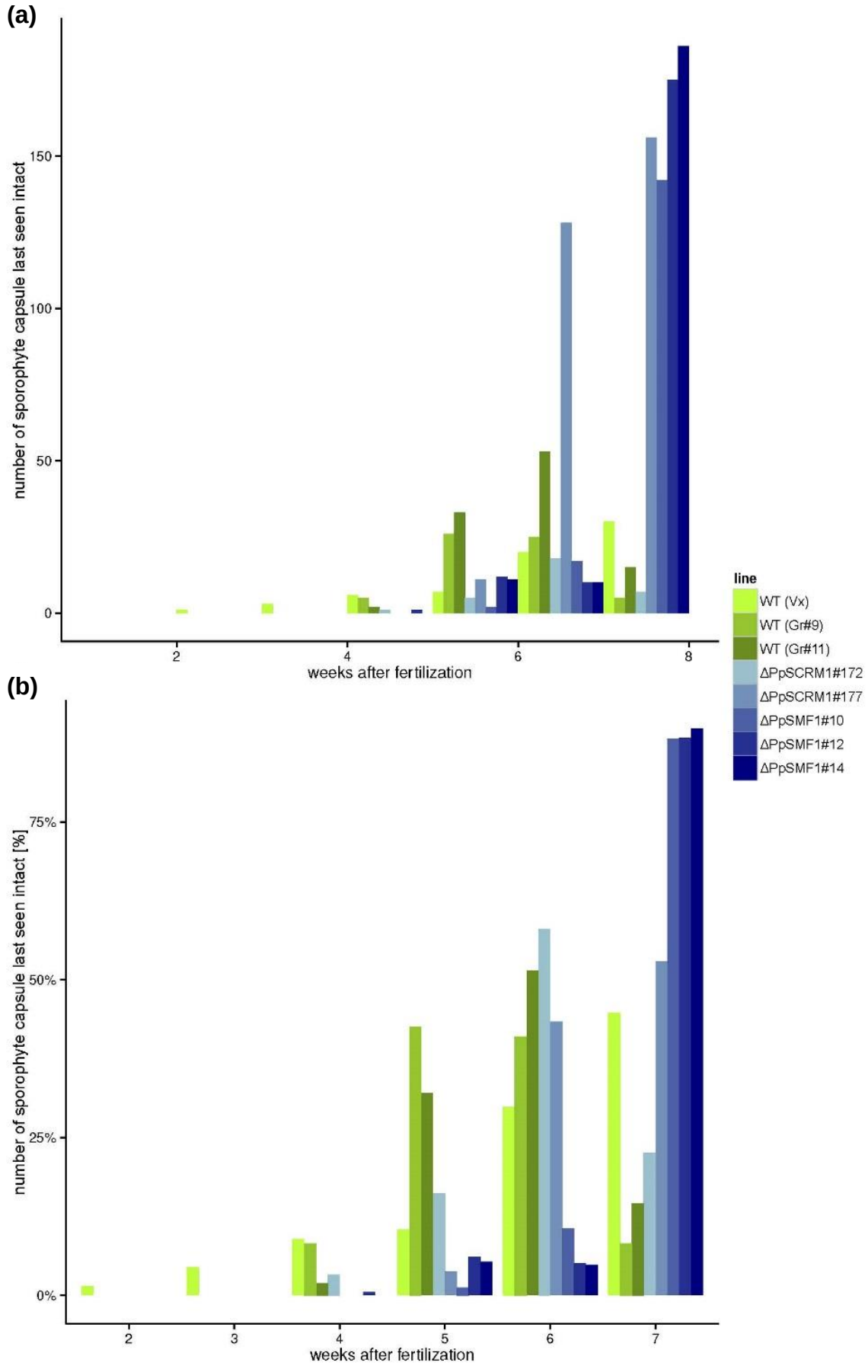

(a)
number of sporophyte capsule last seen intact
150
100
50
0
2
4
6
8
weeks after fertilization
(b)
number of sporophyte capsule last seen intact [%]
75%
50%
25%
0%
2
3
4
5
6
7
weeks after fertilization
line
WT (Vx)
WT (Gr#9)
WT (Gr#11)
ΔPpSCRM1#172
ΔPpSCRM1#177
ΔPpSMF1#10
ΔPpSMF1#12
ΔPpSMF1#14

Figure S14. Loss of *PpSMF1* and the *PpSCRM1* gene functions increases number of capsules intact compared to WT. Developmental time series of sporophyte dehiscence. Individual spore capsules were tracked and the week where an individual capsule was observed intact for the last time was recorded and analysed. The bar plots depict absolute (a) and relative (b) number of spore capsules that have been recorded as still intact for the WT, Δ*PpSMF1* and Δ*PpSCRM1* lines.

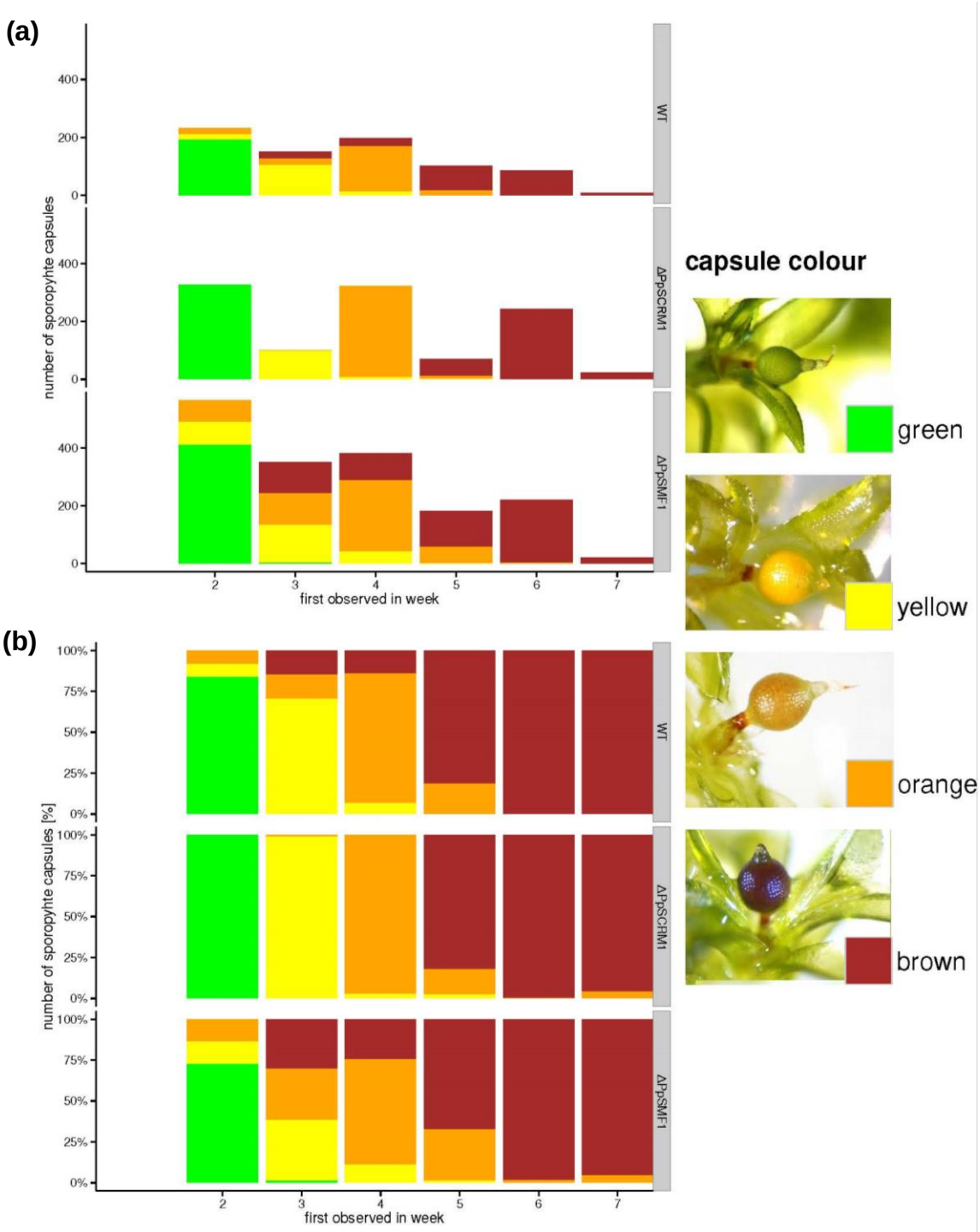


Figure S15. Deletion of *PpSCRM1* results in a delayed transition of capsule colour from green to brown. Developmental time series of sporophyte maturation tracked by changes in capsule colour. The emergence and colour phase transitions of individual maturing spore capsules were tracked. The bar plots depict absolute (a) and relative (b) number of spore capsules that were recorded as green, yellow, orange or brown for the first time in that week for the WT, Δ*PpSMF1* and Δ*PpSCRM1* genotypes in separate panels (grey boxes). Key comprises inset photographs with wild-type examples of the respective colour state.

Supplementary Methods

Phylogenetic analysis of the SMF and SCRM/ICE families of bHLH transcription factors. Members of the SMF and SCRM/ICE gene families in *Arabidopsis thaliana, Populus trichocarpa, Oryza sativa, Sorghum bicolor, Selaginella moellendorffii* and *Physcomitrella patens* were identified based on the results of previous studies of bHLH transcription factors[6-8] and genome-wide gene family definitions available as Hidden Markov Models which were utilized to screen the V3.3 moss proteome (Supplementary Files 1-4)[9]. Previously identified gene family members of *P. patens* were replaced by the more complete gene models from the latest V3.3 release of the moss genome. Protein sequences of identified members were aligned using PASTA and resulting multiple alignments were mapped to coding sequences using the Perl script protal2dna kindly provided by Katja Schuerer and Catherine Letondal. In an iterative process phylogenetic trees were inferred and, together with the tree-sorted multiple alignments and the respective genome browsers, used to identify the evolutionary conserved splice variants and possible haplotype alleles in the case of *Selaginella moellendorffii*. Inference of bootstrapped maximum likelihood trees was performed using RAxML comparing the GTRGAMMAIX and GTRCATIX models. On the final selected trees, speciation and duplication events were inferred manually based on the underlying species tree.

Sporophyte culture conditions. Sporophyte development (Sheffield) was optimised on 42 mm Jiffy 7 peat pellets (Amazon, London) under sterile conditions. Seven day old protonemal tissue grown on BCDAT was homogenised using a Polytron P1200 (KINEMATICA AG, Luzern, Switzerland in 15 ml of sterile H_2O. Dried peat pellets were rehydrated with 40 ml of DH_2O and then autoclaved inside sealed Magenta GA-7 pots (Sigma-Aldrich, Gillingham, UK). Autoclaved pellets were inoculated with 1.5 ml of moss-DH_2O homogenate. Around peat pellets 60 ml of sterile DH_2O was added and the Magentas were sealed using Micropore tape

(3 M) and grown for 8-10 weeks at 25 ºC continuous light. For gametangia induction, samples were then grown at 18 ºC, 100 μmol m^{-2} s^{-1} irradiance, for 10 hours a day, followed by 15 °C for 14 hours in a Sanyo Medicool MPR-161D(H) fitted with Phillips Master TL-D 90 De Luxe 18W/965 fluorescent lamps. After 2-3 weeks, 40 ml of sterile DH_2O water was poured directly onto the moss.

For RT-PCR analysis, ~25 gametophores with at least 15 sporophytes were collected 3 weeks after the application of water. For qPCR analysis immature green sporophyte capsules were picked after 2-3 weeks. For phenotypic analysis, fully expanded mature browning capsules were collected 6 weeks after water application.

Southern Blot method. For Southern blot analysis, gDNA (5 μg) was digested with either *Eco*RI (SMF1 and SMF2 loci) or *Nde*I (SCRM locus) and electrophoretically resolved using a 0.7% agarose gel, transferred to Amersham Hybond-N+ membrane (GE Healthcare, Canada) and UV cross-linked to the membrane (UV Stratalinker, USA). Antibiotic resistance cassette-specific digoxigenin (DIG)-labelled probes were produced by PCR using the NPTII probe primers for the *Δsmf1*, and *Δscrm* lines and the HPTII probe primers for *Δsmf2* lines. Probe sizes were 1221 bp for NPTII, and 883 bp for HPTII. Probe digoxigenin (DIG)-labelling, Southern hybridization and immunological detection were performed using the DIG High Prime DNA Labelling and Detection Starter Kit II (Roche, Germany) according to the manufacturer's instructions.

Yeast 2-Hybrid Assays. The Yeast 2-Hybrid assay was performed using the Matchmaker Gold System (Clontech). Bait (PpSMF1) and prey (PpSCRM1) constructs were cloned by recombination in the vectors pGBKT7 and pGADT7-Rec, respectively, by co-transforming PCR amplicons and vector into yeast strains Y2HGold and Y187. Recombinants were identified by colony-PCR and verified as being in-frame by sequencing. The strains were mated

and spread on selection medium (SD/-Trp/-Leu). Diploids were streaked onto QDO medium (-Trp, -Leu, -His, -Ade) containing X-α-Gal (40μg.ml^{-1}) and Aureobasidin A (125ng.ml^{-1}). See Supplemental Table S1 online for the primer sequences used. The control plasmids, pGADT7-T and pGBKT7-53 were supplied in the Matchmaker Gold System.

Transient expression by microprojectile bombardment. For BiFC experiments, cDNAs of *PpSMF1* and *PpSCRM1* were amplified by PCR (Supplementary Table 1 for the primers) and cloned into plasmids containing the N- and C-terminal halves of YFP (pDH51-GW-YFPn and pDH51-GW-YFPc, respectively) to create *PpSMF1::YFPn/YFPc* and *PpSCRM1::YFPn/YFPc*. Equal quantities of each *YFP* construct were mixed and delivered to *Allium cepa* epidermal tissue by microprojectile bombardment, using a PDS1000 Biolistic system (Bio-Rad, Hemel Hempstead, UK). Plasmid DNA (1 μg per shot) was bound to tungsten microprojectiles (M17, Bio-Rad)[10] for delivery using a 1100 psi rupture disc and a distance of 6 cm from the stopping screen. Following bombardment, plant tissue was incubated for 24 h at 25 ºC in the dark prior to microscopic examination.

Sporophyte maturation rate analysis and lysis experiments. For sporophyte maturation rate analysis and lysis experiments 3 to 5 mature spore capsules were collected from sterile Knop medium plates of WT and knock-out mutants. Spore capsules were placed into 5 ml of liquid Knop medium[11,12], lysed and 500 μl of Knop medium was then pipetted onto fresh individual Knop plates and sealed with Micropore tape and Parafilm 'M'Laboratory Film (Pechiney Plastic Packaging, Chicago). Spores were grown for 10 days under continuous light at 25 ºC. Individual plants were isolated and transferred to new plates and grown for 16 h a day at 25 ºC, with an irradiance of 140 μmol m^{-2} s^{-1} for 5 weeks with both Micropore tape and Parafilm affixed. Plants were then moved to the above mentioned customised Sanyo Medicool growth

chamber set to 15 ºC, 20 µmol m^{-2} s^{-1}, for 10 hours per day and grown for 4 weeks. Plants were then flooded with 10 ml of sterile H_2O and resealed with only Parafilm. After 5 days water was removed and plates were resealed with Parafilm. Weekly rotation and checks were performed to check spore capsule colour and subsequent lysis. For each line, 5 replicate plates were used made with 7 to 10 individual plants analysed per plate.

Spore germination experiments. For spore germination experiments, spore capsules were harvested and air dried for 3–5 d. The capsules were surface-sterilised with 4 % bleach (4.5 % available chlorine) solution and washed with sterile water. Capsules were crushed in 1 ml of sterile water to create a spore suspension that was plated onto BCDAT routine basal medium supplemented with sterile 10 mM calcium chloride. Using a dissection microscope (Leica, Solms, Germany) the number of germinated spores was counted each day for 5 d after the spores were plated. The percentage of spores which germinated was calculated for each day (three replicates per *P. patens* line). A minimum of 200 spores per replicate were included in the analysis.

Assessment of the ploidy of the mutant lines via FCM. The ploidy level of the mutants was determined via Flow cytometry measurements (FCM) of protonema material according to Schween *et al.*[2]. Measurements were performed with a CyFlow Space system (Partec GmbH) which was kindly provided by Prof. Lutz Hein, Institute of Experimental and Clinical Pharmacology and Toxicology of the University of Freiburg. DAPI-stained nuclei were excited with an UV LED and fluorescence was analysed with the FlowMax Software. The haploid stage for *Physcomitrella* protonema is indicated by a main peak at 200, representing nuclei in the G2 phase of the cell cycle. Polyploidisation is indicated by multiplication of the DAPI signals.

Statistical analysis of sporophyte development. Visual and statistical analysis of the data was performed using R utilizing several analysis packages[13-19]. PostHoc hypothesis testing to assess differences in sporophyte development was carried out using simultaneous inference in general parametric models as implemented in the R multcomp package[13]. The performance of Gaussian and Poisson models were compared based on AIC and the overall variability explained by the models. Repeated measurements of capsules dehiscence (Figure 4) was modelled for the combined data and for each genetic background separately using a nested error term correcting for the lab, genetic background, line, plate and capsule with a binomial model implemented using glmer from the lme4 package[19].Unless stated otherwise, significance was assessed and reported at 99.9% confidence level.

SEM preparation details. For SEM examination fresh spore capsules were first fixed in 3 % glutaraldehyde in 0.1 M phosphate buffer for 4 h at 4 °C and then followed standard procedures. Specimens were washed in 0.1 M phosphate buffer twice (15 min each time) at 4 °C. Secondary fixation was carried out in 2 % aqueous osmium tetroxide for 60 min at room temperature and wash step repeated. Specimens were dehydrated in a graded series of ethanol at room temperature (75 % ethanol for 15 min, 95 % ethanol for 15 min, 100 % ethanol for 15 min × 2, 100 % ethanol dried over anhydrous copper sulphate for 15 min). Specimens were critical point dried and mounted on 12.5-mm diameter specimen stubs using adhesive carbon tape. Samples were coated with 25 nm of gold in an Edwards (Crawley, UK) S150B sputter coater and viewed and imaged with a Philips (Amsterdam, Netherlands) XL-20 microscope at an accelerating voltage of 10 KV.

Technovit embedding and sectioning. Mature orange-to-brown 6-7 week old sporophyte capsules were collected and cleared using an ethanol dehydration series. Cleared capsules were imbedded using Technovit 7100 ® (TAAB, www.taab.co.uk) and mounted using

Technovit 3040 ® (TAAB, www.taab.co.uk). Sample sections of 5-10µm were acquired using a Leica RMZ145 microtome and subsequently stained using Toluidine blue for approximately 30 seconds, rinsed in ddH_2O, and mounted in Eukitt® quick-hardening mounting medium (Sigma-Aldrich, UK).

Further Discussion

Phylogenetic analysis of the *SMF* and *SCRM*/*ICE* families of bHLH transcription factors

The inferred phylogenetic tree of the *SMF* gene family (Figure 1d) suggests the following evolutionary scenario: The last common ancestor (LCA) of the land plants included in this study already had a copy of a *FAMA*-like gene. We suggest such a gene may have had some *MUTE*-like functionality. Duplication events then occurred early in the evolution of flowering plants which lead to the formation of the functionally divergent *SPCH* and *MUTE* gene families. The lycophyte *Selaginella moellendorffii* harbours an additional branch of the *FAMA* family that could be linked to the existence of an additional type of stoma found in the leaf margins of these plants[20,21], but requires broader taxonomic sampling for clarification. Inspection of the multiple sequence alignment and the genomic context of the second *FAMA* clade in *Selaginella* (blue node in Figure 1d) suggest that the two loci represent remnant alleles resulting from the sequencing of two mixed haplotypes[22]. Like the additional copies of *FAMA* and *SPCH* in *Populus trichocarpa* and *SPCH* in *Oryza sativa* and *Sorghum bicolor* possibly resulting from ancestral (genome) duplication events in *Populus* and the grasses, the two *FAMA* co-orthologs initially named *PpSMF1* and *PpSMF2* are clear inparalogues that arose from an ancestral duplication event. In line with the qualitatively similar but quantitatively distinct expression profiles and the observed phylogenetic distance, we can assume expressional divergence but not divergence of protein function in the two '*PpFAMA'* (*PpSMF1* and

PpSMF2) inparalogues (Supp. Info. Fig. 1). From a phylogenetic perspective the latter should be renamed to *PpFAMA1* and *PpFAMA2*.

With a total of three duplication events, the phylogenetic tree of the *SCRM/ICE* family of bHLH transcription factors suggests an expansion of this clade in the moss lineage, with four *SCRM/ICE* co-orthologues in the *Physcomitrella patens* genome. All studied flowering plants have two inparalogous copies that arose by single, separate duplication events in the grass/monocot ancestor and along the lineages leading to *Arabidopsis* and poplar, respectively. The lycophyte *Selaginella* most likely has only one copy, as the second locus again appears to be of haplotypic origin. Expression profiles across the developmental stages of the moss (Supp. Info. Fig. 2) provide evidence for expressional divergence of the moss *SCRM/ICE* co-orthologues. Although similarities among the overall spatial expression patterns of the inparalogues could result in the milder dehiscence phenotype observed in *ΔPpSCRM1* mutants, there are specific differences. While *PpSCRM1* is clearly the most abundant inparalogue in mature and developing sporophytes which is also active in most gametophytic tissues, the other *SCRM* inparalogues show slightly divergent patterns either with higher levels in spores (Pp3c1_20960V3.1 and Pp3c8_18070V3.1) or in archegonia and early sporophytes (Pp3c2_16410V3.3). Expansion of *SCMR/ICE* in the moss lineage and the expression patterns is indicative of additional functions of this bHLH family in the gametophytic generation.

Bimolecular fluorescence complementation (BiFC) results suggest that a PpSMF1-PpSCRM1 heterodimer can form *in-vivo* (**Figure 3**, Supp. Info. Fig. 9). Equivalent interactions have been seen between vascular land plant orthologues, implying that the protein-protein interactions and the individual peptides are highly conserved[3]. Our *in-silico* data suggests that an Ebox binding domain (EBD) in PpSMF1 and PpSCRM1, a DNA binding domain in PpSCRM1, and coiled-coil domains in both peptides, share a high degree of sequence similarity with their *A. thaliana* counterparts (Supp. Info. Fig. 10). This further strengthens the idea of the conservation of a heterodimer and not just individual peptides.

In *Arabidopsis*, the EBD in FAMA is critical for allowing the final differentiation step of stomatal development to occur, but this domain is less important in SPCH and MUTE function earlier in stomatal development[23]. Interestingly, in PpSMF1 and PpSMF2 the EBD domains (including the previously identified H-E-R residues[23] and surrounding residues) are identical to FAMA which suggests similar regulatory targets (Supp. Info. Fig. 10). We found that, like with *Arabidopsis* MUTE but not SPCH or FAMA[3], PpSMF1 homodimerised suggesting it has MUTE-like functionality (Supp. Info. Fig. 10). *PpSMF2* expression is very low in *P. patens* even in the sporophyte (Supp. Info. Fig. 1) and therefore it is unsurprising its absence does not alter the wild-type stomatal phenotype. Because *PpSMF1* complements *Arabidopsis* mutants better than *PpSMF2* in *Arabidopsis*[6], it is probable that PpSMF1 has other (as yet unspecified) domains not present in PpSMF2 that permit better targeting of regulatory elements and or interactions with other molecules. We highlight here that PpSMF1 exhibits properties associated with both MUTE and FAMA[3,6], suggesting that PpSMF1 may function similarly to an ancestral MUTE-FAMA-like bHLH[23].

Comparison of sporophyte development in mutant and wild type plants. We assessed the impact of the loss of *PpSMF1* and *PpSCRM1* gene function on sporophyte development, by undertaking a comprehensive developmental study of spore capsule formation in Δ*PpSMF1* and Δ*PpSCRM1* mutant lines in comparison to WT. In total, we followed the development of 1,630 sporophytes in the 8 lines (WT Gransden background n = 180, 129 capsules; WT Villersexel n = 106; *ΔPpSMF1*: 236, 240, 296 capsules; Δ*PpSCRM1*: 92, 352) for which we recorded capsule formation, colour and dehiscence status weekly until the 7th week after initiation of fertilization. We considered only sporophytes that had already developed as green capsules in week 2 after fertilization (See Supp. Info. Fig. 15 for an example). This was true for a total of 1,123 capsules (WT Gransden 2004 background: 61,

103 capsules on 4, 5 plates; WT Villersexel: 67 capsules on 9 plates; Δ*PpSMF1*: 161, 198, 207 capsules on 8 plates each; Δ*PpSCRM1*: 31 and 295 capsules on 6 and 5 plates, respectively).

Development of *P. patens* sporophytes is accompanied by a change in the colouration of the sporophyte, transitioning from young, green, developing sporophytes to yellow and orange capsules, to brown capsules with mature spores (Supp. Info. Fig. 15). Changes in capsule colour are observed in many other bryophytes and has been attributed to changes in flavonoid content and degree of spore maturation[24-26]. We tested for the effect of the loss of *PpSMF1* and *PpSCRM1* function on this trait by comparing sporophytes of mutant and wild-type plants by recording the week in which a transition of capsule colouring occurred (Supp. Info. Fig. 15). Overall, the Δ*PpSCRM1* mutant lines seem to be delayed in the transition from green to brown capsules as compared to the wild-type and the Δ*PpSMF1* lines. This difference is most pronounced in the timing of browning of the capsules – Δ*PpSCRM1* sporophytes are brown after 6 weeks, while Δ*PpSMF1* and the wild-type controls showed brown capsules after a median of 5 weeks ($P<0.001$) (see Methods for statistical analyses).

We tested whether stomata and the genes coordinating stomatal development are involved in capsule dehiscence in *P. patens* by analysing the timing of capsule rupturing in Δ*PpSCRM1* and Δ*PpSMF1* mutant lines in comparison to the wild-type (**Figure 4**, Supp. Info. Figs. 13 and 14). As indicated by the percentage of open spore capsules following the seven weeks after fertilization (**Figure 4** and Supp. Info. Fig. 14), and the timing of capsule dehiscence as measured by the week a capsule was last recorded as intact (Supp. Info. Fig. 15), the three Δ*PpSMF1* capsules rupture significantly ($P < 0.001$) later than the three wild-type and the two Δ*PpSCRM1* mutant capsules. The majority of capsules are open after a median of 6 and 6.5 weeks in the wild-type and the Δ*PpSCRM1* mutants respectively, but the majority of Δ*PpSMF1* sporophytes are still closed in week 7. Although the difference is not as pronounced

compared to the Δ*PpSMF1* mutants, the majority of Δ*PpSCRM1* capsules also dehisce significantly later than the wild-type ($P < 0.01$). These results clearly support the hypothesis that stomatal function in *P. patens* is directly linked to reproductive success; stomata are required for efficient spore capsule dehiscence. These findings strongly support the proposed role of stomata in the facilitation of water loss in late moss sporophytes[27].

Stomatal function in mosses. Evolution of stomata on the sporophytes of otherwise poikilohydric mosses and hornworts, but their absence on the dominant gametophytic generation remains unexplained. The nutritional dependence of bryophyte sporophytes on the gametophyte suggests that the sporophyte's partial homoiohydry (*i.e.*, stomata, waxy cuticle, stratified tissue structure) has evolved instead as a consequence of water dependence. Nevertheless, moss stomata are assumed to play a direct and indirect role in sporophyte nutrition; to enable sporophytic photosynthesis by gas exchange among the multi-layered tissues in developed green sporophytes with developing spores and by establishing a transpirational gradient necessary for assimilate and water transport from the gametophyte to the expanding and pre-dehiscent capsules[28]. *P. patens* sporophytes differ from those of their Funariaceae relatives (such as *F. hygrometrica*) by having a highly reduced seta; the stalk (in the case of *F. hygrometrica*) elevates spore capsules above the laminar boundary layer to enable aerial spore dispersal. Elevation of the capsule necessitates an increased transpirational gradient which is possibly reflected in the higher number of stomata found in *F. hygrometrica* compared to *P. patens* sporophytes[29].

In contrast to the sporophytes of most mosses which develop complex peristome structures dedicated to the release of the mature spores, the sporophytes of the cleistocarpous moss *P. patens* lack these structures and instead release their spores by rupturing of the capsule walls (Figure 4). Evolutionary loss of the peristome and reduction of the seta are assumed to be the result of secondary reduction[30,31]. Given this apparent reduction of sporophytic traits,

the evolutionary retention of stomata in the cleistocarpous moss *P. patens* appears puzzling. Highly reduced sporophytes of several other moss species have lost their stomata entirely[32,33]. A possible explanation may be found in those relatives of *P. patens* that still maintain a complex sporophyte. In the distantly-related peristomate *Sphagnum* mosses there are indications for the involvement of pseudo-stomata in maturational water loss prior to capsule dehiscence[29]. In the peristomate moss *F. hygrometrica* the stomata open during capsule expansion[34] and, like the stomata of *P. patens* and vascular plants, can control their apertures in response to external stimuli such as ABA and CO_2[35]. After this period, stomata remain open and are assumed to aid the dehydration which promotes the shrinkage of the sporophytic tissues and thus generate the forces required to shed the operculum in the case of *F. hygrometrica* or rupture the capsule in the case of *P. patens*[28,29,36,37].

Supplemental Table S1. Primers used in this study

Name and function	Forward Primer	Reverse Primer
PpSMF1 flanking 71RB	CACCGCATCACATCCATCGAAGGA	GAGCACAAAACTTTCTTCGG
PpSMF1 flanking 71LB	GCATGCGGCCGCGGCGCGCCGGTCCGAGTACAATTGTCC	CGATGCGGCCGCGAGGCCACTCGCCTCTTGAG
PpSMF2 flanking 519 RB	CACCGGATCAAAGCTCATTTTCGC	ATGCGTAGTGGTAGCACTCG
PpSMF2 flanking 519 LB	GCATGCGGCCGCGGCGCGCCTTGACGAGAAATTTGAAGG	CGATGCGGCCGCGATGCCACTTCCTTCATGCA
PpSCRM1 KO	TTTGAATTCGCAGACTGGAGTGCAGGAAA	TTTGAATTCCTCCACACCCCACTGATTCG
SMF1 RT-PCR	ATCCAAGTGGTGACGAAGCA	AGCGGCAATGGAGACTGATT
SMF1 External locus	CCTGTTGGTTGTGATTGCAG	CCTGTTGGTTGTGATTGCAG
SMF1 5’ genome integration	CCAGCAACAGCAACTACTCC	TCCACTAGTTCTAGAGCGGC
SMF1 3’ genome integration	TCCACTAGTTCTAGAGCGGC	TGCAACCAAACTGAGCGAAA
SMF2 RT-PCR	GAGCCAACGGATGACACACA	AAATCGAAGGGTTCCCGAGG
SMF2 External locus	GACGAGAAATTTGAAGGTGG	ATGCGTAGTGGTAGCACTCG
SMF2 5’ genome integration	AGTGTCATGGACTACCCCGA	TAGCTGGGCAATGGAATCCG
SMF2 3’ genome integration	TAGGGTTTCGCTCATGTGTTGA	ATAAGCATGGTTGGCTTGTAGC
SCRM1 RT-PCR	GGACGTTGGACCAGAAGAAA	CGCTTTATTCAGCCTCCTCA
SCRM1 External locus	CCTGTTGGTTGTGATTGCAG	CCACAGAAGAATCGTTGGAG
SCRM1 5’ genome Integration	CGCTGGAGCTAGAGGAAGAG	GCGGCTGAGTGGCTCCTTCA
SCRM1 3’ genome Integration	AAATTATCGCGCGCGGTGTC	TACTCGAGGAGGCTGGGATT
qPCR Large Ribosomal Subunit	GACAGGCACAGGGTATTCCT	ATCTTCCGTCGTGTTGATCC
qPCR Small Ribosomal Subunit	ACGGACATTGCATTTAAGACCT	GTCGATTACCTGTGGAGAAGAC
qPCR ADEPRT	AGTATAGTCTAGAGTATGGTACCG	TAGCAATTTGATGGCAGCTC
qPCR SMF1	AACAGCCGTGATTCCCACAA	GTGGGCACTCCCAGATGAAA
qPCR SMF2	GCCAGGTTACTTCATCCAGAAGGG	AGTTCCCTGACGAATTCAATGGC
qPCR SCRM1	CCCTTGGCTTGGATGTACAACAG	CGTCCACATCTTTGGCCTCTG
pYFP-GS-PpSCRM1 primers	GCGTCGCAATCCAATTCC	CTACTGCAAAGAGTGCAAATCACA
pYFP-GS-PpSMF1 primers	GACTACCCCGAGACTGTGC	TCAGAATTGCAGAGAGTGTAGGG
pPpSCRM1-GS-YFP primers	ATGGCGTCGCAATCCAATTCC	CTGCAAAGAGTGCAAATCACATCC

pPpSMF1-GS-YFP primers	ATGGACTACCCCGAGACTGTGC	GAATTGCAGAGAGTGTAGGG
pGAD-SCRM1 Y2H primers	CGTACCAGATTACGCTCATATGAGTACTATGGCGTCGCAATCCAATTCC	GATGGATCCCGTATCGATGCCCACCCTAAGTACTCTGCAAAGAGTGCAAATCACATCC
pGBK-SMF1 Y2H primers	GAAGCTGATCTCAGAGGAGGACCTGAGTACTATGGACTACCCCGAGACTGTGC	CGCTGCAGGTCGACGGATCCCCGGGTCAAGTACTGAATTGCAGAGAGTGTAGGGAGCT
NPTII probe primers (SMF1 & SCRM)	AGAGAGATCTAGACCCCTACTCCAAAAATGTCAAA	GTCAAGAAGGCGATAGAAGGCGATG
HPTII probe primers (SMF2)	AGGGCGAAGAATCTCGTGCTTTCAG	TACTTCTACACAGCCATCGGTCCAG

Supplementary Table S2: International Moss Stock Center (http://www.moss-stock-center.org) accession numbers of plants used in this work.

Plant	IMSC No.
WT (Vx)	40703
ΔPpSMF1-12	40701
ΔPpSMF1-14	40702
WT (GrD12)	40707
ΔPpSMF2-17	40704
ΔPpSMF2-18	40705
WT (Gr04)	40373
ΔPpSCRM1-172	40663
ΔPpSCRM1-177	40664

Supplemental References

1 Ortiz-Ramirez, C. *et al.* A Transcriptome Atlas of Physcomitrella patens Provides Insights into the Evolution and Development of Land Plants. *Mol Plant* **9**, 205-220, doi:10.1016/j.molp.2015.12.002 (2016).

2 Hohe, A. *et al.* An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, Physcomitrella patens. *Curr Genet* **44**, 339-347, doi:10.1007/s00294-003-0458-4 (2004).

3 Kanaoka, M. M. *et al.* SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to Arabidopsis stomatal differentiation. *Plant Cell* **20**, 1775-1785, doi:10.1105/tpc.108.060848 (2008).

4 Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191, doi:10.1093/bioinformatics/btp033 (2009).

5 Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113, doi:10.1186/1471-2105-5-113 (2004).

6 MacAlister, C. A. & Bergmann, D. C. Sequence and function of basic helix-loop-helix proteins required for stomatal development in Arabidopsis are deeply conserved in land plants. *Evolution & Development* **13**, 182-192, doi:10.1111/j.1525-142X.2011.00468.x (2011).

7 Pires, N. & Dolan, L. Origin and Diversification of Basic-Helix-Loop-Helix Proteins in Plants. *Molecular Biology and Evolution* **27**, 862-874, doi:10.1093/molbev/msp288 (2010).

8 Richardt, S. *et al.* Microarray analysis of the moss Physcomitrella patens reveals evolutionarily conserved transcriptional regulation of salt stress and abscisic acid signalling. *Plant Mol Biol* **72**, 27-45, doi:10.1007/s11103-009-9550-6 (2010).

9 Zimmer, A. D. *et al.* Reannotation and extended community resources for the genome of the non-seed plant Physcomitrella patens provide insights into the evolution of plant gene structures and functions. *BMC Genomics* **14**, 1-20, doi:10.1186/1471-2164-14-498 (2013).

10 Sanford, J. C., Smith, F. D. & Russell, J. A. Optimizing the biolistic process for different biological applications. *Methods in Enzymology* **217**, 483-509 (1993).

11 Reski, R. & Abel, W. O. Induction of budding on choloronemata and caulonemata of the moss, *Physcomitrella patens*, using isopentenyladenine. *Planta* **165**, 354-358, doi:10.1007/bf00392232 (1985).

12 Egener, T. *et al.* High frequency of phenotypic deviations in Physcomitrella patens plants transformed with a gene-disruption library. *BMC plant biology* **2**, 6, doi:10.1186/1471-2229-2-6 (2002).
13 Hothorn, T., Bretz, F. & Westfall, P. Simultaneous inference in general parametric models. *Biom J* **50**, 346-363, doi:10.1002/bimj.200810425 (2008).
14 Kleiber, C. & Zeileis, A. in *Use R!* 1 online resource (x, 221 pages) (Springer, New York, 2008).
15 Wickham, H. in *Ggplot2 elegant graphics for data analysis.* 1 online resource (viii, 212 p.) (Springer, Dordrecht ; New York, 2009).
16 Wickham, H. Reshaping Data with the reshape Package. *Journal of Statistical Software* **1** (2007).
17 R: A language and environment for statistical programming (2015).
18 Scales: Scale functions for visualisations (2015).
19 Bates, D., Machler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1-48 (2015).
20 Youguang, Y. & Tan, B. C. The non-functional stomata on the leaf margin of *Selaginellla*. *Philippine Journal of Science* **142**, 245-248 (2013).
21 Valdespino, I. A. Novelties in Selaginella (Selaginellaceae - Lycopodiophyta), with emphasis on Brazilian species. *PhytoKeys* **57**, 93-133, doi:10.3897/phytokeys.57.6489 (2015).
22 Banks, J. A., Nishiyama, T. & Hasebe, M. The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332**, 960-963 doi:10.1126/science.1203810 (2011).
23 Davies, K. A. & Bergmann, D. C. Functional specialization of stomatal bHLHs through modification of DNA-binding and phosphoregulation potential. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 15585-15590, doi:10.1073/pnas.1411766111 (2014).
24 Singer, S. D. & Ashton, N. W. Revelation of ancestral roles of KNOX genes by a functional analysis of Physcomitrella homologues. *Plant Cell Rep* **26**, 2039-2054, doi:10.1007/s00299-007-0409-5 (2007).
25 Glime, J. M. in *Bryophyte ecology* Vol. 1 (Michigan Technological University and the International Association of Bryologists, 2013).
26 Markham, K. R., Moore, N. A. & Porter, L. J. Changeover in flavonoid pattern accompanying reproductive structure formation in a bryophyte. *Phytochemistry* **17**, 911-913, doi:http://dx.doi.org/10.1016/S0031-9422(00)88645-8 (1978).
27 Beerling, D. J. & Franks, P. J. Evolution of stomatal function in 'lower' land plants. *New Phytologist* **183**, 921-925, doi:10.1111/j.1469-8137.2009.02973.x (2009).
28 Merced, A. & Renzaglia, K. Developmental changes in guard cell wall structure and pectin composition in the moss Funaria: implications for function and evolution of stomata. *Annals of Botany* **114**, 1001-1010, doi:10.1093/aob/mcu165 (2014).
29 Duckett, J. G., Pressel, S., P'ng, K. M. Y. & Renzaglia, K. S. Exploding a myth: the capsule dehiscence mechanism and the function of pseudostomata in Sphagnum. *New Phytologist* **183**, 1053-1063, doi:10.1111/j.1469-8137.2009.02905.x (2009).
30 Beike, A. K. *et al.* Molecular evidence for convergent evolution and allopolyploid speciation within the Physcomitrium-Physcomitrella species complex. *BMC Evolutionary Biology* **14**, 1-20, doi:10.1186/1471-2148-14-158 (2014).
31 Liu, Y., Budke, J. M. & Goffinet, B. Phylogenetic inference rejects sporophyte based classification of the Funariaceae (Bryophyta): rapid radiation suggests rampant homoplasy in sporophyte evolution. *Mol Phylogenet Evol* **62**, 130-145, doi:10.1016/j.ympev.2011.09.010 (2012).
32 Hedderson, T. A. & Zander, R. H. Vrolijkheidia circumscissa (Pottiaceae), a new moss genus and species from the Succulent Karoo of South Africa. *Journal of Bryology* **30**, 143-146, doi:10.1179/174328208x282157 (2008).

33 Egunyomi, A. On the stomata of some tropical african mosses. *Lindbergia* **8**, 121-124 (1982).
34 Garner, D. L. B. & Paolillo, D. J. A time-course of sporophyte development in *Funaria hygrometrica* Hedw. *Bryologist* **76**, 356-360 (1973).
35 Chater, C. *et al.* Regulatory Mechanism Controlling Stomatal Behavior Conserved across 400 Million Years of Land Plant Evolution. *Current Biology* **21**, 1025-1029, doi:10.1016/j.cub.2011.04.032 (2011).
36 Garner, D. L. B. & Paolillo, D. J. On the functioning of stomates in *Funaria*. *Bryologist* **76**, 423-427, doi:10.2307/3241726 (1973).
37 Budke, J. M., Goffinet, B. & Jones, C. S. The cuticle on the gametophyte calyptra matures before the sporophyte cuticle in the moss *Funaria hygrometrica* Funariaceae. *American Journal of Botany* **99**, 14-22, doi:10.3732/ajb.1100311 (2012).