The ABCG1 transporter is required for terpene biosynthesis in the lumen of *Marchantia polymorpha* oil bodies.

Edith C. F. Forestier, Paola Asprilla Carabali, Ignacy Bonter, Eftychios Frangedakis and Jim Haseloff.

ABSTRACT

Marchantia polymorpha oil bodies (OBs) are specialised cell structures housing a diverse array of C15-terpenes, called sesquiterpenes. These compounds are known for their roles as herbivore repellents, yet the enzymes responsible for the biosynthesis of their precursors (C5 isoprenoid units) remain poorly characterized. Discrepancies remain between enzyme localizations suggested by computational predictors and those observed in earlier experimental studies, complicating our understanding of terpene biosynthesis. We investigated the localization of isoprenoid biosynthetic enzymes using translational and transcriptional reporters, coupled with confocal microscopy. Most enzymes localized as predicted (*e.g.*, cytosol, chloroplast and the endoplasmic reticulum), and OB cells were identified as the primary sites of terpene biosynthesis.

To explore OBs as potential storage sites for terpenes, we attempted to produce exogenous but easily identifiable compounds in *Marchantia*, such as the diterpene taxadiene and the triterpene β-amyrin. Targeting to OB cells resulted in measurable amounts of these compounds, but their yields remained unaffected by the over-expression of key precursor genes, underscoring challenges in redirecting metabolic flux.

We hypothesized that precursor flux in *Marchantia* primarily supports the synthesis of endogenous sesquiterpenes, which are then transported into the OB lumen. To investigate this process, we focused on ABCG1, an ABC transporter previously reported to localize at the OB membrane. CRISPR-mediated disruption of ABCG1 resulted in a dramatic reduction in sesquiterpene levels, supporting its role in metabolite accumulation within OBs.

Interestingly, a translational reporter of *FTPSL2*, a known endogenous sesquiterpene synthase in *Marchantia*, localized within the OB lumen, suggesting that ABCG1 may instead transport the sesquiterpene precursor farnesyl pyrophosphate (FPP). These

35 findings identified the first potential FPP transporter, an important advancement in our
36 understanding of terpene biosynthesis and storage in *Marchantia polymorpha*.
37
38 **INTRODUCTION.**
39
40 In recent years, *Marchantia polymorpha* has emerged as a valuable model plant[1] for
41 studying terpene biosynthesis and accumulation within specialized storage structures
42 known as oil bodies (OBs)[2]. The oil bodies in *Marchantia* are membrane-bound
43 compartments that primarily house terpenes[3] and bisbibenzyls[4], a class of
44 phenylpropanoid derivatives; the mixture confers anti-feedant properties against
45 herbivores[5]. Terpenes are a diverse class of natural products derived from the
46 condensation of isoprene (C_5H_8) units[6], found across many organisms but
47 predominantly in the plant kingdom. Beyond deterring herbivores as observed in
48 *Marchantia*, they serve other vital roles in plants, such as preventing fungal infections
49 or attracting pollinators. Additionally, their broad range of bioactive properties makes
50 them valuable in medicine, agriculture, food and other industrial applications.
51 The sequestration of bisbibenzyls and terpenes into OBs is critical for protecting
52 *Marchantia* from the potential toxicity of these metabolites when released into other
53 tissues. The unique compartmentalization offered by OBs presents a promising
54 avenue for metabolic engineering, as it provides a natural reservoir for accumulating
55 metabolites that might otherwise be harmful to the plant.
56 Recent developmental studies have provided significant insights into the mechanisms
57 governing OB cell fate determination and formation, highlighting key transcription
58 factors (TFs) involved in these processes. Specific TFs, such as ERF13[7], C1HDZ[5],
59 TGA[8] and MYB2[9,10] have been identified as crucial regulators of OB cell differentiation
60 and maturation[11]. These discoveries have expanded our understanding of how OB
61 cells develop and contribute to the plant's metabolic capacity.
62 Manipulating these TFs has further revealed the delicate balance between OB cell
63 proliferation and plant fitness. For instance, CRISPR-mediated mutation of *TGA*[8] or
64 gain-of-function mutants of *ERF13* led to a dramatic increase in OB cell numbers per
65 plant[7]. However, these alterations were accompanied by reduced growth rates or
66 morphological defects, suggesting that OB cell number is tightly regulated in
67 *Marchantia* to avoid compromising overall plant health and fitness.

Early studies on *Marchantia's* terpene biosynthetic pathway, which were based on immunolocalization of isoprenoid enzymes, hypothesized that key enzymatic steps might occur at the oil body (OB) membrane or within the OB lumen[3]. Today, with the availability of whole-genome sequencing and detailed gene annotations informed by sequence homology with other species, our understanding of the early steps of isoprenoid biosynthesis has significantly improved[12,13] (Table 1 and S1). The mevalonate (MVA) and methylerythritol phosphate (MEP)[14–16] pathways are the two primary metabolic routes responsible for producing the precursors — isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) — to various classes of terpenes (Table 1). The MVA pathway, comprising seven enzymatic steps[17] and typically operating in the cytosol[18], endoplasmic reticulum[19] (ER) and peroxysome[20], generally supplies precursors for sesquiterpenes and triterpenes[21,22], while the MEP pathway, with eight enzymatic steps and primarily localized in plastids[23], commonly provides precursors for monoterpenes and diterpenes[21,22] (Table 1 and S1). In *Marchantia*, sesquiterpenes are the dominant terpene class found in OBs[2], produced by several microbial- and fungal-type terpene synthases[24]. Additionally, *Marchantia* produces the monoterpene limonene, likely synthesized through the cis-prenyl precursor neryl pyrophosphate (NPP) rather than the trans form geranyl pyrophosphate (GPP), as suggested by Kumar and co-workers[25]. However, it remains unclear whether limonene is specifically stored within OBs.

Examination of genome data[12,13] along with computational predictive tools for enzyme localization such as DeepLoc 2.1[26], enabled us to identify putative precursor genes involved in prenyl phosphate synthesis for terpene production (Table 1). However, discrepancies remain between earlier theories suggesting that terpene biosynthetic enzymes are localized to OB membranes and more recent predictions that place MVA enzymes in the cytosol, ER and peroxisome, while MEP enzymes in chloroplasts (Table 1). These discrepancies prompted us to re-examine the localization of key enzymes using modern techniques such as confocal microscopy combined with translational reporters[7]. This approach allowed us to identify the specific cells expressing these genes and determine the subcellular localization of their corresponding proteins.

Building on our findings regarding terpene precursor gene expression and localization, we then explored the potential for metabolic engineering in *Marchantia polymorpha*'s OBs. Specifically, we aimed to produce two valuable terpenes within OBs: the

102 diterpene taxadiene[27], a precursor to the valuable anti-cancer compound Taxol®[28,29],
103 and the triterpene β-amyrin[30], a precursor to compounds such as glycyrrhizin[31], a
104 sweet-tasting constituent of liquorice. These compounds were selected due to their
105 distinct chemical signatures, which facilitate easy detection in chromatographic
106 analyses given the minimal background of these terpene sub-classes in *Marchantia*.
107 To further evaluate the metabolic capacity of *Marchantia* and investigate the transport
108 of sesquiterpenes, we also examined the production of the sesquiterpene amorpha-
109 4,11-diene[32] alongside ABCG1, which we initially hypothesized to transport
110 sesquiterpenes into OBs. Understanding the precise function of this ABC transporter
111 is particularly significant, as prior studies have demonstrated its high and exclusive
112 expression in OB cells[33] and its clear localization to the OB membrane using a
113 translational reporter[7]. This study sheds light on the role of ABCG1 in the terpene
114 accumulation in OBs, addressing a critical gap in our understanding of terpene
115 metabolism in *Marchantia polymorpha*.
116
117 **MAIN**
118
119 **1. Translational and transcriptional reporters of terpenoids pathway genes and**
120 **promoters.**
121 To determine the subcellular localization of key enzymes involved in *Marchantia*'s
122 terpene biosynthesis, we generated constructs of translational reporters, in which
123 each gene's native promoter drove its respective coding sequence fused to the
124 fluorescent tag *mVenus*[34]. In cases where translational reporters were unavailable due
125 to technical challenges, transcriptional reporters were used instead to determine cell-
126 type expression. These consisted of *Marchantia*'s native promoter driving mVenus with
127 a nuclear localisation signal[35] (mVenus-N7), enabling us to investigate the spatial and
128 temporal expression of key genes. Promoter lengths were selected based on Assay
129 for Transposase-Accessible Chromatin with Sequencing (ATAC-seq)[36] data from
130 version 6.1 of the Tak accession in the *Marchantia.info* database[13,24] (Table S2). ATAC-
131 seq identifies regions of open chromatin, which are more accessible to transcription
132 factors and associated with active gene regulation. This approach allowed us to select
133 promoter regions likely to drive gene expression in their native cellular contexts. To
134 delineate cell boundaries, each construct also included a fluorescent membrane

135 marker, *mScarlet*[37], driven by the strong, constitutive *ubiquitin-conjugating enzyme E2*
136 gene promoter and 5'UTR from *Marchantia polymorpha* ($_{Prom5}$*UBE2*)[38], fused to the
137 *Arabidopsis* LTI6b[39] membrane protein.
138 For the identification of different subcellular compartments, such as
139 chloroplasts/plastids, ER, cytosol, Golgi, and mitochondria, we interpreted our findings
140 by comparing the localization patterns observed in our images with described
141 reporters in *Marchantia*[38,40], though distinguishing between ER and cytosol remains
142 challenging under the resolution constraints of our experimental setup.
143 For this study, we selected key genes from the MVA and MEP pathways for their roles
144 as rate-limiting enzymes, their redundancy, or their involvement in the final steps of
145 the pathway. Additionally, we included genes responsible for the formation of linear
146 scaffold precursors for sesquiterpenes, diterpenes, and triterpenes to capture a more
147 comprehensive view of terpene biosynthesis.
148 We first analysed the two annotated *1-deoxy-D-xylulose-5-phosphate synthases*
149 (*MpDXS1 and MpDXS2*) (Table 1), which encode the first committed and rate-limiting
150 enzymes of the MEP pathway[41]. For *MpDXS1*, tagged with mVenus and driven by its
151 native promoter and 5'UTR ($_{Prom5}$*MpDXS1:MpDXS1-mVenus*), mVenus fluorescence
152 was observed in all cell types throughout the plant, with a stronger signal in OB cells
153 (Figure 1A). The protein localized to plastids, displaying a punctate pattern in non-OB
154 cells rather than a homogeneous distribution (Figure 1A). By contrast,
155 $_{Prom5}$*DXS2:DXS2-mVenus* showed a restricted localization pattern, with fluorescence
156 observed in oil body plastids only from Day 0 gemmae to 14-days-old plants (Figure
157 1B and S1D).
158 We examined the final step of the MEP pathway by generating a translational reporter
159 for *(E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase* (*MpHDR*)[42,43] (Table
160 1). We also generated a translational reporter for *(E)-4-hydroxy-3-methylbut-2-enyl*
161 *diphosphate synthase* (*MpHDS*)[42,44] (Table 1). *MpHDS* is adjacent to *MpHDR* and
162 oriented in the opposite direction on the same chromosome, suggesting they share a
163 bidirectional promoter (Figure S2A). Previous studies have highlighted the potential
164 toxic nature of the product of HDS[45,46], which is detoxified in the subsequent enzymatic
165 step by HDR[45,46], making it particularly interesting to study both enzymes together. To
166 our knowledge, this is the first report suggesting that *HDS* and *HDR* genes may be
167 regulated by a shared promoter. Initial experiments revealed that the 2 kb region
168 upstream of *MpHDR* was insufficient to drive expression, whereas the 1.9 kb region

169 upstream of *MpHDS* was adequate (Figures S2B). Expression of a construct driving
170 *MpHDS-eGFP* and *MpHDR-mVenus* bidirectionally (Figure S2C) demonstrated that,
171 like other MEP enzymes, both proteins localized to plastids (Figure 1C and 1D). Their
172 expression was stronger, if not exclusive, to OB cells in Day 0 gemmae (Figure 1C
173 and 1D), before becoming more widespread across tissues at later developmental
174 stages (Figure S1A and S1B). Similarly, GERANYLGERANYL PYROPHOSPHATE
175 SYNTHASE (MpGGPPS) (Table 1), enzyme synthesizing geranylgeranyl
176 pyrophosphate (GGPP), the scaffold for diterpenes[47,48], localized to plastids in every
177 cell types, but with a more intense signal in OB cells (Figure 1E).
178 For the MVA pathway and sesquiterpene synthesis, we examined the rate-limiting
179 enzyme 3-HYDROXY-3-METHYLGLUTARYL-COA REDUCTASE (MpHMGR)[49]
180 (Table 1). Expression of $_{Prom5}$*MpHMGR:MpHMGR-mVenus* revealed a restricted
181 localisation to oil body cells in gemmae (Figure 1G) and 14-day-old plants (Figure
182 S1D). The enzyme appeared to display a dual localization pattern: a punctate
183 distribution around the OB, indicative of Golgi, and a web-like pattern surrounding the
184 nucleus consistent with ER (Figure 1G). As a key enzyme downstream the MVA
185 pathway, we analysed farnesyl pyrophosphate synthase (MpFPS) (Table 1), which
186 forms farnesyl pyrophosphate (FPP)[50], the linear precursor of all sesquiterpenes.
187 MpFPS also exhibited specific expression to OB cells, even after 14 days (Figure 1F
188 and S1E). Contrary to its predicted mitochondrial localisation by DeepLoc 2.1[26] (Table
189 1), the enzyme appeared predominantly cytosolic, as indicated by its uniform signal
190 surrounding the OB and delineating the plastids (Figure 1F).
191 We further investigated the expression pattern[51] of *1-deoxy-D-xylulose 5-phosphate*
192 *reductoisomerases*[52,53] (*MpDXR1* and *MpDXR2*) (Table 1) using transcriptional
193 reporters. $_{Prom5}$*MpDXR1:mVenus-N7* showed ubiquitous expression (Figure 2A),
194 whereas $_{Prom5}$*DXR2:mVenus-N7* showed specificity to OB cells in Day 0 gemma
195 (Figure 2B). Similarly, the expression pattern of the promoter for MEVALONATE
196 DIPHOSPHATE DECARBOXYLASE (MpMVD)[54] (Table 1), the enzyme responsible
197 for the last committed step of the MVA pathway, also exhibited specificity to OB cells,
198 even in 14-day-old thalli (Figure 2C).
199 Finally, given the apparent exclusivity of some MVA biosynthetic steps and MpFPS to
200 OB cells — key intermediates in phytosterol biosynthesis, essential structural
201 components of plant membranes[55] — we examined the transcriptional reporter of the
202 *squalene synthase*[56] (*MpSQS3*) promoter. The enzyme fuses two FPP molecules to

203 form squalene, the precursor of triterpenes and phytosterols. Among the three
204 squalene synthase-like genes annotated in the *Marchantia* genome (*MpSQS1-3*)
205 (Table 1 and S1), we selected *MpSQS3* based on its higher protein sequence
206 homology (60%) with *Arabidopsis thaliana* SQUALENE SYNTHASE. In contrast,
207 MpSQS1 showed greater similarity (73%) to *Arabidopsis* PHYTOENE SYNTHASE,
208 suggesting a role in carotenoid biosynthesis, while MpSQS2 shared 52% homology
209 with an uncharacterized terpenoid synthase. Plants expressing the
210 $_{Prom5}$*SQS3:mVenus-N7* constructs show a mVenus signal in meristematic regions in
211 Day 0 gemmae, followed by ubiquitous expression in later developmental stages
212 (Figure 2D). These findings indicate that sterol biosynthesis is not confined to OB cells
213 but occurs broadly throughout the plant.
214 Overall, our findings provide a comprehensive map of the subcellular localization and
215 expression patterns of key terpene biosynthetic enzymes, offering insights into the
216 spatial organization of biosynthetic intermediates and terpene scaffolds in *Marchantia*
217 *polymorpha*.
218
219 **2. Exogenous production of taxadiene and β-amyrin in *Marchantia polymorpha***
220 **whole plants *vs*. oil body cells.**
221 Given the predominant observed expression of terpene precursor enzymes in OB
222 cells, we aimed to harness the metabolic potential of these specialised structures. As
223 proof of concept, we attempted to produce the diterpene taxadiene[27], and the
224 triterpene β-amyrin[30].
225 We generated constructs to express *Taxus baccata taxadiene synthase* (*TXS*)[57] or
226 *Talinum paniculatum β-amyrin synthase*[58] (*β-AS*), driven by either the constitutive
227 $_{Prom5}$*UBE2* or the oil body-specific Mp*R2R3-MYB2* promoter ($_{Prom5}$*MYB2*)[35], enabling
228 expression throughout the entire plant or specifically within OB cells, respectively. TXS
229 catalyzes the cyclization of the linear diterpene precursor GGPP into taxadiene, while
230 β-AS converts the linear triterpene precursor 2,3-oxidosqualene into β-amyrin.
231 In plants expressing $_{Prom5}$*UBE2:TXS*, we detected a prominent new peak in hexane
232 extracts (Figure 3A), corresponding to taxadiene based on its mass spectrum (MS),
233 as described in prior publications[59,60]. While no commercial taxadiene standards were
234 available, we identified the compound by its characteristic ion at m/z 122 and the
235 molecular ion [M+] at m/z 272 (Figure S3A). Quantification of taxadiene in four
236 independent transformants expressing either $_{Prom5}$*UBE2:TXS* or $_{Prom5}$*MYB2:TXS*

237 revealed significantly higher levels of taxadiene in whole plants compared to those
238 with expression restricted to OB cells (Figure 3D). Specifically, $_{Prom5}$*UBE2:TXS* lines
239 produced approximately 21 µg of taxadiene per g of fresh weight (FW), whereas
240 $_{Prom5}$*MYB2:TXS* lines produced only 0.7 µg/g FW, dodecane equivalent. The correct
241 targeting of TXS protein in OB cells was confirmed using a $_{Prom5}$*MYB2:TXS-mVenus*
242 fusion construct, which showed plastid localization (Figure 3C).
243 Similarly, plants expressing $_{Prom5}$*UBE2:β-AS* produced a new compound, identified as
244 β-amyrin through its MS profile and comparison with an authentic standard (Figure 3B
245 and S3B). The $_{Prom5}$*UBE2:β-AS* lines produced an average of 56 µg of β-amyrin per g
246 of fresh weight (coprostanol equivalent), whereas lines expressing the enzyme
247 specifically in OB cells yielded an average of 1.4 µg/g FW (Figure 3D). In MYB2-driven
248 *β-AS-mVenus* lines, the protein localized specifically in the OB cells and the cytosol
249 (Figure 3C).
250 To investigate whether precursor flux limitations can constrain the production of
251 taxadiene and β-amyrin in oil bodies, we co-expressed *TXS* and *β-AS* with key rate-
252 limiting enzymes of the MEP and MVA pathways (*i.e DXS* and *HMGR*, respectively),
253 as our work demonstrated that many of these enzymes appeared to be highly or
254 specifically expressed to OB cells.
255 For taxadiene production via the MEP pathway, we co-expressed *MpDXS1* or
256 *MpDXS2* with *TXS*. Additionally, to mitigate potential homology-dependent gene
257 silencing associated with overexpressing native *Marchantia* genes[61], we expressed
258 the homologous *PpDXS1* and *PpDXS2* genes from *Physcomitrium patens,* which have
259 recently been re-annotated as *PpDXS1A* and *PpDXS1D*[62]. To further enhance
260 precursor flux, we introduced *MpHDR* and *MpGGPPS*, two key enzymes previously
261 shown to increase GGPP flux for diterpene production in *Nicotiana benthamiana*[63,64].
262 For β-amyrin production through the MVA pathway[65], we overexpressed the *HMGR*
263 genes from *Marchantia* (*MpHMGR*) and *Physcomitrium* (*PpHMGR*), alongside
264 truncated versions designed to remove their negative feedback regulatory domain[66,67],
265 thereby increasing the pool of terpene precursors and allowing unrestrained
266 production of β-amyrin. These truncated variants (*Mp-tHMGR* and *Pp-tHMGR*) lack
267 the first 148 and 146 amino acids, respectively. All genes were expressed under the
268 MYB2 promoter to specifically target OB cells.
269 Quantitative analysis of four independent transformants showed no significant
270 increase in taxadiene or β-amyrin levels upon overexpression of these precursor

genes (Figure 3E). We confirmed the correct targeting to OB cells and subcellular localization of *Marchantia DXS*, *HDR*, *GGPPS*, *HMGR*, and *tHMGR* by fusing these genes with mVenus under the control of the MYB2 promoter (Figure S4). DXSs, HDR, and GGPPS enzymes specifically localized to plastids (Figure S4A, S4B, S4C and S4D). However, despite co-localization with TXS, this did not result in increased taxadiene levels. Over-expression of the precursor genes in the OB cells didn't increase either the level of main endogenous sesquiterpenes (Dataset 1 and Dataset 2), which were tentatively identified using Kovats retention indices[68] (Figure S5 and Table S3). While tHMGR-mVenus appeared to localize in the ER and Golgi apparatus of OB cells (Figure S4E), the native version driven by the MYB2 promoter remained undetectable. To address this, we re-cloned *MpHMGR* under the 2x35S promoter and successfully detected the protein, which accumulated exclusively to the ER of OB cells (Figure S4F). Interestingly, when tHMGR was expressed under the 2x35S promoter, the fusion protein showed expression across all cells (Figure S4G), suggesting that this regulatory domain may play a role in restricting localization to OB cells.

Overall, these results demonstrate that overexpression of selected rate-limiting enzymes was insufficient to significantly enhance terpene yields in *Marchantia* OB cells, suggesting that additional regulatory mechanisms may play critical roles in terpene biosynthesis and accumulation in these specialized cells.

3. Accumulation of terpenes in oil bodies may require a specific transporter.

Despite the high specificity and/or strong expression of terpene biosynthetic pathway genes in OB cells, the over-expression of key enzymes did not lead to increased levels of exogenous terpenes. Moreover, these compounds may not naturally accumulate in OBs: taxadiene may primarily localize to plastids, while β-amyrin could accumulate in the cytosol. We hypothesized that precursor flux in *Marchantia* predominantly supplies sesquiterpenes biosynthesis and that a specific transporter may be required for their accumulation inside OBs. Previous studies demonstrated high and exclusive expression of the *ABCG1* gene in oil body cells[33], with its protein localized to the OB membrane when expressed under its native promoter, and to the plasma membrane of other cells when driven by a non-OB-specific promoter[7].

To investigate whether ABCG1 could facilitate the accumulation of any sesquiterpenes in *Marchantia*'s OBs, we expressed *Artemisia annua amorphadiene synthase* (*AMS*), a sesquiterpene synthase that converts FPP into amorpha-4,11-diene[32]. Cyclisation of

305 FPP into this sesquiterpene likely occurs via the bisabolyl cation[69], similar to the
306 endogenous β-chamigrene[70], therefore we expected this structural similarity to enable
307 amorpha-4,11-diene transport into OBs. We targeted *AMS* expression to either the
308 whole plant ($_{Prom5}$*UBE2*) or specifically in OB cells ($_{Prom5}$*MYB2*). Additionally, we co-
309 expressed *AMS* and *ABCG1* in OB cells, alongside *HMGR*, *DXS2*, and *FPS* to
310 increase precursor availability and channel metabolic flux specifically toward
311 sesquiterpenes, with FPS recognized as a rate-limiting enzyme in their biosynthesis[71].
312 Amorpha-4,11-diene was detected only in transgenic lines expressing the construct
313 $_{Prom5}$*UBE2:AMS* (Figure S6), with an average yield of 0.75 µg/g FW in six out of eight
314 independent transformants (Figure 4A, Dataset 3). No amorpha-4,11-diene was
315 detected when expression was targeted to OB cells, either with or without precursor
316 genes and *ABCG1*, possibly due to insufficient yield and/or co-elution with the β-
317 chamigrene peak (Figure S6). Localization studies of *AMS-mVenus* and *MpFPS-*
318 *mVenus* driven by the MYB2 promoter showed that both proteins localized to the
319 cytosol of OB cells (Figure 4B).
320 We measured endogenous sesquiterpene levels in eight independent transformants
321 for each gene combination. Lines expressing $_{Prom5}$*MYB2:AMS* alone produced
322 relatively low yields, but co-expression of precursor genes significantly increased
323 sesquiterpene levels (Figure 4C, Dataset 3). Among combinations without *ABCG1*,
324 *FPS* + *HMGR* boosted sesquiterpene levels the most, with a 2.8-fold average increase
325 for the major sesquiterpenes compared to a 1.8-fold increase with *FPS* and *DXS2*.
326 The over-expression of *ABCG1* alongside all precursor genes led to the most
327 significant increases in endogenous sesquiterpene levels, with a 3-fold increase in cis-
328 thujopsene, a 3.2-fold increase in β-chamigrene and γ-cuprene, and a 3.6-fold
329 increase in thujopsan-2α-ol (Figure 4C, Dataset 3). The observed increase in
330 endogenous sesquiterpene levels when *HMGR* was expressed under the MYB2
331 promoter also suggests that HMGR protein is likely localized to OB cells, despite its
332 apparent absence in localization studies using the mVenus tag.
333 To further validate the transport role of ABCG1, we generated CRISPR lines to disrupt
334 its function. We obtained six independent transformants with early stop codons in the
335 *ABCG1* DNA sequence (Figure 5A). Terpene analysis revealed a dramatic reduction
336 or complete absence of sesquiterpenes in *abcg1* mutants compared to Cas9-only
337 controls (Figures 5B and 5C). Notably, this reduction was specific to sesquiterpenes,

338 as the levels of phytosterols remained unaffected (Figure 5B and 5C). Given the
339 apparent differences in peak abundance observed in the overlaid chromatograms
340 (Figure 5B), we quantified additional compounds alongside neophytadiene and phytol,
341 such as (*Z*)-1,3-phytadiene and (*E*)-1,3-phytadiene, phytol derivatives identified in
342 prior studies[72,73] (Figure 5B). Quantification revealed no significant changes in these
343 compounds when comparing the five Cas9 controls and six *abcg1* mutant lines (Figure
344 S7). These findings suggest that ABCG1 plays a critical role in the accumulation of
345 endogenous sesquiterpenes while having no effect on other terpene sub-classes or
346 on the accumulation of exogenous sesquiterpenes.
347

348 **4. Translational reporter analysis of *Marchantia* endogenous sesquiterpene**
349 **synthase suggests transport of FPP into OBs.**

350 As part of the broader screening of translational reporters described above, we initially
351 investigated one fungal-type terpene synthase-like protein (MpFTPSLs) from a group
352 of seven homologs located on the same chromosome and believed to produce the
353 majority of endogenous sesquiterpenes in *Marchantia*[25,74]. Among these, we studied
354 *MpFTPSL2* (*Mp6g04320*), as this sesquiterpene synthase was previously
355 characterized by Kumar and co-workers[25].
356 Interestingly, the mVenus signal began to appear in 14-day-old plants expressing
357 $_{Prom5}$*FTPSL2:FTPSL2-mVenus*, with no detectable signals observed at earlier
358 developmental stages, such as gemmae or young gemmalings. The signal was
359 localized to OB cells in only one of the four thallus lobes in two independent lines, and
360 specifically within the OB itself (Figure 6A). This raised the possibility that FTPSL2
361 enzymatic activity depends on FPP transport into the OB. These serendipitous
362 observations were consistent across multiple OB cells in both independent lines,
363 supporting the specific targeting of FTPSL2 and its potential functional role.
364 Intriguingly, DeepLoc 2.1[26] predicted the enzyme to be cytosolic but to also contain a
365 nuclear localisation signal, suggesting that FTPSL2 may localize to a compartment
366 not accounted for by this computational tool.
367 To further investigate the contributions of the promoter and CDS in governing the
368 observed localization, and to assess whether the *FTPSL2* promoter could offer a more
369 effective strategy for metabolic experiments, we generated lines expressing the
370 *FTPSL2* CDS under the OB-specific $_{Prom5}$*MYB2* promoter and lines expressing the
371 *AMS* CDS with $_{Prom5}$*FTPSL2*. Interestingly, $_{Prom5}$*MYB2:FTPSL2* directed the protein to

372 the cytosol of OB cells, rather than the OB lumen (Figure 6B), while AMS-mVenus
373 expressed under the *FTPSL2* promoter didn't provide any detectable signal from
374 gemma to 14-day-old plants (Figure 6C). These results suggest that both the promoter
375 and CDS play critical roles in determining the localization of terpene synthases,
376 supporting the validity of our initial strategy while offering insights for further
377 optimization.
378 Finally, considering the strong impact of *ABCG1* disruption on sesquiterpene
379 accumulation and the fact that FPP is a substrate for OB lumen-localized
380 sesquiterpene synthases, these findings suggest that ABCG1 may function as an FPP
381 transporter, facilitating its entry into OBs.
382
383 **DISCUSSION**
384
385 **Translational and transcriptional reporters studies to re-assess sub-localization**
386 **of isoprenoid biosynthetic enzymes and cell type expression specificity.**
387 For each translational or expression pattern reporter, we selected promoter lengths
388 based on a single ATAC-seq[36] peak, ensuring coverage of the putative core promoter
389 and upstream regulatory regions. In some cases, promoters (including the 5'UTR)
390 were shorter than the typically advised 2.5 kb length[35,38], balancing the need for
391 regulatory coverage with constraints related to sequence synthesis. For instance, a
392 promoter as short as 970 bp like $_{Prom5}$*GGPPS* was sufficient to drive strong and
393 ubiquitous signal in chloroplasts/plastids. However, in other instances such as
394 $_{Prom5}$*DXR2*, promoter functionality required extending the length from 1.7 kb to 2.7 kb
395 to achieve detectable expression.
396 Interestingly, $_{Prom5}$*FTPSL2* required the inclusion of three ATAC-seq peaks to become
397 functional, although the precise conditions influencing the activity of the 2.2 kb
398 promoter warrant further investigation. Notably, another study using a 4.1 kb
399 $_{Prom5}$*FTPSL2* was able to detect a dTomato signal targeted to the nucleus and showed
400 specificity to the OB cells[74]. While it is tempting to hypothesize that altering promoter
401 lengths could include or exclude positive or negative regulatory elements affecting
402 cell-type specificity, further systematic studies are needed to draw concrete
403 conclusions about the relationship between promoter structure and regulatory
404 outcomes.

405 Despite potential variability in cell-type detection due to differences in promoter length,
406 our findings robustly confirmed the subcellular localization of enzymes involved in the
407 early steps of terpene biosynthesis. Enzymatic steps of the MEP pathway and
408 diterpene biosynthesis, including DXS, HDS, HDR, and GGPPS, were localized to
409 plastids/chloroplasts, consistent with their predicted roles in producing isoprenoid
410 precursors for plastid-derived terpenes. In contrast, steps of the MVA pathway and
411 sesquiterpene biosynthesis, represented by HMGR and FPS, were localized to the
412 ER, Golgi, and cytosol, supporting their involvement in cytosolic terpene precursor
413 production. With the exception of FPS, all studied enzymes localized as predicted by
414 DeepLoc2.1[26]. However, our experimental conditions allowed us to localise steady-
415 state accumulation of proteins and did not account for any dynamic behavior of protein
416 trafficking[75]. Our findings corroborate earlier studies[3] that reported high or exclusive
417 expression of isoprenoid biosynthetic enzymes in OB cells. Given that OBs are formed
418 through redirection of the secretory pathway[7], some enzymes like HMGR – which
419 localizes to both the ER and Golgi depending on the OB cell or plant being imaged –
420 may transit to other compartments during OB development, supporting the possibility
421 of their transient presence around the OB membrane[3].
422 The subcellular localization of FTPSL2 remains an intriguing aspect of our findings.
423 Our results demonstrate that promoter choice influences its localization, with
424 $_{Prom5}$*FTPSL2* directing the protein inside the OBs, whereas $_{Prom5}$*MYB2* results in
425 cytosolic localization within the same cell type. A similar phenomenon occured with
426 the ABCG1 protein, which localized to the OB membrane when regulated by its native
427 promoter, but redirected to the plasma membrane under the control of a non-OB-
428 specific promoter[7]. In this case, the change in cell type and the absence of an OB
429 membrane in non-OB cells may have led ABCG1 to associate with the default
430 membrane compartment. Given that our mScarlet membrane marker delineates both
431 the OB and plasma membranes, these structures may share common biophysical or
432 molecular characteristics. A comparable promoter-dependent influence on localization
433 has been observed in yeast, where researchers demonstrated that promoter
434 sequences encode information affecting mRNA localization and translation efficiency,
435 potentially through co-transcriptional recruitment of RNA-binding proteins,
436 transcription factors, or cis-regulatory elements[76]. While our study did not directly
437 investigate mRNA localization, our findings provide further evidence that promoter
438 sequences play a broader role in determining protein localization within a same cell,

439 highlighting their previously underappreciated regulatory influence beyond
440 transcription.
441 Nevertheless, our results confirm the specialized role of OB cells in terpene
442 biosynthesis and highlight the importance of carefully selecting promoters to ensure
443 accurate expression and localization of biosynthetic enzymes.
444
445 **Pathway specificity in terpene synthesis: MVA and MEP contributions.**
446 Our study suggests that the MVA pathway predominantly provides precursors for
447 sesquiterpene synthesis in *Marchantia polymorpha*, supported by the apparent
448 exclusive expression of HMGR and MVD in OB cells and the significant increases in
449 endogenous sesquiterpene levels observed when *HMGR* and *FPS* were co-
450 expressed. Notably, FPS – also a potential OB-specific enzyme – emerged as a critical
451 limiting step in sesquiterpene biosynthesis, consistent with previous findings in
452 *Nicotiana tabacum*[77].
453 The MEP pathway's contribution to sesquiterpene synthesis cannot be ruled out, given
454 the specific expression of *DXS2* and *DXR2* in OB cells and the fact that a previous
455 study using GC-MS analysis of physically extracted OB content detected
456 sesquiterpenes[2] as the only sub-class of terpenes. If OBs also store monoterpenes
457 like limonene[25] or specific diterpenes, their levels remained undetectable or
458 unchanged despite the targeted expression of precursor supply genes. Furthermore,
459 over-expression of key MEP genes and *GGPPS* had no effect on taxadiene yield,
460 despite co-localization of these enzymes with TXS in plastids. This suggests that a
461 significant portion of the precursors is redirected toward sesquiterpene synthesis
462 instead. Supporting this, co-expression of *DXS2* with *FPS*, *HMGR*, and *ABCG1* mildly
463 increased the sesquiterpene levels compared to the same combination of genes
464 without *DXS2*, demonstrating a modest effect of *DXS2* on sesquiterpene yield. Dual
465 contribution from the MVA and MEP pathways in *Marchantia polymorpha* would
466 parallel findings in *Artemisia annua*, where inhibitor assays demonstrated a mixed
467 origin of precursors[78].
468 In contrast, *DXS1* and *DXR1* are ubiquitously expressed and likely play broader roles
469 in synthesizing essential metabolites, such as phytyl diphosphate for chlorophyll
470 production. Supporting this, ChIP-seq studies in *Marchantia* have shown that the
471 chloroplast biogenesis regulator GLK binds to the promoter of *DXR1*[79], further

472 highlighting its role in primary metabolism rather than specialized metabolite
473 production.
474 The specific localisation of HMGR, MVD, and FPS protein fusions in OB cells raises
475 intriguing questions about how essential metabolites, such as sterols, are produced in
476 non-OB cells. Despite the ubiquitous expression of squalene synthase *SQS3*, which
477 catalyzes the formation of squalene as a precursor to sterols, the source of precursors
478 for sterol synthesis in non-OB tissues remains unclear. Notably, sterol levels in *abcg1*
479 mutants were unaffected, suggesting that OBs are not the primary site of sterol
480 synthesis or storage for the rest of the plant. This raises the possibility that other, as-
481 yet-unannotated MVA enzymes may supply precursors for sterol biosynthesis in non-
482 OB cells. Alternatively, it is plausible that the genes studied here are expressed
483 ubiquitously but accumulate at levels too low to be detected in other cell types with
484 our confocal microscopy studies.
485 These findings highlight the complexity of precursor flux and compartmentalization in
486 *Marchantia*, accentuating the need to further investigate how primary metabolites
487 synthesis is maintained across tissues.
488
489 **ABCG1 may be a FPP transporter, the first of its kind.**
490 Our study demonstrated that exogenous terpenes, such as amorpha-4,11-diene,
491 taxadiene, and β-amyrin, could be produced throughout the whole plant. However,
492 engineering their production within OB cells was more challenging. We initially
493 hypothesized that ABCG1 transporter could function as a broad-spectrum terpene
494 transporter within OBs, a theory supported by the sharp reduction in endogenous
495 sesquiterpene levels observed in *abcg1* mutants. However, the detection of an
496 endogenous sesquiterpene synthase (FTPSL2) within the OB lumen suggests a more
497 specific role for ABCG1: the transport of FPP into OBs. Combined with insights from
498 our metabolic engineering experiments, this model implies that precursor flux from
499 both MVA and MEP pathways is predominantly directed toward FPP production, which
500 is then transported into OBs for processing by endogenous sesquiterpene synthases.
501 This would explain why co-expression of different *HMGR* versions with *β-AS* did not
502 increase β-amyrin yield: little FPP remains in the cytosol, limiting its availability for β-
503 amyrin biosynthesis.
504 The localization of FTPSL2 inside OBs provides the strongest evidence that FPP must
505 be transported into this compartment. However, the conditions under which FTPSL2-

506 mVenus was detected inside OBs remain unclear, as expression was not observed
507 throughout the plant. One possibility is that this striking observation occurred in
508 response to unidentified abiotic factors, during which OB cells in a particular region of
509 the plant entered an “oil body phase”, redirecting the secretory pathway toward OBs,
510 as hypothesized by Kanazawa et al[7]. Additional experiments identifying the signals
511 that drive terpene synthase localization within OBs would provide a more complete
512 understanding of terpene biosynthesis in *Marchantia*.
513 Considering this precursor transport, successful taxadiene accumulation would
514 require GGPPS to be targeted outside the plastids, a transporter to facilitate GGPP
515 entry into OBs, and the engineering of TXS to localize within OBs. Similarly, producing
516 amorpha-4,11-diene or β-amyrin in OBs would necessitate targeting their respective
517 biosynthetic enzymes – and SQS3 in the case of β-amyrin – to this compartment. It is
518 likely that the endogenous sesquiterpene synthases in *Marchantia* contain specific
519 amino acid motifs that direct them to OBs – features absent in exogenous enzymes.
520 This highlights the need for advanced synthetic biology approaches to engineer
521 enzymes with the necessary targeting signals for OB localization, particularly if the
522 target compounds prove deleterious when produced in all tissues.
523 We propose a preliminary map of terpene biosynthesis in *Marchantia* OB cells (Figure
524 7). This map includes key steps in primary metabolism, emphasizing the role of certain
525 terpene precursors in vital plant functions, such as FPP for sterols and GGPP for
526 chlorophyll and gibberellin derivative GA_{12}[80], although their specific presence in OB
527 cells has yet to be confirmed (Figure 7).
528 Overall, our work provides the first evidence for a prenyl diphosphate transporter and
529 confirms that the OB lumen is a catalytic compartment for the final stages of terpene
530 synthesis, rather than a storage site for the products. This study lays a foundation for
531 uncovering similar mechanisms in other plant species, potentially transforming our
532 understanding of metabolite compartmentalization and enabling new approaches in
533 plant synthetic biology. Direct biochemical validation of ABCG1-mediated FPP
534 transport will be an important next step. Future work investigating ABCG1-FPP
535 interactions will be critical to defining its substrate specificity and transport mechanism.
536 Moreover, given the apparent plastid-to-cytosol flux of isoprenoid precursors in
537 *Marchantia* OB cells, this discovery may pave the way for identifying transporters
538 responsible for the movement of IPP, DMAPP, or other intermediates across the plastid
539 membrane.

540

541 **MATERIAL AND METHODS.**

542

543 **Plant material and growth conditions.**

544 *Marchantia polymorpha* subspecies *ruderalis* accessions Cam-1 (male) and Cam-2
545 (female) were used in this study[81]. Plants were cultivated on solid 0.5× Gamborg B-5
546 basal medium (#G398, PhytoTech Labs, Lenexa, Kansas, USA) adjusted to pH 5.7–
547 5.8 and solidified with 1.2% (w/v) agar micropropagation grade (#A296, PhytoTech
548 Labs). They were maintained under continuous light at 22°C with a light intensity of
549 150 µmol/m²/s. For general propagation of the lines, plants were grown in 94 × 16 mm
550 Petri dishes (#633181, Greiner Bio-One, Kremsmünster, Austria). For imaging
551 purposes, plants were cultivated on gridded 65 × 14.5 mm Gosselin™ Petri dishes
552 (#BB64-01, Corning, Corning, NY, USA). Several gemmae, originating from four
553 independent transformants, were grown on these plates and imaged at developmental
554 stages corresponding to Day 3, Day 5, Day 7, and Day 14 after germination.
555 For gemma cup production and the collection of material for terpene analysis, the
556 media were supplemented with 0.5% (w/v) sucrose. Plants destined for terpene
557 extraction and quantification were grown in 100 × 25 mm Petri dishes (#D943,
558 Phytotech Labs) containing 50 mL of medium.
559 Spore production was carried out in Microbox micropropagation containers (SacO2)
560 under long-day conditions (16 h light/8 h dark) with light supplemented by far-red
561 illumination, following previously established protocol[38].

562

563 **Plasmid construction for over-expression studies and CRISPR-based mutant**
564 **generation.**

565 DNA fragments used to generate mVenus and metabolic constructs were synthesized
566 by Genewiz (Azenta, Burlington, Massachusetts, USA) either as linear fragments or
567 in pUAP1 vector[82]. Linear fragments included overhangs compatible with LguI
568 (#ER1931, ThermoFisher Scientific, Waltham, Massachusetts, USA) cloning into L0
569 acceptor vectors following the Loop protocol[38]. Linear fragments and sequences
570 synthesized in pUAP1 vectors, contained overhangs enabling direct cloning into L1 or
571 pBy01[35] vectors using BsaI-HF ®v2 enzyme (#R3733S, New England Biolabs,
572 Ipswitch, Massachusetts, USA).

573 *MpFPS* and *MpHDS* sequences were cloned from the *Marchantia* transcriptome as
574 they didn't require to remove the internal restriction sites of BsaI and LguI,
575 incompatible with the Loop cloning system. Total RNA was first extracted from
576 *Marchantia* tissue using the RNeasy Plant Mini Kit (#74904, Qiagen, Hilden, Germany)
577 and reverse-transcribed into cDNA with Superscript IV (#18090010, Invitrogen,
578 Waltham, Massachusetts, USA) using random hexamer oligos (#N8080127,
579 Invitrogen) following the protocol from the manufacturer. Target sequences were
580 subsequently amplified from the cDNA using VeriFi® Polymerase Mix (#PB10.43-01,
581 PCR Biosystem, London, UK). PCR products were purified using the QIAquick PCR
582 Purification Kit (#28104, Qiagen) before being used in the cloning steps.
583 DNA sequences exogenous to *Marchantia* were domesticated when necessary and
584 codon-optimised using the Genewiz codon-optimisation tool prior to ordering. NCBI
585 accession numbers for each exogenous genes are as follows: *TXS* (AY424738), *β-AS*
586 (MG492000), *PpDXS1* (XM_024524883), *PpDXS2* (XM_024533934), *PpHMGR*
587 (XM_024507706) and *AMS* (AY006482). Certain sequences were re-amplified to
588 remove stop codons and introduce the appropriate overhangs for fusion with mVenus
589 or eGFP tags. Primers used in this study were ordered from IDT (Integrated DNA
590 Technology, Coralville, Iowa, USA) (Table S4).
591 Vectors, promoters, terminators, tags, and pre-assembled cassettes used to generate
592 the constructs were sourced from the OpenPlant toolkit[38] or adapted from Romani et
593 al[35]. Final plasmids (L2, L3, or pBy01) were sequenced at Plasmidsaurus (Eugene,
594 Oregon, USA) to confirm proper assembly and integration of all transcription units.
595 Details of the sequences and vectors syntax are available in Dataset 5.
596 The guide RNA sequence used to generate *abcg1* mutants was designed using the
597 CasFinder tool[83]. This guide was ordered as primers with appropriate overhangs, and
598 cloned into a single-guide acceptor vector following the protocol of Sauret-Guëto et
599 al[23].
600
601 ***Agrobacterium*-mediated transformation of *Marchantia* spores and selection of**
602 **independant transformants.**
603 *Marchantia* spores were sterilized and transformed following previously described
604 protocols, with slight modifications. Briefly, a single archegoniophore, dried and stored
605 in silica beads, was mixed in 1.5 mL of a chlorinated solution consisting of one Milton

606 sterilizing tablet (troclosene sodium; Boots, UK) dissolved in 10 mL of sterile water.
607 Spores were released into the solution, filtered through a 40 µm cell strainer (#542040,
608 Greiner Bio-One) and incubated in the sterilizing solution for 30 minutes. After
609 sterilization, spores were harvested by centrifugation, resuspended in sterile water,
610 and spread on Gamborg medium agar plates (without sucrose).
611 After 5 days of germination, the sporelings were co-cultured with *Agrobacterium*
612 *tumefaciens* GV3101 carrying the constructs of interest. Transformation of
613 *Agrobacterium* with the constructs was performed using the freeze thaw method[84].
614 Sporelings and *Agrobacterium* were co-cultured for 2 days in a 4 mL solution of 0.5×
615 Gamborg B-5 plus supplements like previously described[38], and incubated at 22°C
616 under continuous shaking and light. Following co-culture, sporellings were collected
617 on 70 µm cell strainers (#542070, Greiner Bio-One), rinsed thoroughly with sterile
618 water, and plated on 90 mm Petri dishes containing 0.5× Gamborg B5 medium
619 supplemented with 100 µg/mL cefotaxime (#BIC0111, Apollo Scientific, Bredbury, UK)
620 to eliminate *Agrobacterium* and 20 µg/mL hygromycin (#10687010, Invitrogen) to
621 select for transformants.
622 After two weeks on the first selective plates, eight transformants were typically
623 transferred to a second plate containing the same selection agents to confirm
624 successful transformation and eliminate residual *Agrobacterium*. For lines expressing
625 fluorescent proteins (mScarlet, mVenus or eGFP), transformants were pre-selected
626 under a Leica stereo microscope (#MDG41, Leica, Wetzlar, Germany) using suitable
627 filter sets to confirm the presence of fluorescence signals. For lines expressing
628 exogenous terpenes, eight plants were randomly chosen, grown for one month on the
629 second selective plates, and subsequently extracted to confirm the presence of the
630 terpene of interest. The four best lines producing the desired terpene were then
631 transferred to sucrose-supplemented plates to quantify endogenous and exogenous
632 terpene levels.
633 For lines co-expressing *AMS* and precursor supply genes, a single vector carrying up
634 to six transcription units could not be constructed due to cloning or vector size
635 limitations. Attempts to co-transform using two *Agrobacterium* strains carrying distinct
636 plasmids — one with hygromycin B as the selection agent and the other with
637 chlorosulfuron — resulted in an insufficient number of transformants. Therefore,
638 transformants were selected based on hygromycin B resistance for *AMS* (with or
639 without *ABCG1*) on one vector, and the presence of an mScarlet-Lti6b fluorescent

640 signal from the second vector carrying *MpFPS*, *MpHMGR*, and/or *MpDXS2*. All eight
641 independent lines were retained to measure amorphadiene and endogenous terpene
642 level.
643 To genotype CRISPR-generated *abcg1* mutants and Cas9-only control lines, genomic
644 DNA was extracted using the InnuPrep Plant DNA kit (#845-KS-1060050, IST
645 Innuscreen, Berlin, Germany). For *abcg1* lines, a 500 bp DNA fragment spanning the
646 gRNA target site was amplified using Q5® High-Fidelity DNA Polymerase (#M0491S,
647 New England Biolabs). The resulting PCR products were purified and sequenced by
648 Genewiz to confirm the disruption or not of the *ABCG1* coding sequence. The
649 presence of the Cas9 cassette in control lines was confirmed by amplification with
650 specific primers (listed in Table S4) and subsequent sequencing. A total of 16 lines
651 were genotyped for the abcg1 ones, and eight for the Cas9 controls.
652
653 **Laser scanning confocal microscopy.**
654 For transgenic lines expressing fluorescent proteins, images were acquired on an
655 upright Leica SP8X confocal microscope equipped with a 460–670 nm
656 supercontinuum white light laser, two continuous wavelength laser lines of 405 nm and
657 442 nm and a five-channel spectral scanhead (four hybrid detectors and one
658 photomultiplier). Imaging was conducted using either a 25× water immersion (Fluotar
659 VISIR 25x/0.95 WATER) or a 40× water immersion objective (HC PL APO CS2
660 40×/1.10 WATER) for magnifying the subcellular localisations of fluorescent proteins.
661 Excitation laser wavelength and fluorescence emission bandwidth windows were as
662 follows: 515 nm and 525–550 nm (for mVenus); 570 nm and 591-621 nm (for
663 mScarlet); 442 nm and 645–664 nm (for chlorophyll autofluorescence). Each channel
664 was acquired separately using a hybrid detector, with a sequential scan for each
665 channel performed on the Leica LAS X software. To image sequentially the mVenus,
666 eGFP and mScarlet fluorescences on the translational reporter lines carrying the
667 bidirectional construct of *HDS-eGFP* and *HDR-mVenus*, excitation and emission
668 bandwidth windows were adjusted as follow to prevent signal crosstalk: 487 nm and
669 497-513 nm (for eGFP); 521 nm and 530-550 nm (for mVenus); 570 nm and 588-598
670 nm (for mScarlet).
671 Day 0 gemmae were imaged by mounting them on a glass slide with perfluorodecalin[85]
672 (#130040250, ThermoFisher) and covering them with a glass coverslip. For plants
673 aged 3 to 14 days, a 15 x 16 mm gene frame (#AB0577, ThermoFisher) was placed

674 on the glass slide to prevent compression of the plant material under the coverslip.
675 Whole gemmae and overviews of the meristematic region in older plants were imaged
676 using Z-stack scans, which were processed in Fiji[86] to generate maximum intensity
677 projections of the Z-stacks. High magnification images were acquired as single scans
678 to better display subcellular features.
679
680 **Extraction of terpenes and analysis by gas chromatography-mass spectrometry**
681 **(GC-MS).**
682 Approximately 200 mg of frozen tissue from 1-month-old *Marchantia polymorpha*
683 plants was extracted with 1 mL of cold methanol containing 5 mM NaCl to quench
684 enzymatic activity, following the method of Kumar and co-workers[25]. To enable
685 quantification, 5 µg of dodecane (#297879, Sigma-Aldrich) was included as an internal
686 standard. The tissue was disrupted in the solvent using a 3 mm stainless steel ball
687 (#2205, Durston, High Wycombe, UK) and homogenized with a TissueLyser II
688 (Qiagen). Samples were subsequently agitated on a Vibrax® shaker (#0002819002,
689 IKA, Staufen, Germany) at 2000 rpm for two hours, and the resulting extracts were
690 centrifuged to remove plant debris. The methanolic phase was then extracted once
691 with hexane to isolate non-polar and medium-polar terpenes from the upper organic
692 layer.
693 Hexane extracts (200 µL) were analyzed on a Trace 1300 gas chromatograph
694 (ThermoFisher) coupled to an ISQ 700 mass spectrometer (ThermoFisher) with a
695 Zebron CD-5MS column (30 m × 0.25 mm × 0.25 µm; ThermoFisher). A splitless
696 injection (1 µL) was conducted at 230°C, and the GC-MS oven parameters were
697 adapted from Kumar's work[25] as follows: the oven temperature was initially set to 70°C
698 and held for 3 min, followed by a ramp of 20°C/min to 90°C, a second ramp of 3°C/min
699 to 180°C, a third ramp of 5°C/min to 240°C, and a final ramp of 20°C/min to 300°C,
700 with a 6-min hold at this temperature. The MS began data acquisition after a 5.5-
701 minute solvent delay, with the transfer line and ion source temperatures set at 250°C
702 and 270°C, respectively. Scanning was conducted in full-scan mode (scan time: 0.17
703 s) over a mass range of 40–600 atomic mass units. Helium was used as the carrier
704 gas, with a flow rate of 1.2 mL/min for sesquiterpenes in the taxadiene and β-amyrin
705 datasets. For amorphadiene detection, the carrier flow rate was reduced to 0.9
706 mL/min. Chromatograms were processed using Chromeleon™ software
707 (ThermoFisher), and terpene quantification was performed relative to the internal

708 standard dodecane. Tentative identification of the major sesquiterpenes was achieved
709 by comparing mass spectra to published data and by calculating retention indices
710 using a single sample run with a C8–C40 alkane standard mixture (#40147-U,
711 Superlco, Bellefonte, Pennsylvania, USA) as described by Adams[68]. Note that the
712 endogenous terpene composition in CAM accessions may slightly differ from that
713 reported for other accessions, such as Tak, in previous studies. The authentic standard
714 of amorpha-4,11-diene was provided by Tomasz Czechowski and Ian Graham
715 (University of York).
716 To achieve more accurate quantification of β-amyrin and sterols, the extraction
717 protocol and GC-MS method were modified. The same amount of frozen material was
718 extracted with cold methanol supplemented with 10 µg of coprostan-3-ol (#C7578,
719 Sigma-Aldrich) as an internal standard. The samples were extracted twice with hexane
720 to obtain a total hexane volume of 1.5 mL, which was then dried under nitrogen flow
721 in a Genevac™ concentrator EZ-2 (#EZ3P-23050-NN0, Genevac Ltd, Ipswich, UK).
722 Extracts were derivatized with 50 µL 1-trimethylsilyl-imidazole (TMS; #92718,
723 Superlco) and resuspended in 300 µL hexane. For GC-MS analysis, 200 µL of the
724 derivatized samples was transferred into vials for direct injection. The method used for
725 sterol/triterpene analysis was as follows: 130°C hold for 2 min; ramp of 30°C/min until
726 220°C, then 2°C/min until 300°C, hold for 10 min. Solvent delay of 10 minutes before
727 acquiring the MS data. The authentic standard of β-amyrin was provided by James
728 Reed and Anne Osbourne (John Innes Center, Norwich, UK).
729 For terpene quantification in *abcg1* CRISPR and Cas9 control lines, extractions were
730 performed on genotyped plants from the second selection stage, as loss of the
731 CRISPR genotype was observed during propagation from gemmae. Given this
732 observation, we considered the possibility of chimerism[87] and extracted as much
733 biomass as was reasonably available from individual plants. This approach aimed to
734 minimize variability, homogenize the material, and ensure accurate results.
735 Box plots showing individual data points for each quantified terpene were created
736 using PRISM software version 10 (GraphPad, La Jolla, California, USA), and the
737 results of statistical comparisons were integrated into the visualizations. For lines
738 expressing exogenous terpene synthases, statistical analyses were performed using
739 one-way analysis of variance (ANOVA) to assess differences across gene
740 combinations. Homogeneity of variances was evaluated using the Brown-Forsythe
741 and Bartlett's tests, with Dunnett's multiple comparison test applied post-hoc to

742 compare each gene combination to the respective control ($_{Prom5}$*MYB2:β-AS*, *TXS*, or
743 *AMS*). The significance threshold (alpha) was set at 0.05. Four biological replicates
744 expressing the constructs were analyzed to quantify both exogenous (β-amyrin and
745 taxadiene) and endogenous compounds (n=4), while eight biological replicates
746 expressing the constructs were analyzed for amorphadiene lines, focusing on the
747 quantification of endogenous terpenes since amorphadiene was not detected (n=8).
748 For comparisons of terpene levels between *abcg1* mutants (n = 6) and Cas9 controls
749 (n = 5), Welch's t-test was used when variances were unequal, and unpaired t-test
750 was used when variances were equal, as determined by PRISM. Both tests accounted
751 for differences in sample size between the two groups. Asterisks were added to the
752 box plots to indicate statistically significant differences (P<0.05) between *abcg1*
753 mutants and Cas9 control lines, as determined by Welch's or unpaired t-tests.
754
755 **REFERENCES**
756
757 1. Bowman, J. L. *et al.* The renaissance and enlightenment of Marchantia as a
758 model system. *Plant Cell* **34**, 3512–3542 (2022).
759 2. Tanaka, M. *et al.* Direct evidence of specific localization of sesquiterpenes and
760 marchantin A in oil body cells of Marchantia polymorpha L. *Phytochemistry* **130**, 77–
761 84 (2016).
762 3. Suire, C. *et al.* Cellular Localization of Isoprenoid Biosynthetic Enzymes in
763 *Marchantia polymorpha* . Uncovering a New Role of Oil Bodies. *Plant Physiology* **124**,
764 971–978 (2000).
765 4. Jensen, S. *et al.* Marchantin A, a macrocyclic bisbibenzyl ether, isolated from
766 the liverwort Marchantia polymorpha, inhibits protozoal growth in vitro. *Phytomedicine*
767 **19**, 1191–1195 (2012).
768 5. Romani, F. *et al.* Oil Body Formation in Marchantia polymorpha Is Controlled
769 by MpC1HDZ and Serves as a Defense against Arthropod Herbivores. *Current Biology*
770 **30**, 2815-2828.e8 (2020).
771 6. Ruzicka, L. The isoprene rule and the biogenesis of terpenic compounds.
772 *Experientia* **9**, 357–367 (1953).
773 7. Kanazawa, T. *et al.* The liverwort oil body is formed by redirection of the
774 secretory pathway. *Nat Commun* **11**, 6152 (2020).
775 8. Gutsche, N. *et al.* Mp *TGA* , together with Mp *NPR* , regulates sexual

776 reproduction and independently affects oil body formation in *Marchantia polymorpha*.
777 *New Phytologist* **241**, 1559–1573 (2024).
778 9. Kubo, H. *et al.* Biosynthesis of riccionidins and marchantins is regulated by
779 R2R3-MYB transcription factors in Marchantia polymorpha. *J Plant Res* **131**, 849–864
780 (2018).
781 10. Kubo, H., Sunhwa, K., Teramori, H. & Takanashi, K. MpR2R3-MYB2 is a key
782 regulator of oil body formation in Marchantia polymorpha. *Planta* **260**, 68 (2024).
783 11. Romani, F. *et al.* Liverwort oil bodies: diversity, biochemistry, and molecular cell
784 biology of the earliest secretory structure of land plants. *Journal of Experimental*
785 *Botany* **73**, 4427–4439 (2022).
786 12. Ishizaki, K., Nishihama, R., Yamato, K. T. & Kohchi, T. Molecular Genetic Tools
787 and Techniques for Marchantia polymorpha Research. *Plant and Cell Physiology* **57**,
788 262–270 (2016).
789 13. Montgomery, S. A. *et al.* Chromatin Organization in Early Land Plants Reveals
790 an Ancestral Association between H3K27me3, Transposons, and Constitutive
791 Heterochromatin. *Current Biology* **30**, 573-588.e7 (2020).
792 14. Rohmer, M., Knani, M., Simonin, P., Sutter, B. & Sahm, H. Isoprenoid
793 biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl
794 diphosphate. *Biochemical Journal* **295**, 517–524 (1993).
795 15. Lichtenthaler, H. K., Schwender, J., Disch, A. & Rohmer, M. Biosynthesis of
796 isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent
797 pathway. *Febs Letters* **400**, 271–274 (1997).
798 16. Rohmer, M. The discovery of a mevalonate-independent pathway for
799 isoprenoid biosynthesis in bacteria, algae and higher plants†. *Nat. Prod. Rep.* **16**, 565–
800 574 (1999).
801 17. Bouvier, F., Rahier, A. & Camara, B. Biogenesis, molecular regulation and
802 function of plant isoprenoids. *Progress in Lipid Research* **44**, 357–429 (2005).
803 18. Rodríguez-Concepción, M. & Boronat, A. Breaking new ground in the regulation
804 of the early steps of plant isoprenoid biosynthesis. *Current Opinion in Plant Biology*
805 **25**, 17–22 (2015).
806 19. Leivar, P. *et al.* Subcellular Localization of Arabidopsis 3-Hydroxy-3-
807 Methylglutaryl-Coenzyme A Reductase. *Plant Physiology* **137**, 57–69 (2005).
808 20. Simkin, A. J. *et al.* Peroxisomal localisation of the final steps of the mevalonic
809 acid pathway in planta. *Planta* **234**, 903–914 (2011).

810 21. Chappell, J. Biochemistry and Molecular Biology of the Isoprenoid Biosynthetic
811 Pathway in Plants. *Annual Review of Plant Biology* **46**, 521–547 (1995).
812 22. KUZUYAMA, T. Mevalonate and Nonmevalonate Pathways for the
813 Biosynthesis of Isoprene Units. *Bioscience, Biotechnology, and Biochemistry* **66**,
814 1619–1627 (2002).
815 23. Hsieh, M.-H., Chang, C.-Y., Hsu, S.-J. & Chen, J.-J. Chloroplast localization of
816 methylerythritol 4-phosphate pathway enzymes and regulation of mitochondrial genes
817 in ispD and ispE albino mutants in Arabidopsis. *Plant Mol Biol* **66**, 663–673 (2008).
818 24. Bowman, J. L. *et al.* Insights into Land Plant Evolution Garnered from the
819 Marchantia polymorpha Genome. *Cell* **171**, 287-304.e15 (2017).
820 25. Kumar, S. *et al.* Molecular Diversity of Terpene Synthases in the Liverwort
821 Marchantia polymorpha. *Plant Cell* tpc.00062.2016 (2016) doi:10.1105/tpc.16.00062.
822 26. Ødum, M. T. *et al.* DeepLoc 2.1: multi-label membrane protein type prediction
823 using protein language models. *Nucleic Acids Research* **52**, W215–W220 (2024).
824 27. Koepp, A. E. *et al.* Cyclization of Geranylgeranyl Diphosphate to Taxa-
825 4(5),11(12)-diene Is the Committed Step of Taxol Biosynthesis in Pacific Yew *.
826 *Journal of Biological Chemistry* **270**, 8686–8690 (1995).
827 28. Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P. & McPhail, A. T. Plant
828 antitumor agents. VI. Isolation and structure of taxol, a novel antileukemic and
829 antitumor agent from Taxus brevifolia. *J. Am. Chem. Soc.* **93**, 2325–2327 (1971).
830 29. Suffness, M. *TAXOL®: Science and Applications*. (CRC Press, London, 2021).
831 doi:10.1201/9780138737368.
832 30. Kushiro, T., Shibuya, M. & Ebizuka, Y. β-Amyrin synthase. *European Journal*
833 *of Biochemistry* **256**, 238–244 (1998).
834 31. Fiore, C., Eisenhut, M., Ragazzi, E., Zanchin, G. & Armanini, D. A history of the
835 therapeutic use of liquorice in Europe. *J Ethnopharmacol* **99**, 317–324 (2005).
836 32. Bouwmeester, H. J. *et al.* Amorpha-4,11-diene synthase catalyses the first
837 probable step in artemisinin biosynthesis. *Phytochemistry* **52**, 843–854 (1999).
838 33. Wang, L. *et al.* The maturation and aging trajectory of Marchantia polymorpha
839 at single-cell resolution. *Developmental Cell* **58**, 1429-1444.e6 (2023).
840 34. Kremers, G.-J., Goedhart, J., van Munster, E. B. & Gadella, T. W. J. Cyan and
841 Yellow Super Fluorescent Proteins with Improved Brightness, Protein Folding, and
842 FRET Förster Radius,. *Biochemistry* **45**, 6570–6580 (2006).
843 35. Romani, F. *et al.* The landscape of transcription factor promoter activity during

844 vegetative development in Marchantia. *Plant Cell* **36**, 2140–2159 (2024).
845 36. Karaaslan, E. S. *et al.* Marchantia TCP transcription factor activity correlates
846 with three-dimensional chromatin structure. *Nat. Plants* **6**, 1250–1261 (2020).
847 37. Bindels, D. S. *et al.* mScarlet: a bright monomeric red fluorescent protein for
848 cellular imaging. *Nat Methods* **14**, 53–56 (2017).
849 38. Sauret-Güeto, S. *et al.* Systematic Tools for Reprogramming Plant Gene
850 Expression in a Simple Model, *Marchantia polymorpha*. *ACS Synth. Biol.* **9**, 864–882
851 (2020).
852 39. Cutler, S. R., Ehrhardt, D. W., Griffitts, J. S. & Somerville, C. R. Random
853 GFP::cDNA fusions enable visualization of subcellular structures in cells of
854 *Arabidopsis* at a high frequency. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3718–3723 (2000).
855 40. Minamino, N. *et al.* RAB GTPases in the Basal Land Plant Marchantia
856 polymorpha. *Plant and Cell Physiology* **59**, 850–861 (2018).
857 41. Lichtenthaler, H. K. THE 1-DEOXY-D-XYLULOSE-5-PHOSPHATE PATHWAY
858 OF ISOPRENOID BIOSYNTHESIS IN PLANTS. *Annual Review of Plant Biology* **50**,
859 47–65 (1999).
860 42. Rodríguez-Concepción, M. & Boronat, A. Elucidation of the Methylerythritol
861 Phosphate Pathway for Isoprenoid Biosynthesis in Bacteria and Plastids. A Metabolic
862 Milestone Achieved through Genomics. *Plant Physiology* **130**, 1079–1089 (2002).
863 43. Hsieh, M.-H. The Arabidopsis IspH Homolog Is Involved in the Plastid
864 Nonmevalonate Pathway of Isoprenoid Biosynthesis. *PLANT PHYSIOLOGY* **138**,
865 641–653 (2005).
866 44. Seemann, M., Tse Sum Bui, B., Wolff, M., Miginiac-Maslow, M. & Rohmer, M.
867 Isoprenoid biosynthesis in plant chloroplasts via the MEP pathway: Direct
868 thylakoid/ferredoxin-dependent photoreduction of GcpE/IspG. *FEBS Letters* **580**,
869 1547–1552 (2006).
870 45. Li, Q. *et al.* Balanced activation of IspG and IspH to eliminate MEP intermediate
871 accumulation and improve isoprenoids production in Escherichia coli. *Metabolic*
872 *Engineering* **44**, 13–21 (2017).
873 46. Forestier, E. C. F., Brown, G. D., Harvey, D., Larson, T. R. & Graham, I. A.
874 Engineering Production of a Novel Diterpene Synthase Precursor in Nicotiana
875 benthamiana. *Frontiers in Plant Science* **12**, 8 (2021).
876 47. Laferrière, A. & Beyer, P. Purification of geranylgeranyl diphosphate synthase
877 from *Sinapis alba* etioplasts. *Biochimica et Biophysica Acta (BBA) - Protein Structure*

878 *and Molecular Enzymology* **1077**, 167–172 (1991).
879 48. Kuntz, M. *et al.* Identification of a cDNA for the plastid-located geranylgeranyl
880 pyrophosphate synthase from Capsicum annuum: correlative increase in enzyme
881 activity and transcript level during fruit ripening. *The Plant Journal* **2**, 25–34 (1992).
882 49. Bach, T. J. Hydroxymethylglutaryl-CoA reductase, a key enzyme in phytosterol
883 synthesis? *Lipids* **21**, 82–88 (1986).
884 50. Poulter, C. D., Argyle, J. C. & Mash, E. A. Farnesyl pyrophosphate synthetase.
885 Mechanistic studies of the 1'-4 coupling reaction with 2-fluorogeranyl pyrophosphate.
886 *Journal of Biological Chemistry* **253**, 7227–7233 (1978).
887 51. Tse, S. W. *et al.* Optimizing Promoters and Subcellular Localization for
888 Constitutive Transgene Expression in Marchantia polymorpha. *Plant and Cell*
889 *Physiology* **65**, 1298 (2024).
890 52. Kuzuyama, T., Takahashi, S., Watanabe, H. & Seto, H. Direct formation of 2-
891 C-methyl-d-erythritol 4-phosphate from 1-deoxy-d-xylulose 5-phosphate by 1-deoxy-
892 d-xylulose 5-phosphate reductoisomerase, a new enzyme in the non-mevalonate
893 pathway to isopentenyl diphosphate. *Tetrahedron Letters* **39**, 4509–4512 (1998).
894 53. Takahashi, S., Kuzuyama, T., Watanabe, H. & Seto, H. A 1-deoxy-D-xylulose
895 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-
896 phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis.
897 *Proceedings of the National Academy of Sciences of the United States of America* **95**,
898 9879–9884 (1998).
899 54. Bloch, K., Chaykin, S., Phillips, A. H. & De Waard, A. Mevalonic Acid
900 Pyrophosphate and Isopentenylpyrophosphate. *Journal of Biological Chemistry* **234**,
901 2595–2604 (1959).
902 55. Schaller, H. The role of sterols in plant growth and development. *Prog Lipid Res*
903 **42**, 163–175 (2003).
904 56. Mookhtiar, K. A., Kalinowski, S. S., Zhang, D. & Poulter, C. D. Yeast squalene
905 synthase. A mechanism for addition of substrates and activation by NADPH. *Journal*
906 *of Biological Chemistry* **269**, 11201–11207 (1994).
907 57. Besumbes, Ó. *et al.* Metabolic engineering of isoprenoid biosynthesis in
908 Arabidopsis for the production of taxadiene, the first committed precursor of Taxol.
909 *Biotechnology and Bioengineering* **88**, 168–175 (2004).
910 58. Afifah, I. Q., Wibowo, I. & Faizal, A. A newly identified β-amyrin synthase gene
911 hypothetically involved in oleanane-saponin biosynthesis from Talinum paniculatum

912 (Jacq.) Gaertn. *Heliyon* **9**, e17707 (2023).
913 59. Hasan, Md. M. *et al.* Metabolic engineering of Nicotiana benthamiana for the
914 increased production of taxadiene. *Plant Cell Rep* **33**, 895–904 (2014).
915 60. Li, J. *et al.* Chloroplastic metabolic engineering coupled with isoprenoid pool
916 enhancement for committed taxanes biosynthesis in Nicotiana benthamiana. *Nat*
917 *Commun* **10**, 4850 (2019).
918 61. Park, Y.-D. *et al.* Gene silencing mediated by promoter homology occurs at the
919 level of transcription and results in meiotically heritable alterations in methylation and
920 gene activity. *The Plant Journal* **9**, 183–194 (1996).
921 62. Horn, A., Lu, Y., Astorga Ríos, F. J., Toft Simonsen, H. & Becker, J. D.
922 Transcriptional and functional characterization in the terpenoid precursor pathway of
923 the early land plant *Physcomitrium patens*. *Plant Biol J* plb.13741 (2024)
924 doi:10.1111/plb.13741.
925 63. Brückner, K. & Tissier, A. High-level diterpene production by transient
926 expression in Nicotiana benthamiana. *Plant methods* **9**, 46 (2013).
927 64. Forestier, E. C. F. *et al.* Developing a *Nicotiana benthamiana* transgenic
928 platform for high-value diterpene production and candidate gene evaluation. *Plant*
929 *Biotechnol J* **19**, 1614–1623 (2021).
930 65. Schaller, H. *et al.* Expression of the Hevea brasiliensis (H.B.K.) Mull. Arg. 3-
931 Hydroxy-3-Methylglutaryl-Coenzyme A Reductase 1 in Tobacco Results in Sterol
932 Overproduction. *Plant Physiology* **109**, 761–770 (1995).
933 66. Chin, D. J. *et al.* Nucleotide sequence of 3-hydroxy-3-methyl-glutaryl coenzyme
934 A reductase, a glycoprotein of endoplasmic reticulum. *Nature* **308**, 613–617 (1984).
935 67. Chappell, J., Wolf, F., Proulx, J., Cuellar, R. & Saunders, C. Is the Reaction
936 Catalyzed by 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase a Rate-Limiting
937 Step for Isoprenoid Biosynthesis in Plants? *Plant Physiology* **109**, 1337–1343 (1995).
938 68. Adams, R. P. *Identification of Essential Oil Components by Gas*
939 *Chromatography Mass Spectroscopy*. (Allured Publishing Corporation, Carol Stream,
940 Ill, 2007).
941 69. Picaud, S. *et al.* Amorpha-4,11-diene synthase: mechanism and
942 stereochemistry of the enzymatic cyclization of farnesyl diphosphate. *Arch Biochem*
943 *Biophys* **448**, 150–155 (2006).
944 70. Cane, D. E. 2.06 - Sesquiterpene Biosynthesis: Cyclization Mechanisms. in
945 *Comprehensive Natural Products Chemistry* (eds. Barton, S. D., Nakanishi, K. & Meth-

946 Cohn, O.) 155–200 (Pergamon, Oxford, 1999). doi:10.1016/B978-0-08-091283-
947 7.00039-4.
948 71. Cunillera, N. *et al. Arabidopsis thaliana* Contains Two Differentially Expressed
949 Farnesyl-Diphosphate Synthase Genes (∗). *Journal of Biological Chemistry* **271**,
950 7774–7780 (1996).
951 72. Grossi, V. & Rontani, J.-F. Photosensitized oxygenation of phytadienes.
952 *Tetrahedron Letters* **36**, 3141–3144 (1995).
953 73. Grossi, V. *et al.* Formation of phytadienes in the water column: myth or reality?
954 *Organic Geochemistry* **24**, 833–839 (1996).
955 74. Takizawa, R. *et al.* Fungal-Type Terpene Synthases in Marchantia polymorpha
956 Are Involved in Sesquiterpene Biosynthesis in Oil Body Cells. *Plant and Cell*
957 *Physiology* **62**, 528–537 (2021).
958 75. Crook, O. M. *et al.* Inferring differential subcellular localisation in comparative
959 spatial proteomics using BANDLE. *Nat Commun* **13**, 5948 (2022).
960 76. Zid, B. M. & O'Shea, E. K. Promoter sequences direct cytoplasmic localization
961 and translation of mRNAs during starvation in yeast. *Nature* **514**, 117–121 (2014).
962 77. Daudonnet, S., Karst, F. & Tourte, Y. Expression of the farnesyldiphosphate
963 synthase gene of Saccharomyces cerevisiae in tobacco. *Molecular Breeding* **3**, 137–
964 145 (1997).
965 78. Nair, P. *et al.* Differentially Expressed Genes during Contrasting Growth Stages
966 of Artemisia annua for Artemisinin Content. *PLoS ONE* **8**, e60375 (2013).
967 79. Yelina, N. E. *et al.* Streamlined regulation of chloroplast development in the
968 liverwort Marchantia polymorpha. 2023.01.23.525199 Preprint at
969 https://doi.org/10.1101/2023.01.23.525199 (2024).
970 80. Sun, R. Biosynthesis of gibberellin-related compounds modulates far-red light
971 responses in the liverwort Marchantia polymorpha.
972 81. Delmans, M., Pollak, B. & Haseloff, J. MarpoDB: An Open Registry for
973 Marchantia Polymorpha Genetic Parts. *Plant and Cell Physiology* **58**, e5 (2017).
974 82. Patron, N. J. *et al.* Standards for plant synthetic biology: a common syntax for
975 exchange of DNA parts. *New Phytologist* **208**, 13–19 (2015).
976 83. Aach, J., Mali, P. & Church, G. M. CasFinder: Flexible algorithm for identifying
977 specific Cas9 targets in genomes. 005074 Preprint at https://doi.org/10.1101/005074
978 (2014).
979 84. Höfgen, R. & Willmitzer, L. Storage of competent cells for Agrobacterium

980 transformation. *Nucleic Acids Res* **16**, 9877 (1988).
981 85. Littlejohn, G. R., Gouveia, J. D., Edner, C., Smirnoff, N. & Love, J.
982 Perfluorodecalin enhances in vivo confocal microscopy resolution of Arabidopsis
983 thaliana mesophyll. *New Phytologist* **186**, 1018–1025 (2010).
984 86. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis.
985 *Nat Methods* **9**, 676–682 (2012).
986 87. Song, G.-Q., Urban, G., Ryner, J. T. & Zhong, G.-Y. Gene Editing Profiles in 94
987 CRISPR-Cas9 Expressing T0 Transgenic Tobacco Lines Reveal High Frequencies of
988 Chimeric Editing of the Target Gene. *Plants* **11**, 3494 (2022).
989

	Name	Abbreviation	Putative localisation (DeepLoc 2.1)
MEP	1-Deoxy-D-xylulose-5-phosphate synthase 1	*MpDXS1*	Plastids
	1-Deoxy-D-xylulose-5-phosphate synthase 2	*MpDXS2*	Plastids
	1-Deoxy-D-xylulose 5-phosphate reductoisomerase 1	*MpDXR1*	Plastids
	1-Deoxy-D-xylulose 5-phosphate reductoisomerase 2	*MpDXR2*	Plastids
	4-Diphosphocytidyl-2-*C*-methyl-*D*-erythritol synthase	*MpMCT*	Plastids
	4-Diphosphocytidyl-2-*C*-methyl-*D*-erythritol 2-phosphate kinase	*MpCMK*	Plastids
	2-*C*-methyl-*D*-erythritol 2,4-cyclodiphosphate synthase	*MpMDS*	Plastids
	1-Hydroxy-2-methyl-2-*(E)*-butenyl 4-diphosphate synthase	*MpHDS*	Plastids
	1-Hydroxy-2-methyl-2-*(E)*-butenyl 4-diphosphate reductase	*MpHDR*	Plastids
	Isopentenyl diphosphate isomerase 1	*MpIDI1*	Plastids
	Isoprenyl (geranyl) diphosphate synthase 5	*MpIDS5/MpGPS*	Mitochondria
	Cis-prenyltransferase 5 / Neryl pyrophophate synthase	*MpCPT5*	Plastids
	Isoprenyl (geranylgeranyl) diphosphate synthase 2	*MpIDS2/MpGGPPS*	Plastids
MVA	Acetoacetyl-coA thiolase 2	*MpACT2*	Peroxisome
	3-Hydroxy-3-methylglutaryl-coA synthase	*MpHMGS*	Cytoplasm
	3-Hydroxy-3-methylglutaryl-coA reductase	*MpHMGR*	Endoplasmic reticulum
	Mevalonate-5-kinase	*MpMK*	Cytoplasm
	Phosphomevanolate kinase	*MpPMK*	Cytoplasm
	Mevanolate diphosphate decarboxylase.	*MpMVD*	Cytoplasm
	Isopentenyl diphosphate isomerase 2	*MpIDI2*	Cytoplasm
	Farnesyl pyrophosphate synthase	*MpFPS*	Mitochondria
	Squalene synthase 3	*MpSQS3*	Endoplasmic reticulum
	Squalene/phytoene synthase 1	*MpSQS1*	Plastids
	Squalene/phytoene synthase 2	*MpSQS2*	Mitochondria

Table 1. Overview of MVA, MEP, and terpene scaffold biosynthetic genes in *Marchantia polymorpha*.
This table lists key genes studied or referenced in this work, including those involved in the methylerythritol phosphate (MEP) and mevalonate (MVA) pathways, as well as enzymes responsible for synthesizing terpene scaffolds for different sub-classes: monoterpenes (IDS5/GPS, CPT5), diterpenes (IDS2/GGPPS), sesquiterpenes (FPS), and triterpenes/sterols (SQS3). Abbreviations correspond to gene names, and their putative protein localizations were predicted using DeepLoc 2.1. Detailed annotations and functional descriptions for each gene are provided in Table S1.

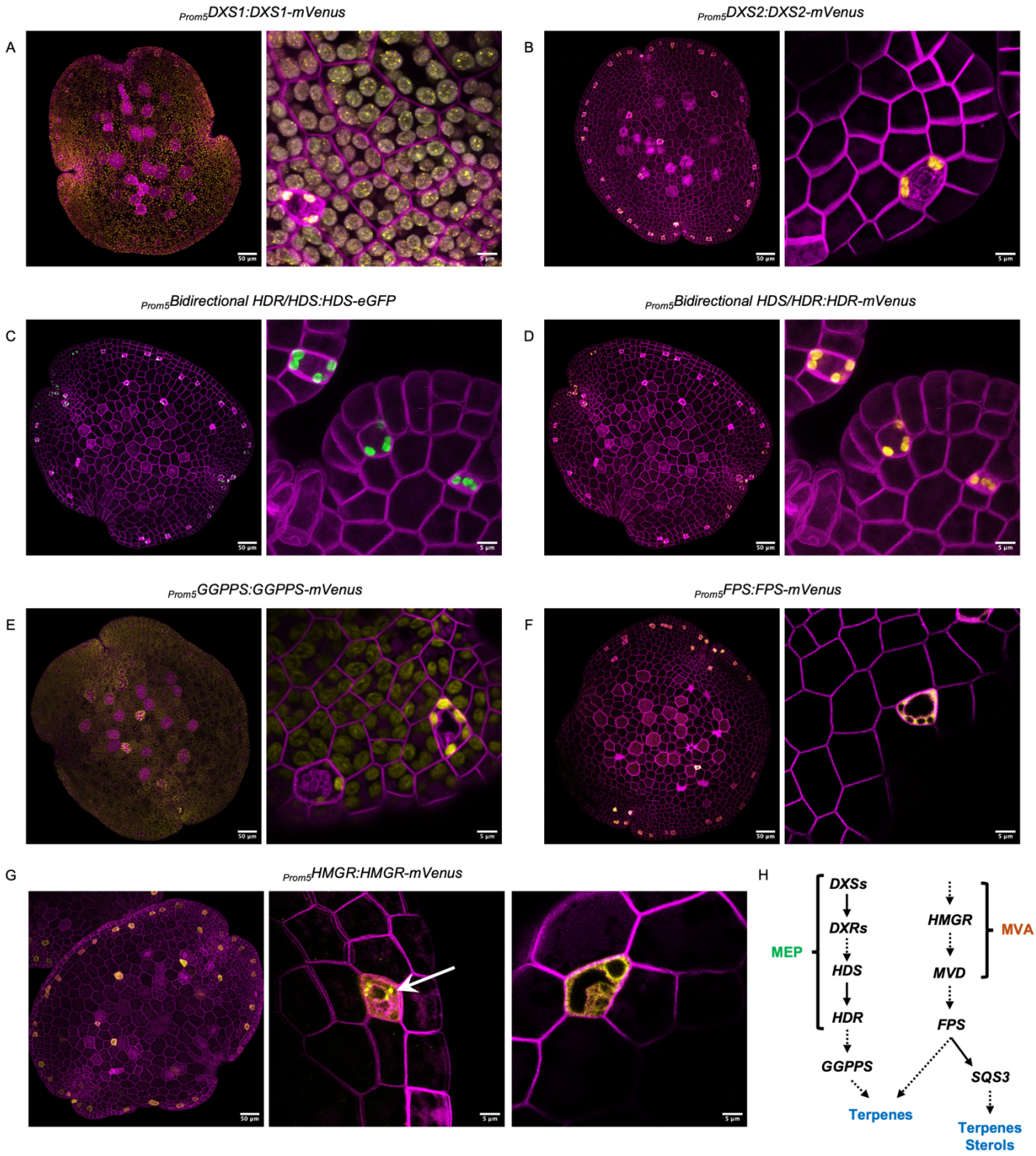


Figure 1. Confocal imaging of translational reporters for key *Marchantia polymorpha* isoprenoid biosynthetic genes.
(A) $_{Prom5}$*DXS1:DXS1-mVenus*, (B) $_{Prom5}$*DXS2:DXS2-mVenus*, (C) $_{Prom5}$*Bidirectional HDR:HDS:HDS-eGFP*, (D) $_{Prom5}$*Bidirectional HDS:HDR:HDR-mVenus*, (E) $_{Prom5}$*GGPPS:GGPPS-mVenus*, (F) $_{Prom5}$*FPS:FPS-mVenus*, and (G) $_{Prom5}$*HMGR:HMGR-mVenus*. Each construct is represented by two images: the full gemmae at day 0 (left panels, scale bar: 50 µm) and a higher magnification view showing subcellular localization (right panels, scale bar: 5 µm). (H) Simplified biosynthetic pathway scheme highlighting enzymatic steps analyzed in this study. The *mVenus* signal (yellow) or eGFP (green) for HDS (C) indicates the subcellular localization of the corresponding enzymes, while *mScarlet* (purple) delineates the cellular boundaries. For the magnified image in (A), chloroplast/plastid autofluorescence (grey) is included to demonstrate sublocalization of DXS1 in the plastids. Panel (G) includes one whole-gemma image and two high magnification images to illustrate the localization of HMGR-mVenus to the ER and Golgi apparatus. White arrow shows the localizations to Golgi.

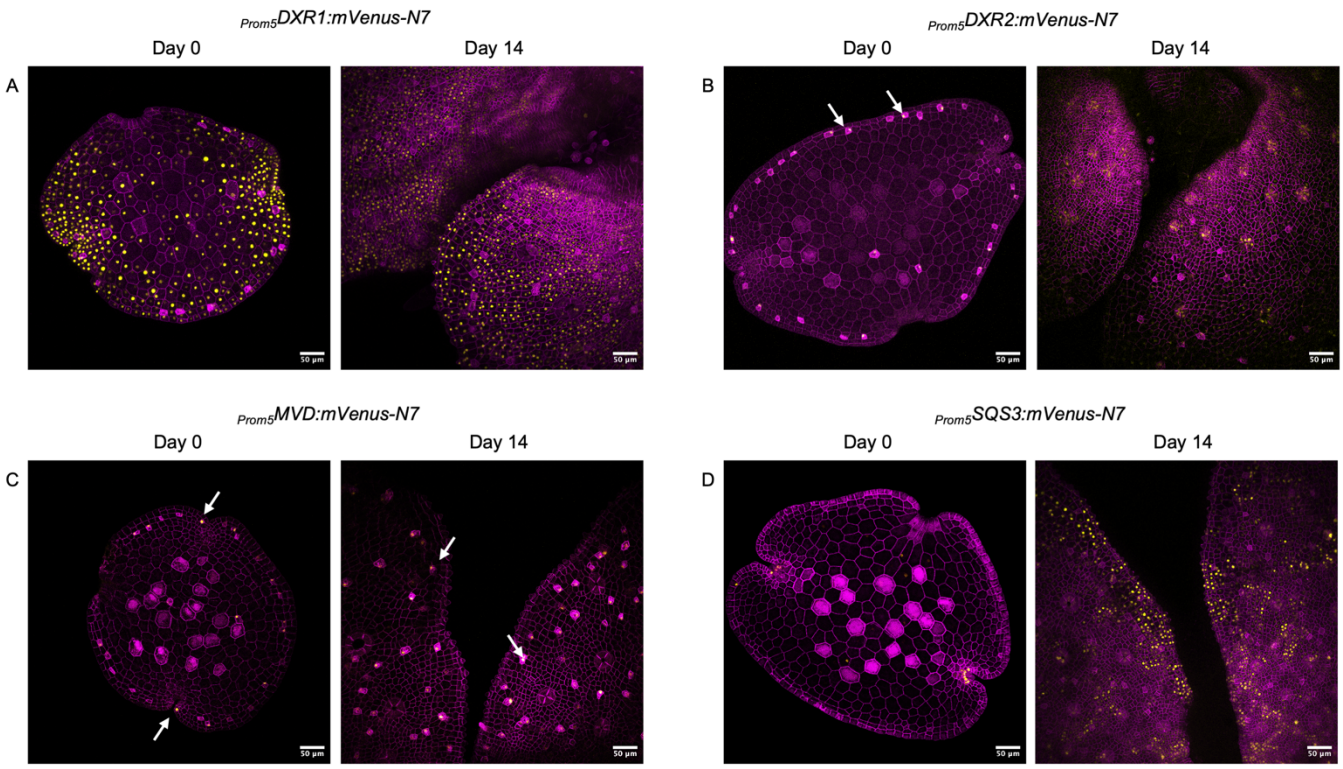


Figure 2. Confocal imaging of translational reporters for additional isoprenoid biosynthetic genes in *Marchantia polymorpha.*
(A) *Prom5DXR1:mVenus-N7*, (B) *Prom5DXR2:mVenus-N7*, (C) *Prom5MVD:mVenus-N7*, and (D) *Prom5SQS3:mVenus-N7*. The mVenus signal (yellow) localized to the nucleus indicates the activation of these promoters in specific cell types, while mScarlet fluorescence (purple) delineates cellular boundaries. Each construct is represented by two panels: day 0 gemmae (left, scale bar: 50 µm) and the meristem area of 14-day-old plants (right, scale bar: 50 µm). White arrows indicate mVenus signals detected in oil body cells.

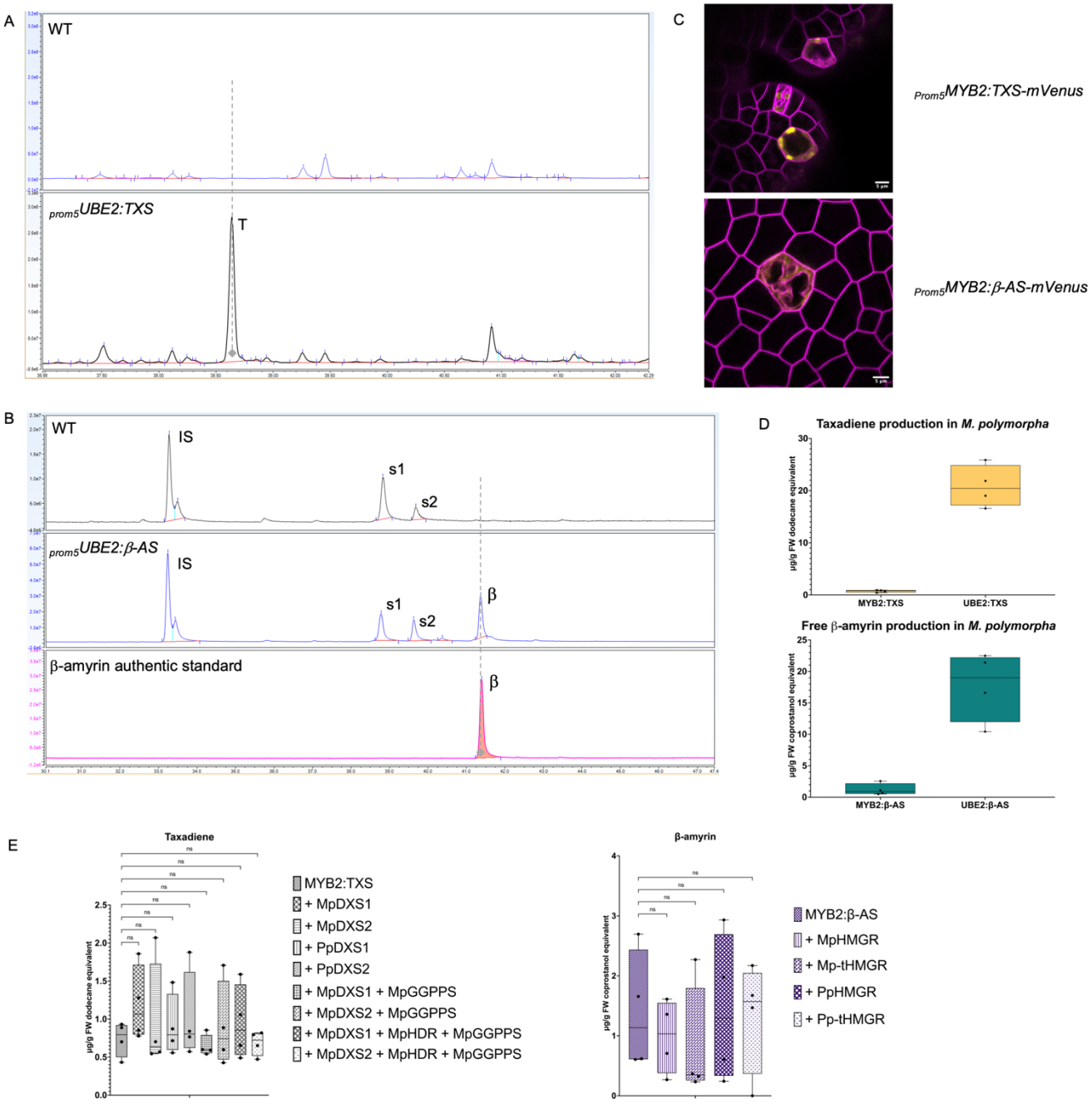


Figure 3. Production of taxadiene and β-amyrin in *Marchantia polymorpha* and subcellular localization of their synthases.
(A) Total ion chromatograms (TICs) of terpene extracts from wild-type (WT, upper panel) and transgenic plants expressing *Prom5UBE2:TXS* (*taxadiene synthase*, lower panel), highlighting the taxadiene peak (T). (B) TICs of trimethylsilyl-derivatized terpene extracts, showing WT (upper), a transgenic line expressing *Prom5UBE2:β-AS* (*β-amyrin synthase,* middle), and an authentic β-amyrin standard (lower). Peaks: IS (internal standard: coprostan-3-ol), s1 (campesterol), s2 (stigmasterol), and β (β-amyrin).
(C) Subcellular localization of TXS and β-AS fused to mVenus under the *Prom5MYB2* promoter, showing plastid localization for TXS (upper panel) and ER or cytosol localization for β-AS (lower panel). mScarlet fluorescence (purple) delineates cellular boundaries. Scale bar: 5 µm. (D) Quantification of taxadiene (upper graph) and β-amyrin (lower graph) in plants expressing *Prom5UBE2:TXS* and *Prom5UBE2:β-AS* respectively (whole plant expression) or *Prom5MYB2:TXS* and *Prom5MYB2:β-AS* respectively (oil body cell-specific expression). Data represent four independent lines per condition.
(E) Quantification of taxadiene (left) and β-amyrin (right) in transgenic lines over-expressing precursor genes. For taxadiene, combinations included *MpDXS1*, *MpDXS2*, *PpDXS1*, *PpDXS2*, *MpHDR*, and *MpGGPPS*. For β-amyrin, *MpHMGR*, *Mp-tHMGR*, *PpHMGR*, and *Pp-tHMGR* were tested. Box plots represent data points from independent transformants. One-way ANOVA with Dunnett's multiple comparison tests ($p < 0.05$) showed no significant differences across gene combinations.

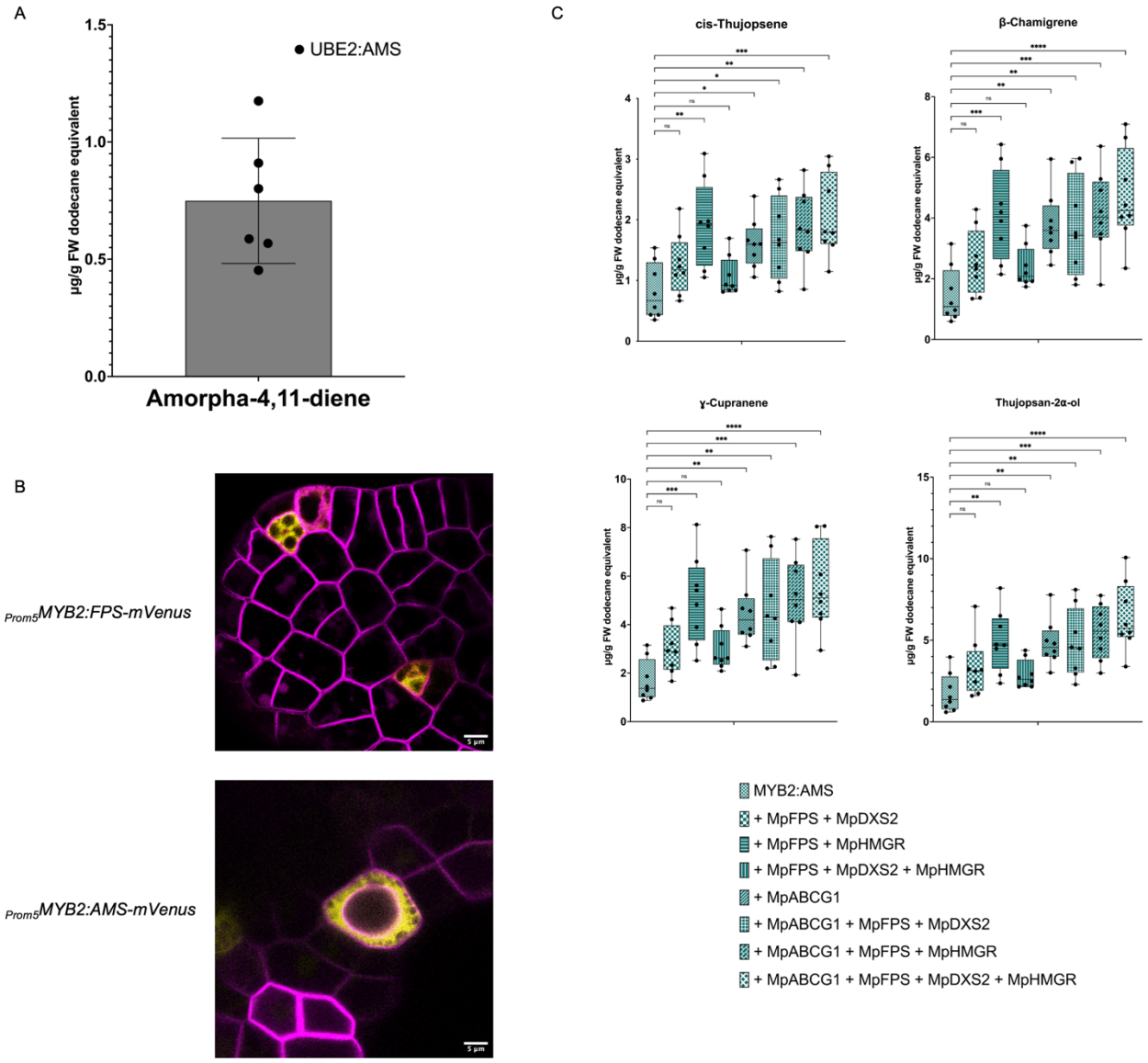
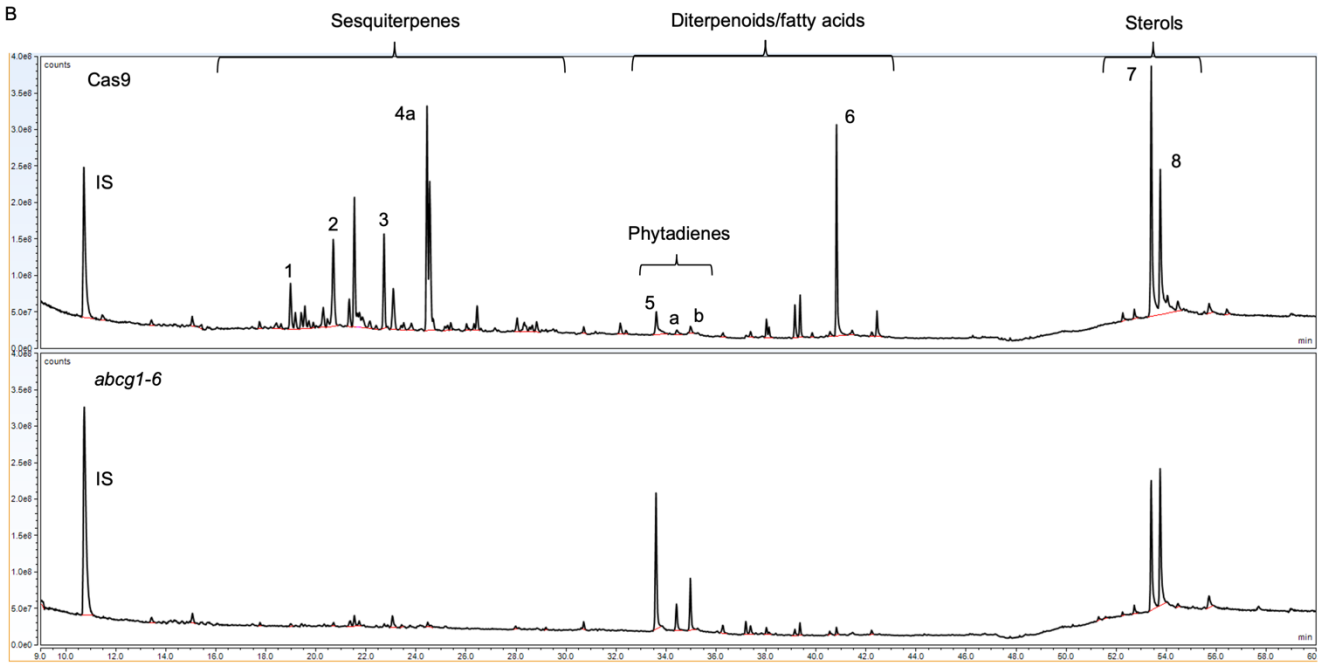


Figure 4. Expression of amorphadiene synthase (*AMS*) in *Marchantia polymorpha* with or without precursor supply genes and the ABCG1 transporter, and its effect on endogenous sesquiterpene levels.
(A) Quantification of amorpha-4,11-diene levels in plants expressing *Artemisia annua* AMS under the *UBE2* promoter. The compound was detected in six independent transformants (mean ± SD; n = 6). (B) Subcellular localization of FPS and AMS proteins when expressed under the oil-body-specific *MYB2* promoter. mVenus fluorescence (yellow) indicates protein localization, with signals observed in oil body cells, while mScarlet fluorescence (purple) marks cell boundaries. Scale bars: 5 µm. (C) Quantification of endogenous sesquiterpenes, including cis-thujopsene, β-chamigrene, γ-cuprenene, and thujopsan-2α-ol, in lines expressing *AMS* with or without precursor supply genes (*MpFPS*, *MpDXS2*, and/or *MpHMGR*) and with or without *ABCG1*, under $_{Prom5}$*MYB2*. Box plots represent data points from eight independent transformants (n = 8), with significance determined by one-way ANOVA followed by Dunnett's multiple comparison test ($p < 0.05$).

A

WT *ABCG1* 301- CAGATGAACCCAGAGGAGGATAATGAGAAGTTTCTCGCCAAGCTAA
abcg1_2 301- CAGATGAACCCAG __GAGGATAATGAGAAGTTTCTCGCCAAGCTAA
abcg1_5 301- CAGATGAACCCAGTCCCACTATGCAGATACGTGAGTATATGTCGGATATGTACCACTATGCAGAGGAGGATAATGAGAAGTTTCTCGCCAAGCTAA
abcg1_6 301- CAGATGAACCCAG __GAGGATAATGAGAAGTTTCTCGTAAGAGAACCTGCTTCATTATTTATAATGAGAAGTTTCTCGCCAAGCTAA
abcg1_7 301- CAGATGAACCTTCTCCATTTCAAGATACCTTTTTCATGTTATTGAATTGGTGAGTTCAATGAATTGGTTAATTTCAGGAGGATAATGAGAAGTTTCTCGCCAAGCTAA
abcg1_10 301- CAGATGAAC ____AGGAGGATAATGAGAAGTTTCTCGCCAAGCTAA
abcg1_11 301- CAGATGAACTGGGTTTCTCGCCAAGCTAAGATCCCATGATGAGAAGTTTCTCGCCAAGCTAA



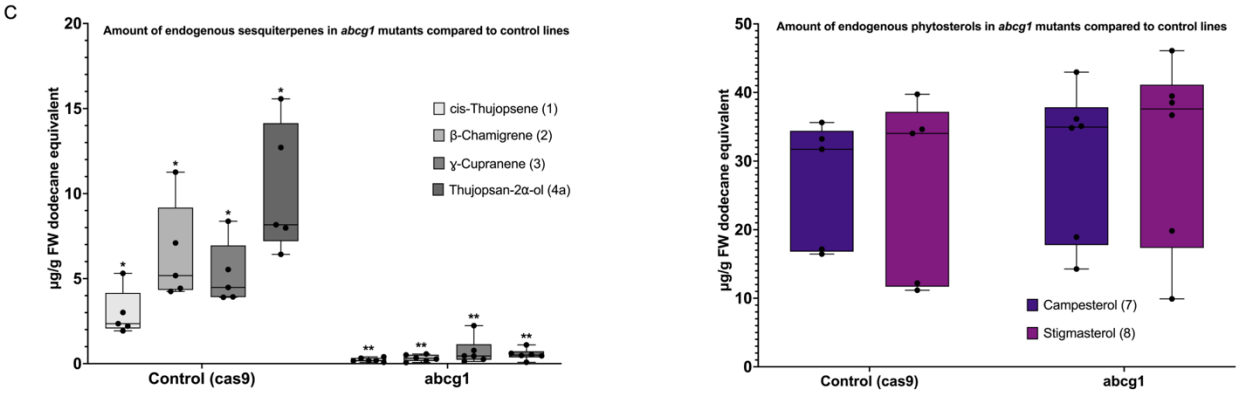


Figure 5. Analysis of endogenous terpene profiles in CRISPR-generated abcg1 mutants compared to Cas9 controls.
(A) DNA sequence alignment of the *abcg1* genomic region in wild-type (WT) and CRISPR-generated *abcg1* mutant lines. Dashes indicate deletions, red sequences represent insertions, and purple nucleotides correspond to premature stop codons introduced by the deletions and/or insertions. (B) TICs of terpene extracts from a Cas9 control (upper chromatogram) and abcg1 mutant line (lower chromatogram). Peaks corresponding to quantified compounds are labeled: sesquiterpenes (1: cis-thujopsene, 2: β-chamigrene, 3: γ-cuprenene, 4a: thujopsan-2α-ol), fatty acids/diterpenoids (5: neophytadiene, 6: phytol, a: (*Z*)-1,3-phytadiene, b: (*E*)-1,3-phytadiene), and phytosterols (7: campesterol, 8: stigmasterol). (C) Quantification of endogenous terpene levels in *abcg1* mutants (n=6) compared to Cas9 controls (n=5). The left graph shows the amounts of the four major sesquiterpenes (1 to 4a) labelled in panel B. The right graph displays the levels of phytosterol (7 and 8) labelled in panel B. Box plots display individual data points, with bars representing the median and interquartile range. Statistical significances are denoted by asterisks (one-way ANOVA with Dunnett's multiple comparison tests ($p < 0.05$)).

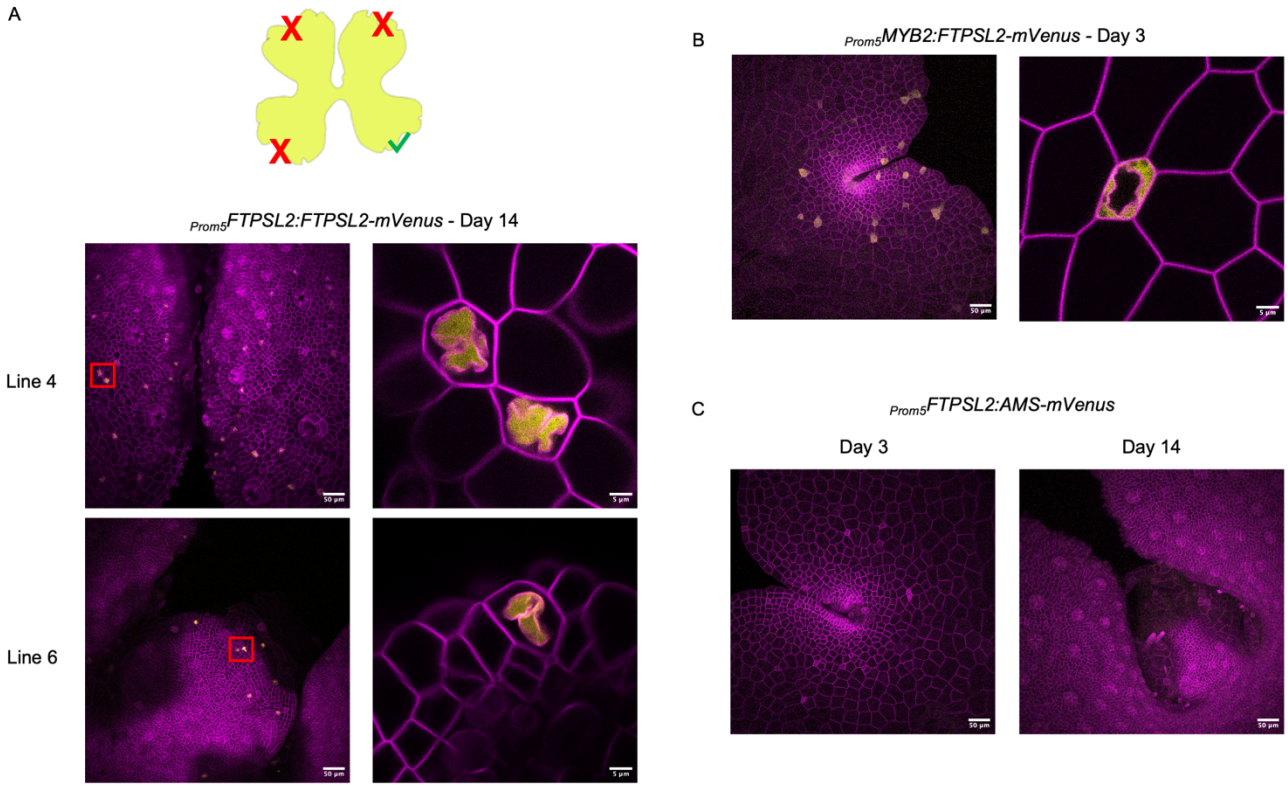


Figure 6. Localization of the endogenous sesquiterpene synthase FTPSL2 and its promoter-driven constructs in *Marchantia polymorpha*.
(A) Confocal imaging of the translational reporter *Prom5FTPSL2:FTPSL2-mVenus* in two independent 14-day-old plants (lines 4 and 6). Left panels show the meristematic region (scale bar: 50 µm), with red squares indicating the areas magnified in the right panels (scale bar: 5 µm). An illustration of a 14-day-old *Marchantia* thallus (adapted from Marchantia.info) highlights that the mVenus signal was detected in only one of the four thallus lobes (green checkmark) and absent in the others (red X). (B) Subcellular localization of FTPSL2-mVenus driven by the oil-body-specific MYB2 promoter. The left panel shows a broader region of the meristem (scale bar: 50 µm), while the right panel magnifies an oil body cell (scale bar: 5 µm). (C) Expression of AMS-mVenus driven by the FTPSL2 promoter in 3-day-old (left) and 14-day-old plants (right) revealed no detectable mVenus fluorescence.

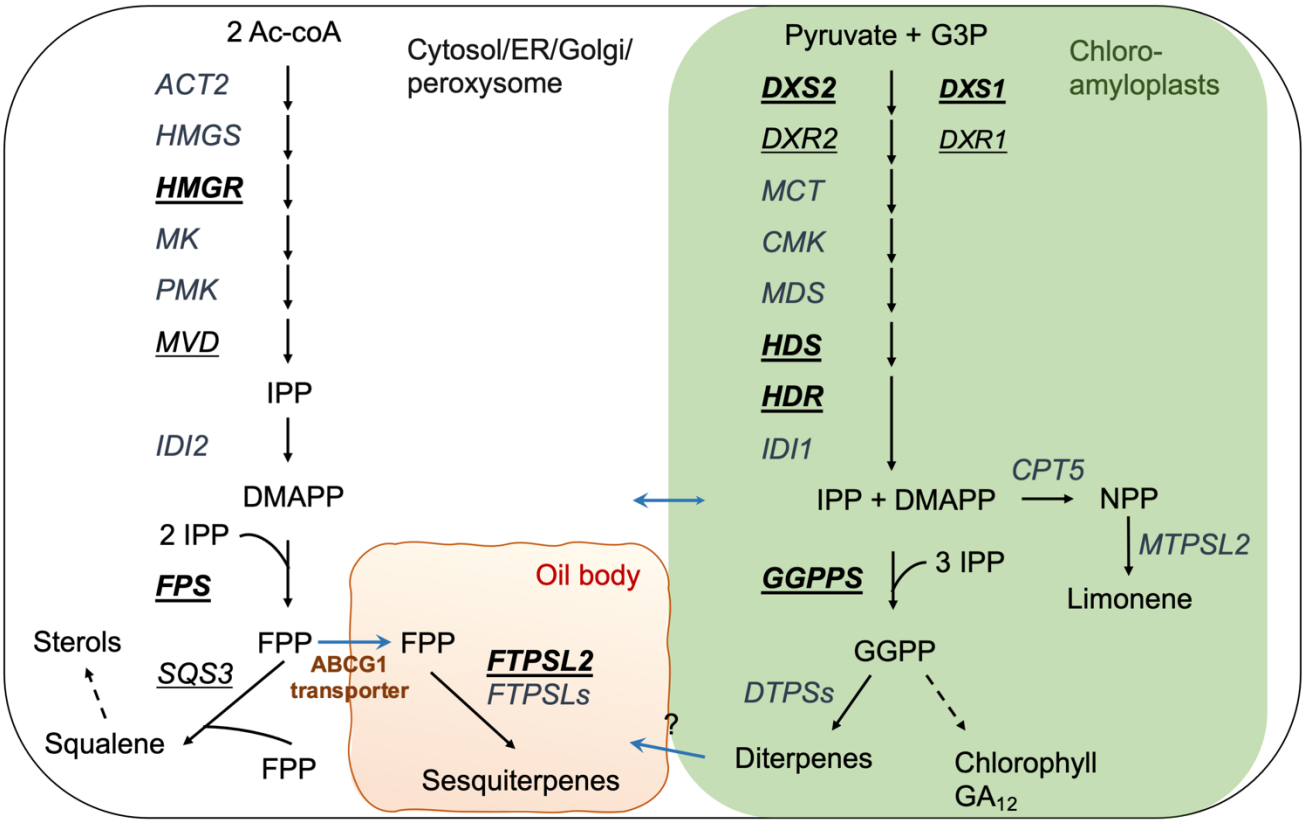


Figure 7. Overview of the isoprenoid precursor pathways in *Marchantia polymorpha* oil body cells based on previous works and findings from this study.
Biosynthetic scheme illustrating the mevalonate (MVA) and methylerythritol phosphate (MEP) pathways in *Marchantia polymorpha*. Enzymes shown in bold and underlined were localized using translational reporters confirming their subcellular localization. Enzymes shown in underlined text only were analyzed with promoter expression constructs, indicating activity in oil body cells. Key precursor flows toward sterols, diterpenes, and other metabolites are also indicated. Details of enzymatic steps and annotations can be found in Table 1 and Table S1. Abbreviations: Ac-CoA, acetyl coenzyme A; G3P, glyceraldehyde-3-phosphate.