Production of human papillomavirus type 16 virus-like particles in Physcomitrella photobioreactors

Paul Alexander Niederau[1], Maria Caroline Weilguny[1,2], Sarah Chamas[1], Caitlin Elizabeth Turney[1], Juliana Parsons[1], Marta Rodríguez-Franco[3], Sebastian N.W. Hoernstein[1], Eva L. Decker[1], Henrik Toft Simonsen[2,4], Ralf Reski[1,5,6 *]

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

[2] Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

[3] Cell Biology, Faculty of Biology, University of Freiburg, Schänzlestr. 1, 79104 Freiburg, Germany

[4] Laboratoire de Biotechnologies Végétales appliquées aux Plantes Aromatiques et Médicinales, CNRS, UMR 5079, Université Jean Monnet, 42023 Saint-Etienne Cedex 2, France

[5] Signalling Research Centre BIOSS and CIBSS, University of Freiburg, Schänzlestr. 18, 79104 Freiburg, Germany

[6] Cluster of Excellence *liv*MatS @ FIT – Freiburg Centre for Interactive Materials and Bioinspired Technologies, University of Freiburg, Georges-Köhler-Allee 105, 79110 Freiburg, Germany

*Corresponding author: ralf.reski@biologie.uni-freiburg.de

ORCID:

P.A.N.	0009-0004-0194-8570
M.C.W.	0009-0009-9440-6872
S.C.	0009-0007-5704-5639
C.E.T.	0009-0001-3113-4648
J.P.	0000-0001-6261-2342
M.R.F.	0000-0003-1183-2075
S.N.W.H.	0000-0002-2095-689X
E.L.D.	0000-0002-9151-1361
H.T.S.	0000-0003-3070-807X
R.R.	0000-0002-5496-6711

Author E-mails:

P.A.N.	alexander.niederau@biologie.uni-freiburg.de
M.C.W.	mariaweilguny@gmail.com
S.C.	sarah-chamas@web.de
C.E.T.	caitlin.turney@email.uni-freiburg.de
J.P.	juliana.parsons@biologie.uni-freiburg.de
M.R.F.	marta.rodriguez@biologie.uni-freiburg.de
S.N.W.H.	sebastian.hoernstein@biologie.uni-freiburg.de
E.L.D.	eva.decker@biologie.uni-freiburg.de
H.T.S.	henrik.toft.simonsen@univ-st-etienne.fr
R.R.	ralf.reski@biologie.uni-freiburg.de

Abstract

Human papillomavirus (HPV) infections are the main causal agent of cervical and other anogenital cancers. Modern HPV vaccines are based on virus-like particles (VLPs), which are self-assembling nanoparticles composed of viral structural proteins. VLPs mimic native virions but lack viral DNA, rendering them non-infectious while retaining high immunogenicity. Despite available vaccines, the majority of cervical cancer deaths are due to HPV infections, particularly in low- and middle-income countries where inequalities in access to vaccination, screening, and treatment persist. While current HPV vaccines are produced in yeast or insect cells, alternative production platforms are explored, focusing on cost-effectiveness and scalability. Here, we used the moss plant Physcomitrella (*Physcomitrium patens*) to produce HPV-16 VLPs *via* expression of the major structural protein L1. We characterized Physcomitrella chloroplast transit peptides, which we used for targeting of L1 to chloroplasts, leading to higher recombinant protein accumulation compared to nuclear or cytoplasmic localization. We confirmed subcellular localization with confocal laser scanning microscopy and found L1 to accumulate within the chloroplast stroma. Production in 5 L-photobioreactors yielded more than 0.3 mg L1 per g fresh weight. We established a purification protocol for moss-produced L1 using a combination of ammonium sulphate precipitation and cation exchange chromatography. Purified samples were subjected to a controlled dis- and reassembly, ultimately yielding fully assembled HPV-16 L1 VLPs. This is the first report of production, purification and assembly of VLPs in a non-vascular plant.

Keywords: bryophyte, bioreactor, molecular farming, nanoparticle, plant biotechnology, vaccine candidate

Key message: First production of virus-like particles as a vaccine candidate in a non-vascular plant.

Word count: 6809

68 # Introduction

69 Human papillomavirus (HPV) infections are a major source of cancer, causing approximately
70 690,000 cases annually (zur Hausen et al., 1981; zur Hausen, 2002; de Martel et al., 2020). The most
71 prevalent type is cervical cancer, accounting for about 80% of HPV-attributable cancers. Other HPV-
72 related cancers include anogenital and oropharyngeal cancers (Wu et al., 2024). The burden of HPV
73 infection is not evenly distributed, with the highest prevalence in low- and middle-income countries, as a
74 result of inequities in access to vaccination, screening, and treatment services (World Health Organisation,
75 2024a).

76 HPV belongs to the *Papillomaviridae*, comprising nearly 200 different types, which are commonly
77 transmitted sexually (Ljubojevic & Skerlev, 2014). The majority of HPV-attributable cancers are linked to
78 two types in particular, namely HPV-16 and HPV-18 (de Martel et al., 2020; Wu et al., 2024). HPV is a non-
79 enveloped double stranded DNA virus of 8000 bp, encoding six early regulatory proteins (E1, E2, E4 – E7)
80 and two late structural proteins (L1, L2) (Roden & Stern, 2018). L1 proteins multimerize into pentamers,
81 the capsomers, and L1 capsomers and L2 monomers are imported into the nucleus *via* nuclear localization
82 signals (NLS), where HPV virions assemble (Cerqueira & Schiller, 2017). While the role of L2 in virion
83 assembly differs from type to type, L1 can self-assemble into virus-like particles (VLPs) in the absence of
84 L2 (Kirnbauer et al., 1992). The final HPV-16 virion is composed of 360 L1 monomers, organized in 72
85 pentamers, and 72 L2 monomers and is reinforced by inter-pentameric disulfide bonds (Baker et al., 1991;
86 Buck et al., 2005). Because the viral life cycle depends on differentiated epithelial cells, it is difficult to
87 obtain large quantities of virions. Therefore, today's HPV vaccines are based on L1 VLPs.

88 VLPs are nanoparticles consisting solely of viral structural proteins (Marsian & Lomonossoff, 2016;
89 Nooraei et al., 2021). Therefore, they cannot replicate and infect cells. These nanoparticles resemble
90 native viruses in size, shape, and surface structure, presenting highly repetitive immunogenic epitopes.
91 Although VLP-based HPV vaccines are currently produced in microbes (World Health Organization,
92 2024b), plant-based expression systems are gaining attention in this respect (Scotti & Rybicki, 2013;
93 Rybicki, 2020). In *Nicotiana benthamiana*, expression of L1 and L2 in combination with a reporter gene
94 generated pseudovirions (PsVs), which could substitute mammalian cell culture-derived PsVs in
95 neutralisation assays (Lamprecht et al., 2016; Adams et al., 2023). Alternatively, Tobacco Mosaic Virus
96 VLPs produced in *N. benthamiana* were used for encapsidation and targeted delivery of mRNA encoding
97 oncogenic E6, a crucial target in therapeutic HPV vaccine development (Karan et al., 2024).

98 While first attempts to produce L1-based VLPs in plants had low yields (Biemelt et al., 2003;
99 Varsani et al., 2003), considerable improvements have been achieved, e.g. by codon optimization and
100 subcellular targeting strategies in transiently transformed tobacco (Maclean et al., 2007). A human codon-
101 optimized version (62% GC content) resulted in highest yields as compared to *N. benthamiana* optimized
102 (49% GC) and the native sequence (37% GC), presumably not due to the GC content but due to inhibitory
103 5' sequences, which act on transcription and mRNA processing (Maclean et al., 2007; Hitzeroth et al.,
104 2018). Further, these authors found chloroplast-targeted L1 yield highest at 0.5 mg / g fresh weight (FW)
105 followed by cytoplasmic L1 (without the C-terminal 22 aa NLS), whereas L1 targeted to the ER was not
106 detectable. Similarly, Zahin et al. (2016) reported 0.25 mg L1 / g FW *via* chloroplast targeting while leaving
107 the NLS intact. Although tobacco remains a prominent option, other plants are increasingly explored for
108 the efficient production of biopharmaceuticals.

109 The moss Physcomitrella has emerged as an alternative for biopharmaceutical production,
110 offering several advantages over traditional systems (Reski et al., 2015; Reski, 2018; Decker & Reski, 2020).
111 Its haploid dominant phase and high rate of homologous recombination enable precise gene targeting
112 (Strepp et al., 1998; Wiedemann et al., 2018; Rempfer et al., 2022). This unique feature facilitated the
113 modification of the N-glycosylation pathway, eliminating plant-specific modifications, such as β1,2-
114 xylosylation, α1,3-fucosylation, and Lewis A epitopes (Koprivova et al., 2004; Parsons et al., 2012; Decker
115 et al., 2014). Moreover, enzymes necessary for sialic acid synthesis, activation, and linkage to protein N-
116 glycans were introduced (Bohlender et al., 2020). These modifications have resulted in more
117 homogeneous and human-like N-glycan structures on recombinant proteins. The potential of
118 Physcomitrella as a biopharmaceutical production platform is demonstrated by several products in
119 development. For example, moss-produced α-galactosidase A has successfully completed clinical phase
120 Ib (Shen et al., 2016; Hennermann et al., 2019). Other moss-produced proteins include human factor H
121 (FH) and synthetic FH-related multitarget regulators MFHR1 and MFHR13, which are potential treatments
122 for complement-related disorders (Büttner-Mainik et al., 2011; Michelfelder et al., 2017, 2018; Top et al.,
123 2019; Ruiz-Molina et al., 2022a). Importantly, Physcomitrella cultivation in photobioreactors is tightly
124 controlled (Hohe et al., 2002), and thus compliant with Good Manufacturing Practice (GMP) standards,
125 ensuring high-quality production of recombinant proteins (Reski et al., 2018; Decker & Reski, 2020). Plants
126 in general (Rosales-Mendoza et al., 2016) and Physcomitrella in particular (Rosales-Mendoza et al., 2014)
127 are considered promising production hosts for innovative vaccines. The first moss-produced vaccine
128 candidate, an ENV-derived multi-epitope HIV chimeric protein, was immunogenic in mice (Orellana-

129 Escobedo et al., 2015). However, until now, there was no record of VLP-production in Physcomitrella, or
130 any other non-vascular plant.

131 Here, we analysed the expression of HPV-16 L1 in Physcomitrella. We characterized five
132 Physcomitrella chloroplast transit peptides and investigated the effect of subcellular L1 localization on
133 recombinant protein yields. We further established L1 production in photobioreactors and subsequent
134 purification of the product. Finally, the assembly of moss-produced HPV-16 L1 to VLPs was demonstrated
135 with transmission electron microscopy.

136

137 Materials and methods

138 Plant material

139 Physcomitrella (new species name: *Physcomitrium patens* (Hedw.) Mitt.; IMSC acc. no. 41269),
140 was cultivated according to Reski & Abel (1985) and Decker et al. (2015).

141

142 Optimization of HPV-16 L1 coding sequence

143 The CDS of h(human codon-optimized)L1 (GenBank acc. no. DQ067889, Maclean et al., 2007) was
144 used for optimization using physCO (https://www.plant-biotech.uni-freiburg.de/tools/physco/, Top et al.,
145 2021). Putative alternative splicing motifs AGGT were identified *via* search function. MicroRNA-binding
146 sites were predicted *via* psRNATarget with the library of 280 published Physcomitrella microRNAs and the
147 default settings *Schema V2* (2017 release) (http://plantgrn.noble.org/psRNATarget/, Dai & Zhao, 2011;
148 Dai et al., 2018). Alternative splice sites and microRNA-binding sites were removed and substituted by
149 synonymous codons based on http://www.kazusa.or.jp/codon/. Sequence editing was done in Benchling
150 (https://www.benchling.com/).

151

152 Expression vectors

153 Chloroplast transit peptides (CTPs) were selected from a list of proteins with experimentally
154 confirmed cleavage of a predicted CTP (Hoernstein et al., 2024). CTP-eGFP vectors were constructed based
155 on a vector containing an eGFP CDS with an N-terminal linker (sequence: GGGGGA) under control of the
156 PpActin5 promoter (Weise et al., 2006) and nos terminator (Hoernstein et al., 2023). CTP candidates were
157 amplified from Physcomitrella cDNA and inserted at the N-terminus of the linker using Gibson Assembly

158 (Gibson et al., 2009). The Citrine-L1 fusion constructs are based on a vector containing the PpActin5
159 promoter, nos terminator and citrine CDS (Wiedemann et al., 2018). The CDS of pL1, pL1Δ22 and CTP-pL1
160 were amplified using Phusion™ polymerase (Thermo Fisher Scientific, Waltham, USA), cloned into the
161 pJet1.2 backbone (Thermo Fisher Scientific) and ligated into the expression vector using T4 Ligase (Thermo
162 Fisher Scientific), *via Xho*I and *BamH*I sites. The pL1 expression vectors were constructed based on a vector
163 containing the PpActin5 promoter, nos terminator and hpt cassette (Ruiz-Molina et al., 2022a). The pL1
164 sequence was synthesized by Genewiz (Leipzig, Germany). For constructs pL1 and pL1Δ22, the pL1 CDS
165 was amplified using Phusion™ polymerase (Thermo Fisher Scientific) and cloned as described above. For
166 construct CTP-pL1, the pL1 and CTPc5 CDS were amplified and inserted into the backbone using Gibson
167 Assembly. Vectors were sequenced by Eurofins Genomics (Ebersberg, Germany). Vector maps are
168 compiled in **Supplementary Figure S1**. Primers were synthesized by Eurofins Genomics and are compiled
169 in **Supplementary Table S1**.
170

171 ## Protoplast preparation and transformation

172 Physcomitrella protoplasts were prepared and transformed based on existing protocols (Hohe &
173 Reski, 2002; Schween et al., 2003; Hohe et al., 2004; Decker et al., 2015). Transformed protoplasts were
174 incubated at 22 °C for 24 h in the dark and subsequently for 72 h at a 16/8-h light/dark photoperiod and
175 light intensity of 50–70 µmol m^{-2} s^{-1} before plating on KnopME + 25 mg/L hygromycin agar plates covered
176 with cellophane for two weeks. Surviving colonies were transferred to non-selective conditions for two
177 weeks before a second selection on KnopME + 50 mg/L hygromycin.
178

179 ## Fluorescence microscopy

180 Protoplasts were analysed 2-3 days after transfection with an AxioCam MRc5 camera on an
181 Axioplan2 stereo microscope (Zeiss, Oberkochen, Germany). Pictures were taken with a 40x objective or
182 a 63x objective with immersion oil. A FITC filter was used for eGFP, a rhodamine filter for chlorophyll
183 autofluorescence, and a FITC_LP filter for the combination of the two. Images were edited with GIMP
184 2.10.22 and ImageJ. For confocal microscopy, an AXIO Observer.Z1 Inverted Fluorescence Motorized
185 Phase Contrast Microscope Pred 7 with the objective Plan-Apochromat 63x/1.4 Oil DIC M27 (Zeiss) was
186 used. EGFP signals were detected at 488 / 509 nm, and chlorophyll autofluorescence at 587 / 610 nm.
187 Images were edited in ZEN 3.5 (Blue edition, ZEN Lite), ImageJ, and GIMP 2.10.22.

188 For analysis of subcellular targeting, an inverted confocal laser scanning microscope LSM 880 with
189 the objective LD LCI Plan-Apochromat 40x/1.2 H_2O autocorr and FastAiryScan mode (Zeiss) was used.
190 Citrine signals were detected at 488 / 516 nm, and chlorophyll autofluorescence at 561 / 595 nm. Images
191 were edited in ZEN 3.9 (Blue edition) and processed *via* deconvolution using default setting and
192 orthogonal projection using maximum intensity projection.
193

194 ## Biomass accumulation

195 Flasks containing 30 mL KnopME medium were inoculated at a density of 60 mg DW/L and
196 cultured for 4 weeks at the conditions described above. Biomass DW was measured by filtering 3x 10 mL
197 of culture through miracloth and subsequent drying at 105 °C for 2 hours. Biomass accumulation was
198 analysed by linear regression analysis. Slopes were compared using ANCOVA and post hoc multiple
199 comparisons test (TukeyHSD). The analysis was conducted in RStudio (version 4.2.2) and GraphPad Prism
200 (version 8.1).
201

202 ## Bioreactor operation

203 Stirred tank bioreactors (Applikon, Schiedam, Netherlands) were inoculated at 60 mg DW/L moss
204 density in 5 L KnopME. The light intensity was set to 160 µmol m^{-2} s^{-1} and increased to 350 µmol m^{-2} s^{-1}
205 after 4 days. At day 4, auxin 1-Naphtalene acetic acid (NAA, Sigma, St. Louis, USA) dissolved in 0.5 M KOH
206 was added to a final concentration of 5 µM. Throughout the bioreactor run, the pH was kept at pH 5.8,
207 temperature at 22 °C, and aeration at 0.3 vvm with 2% CO_2. The culture was continuously agitated with a
208 pitched 3 blade impeller at 500 rpm.
209

210 ## Protein extraction, purification and VLP assembly

211 For small scale extraction, 30 – 60 mg FW of vacuum-filtered moss material was combined with a
212 glass and a steel bead (each 3 mm in diameter) in a 2 mL reaction tube, frozen in liquid nitrogen and the
213 tissue lysed using a tissue lyser for 2 min at 28 Hz. Homogenized material was resuspended in PBS (2.7
214 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.137 M NaCl, pH 7.2), 0.01% Tween20, and protease inhibitor
215 (P9599, Sigma) in a ratio 3 mL/g. Samples were placed in an ice-cold sonication bath for 15 min,
216 centrifuged for 30 min at 14,000 x g and 4 °C, and the supernatant taken off for further analysis.
217 For large scale protein extraction, 5 g FW of vacuum-filtered moss material were frozen in liquid
218 nitrogen and ground to a fine powder. The material was combined with 22 mL PBS and 0.01% Tween20,

219 50 mM DTT and 100 µL protease inhibitor. The mixture was homogenized using an ULTRA-TURRAX (IKA,
220 Staufen, Germany) at 10,000 rpm for 10 min on ice. Next, the extract was treated with a Q500 sonicator
221 using a CL-334 converter (QSonica, Newtown, USA) at amplitude 55%, 10s on & 40s off for 20 min at 8 °C.
222 Debris was removed by two centrifugations at 4500 x g for 10 min and 20,000 x g for 20 min, both at 4 °C.
223 The sample was combined with solid $(NH_4)_2SO_4$ to a final concentration of 45% and proteins precipitated
224 under stirring for 1 h at 8 °C. Precipitated proteins were pelleted by centrifugation at 12,000 x g for 10
225 min at 8 °C. Pellets were dissolved in 4 mL CaEx chromatography binding buffer (2.7 mM KCl, 10 mM
226 Na_2HPO_4, 1.8 mM KH_2PO_4, 0.068 M NaCl, 0.01% Tween 20, 5% Glycerol, 5 mM DTT, pH 6.8) and dialysed
227 (20k MW cut off, 88405, Thermo Fisher Scientific) overnight against the same buffer at 8 °C. The sample
228 was diluted to a final volume of 10 mL using binding buffer, and passed through a 0.22 µm filter (Roth,
229 Karlsruhe, Germany) before loading onto a 1 mL HiTrap SP HP cation exchange column (Cytiva,
230 Marlborough, USA), connected to an ÄKTA™ start (Cytiva). The column was washed with 5 column
231 volumes (CV) of 0.3 M NaCl (based on binding and elution buffer) and L1 eluted *via* 4 CV of elution buffer
232 (2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 1 M NaCl, 0.01% Tween 20, 5% Glycerol, 5 mM DTT, pH
233 6.8)

234 Eluted fractions containing L1 were pooled and dialysed against disassembly buffer (22.7 mM KCl,
235 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.068 M NaCl, 0.05% Tween20, 5 mM DTT, pH 8.5) overnight at 8 °C
236 with 3x buffer exchange. Samples were concentrated 5x using a vacuum concentrator (Concentrator plus,
237 Eppendorf, Hamburg, Germany). For reassembly, samples were dialysed against reassembly buffer (22.7
238 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 1 M NaCl, 0.05% Tween20, pH 6) overnight at 8 °C with 3x
239 buffer exchange.

240

241 SDS-PAGE and Western blot

242 SDS-PAGE and Western blots were performed based on Bohlender et al. (2022). Two systems
243 were used for protein separation, either 7.5% polyacrylamide gels (Mini-PROTEAN® TGX™ Precast Gels,
244 Rio-Rad, Munich, Germany) in TGS buffer at 120 V, or 4 – 12% Bis-Tris gels (mPAGE™ Precast Gels, Merck,
245 Darmstadt, Germany) in MOPS buffer at 150 V. As first antibody served 1:5000 diluted CamVir1 (Thermo
246 Fisher Scientific), as second antibody peroxidase-linked rabbit anti-mouse secondary antibody (Cytiva)
247 diluted 1:12,500. For dot blots, protein samples were blotted on a PVDF membrane (Cytiva) using a
248 filtration manifold kit (SRC-96/1, Schleicher&Schuell, Dassel, Germany). Antibody H16.V5

249 (AntibodySystems, Schiltigheim, France) was diluted 1:5000. Dot blots were developed as described for
250 Western blots. Yeast-produced HPV-16 L1 (Abcam, Cambridge, UK) was used as a positive control.
251

252 ## L1 quantification

253 The ELISA protocol was based on Studentsov et al. (2002). In short, 96-well microtiter plates
254 (Greiner, Kremsmünster, Austria) were coated with 1:500 CamVir1 antibody (Thermo Fisher Scientific) in
255 PBS overnight at 8 °C. Wells were washed twice with washing buffer (1x PBS, 0.05% Tween20) and blocked
256 with blocking buffer (1x PBS, 1% polyvinyl alcohol MW 47,000, 0.05% Tween20) for 3 h at 37 °C. After 3x
257 washing, samples were diluted in blocking buffer and incubated in the wells for 1.5 h at 37 °C. Wells were
258 washed 3x before adding 1:1000 diluted polyclonal anti-HPV-16 L1 antibody from rabbit (Cusabio,
259 Houston, USA) for 1 h at 37 °C. After 5x washing, wells were incubated with 1:1000 diluted anti-rabbit
260 HRP-conjugated antibody NA934 (Cytiva) for 1 h at 37°C. Wells were washed 3x before adding 100 µL TMB
261 substrate (Thermo Fisher Scientific). The reaction was stopped with 50 µL of 2 M H_2SO_4. Absorbance was
262 measured at 450 / 595 nm for reference. The signal of a wildtype sample was subtracted to account for
263 background signal.
264

265 ## Transmission electron microscopy

266 TEM protocols are based on Abel et al. (1989). Negative staining of purified VLPs and moss extracts
267 was performed using 300 Mesh glow discharged formvar/carbon-coated copper grids (Electron
268 Microscopy Sciences, Hatfield, USA). Five µL of sample were applied on a grid and incubated for 5 min at
269 RT. Grids were washed 3x with water and stained with 2% (w/v) uranyl acetate. After 30 sec, uranyl
270 acetate was removed with a filter paper, grids were air dried, and inspected with a Hitachi HT7800 TEM
271 (Tokyo, Japan) coupled to a Xarosa CMOS camera (Emsis, Münster, Germany).

272 # Results

273 ## Adaptation of L1 CDS for Physcomitrella-optimized expression

274 The CDS of HPV-16 L1 was optimized for expression in Physcomitrella. The goal was to adapt the
275 CDS regarding codon usage and preventing potential degradation through microRNAs and splicing. A
276 human codon-optimized version (hL1, **Supplementary Figure S2**) was used as a basis for optimization
277 since it had led to higher yields than plant-optimized versions or the native viral sequence in *Nicotiana*
278 *benthamiana* (Maclean et al., 2007).

279 For codon optimization, physCO (Top et al., 2021) was used, which changed nine out of 505
280 codons and slightly increased the GC content from 62.25 % to 62.85 % (**Supplementary Figure S2**). Next,
281 the CDS was scanned for the motif AGGT, which is the most common motif at exon-intron borders for
282 alternative splicing in Physcomitrella (Top et al., 2021). Transformation of a heterologous cDNA containing
283 the AGGT motif led to alternative splicing in Physcomitrella, a process referred to as heterosplicing (Top
284 et al., 2021). Scanning the physCO-optimized hL1 sequence yielded 5 AGGT-motifs which were replaced
285 by synonymous codons (**Supplementary Figure S2**). MicroRNAs bind complementary sequences in the
286 mRNA and thereby cause inhibition of translation or degradation of mRNA (Jones-Rhoades et al., 2006;
287 Khraiwesh et al., 2010). One microRNA binding site was predicted within the CDS using psRNATarget (Dai
288 & Zhao, 2011; Dai et al., 2018). Therefore, we changed three bases within the predicted binding region
289 using synonymous codon substitutions. The final Physcomitrella-optimized CDS was named pL1
290 (**Supplementary Figure S2**) and differs from hL1 by 17 out of 1515 nucleotides without changing the aa
291 sequence.
292

293 ## Characterization of chloroplast transit peptides

294 Different constructs were designed to target L1 to specific cellular compartments, including
295 chloroplasts. Fusing chloroplast transit peptides (CTPs) to the L1 CDS led to increased protein yields in
296 tobacco as compared to constructs lacking CTPs (Maclean et al., 2007; Zahin et al., 2016). To assess
297 whether this is true for L1 production in Physcomitrella, a selection of Physcomitrella-derived CTPs was
298 characterized.
299 Initially, the CTP of FtsZ1-1 (Gremillon et al., 2007) was considered. However, cleavage site
300 analysis *via* TargetP-2.0 (Armenteros et al., 2019) predicted 37 aa of this CTP to remain at the N-terminus
301 of L1 after chloroplast import. This CTP was rejected because the additional aa could change L1 properties
302 and hamper VLP assembly. A list of alternative CTPs was derived from the Physcomitrella N-terminome
303 (Hoernstein et al., 2024) (**Supplementary Table S2**). Here, experimentally detected protein N-termini
304 were compared with TargetP-2.0 predictions for cleavable N-terminal targeting sequences. Five CTPs were
305 selected where an exact match between prediction and experimentally validated N-terminus was
306 observed across different experiments (Hoernstein al., 2024), namely CTPc5, CTPc10.1, CTPc10.2, CTPc21,
307 CTPc22 (**Table 1**). Their sequences were fused *in silico* to the N-terminus of L1 and analysed with TargetP-
308 2.0 to validate the correct cleavage of the fusion.

309 Our *in-silico* analysis indicated CTPc22 to end with a methionine, which could introduce an
310 alternative translation initiation site, resulting in translation of L1 without the CTP (**Table 1**). Therefore,
311 an alternative version lacking the C-terminal methionine (CTPc22noMet) was included. The likelihoods for
312 chloroplast localization were above 70 % for all candidates except CTPc10.1 (**Table 1**), which was rejected.

313 **Table 1: *In-silico* analysis of Physcomitrella-derived chloroplast transit peptides.** CTPs were selected based on mass
314 spectrometry datasets (Hoernstein et al., 2024). *In-silico* predictions of likelihood of chloroplast localization target
315 sequence cleavage were performed with TargetP-2.0.

Name of CTP	Originating from protein model	Predicted cleavage of CTP-L1	Length (aa)	Likelihood of GFP chloroplast localization	Likelihood of L1 chloroplast localization
FtsZ1-1	Pp3c22_4940V3.1	37 aa remain on N-terminus	84	0.95	0.92
CTPc5	Pp3c5_1370V3.1	-	74	0.987	0.943
CTPc10.1	Pp3c10_10490V3.1	-	63	0.456	0.229
CTPc10.2	Pp3c10_2940V3.1	-	70	0.908	0.845
CTPc21	Pp3c21_9980V3.1	-	43	0.956	0.891
CTPc22	Pp3c22_5470V3.1	-	62	0.939	0.739
CTPc22noMet	Pp3c22_5470V3.1	-	61	0.913	0.727

316

317 The remaining five CTPs (CTPc5, CTPc10.2, CTPc21, CTPc22, CTPc22noMet) were fused to the CDS
318 of eGFP (**Supplementary Figure S1**) and expressed in Physcomitrella protoplasts. Analysis *via* fluorescence
319 microscopy confirmed localization of eGFP signals in chloroplasts for CTPc5, CTPc10.2, CTPc22, and
320 CTPc22noMet, but not for CTPc21 (**Supplementary Figure S3**). Of the four remaining CTPs, CTPc5 had the
321 highest likelihood for plastid localization (**Table 1**) and were examined *via* confocal microscopy
322 (**Supplementary Figure S3**). Finally, we selected CTPc5 for the chloroplast-targeting of L1.
323

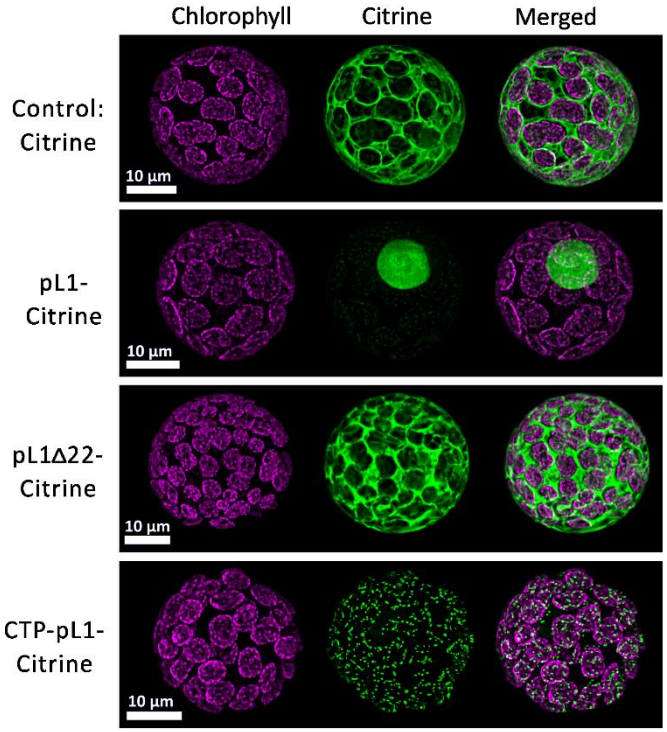
324 Targeting of L1 to nuclei and chloroplasts

325 To evaluate the impact of L1 localization on the yield, we generated three variants (pL1, pL1Δ22,
326 CTP-pL1). pL1 contains the full-length pL1 CDS, which includes the native HPV-16 NLS (Cerqueira & Schiller,

327 2017). For pL1Δ22 the NLS was removed. In CTP-pL1 the Physcomitrella CTPc5 was fused to the 5' of the
328 pL1 CDS, including the NLS.

329 After generation of these CDS variants, we aimed to evaluate L1 targeting to subcellular
330 compartments. The CDS of pL1, pL1Δ22 and CTP-pL1 were fused to the 5' end of Citrine CDS
331 (**Supplementary Figure S1**). Physcomitrella protoplasts were transfected with the plasmids and Citrine
332 localization was analysed using fluorescence microscopy. A vector encoding Citrine without any L1 variant
333 fusion served as a control.

334 Cells transformed with the control vector exhibited predominantly cytoplasmic Citrine signals,
335 whereas diffuse signals also occurred in nuclei (**Figure 1**). For pL1-Citrine, the signal localized to nuclei,
336 demonstrating that the HPV-16 NLS is functional in Physcomitrella. This is further supported by the fact
337 that removal of the NLS in pL1Δ22-Citrine results in cytoplasmic localization. For CTP-pL1-Citrine, the
338 signal is predominantly in chloroplasts, but is not evenly distributed, as seen for CTP-eGFP fusions lacking
339 L1 (**Supplementary Figure S3**). Instead, the pL1-Citrine fusion concentrates in several spots of high
340 fluorescence intensity and <1 µm in diameter. Overall, we confirmed successful L1 targeting to the
341 intended compartments.
342



343
344 **Figure 1: Subcellular localization of L1 variants.** Confocal microscopy images of Physcomitrella protoplasts transformed with
345 constructs encoding (CTP-)pL1(Δ22)-Citrine fusion proteins or Citrine alone. For each sample, chlorophyll autofluorescence in
346 magenta is seen on the left, Citrine signals in the middle and the merged signals on the right. In the control and pL1Δ22-Citrine,

347 the Citrine signals do not enter the chloroplasts and are mainly in the cytoplasm. For pL1-Citrine, the Citrine signals are in the
348 nucleus. In the CTP-pL1-Citrine sample, the Citrine signals form small spots of high intensity in the chloroplasts.
349

350 Characterization of L1 production

351 After validating the subcellular localization, pL1, pL1Δ22 and CTP-pL1 expression constructs were
352 cloned (**Supplementary Figure S1**). The linearized constructs were transferred into Physcomitrella
353 protoplasts and obtained plants were screened for productivity by Western blot and ELISA.
354 Western blot analysis revealed the L1 monomer of 56 kDa (pL1, CTP-pL1) and 54 kDa (pL1Δ22)
355 (**Figure 2**). In addition, pL1 and CTP-pL1 lines exhibited a band at approximately 54 kDa. All lines showed
356 another band at around 45 kDa, even though the signal was relatively weak in CTP lines and more
357 prominent in pL1Δ22 lines. Smaller protein isoforms can originate from degradation or DNA or RNA,
358 resulting in truncated transcripts and, subsequently, truncated L1 isoforms (Murén et al., 2009; Top et al.,
359 2021). The latter possibilities were ruled out by amplifying the L1 CDS *via* PCR from cDNA, which yielded
360 amplicons of the expected sizes, corresponding to the Western blot bands of 56 kDa (pL1, CTP-pL1) and
361 54 kDa (pL1Δ22), but not additional amplicons (**Supplementary Figure S4**). Also, two bands are visible at
362 130 – 180 kDa, of which the highest band at 180 kDa is also present in the yeast-produced L1 control. This
363 pattern of two close-by bands repeats another two times at higher molecular weight outside of the marker
364 range (**Figure 2**). ELISA analysis indicates higher yields in CTP-pL1 lines as compared to pL1 or pL1Δ22 lines.
365 In shaking flasks, the best producer is CTP-pL1 30 with 142 µg L1/g FW (**Figure 2**).
366

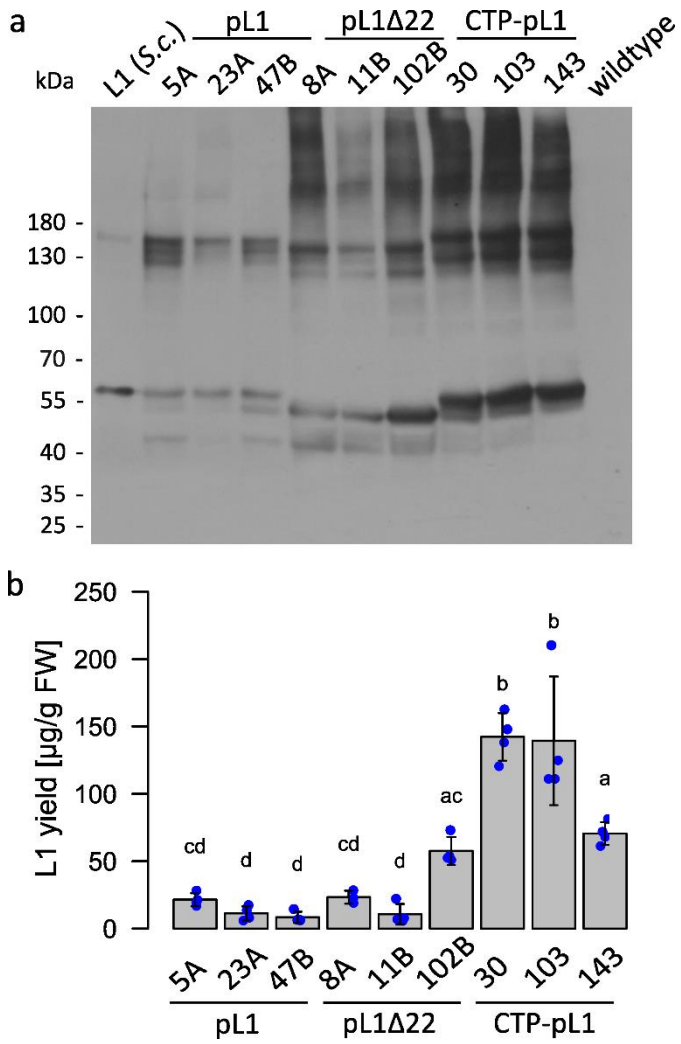


Figure 2: Analysis of HPV-16 L1 production *via* Western blot and ELISA. (**a**) For each of the constructs pL1, pL1Δ22, and CTP-pL1 three lines were tested *via* Western blot for expression of HPV-16 L1. 50 ng *S. cerevisiae*-produced L1 served as a positive control and Physcomitrella wildtype as a negative control. The proteins for the Western blot were separated on a 7.5% TGS polyacrylamide gel under reducing conditions before being transferred to a membrane and developed using CamVir1 antibody (1:5000). (**b**) Production levels of L1 were quantified *via* ELISA. Wildtype background signal was subtracted before analysis *via* one-way ANOVA and post hoc multiple comparisons test (TukeyHSD). Means and standard deviations are calculated based on n ≥ 3 technical replicates from two dilutions. Statistical significance (α = 0.05) is indicated by Compact Letter Display.

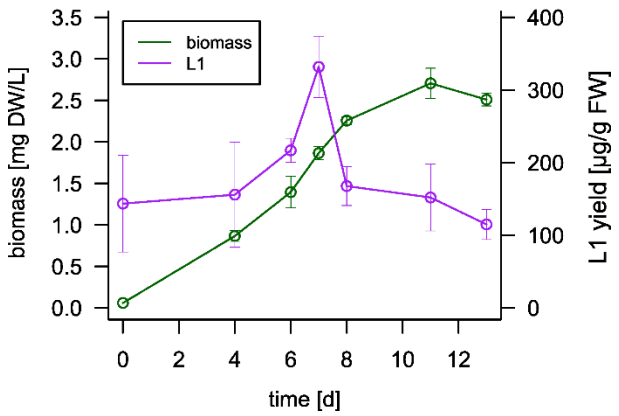
Besides the amount of recombinant protein per g FW biomass, the accumulation of biomass is important in biomanufacturing as high product yields per biomass can be trivial if biomass accumulation is slow. To study if recombinant protein amount or subcellular localization of L1 hampers plant growth, we studied the time course of biomass accumulation of the nine lines in shaking flasks for 4 weeks.

In our conditions, exponential growth could not be measured. This might be due to the culture conditions, to the rather long intervals between sampling, and/or to the intrinsic growth characteristics

382 of Physcomitrella protonema, which grows by apical cell division (Reski, 1998). Therefore, biomass density
383 was fitted using linear regression analysis model (**Supplementary Figure S5**), using the calculated slope to
384 approximate biomass accumulation rates. Comparison of the slopes between the lines pL1, pL1Δ22 and
385 CTP-pL1 showed pL1 lines accumulated biomass faster than pL1Δ22 ($p = <0.0001$) and CTP-pL1 ($p = 0.0001$)
386 lines. If this difference is due to the distinct L1 localization or to lower L1 production in pL1 lines (**Figure**
387 **2**) is unclear. Line pL1 47B showed the highest slope of 255 mg DW/L/week, as compared to line CTP-pL1
388 30 with a slope of 211 mg DW/L/week. However, this difference is negligible considering the yield of pL1
389 47B is 8 µg L1/g FW as to 142 µg L1/g FW for CTP-pL1 30 (**Figure 2**). Consequently, line CTP-pL1 30 was
390 chosen for VLP production.
391

392 ## Upscaled L1 production

393 For production of VLPs, the cultivation of line CTP-pL1 30 was upscaled to 5 L-photobioreactors.
394 The goal was to identify the best time point for harvesting material used for subsequent L1 purification.
395 Biomass increased from initially 60 mg DW/L to around 870 mg DW/L at day 4 (**Figure 3**). L1
396 production remained constant during the first 4 days, in the range of the yield observed for flasks. Upon
397 addition of the phytohormone auxin on day 4, L1 levels increased on day 6 and peaked at day 7 at 332 µg
398 / g FW, 130% higher as compared to day 0. Afterwards, L1 production dropped sharply to levels prior to
399 auxin addition. Biomass continued to accumulate until day 11, after which it dropped slightly. These
400 conditions, including day 7 as the ideal harvesting time point, were used for further L1 production.
401



402
403 **Figure 3: Analysis of biomass accumulation and L1 production of line CTP-pL1 30 in a 5 L-photobioreactor.** The bioreactor was
404 inoculated with 60 mg dry weight (DW)/L and monitored over a course of 13 days. On day 4, 5 µM NAA was added. L1 production
405 peaked at day 7 at 332 µg / g fresh weight (FW). Biomass and L1 measurements are based on n ≥ 3 technical replicates. Production
406 levels of L1 were quantified *via* ELISA.

407

408 ## Two-step purification removes bulk of cellular proteins

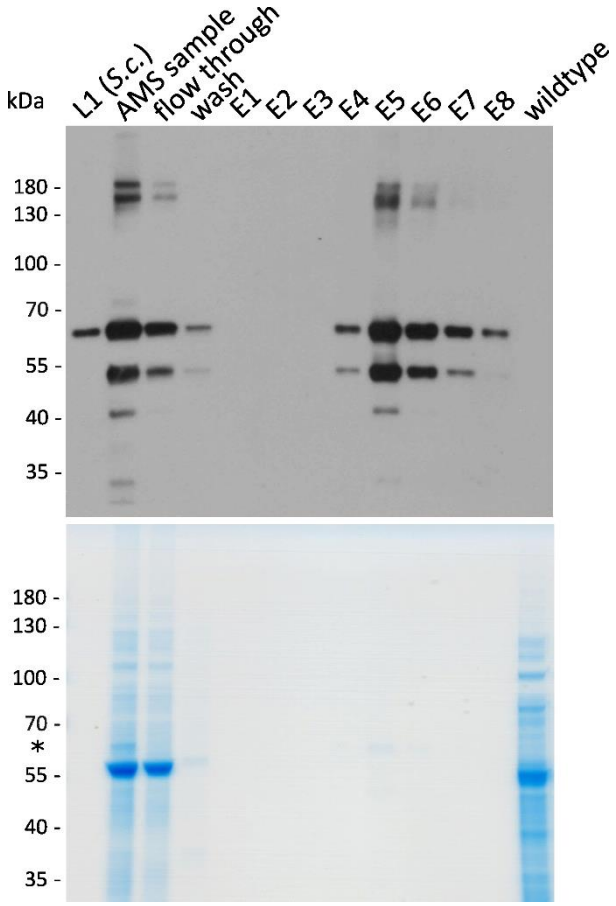
409 Purification of L1 was performed using ammonium-sulphate (AMS) precipitation, followed by
410 cation exchange (CaEx) chromatography. While AMS removes the bulk of cellular proteins, CaEx removes
411 residual proteins and concentrates the sample.

412 We tested the efficacy of 40%, 45% and 50% AMS to precipitate L1 from an extract containing the
413 soluble protein fraction from CTP-pL1 30 plant material produced in the bioreactor. We found 45% AMS
414 to be sufficient to effectively remove L1 from the supernatant, whereas at 40% AMS small amounts of L1
415 remained in the supernatant and at 50% AMS higher amounts of cellular protein were precipitated
416 together with L1 (**Supplementary Figure S6**). In preparation for chromatographic purification, the
417 precipitated proteins were resuspended in CaEx binding buffer and dialysed overnight in binding buffer
418 to remove residual AMS. Cloudiness of samples indicated a reduced solubility of precipitated proteins,
419 and undissolved matter was removed by filtration. The relative amount of L1 and host protein was
420 evaluated *via* Western blots and Coomassie-staining, respectively (**Supplementary Figure S6**). Samples
421 after filtration showed a slight reduction in host protein content, but, except for the sample at 40% AMS,
422 no reduction in L1. Overall, precipitation at 45% AMS was selected, which is in line with purifications of
423 yeast-produced L1 (Park et al., 2008; Kim et al., 2010).

424 Next, CaEx chromatography was used to remove residual cellular proteins. The bulk of cellular
425 proteins did not bind to the column and was detected in the flow through (**Figure 4**, **Supplementary Figure**
426 **S7**). Notably, a large amount of L1 did not bind to the column and was detected in the flow through.
427 Different binding buffer compositions were tested to increase L1 binding to the column, e.g. lowering the
428 pH from 7.2 to 6.8, lowering the concentration of NaCl from 0.137 M to 0.068 M, addition of 5% glycerol,
429 and re-loading of flow-through fraction. Nevertheless, a large part of L1 did not bind to the column and
430 was found in the flow through. During column wash with 5 mL at 0.3 M NaCl, small amounts of
431 contaminating proteins and L1 were washed off the column. For elution, NaCl was increased to 1 M NaCl
432 and 0.5 mL fractions were collected, in total 4 mL. L1 eluted from the column in fractions E4 – E8 (**Figure**
433 **4**), which were pooled for further analysis.

434 In contrast to the confirmation of L1 expression (**Figure 2**) and the establishing of AMS
435 precipitation (**Supplementary Figure S6**), DTT was used throughout the extraction and purification leading
436 to CaEx chromatography. In comparison, the presence of DTT markedly reduces the Western blot signal
437 intensity above 180 kDa (**Figure 4**). We hypothesize that in the absence of DTT, reactive oxygen species in

438 the plant extract cross-link L1 with itself and/or host proteins. Irrespective of the presence or absence of
439 DTT during extraction and purification, two prominent bands in the range of 130 and 180 kDa are present
440 in all Western blots (**Figures 2 and 4, Supplementary Figure S6**).
441



442

443 **Figure 4: Two-step purification of L1 removes bulk of cellular protein.** Analysis of L1 and residual cellular host proteins during
444 purification under reducing conditions *via* Western blot and Coomassie-stained SDS-PAGE (12% Bis-Tris gel). After precipitation
445 with 45% AMS, the filtered sample was subjected to cation exchange chromatography (CaEx) for final purification. Column-bound
446 L1 eluted at 1 M NaCl in fractions E4 – E8. The asterisk in the Coomassie-stained SDS-PAGE marks the band referring to HPV-16
447 L1. 50 ng *S. cerevisiae*-produced L1 served as a positive control and Physcomitrella wildtype as a negative control.
448

449 ## Capsomers and VLPs

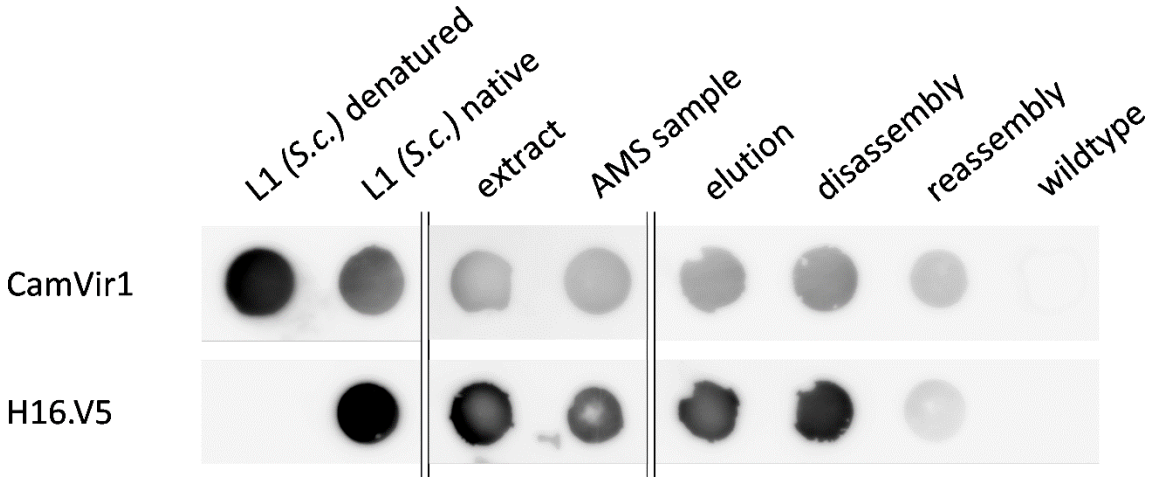
450 Next, we performed a dis- and reassembly protocol on the pooled elution fractions, as it yields
451 more uniform and stable VLPs, which are less prone to aggregation (McCarthy et al., 1998; Mach et al.,
452 2006). First, pooled elution fractions were dialysed at low salt concentrations, high pH and DTT, that
453 results in the disassembly of VLPs or other multi-pentameric L1-based structures into capsomers. Second,

454 samples were dialysed at high salt, low pH and in the absence of reducing agent, that leads to the
455 reassembly of capsomers into VLPs.

456 To assess the conformation of L1 throughout the process, dot blots were performed using
457 different monoclonal antibodies. CamVir1 detects a linear epitope in both denatured and native L1 and
458 served as a control. In contrast, H16.V5 detects only conformational epitopes in both, L1 capsomers and
459 VLPs, but not monomeric or denatured L1.

460 The dot blot showed the presence of L1 in all CTP-pL1 30-derived samples throughout the
461 purification process, as well as yeast-derived denatured and native L1 (**Figure 5**). In comparison, H16.V5
462 detects L1 in the native positive control and in all CTP-pL1 30-derived samples, indicating the presence of
463 at least capsomers, if not VLPs, already in the raw extract. Further, capsomers seem to be stable
464 throughout AMS precipitation, CaEx chromatography, disassembly and reassembly.

465



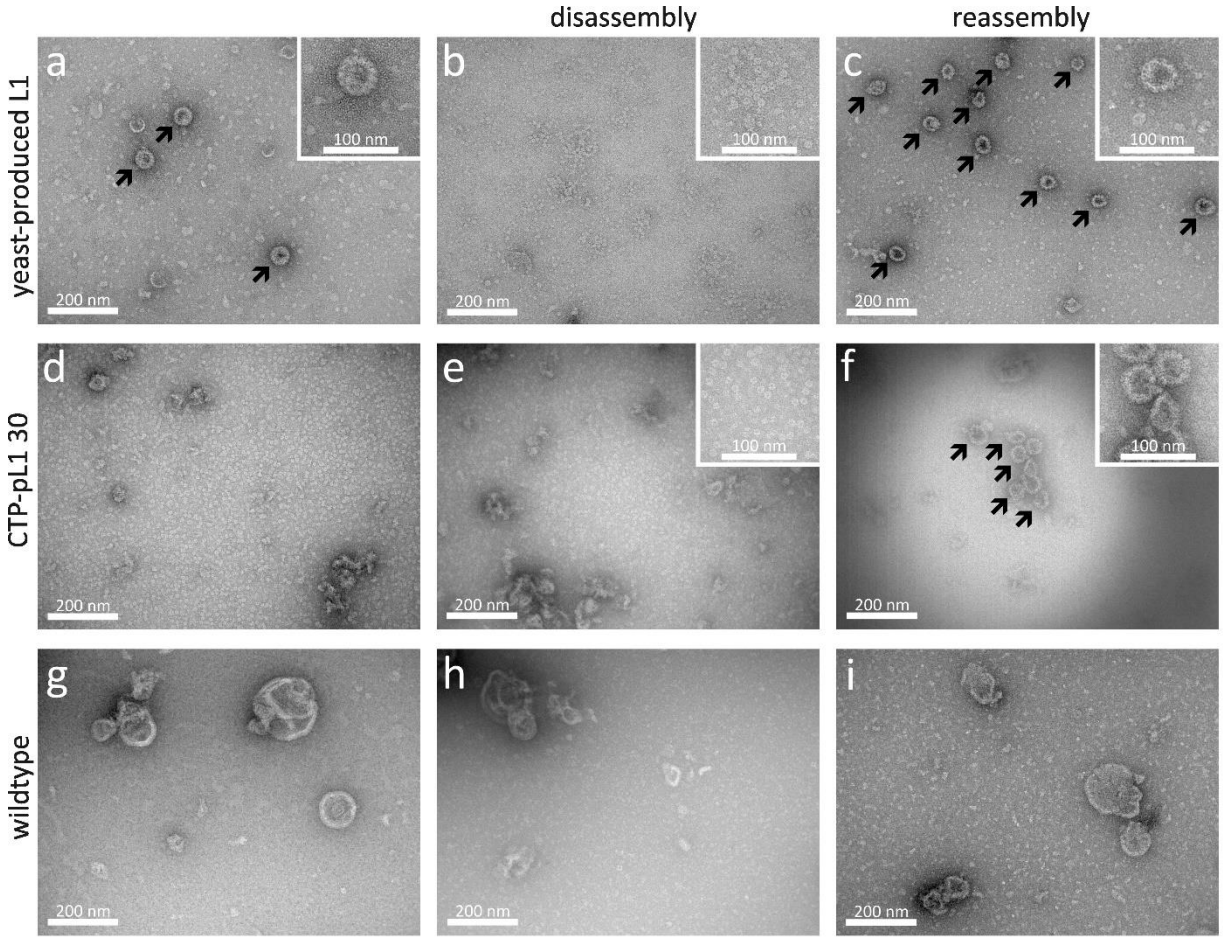
466

467 **Figure 5: Moss-produced HPV-16 L1 assembles into capsomers, which are stable throughout purification.** Dot blots were
468 performed using two different antibodies to analyse the formation of capsomers in yeast-produced L1 and samples from CTP-
469 pL1 30 purification (extract, after AMS precipitation, elution) and dis- and reassembly. The CamVir1 antibody detects a linear
470 epitope in both denatured and native L1. The neutralizing antibody H16.V5 detects specifically conformational epitopes in L1
471 capsomers and VLPs. Positive H16.V5 signals indicate HPV-16 L1 monomers to assemble at least into capsomers, *in vivo* or after
472 extraction, which are stable throughout the purification. The experiment only allows qualitative but not quantitative comparison
473 since sample concentrations were adjusted individually. `L1 denatured´ was incubated at 95 °C for 5 min, all other samples were
474 kept at 4 °C. A wildtype sample was used as a negative control.

475

476 The assembly of L1 capsomers into VLPs was verified with Transmission Electron Microscopy
477 (TEM). As a control, yeast-produced L1 was analysed without prior treatment, after disassembly, and after
478 reassembly. In the untreated sample, particles of ca. 60 nm in size were detected. After disassembly,

479 structures of 5 – 10 nm in size were observed, corresponding to capsomers. VLPs reshape upon
480 reassembly treatment, demonstrating the successful dis- and reassembly of yeast-derived L1 VLPs.
481 Notably, reassembled VLPs are heterogenous in size and shape, despite the controlled dis- and reassembly
482 (**Figure 6**).
483



484
485 **Figure 6: Analysis of dis- and reassembled VLPs from moss and yeast *via* transmission electron microscopy (TEM).** Yeast-derived
486 L1 was analysed untreated (**a**), after disassembly (**b**), and after reassembly (**c**) *via* TEM. CTP-pL130 (**d** – **f**) and wildtype (**g** – **i**)
487 extracts were purified *via* AMS precipitation and CaEx chromatography (**d** & **g**), after disassembly (**e** & **h**), and after reassembly (**f**
488 & **i**). Untreated yeast-derived VLPs (**a**) disintegrate into capsomers upon disassembly (**b**) and reshape into VLPs upon reassembly
489 treatment (**c**). TEM shows the presence of moss-derived HPV-16 L1 capsomers in the elution (**d**) and disassembly sample (**e**) and
490 the assembly into VLPs in the reassembly sample (**f**). Capsomer and VLP-like structures are absent from the wildtype control
491 samples. VLPs are indicated by black arrows.
492
493 To analyse L1 assembly into VLPs, samples were taken from the pooled CaEx chromatography
494 elution fractions, after disassembly, and after reassembly.

495 In the elution sample, capsomeric structures comparable to those observed in the disassembled
496 yeast sample were observed, which are not present in the corresponding wildtype control (**Figure 6**). The
497 abundance of capsomers in the elution fractions was surprising given the fact that the elution buffer has
498 a high ionic strength, which favours reassembly. However, before column loading, the sample was
499 dialysed overnight in binding buffer, which favours disassembly. The exposure to high ionic strength
500 during elution might have been too short to allow VLP assembly. In addition, both binding and elution
501 buffer contained DTT, which induces disassembly (McCarthy et al., 1998; Mach et al., 2006). The
502 disassembled CTP-pL1 30-derived sample depicted capsomers as seen in the elution fraction, and for
503 disassembled yeast-derived L1 (**Figure 6**). **Figure 6f** displays fully assembled, moss-derived HPV-16 L1
504 VLPs. Again, VLPs were heterogenous in size and shape, as seen for reassembled yeast-derived VLPs.
505 Moss-produced VLPs tended to aggregate, which was not observed for reassembled yeast-derived VLPs,
506 but is common for HPV-16 L1 VLPs (Shi et al., 2005). No VLPs were detected in wildtype samples (**Figure**
507 **6**).
508

509 ## Discussion

510 Various strategies were explored for increasing production yields of HPV-16 L1 VLPs in plants, one
511 of which is optimization of codon usage (Maclean et al., 2007; Hitzeroth et al., 2018). The latter found that
512 codon optimization of L1 improves transcription, rather than translation. This is due to RNA motifs in the
513 L1 CDS, which posttranscriptionally regulate the virus life cycle and lead to pre-mRNA degradation (Mori
514 et al., 2006; Zhao & Schwartz, 2008). It is argued, that codon optimization removes those motifs, thereby
515 increasing mRNA and protein levels, irrespective of whether the CDS was optimised for expression in
516 humans, tobacco (Hitzeroth et al., 2018) or Physcomitrella (this study). Our Physcomitrella-optimized CDS
517 has a GC content of 62.85% as compared to hL1 with 62.25% (Maclean et al., 2007) and the wildtype
518 sequence with 38.27% (Varsani et al., 2003). Combined with the removal of microRNA binding sites and
519 heterosplice sites we obtained Physcomitrella-optimized pL1 CDS. We argue that pL1 addresses the issues
520 of L1 intrinsic degradation motifs as well as Physcomitrella-specific posttranscriptional degradation.
521 Besides codon optimization, research focused on localization signals or transit peptides. The most
522 common strategy is sequestration to chloroplasts, either *via* nuclear transformation and posttranslational
523 import into chloroplasts (Maclean et al., 2007; Pineo et al., 2013; Zahin et al., 2016; Hitzeroth et al., 2018;
524 Naupu et al., 2020; Muthamilselvan et al., 2023) or *via* transplastomic lines (Fernández-San Millán et al.,
525 2008; Lenzi et al., 2008). Here, we experimentally characterized five Physcomitrella CTPs, of which we

526 used CTPc5 for sequestration of L1 to chloroplasts. Employing the HPV-16 native NLS, L1 was targeted to
527 nuclei, whereas the removal of the NLS resulted in cytoplasmic L1. Earlier studies confirmed functionality
528 of the NLS in mammalian cells (Leder et al., 2001; Bissa et al., 2015). However, the NLS failed to localize
529 L1 to nuclei in tobacco (Šmídková et al., 2010). Our observation of the HPV-16 native NLS targeting L1 to
530 nuclei is another example for a conserved cross-kingdom functionality of localization signals between
531 Physcomitrella and mammals or human-infecting viruses (Gitzinger et al., 2009).

532 Western blots of L1-producing lines showed a variety of bands, both smaller and larger than the
533 expected 56 kDa or 54 kDa. A comparable banding pattern was observed when expressing L1 in tobacco
534 chloroplasts (Fernández-San Millán et al., 2008). Likewise, yeast-produced L1 exhibited bands with the
535 size of L1 dimers and trimers (Kim et al., 2007; Kim et al., 2010). L1 dimers and trimers are based on inter-
536 capsomeric disulfide bridges which stabilize virions and VLPs (Sapp et al., 1998; Buck et al., 2005; Buck et
537 al., 2013). McCarthy et al. (1998) resolved L1 multimers with DTT. Likewise, Buck et al. (2005)
538 demonstrated absence and presence of L1 dimers and trimers in a comparison of reducing and non-
539 reducing SDS-PAGE and Western blots, respectively. An investigation of cysteine 175 and cysteine 428 L1
540 mutants resulted in the absence of bands at 130 kDa and 180 kDa, thereby linking disulfide bridge-based
541 dimers and trimers to the observed bands (Buck et al., 2005). Surprisingly, reducing agents such as DTT or
542 β-mercaptoethanol did not result in the disappearance of these bands in our and other studies (Kim et al.,
543 2007; Fernández-San Millán et al., 2008; Kim et al., 2010). This indicates a tolerance of L1 dimers and
544 trimers to reducing agents or a different origin of multimerization observed at 130 kDa and 180 kDa.
545 Castells-Graells & Lomonossoff (2021) observed multimerization of *Nudaurelia capensis* omega virus α-
546 peptide produced in tobacco, irrespective of DTT, but not when produced in insect cells. These authors
547 hypothesize that cross-linkage of α-peptide occurs during extraction and purification and occurs in plant-
548 based expression systems. Regarding the observed bands <56 kDa, proteolytic degradation was proposed
549 in other studies observing the same phenomenon (Buck et al., 2005; Fernández-San Millán et al., 2008).
550 In tobacco, a shift from bands above 100 kDa in young leaves to more prominent bands at 45 kDa and 20
551 kDa in older leaves indicated proteolytic degradation of L1 over time (Fernández-San Millán et al., 2008).
552 In our study, bands at 45 kDa are most prominent in pL1Δ22 lines with cytoplasmic L1 and less prominent
553 for plastidic L1. This might result from organelle-specific proteases (Coates at al., 2022).

554 In Physcomitrella, combining the full-length L1 with a CTP significantly increases L1 yield as
555 compared to full length L1 or L1 without NLS, which is in line with findings in tobacco (Maclean et al.,
556 2007; Zahin et al., 2016). In our study, L1 yields above 0.3 mg/g FW were achieved in the photobioreactor.
557 In tobacco, yields of 0.1 – 1.2 mg/g FW (Maclean et al., 2007; Pineo et al., 2013; Zahin et al., 2016;

558 Hitzeroth et al., 2018) and up to 3 mg/g FW in transplastomic lines have been reported (Fernández-San
559 Millán et al., 2008). The success of targeting L1 to chloroplasts over localization to other organelles may
560 have several reasons. One is that plastid L1 may be protected from proteases absent in chloroplasts (van
561 Wijk, 2024). Besides, CTPs also act on the transcriptional level. Upon expression in tobacco, chimeric L1/L2
562 transcripts containing a CTP were more abundant than transcripts lacking a CTP, and consequently yielded
563 more protein (Hitzeroth et al., 2018). So far, it is unclear whether this difference is due to enhanced
564 transcription, increased mRNA half-life or changes in mRNA secondary structure.

565 For purification of L1 from Physcomitrella extracts, a combination of AMS precipitation and CaEx
566 chromatography was employed. We found 45% AMS to be sufficient to efficiently precipitate L1. Studies
567 from yeast (Park et al., 2008; Kim et al., 2010) and tobacco (Zahin et al., 2016) also used 45% AMS, whereas
568 sometimes 50% AMS was used (Kim et al., 2018). CaEx chromatography showed that a large part of L1
569 does not bind to the column. Likewise, Zahin et al. (2016) reported tobacco-produced L1 to bind poorly
570 to CaEx columns. Kim et al. (2010) purified yeast-produced L1 by either CaEx or heparin chromatography.
571 CaEx chromatography showed unbound L1 in the flow through, as observed in our study, whereas no L1
572 was found in the flow through employing heparin chromatography. Pineo et al. (2013) produced L1/L2
573 chimaeras in tobacco and experienced loss of L1 and insufficient removal of contaminating proteins using
574 heparin chromatography. Heparin is a highly sulphated form of heparan sulphate, a proteoglycan found
575 on host cell surfaces and a requirement for HPV virion attachment and infection (Knappe et al., 2007).
576 However, a prerequisite for heparin binding is the assembly of L1 to capsomers or VLPs (Rommel et al.,
577 2005), meaning monomers and not properly folded L1 are lost. Overall, chromatography yields for VLPs
578 differ and are influenced by different factors, including the production host, the used materials and
579 buffers, and purification prior to chromatography, making it difficult to compare purification strategies
580 across publications. Here, we achieved to remove the bulk of cellular proteins from moss extracts, but
581 experienced a substantial loss of L1 in the flow through. To address this, further optimization of
582 purification might be achieved by combinations with other methods, such as heparin chromatography.
583 Our dot blot analysis confirmed the presence of intact capsomers or VLPs in extracts and after AMS
584 precipitation, suggesting heparin chromatography as an option for the second step in purification.

585 TEM analysis of purified and reassembled CTP-pL1 samples shows VLPs of ca. 60 nm in size. These
586 nanoparticles assembled from capsomers as a result of the reassembly treatment. To the best of our
587 knowledge, this is the first report of VLPs produced in a non-vascular plant and of *in vitro* reassembly of
588 HPV-16 L1 VLPs from capsomers produced in any plant system. The VLPs are heterogenous in shape and
589 size. The same is true for yeast-derived VLPs after reassembly, whereas untreated VLPs were not

590 heterogenous. This is surprising, since controlled dis- and reassembly of VLPs can enhance stability and
591 uniformity (McCarthy et al., 1998; Mach et al., 2006; Shi et al., 2007). Besides high ionic strength and low
592 pH buffer, inter-pentameric disulfide bridges positively influence VLP stability and uniformity (Buck et al.,
593 2005). This could be achieved through prolonged incubation of cleared cell lysates before purification or
594 addition of oxidizing agents, which accelerate disulfide bridge formation. However, prolonged incubation
595 causes aggregation of HPV-16 L1 VLPs, as observed here, but could be addressed by improved formulation
596 (Shi et al., 2005).
597 In conclusion, we generated a CDS optimized for expression of HPV-16 L1 in Physcomitrella. We
598 further characterized five CTPs and confirmed the localization of L1 in chloroplasts. In addition, we
599 demonstrated chloroplast localization to yield higher L1 levels than targeting to nuclei or cytoplasm.
600 Cultivation in photobioreactors yielded more than 0.3 mg/g FW, which is in the range of other plant-based
601 expression platforms. We established a purification protocol and demonstrated the successful assembly
602 of moss-produced HPV-16 L1 VLPs. Future efforts should focus on further enhancing purification and
603 formulation of VLPs for immunological studies of this vaccine candidate.
604

605 ## Acknowledgements

606 We thank the Life Imaging Centre (LIC) of the University of Freiburg for support with confocal microscopy.
607 We are grateful for technical assistance by Agnes Novakovic, Britta Rothgänger, Christine Glockner and
608 Richard Haas and to Anne Katrin Prowse for language editing.
609

610 ## Author Contributions

611 PAN acquired funding, designed research, performed experiments, analysed data and wrote the
612 manuscript. MCW and SC performed experiments. CET, MRF and SNWH performed experiments and
613 analysed data. JP and ELD analysed data and helped writing the manuscript. HTS acquired funding and
614 conceptualized the idea. RR designed and supervised research, acquired funding and wrote the
615 manuscript. All authors read and approved the final version of the manuscript.
616

617 ## Conflict of interest disclosure

618 All authors declare no conflict of interest.
619

Data availability statement

Nucleotide sequence data was submitted to GenBank under accession numbers PV029542 (pL1) and PV068547(CTPc5). Moss lines and vectors were submitted to the International Moss Stock Centre (IMSC) under accession numbers 40963 (pL1 5A), 40964 (pL1 23A), 40965 (pL1 47B), 40966 (pL1Δ22 8A), 40967 (pL1Δ22 11B), 40968 (pL1Δ22 102B), 40969 (CTP-pL1 30), 40970 (CTP-pL1 103), 40971 (CTP-pL1 143), 1840 (pAct5_linker_eGFP_35ST_hpt), 2038 (pAct5_CTP_linker_eGFP_35ST_hpt), 1193 (pAct5_Citrine_nosT_hpt), 2249 (pAct5_pL1-Citrine_nosT_hpt), 2250 (pAct5_pL1Δ22-Citrine_nosT_hpt), 2251 (pAct5_CTP-pL1-Citrine_nosT_hpt), 1997 (pAct5_pL1_35ST_hpt), 1998 (pAct5_pL1Δ22_35ST_hpt) and 2048 (pAct5_CTP-pL1_35ST_hpt).

Funding statement

This work was supported by Maria-Sklodowska-Curie Actions Innovative Training Network under the Horizon 2020 programme under Grant Agreement No. 765115 (MossTech), by the German Research Foundation DFG under Germany's Excellence Strategy (CIBSS – EXC-2189 – Project ID 390939984), and the International Network Programme by the Danish Agency for Higher Education and Science (grant number 1113-00016B). The TEM (Hitachi HT7800) was funded by a DFG grant (project number 426849454) and is operated by the University of Freiburg, Faculty of Biology, as a partner unit within the Microscopy and Image Analysis Platform MIAP and the Life Imaging Center LIC, Freiburg. PAN gratefully acknowledges financial support by Studienstiftung des deutschen Volkes.

641 ## Short legends for Supporting Information

642 Supplementary Figure S1 **Vector maps**

643 Supplementary Figure S2 **Generation of Physcomitrella-optimized (p)L1**

644 Supplementary Figure S3 **Characterization of CTPs by analysis of subcellular eGFP localization**

645 Supplementary Figure S4 **Analysis of CDS integrity in L1-producing lines**

646 Supplementary Figure S5 **Comparison of biomass accumulation rates**

647 Supplementary Figure S6 **Ammonium sulphate precipitation**

648 Supplementary Figure S7 **Chromatogram of Cation Exchange purification**

649 Supplementary Table S1 **List of primers**

650 Supplementary Table S2 **Predicted chloroplast transit peptides for six selected proteins in**

651 **Physcomitrella**

652 ## References

653 Abel WO, Knebel W, Koop HU, Marienfeld JR, Quader H, Reski R, Schnepf E, Sporlein B (1989) A
654 cytokinin-sensitive mutant of the moss, *Physcomitrella patens*, defective in chloroplast division.
655 *Protoplasma* 152, 1-13. https://doi.org/10.1007/BF01354234
656
657 Adams A, Hendrikse M, Rybicki EP, Hitzeroth II (2023) Optimal size of DNA encapsidated by plant
658 produced human papillomavirus pseudovirions. *Virology* 580, 88 – 97.
659 https://doi.org/10.1016/j.virol.2023.02.003
660
661 Armenteros JJA, Salvatore M, Emanuelsson O, Wither O, von Heijne G, Elofsson A, Nielsen H (2019)
662 Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance* 2,
663 e201900429. https://doi.org/ 10.26508/lsa.201900429
664
665 Baker TS, Newcomb WW, Olson NH, Cowsert LM, Olson C, Brown JC (1991) Structures of bovine and
666 human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image
667 reconstruction. *Biophysical Journal* 60, 1445–1456. https://doi.org/10.1016/S0006-3495(91)82181-6
668
669 Biemelt S, Sonnewald U, Galmbacher P, Willmitzer L, Müller M (2003) Production of human
670 papillomavirus type 16 virus-like particles in transgenic plants. *Journal of Virology* 77, 9211–9220.
671 https://doi.org/10.1128/JVI.77.17.9211-9220.2003
672
673 Bissa M, Zanotto C, Pacchioni S, Volonté L, Venuti A, Lembo D, De Giuli Morghen C, Radaelli A (2015) The
674 L1 protein of human papilloma virus 16 expressed by a fowlpox virus recombinant can assemble into
675 virus-like particles in mammalian cell lines but elicits a non-neutralising humoral response. *Antiviral*
676 *Research* 116, 67–75. https://doi.org/10.1016/j.antiviral.2015.01.012
677
678 Bohlender LL, Parsons J, Hoernstein SNW, Rempfer C, Ruiz-Molina N, Lorenz T, Rodríguez Jahnke F, Figl
679 R, Fode B, Altmann F, Reski R, Decker EL (2020) Stable protein sialylation in Physcomitrella. *Frontiers in*
680 *Plant Science* 11, 610032. https://doi.org/10.3389/FPLS.2020.610032
681

682 Bohlender LL, Parsons J, Hoernstein SNW, Bangert N, Rodriguez-Jahnke F, Reski R, Decker EL (2022)
683 Unexpected arabinosylation after humanization of plant protein N-glycosylation. *Frontiers in*
684 *Bioengineering and Biotechnology* 10, 838365. https://doi.org/10.3389/fbioe.2022.838365
685
686 Buck CB, Thompson CD, Pang YYS, Lowy DR, Schiller JT (2005) Maturation of papillomavirus capsids.
687 *Journal of Virology* 79, 2839–2846. https://doi.org/10.1128/JVI.79.5.2839-2846.2005
688
689 Buck CB, Day PM, Trus BL (2013) The papillomavirus major capsid protein L1. *Virology* 445, 169-174.
690 https://doi.org/ 10.1016/j.virol.2013.05.038
691
692 Büttner-Mainik A, Parsons J, Jérôme H, Hartmann A, Lamer S, Schaaf A, Schlosser A, Zipfel PF, Reski R,
693 Decker EL (2011) Production of biologically active recombinant human factor H in Physcomitrella. *Plant*
694 *Biotechnology Journal* 9, 373–383. https://doi.org/10.1111/J.1467-7652.2010.00552.X
695
696 Castells-Graells R, Lomonossoff GP (2021) Plant-based production can result in covalent cross-linking of
697 proteins. *Plant Biotechnology Journal* 19, 1095–1097. https://doi.org/10.1111/pbi.13598
698
699 Cerqueira C, Schiller JT (2017) Papillomavirus assembly: An overview and perspectives. *Virus Research*
700 231: 103–107. https://doi.org/10.1016/j.virusres.2016.11.010
701
702 Coates RJ, Young MT, Scofield S (2022) Optimising expression and extraction of recombinant proteins in
703 plants. *Frontiers in Plant Science* 13, 1074531. https://doi.org/10.3389/fpls.2022.1074531
704
705 Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Research* 39,
706 W155-W159. https://doi.org/10.1093/nar/gkr319
707
708 Dai X, Zhuang, Z, Zhao, PX (2018) psRNATarget: a plant small RNA target analysis server (2017 release).
709 *Nucleic Acids Research* 46, W49–W54. https://doi.org/10.1093/NAR/GKY316
710
711 de Martel C, Georges D, Bray F, Ferlay J, Clifford GM (2020) Global burden of cancer attributable to
712 infections in 2018: a worldwide incidence analysis. *The Lancet Global Health* 8, e180–e190.
713 https://doi.org/10.1016/S2214-109X(19)30488-7

714

715 Decker EL, Reski R (2020) Mosses in biotechnology. *Current Opinion in Biotechnology* 61, 21–27.
716 https://doi.org/10.1016/j.copbio.2019.09.021
717

718 Decker EL, Wiedemann G, Reski R (2015) Gene targeting for precision glyco-engineering: Production of
719 biopharmaceuticals devoid of plant-typical glycosylation in moss bioreactors. *Methods in Molecular*
720 *Biology* 1321, 213-224. https://doi.org/10.1007/978-1-4939-2760-9_15
721

722 Decker EL, Parsons J, Reski R (2014) Glyco-engineering for biopharmaceutical production in moss
723 bioreactors. *Frontiers in Plant Science* 5, 346. https://doi.org/10.3389/fpls.2014.00346
724

725 Fernández-San Millán A, Ortigosa SM, Hervás-Stubbs S, Corral-Martínez P, Seguí-Simarro JM, Gaétan J,
726 Coursaget P, Veramendi J (2008) Human papillomavirus L1 protein expressed in tobacco chloroplasts
727 self-assembles into virus-like particles that are highly immunogenic. *Plant Biotechnology Journal* 6, 427–
728 441. https://doi.org/10.1111/j.1467-7652.2008.00338.x
729

730 Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO (2009) Enzymatic assembly of DNA
731 molecules up to several hundred kilobases. *Nature Methods* 6, 343–345.
732 https://doi.org/10.1038/nmeth.1318
733

734 Gitzinger M, Parsons J, Reski R, Fussenegger M (2009) Functional cross-kingdom conservation of
735 mammalian and moss (*Physcomitrella patens*) transcription, translation and secretion machineries. *Plant*
736 *Biotechnology Journal* 7, 73–86. https://doi.org/10.1111/J.1467-7652.2008.00376.X
737

738 Gremillon L, Kiessling J, Hause B, Decker EL, Reski R, Sarnighausen E (2007) Filamentous temperature-
739 sensitive Z (FtsZ) isoforms specifically interact in the chloroplasts and in the cytosol of *Physcomitrella*
740 *patens*. *New Phytologist* 176, 299–310. https://doi.org/10.1111/j.1469-8137.2007.02169.x
741

742 Hennermann JB, Arash-Kaps L, Fekete G, Schaaf A, Busch A, Frischmuth T (2019) Pharmacokinetics,
743 pharmacodynamics, and safety of moss-aGalactosidase A in patients with Fabry disease. *Journal of*
744 *Inherited Metabolic Disease* 42, 527–533. https://doi.org/10.1002/JIMD.12052
745

746 Hitzeroth II, Chabeda A, Whitehead MP, Graf M, Rybicki EP (2018) Optimizing a human papillomavirus
747 type 16 L1-based chimaeric gene for expression in plants. *Frontiers in Bioengineering and Biotechnology*
748 6, 101. https://doi.org/10.3389/fbioe.2018.00101
749
750 Hoernstein SNW, Oezdemir B, van Gessel N, Miniera AA, Rogalla von Bieberstein B, Nilges L, Schweikert
751 Farinha J, Komoll R, Glauz S, Weckerle T, Scherzinger F, Rodriguez-Franco M, Mueller-Schuessele S, Reski
752 R (2023) A deeply conserved protease, acylamino acid-releasing enzyme (AARE), acts in ageing in
753 Physcomitrella and Arabidopsis. *Communications Biology* 6, 61. https://doi.org/10.1038/s42003-023-
754 04428-7
755
756 Hoernstein SNW, Schlosser A, Fiedler K, van Gessel N, Igloi GL, Lang D, Reski R (2024) A snapshot of the
757 Physcomitrella N-terminome reveals N-terminal methylation of organellar proteins. *Plant Cell Reports*
758 43, 250. https://doi.org/10.1007/s00299-024-03329-1
759
760 Hohe A, Reski R (2002) Optimisation of a bioreactor culture of the moss *Physcomitrella patens* for mass
761 production of protoplasts. *Plant Science* 163, 69-74. https://doi.org/10.1016/S0168-9452(02)00059-6
762
763 Hohe A, Decker EL, Gorr G, Schween G, Reski R (2002) Tight control of growth and cell differentiation in
764 photoautotrophically growing moss (*Physcomitrella patens*) bioreactor cultures. *Plant Cell Reports* 20,
765 1135-1140. https://doi.org/10.1007/s00299-002-0463-y
766
767 Hohe A, Egener T, Lucht JM, Holtorf H, Reinhard C, Schween G, Reski R (2004) An improved and highly
768 standardized transformation procedure allows efficient production of single and multiple targeted gene-
769 knockouts in a moss, *Physcomitrella patens*. *Current Genetics* 44, 339-347.
770 https://doi.org/10.1007/s00294-003-0458-4
771
772 Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. *Annual*
773 *Review of Plant Biology* 57, 19-53. https://doi.org/10.1146/annurev.arplant.57.032905.105218
774
775 Karan S, Affonso De Oliveira JF, Moreno-Gonzalez MA, Steinmetz NF (2024) A self-amplifying human
776 papillomavirus 16 vaccine candidate delivered by tobacco mosaic virus-like particle. *ACS Applied Bio*
777 *Materials* 7, 7675 – 7683. https://doi.org/10.1021/acsabm.4c01239

Khraiwesh B, Arif MA, Seumel GI, Ossowski S, Weigel D, Reski R, Frank W (2010) Transcriptional control of gene expression by microRNAs. *Cell* 140, 111-122. https://doi.org/ 10.1016/j.cell.2009.12.023

Kim SN, Jeong HS, Park SN, Kim HJ (2007) Purification and immunogenicity study of human papillomavirus type 16 L1 protein in *Saccharomyces cerevisiae*. *Journal of Virological Methods* 139, 24-30. https://doi.org/ 10.1016/j.jviromet.2006.09.004

Kim HJ, Kim SY, Lim SJ, Kim JY, Lee SJ, Kim HJ (2010) One-step chromatographic purification of human papillomavirus type 16 L1 protein from *Saccharomyces cerevisiae*. *Protein Expression and Purification* 70, 68–74. https://doi.org/10.1016/j.pep.2009.08.005

Kim HJ, Cho SY, Park MH, Kim HJ (2018) Comparison of the size distributions and immunogenicity of human papillomavirus type 16 L1 virus-like particles produced in insect and yeast cells. *Archives of Pharmacal Research* 41, 544–553. https://doi.org/10.1007/s12272-018-1024-4

Kirnbauer R, Booy F, Cheng N, Lowy DR, Schiller JT (1992) Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. *Proceedings of the National Academy of Sciences USA* 89, 12180–12184. https://doi.org/10.1073/pnas.89.24.12180

Knappe M, Bodevin S, Selinka HC, Spillmann D, Streeck RE, Chen XS, Lindahl U, Sapp M (2007) Surface-exposed amino acid residues of HPV16 L1 protein mediating interaction with cell surface heparan sulfate. *Journal of Biological Chemistry* 282, 27913–27922. https://doi.org/10.1074/jbc.M705127200

Koprivova A, Stemmer C, Altmann F, Hoffmann A, Kopriva S, Gorr G, Reski R, Decker EL (2004) Targeted knockouts of *Physcomitrella* lacking plant-specific immunogenic N-glycans. *Plant Biotechnology Journal* 2, 517–523. https://doi.org/10.1111/J.1467-7652.2004.00100.X

Lamprecht RL, Kennedy P, Huddy SM, Bethke S, Hendrikse M, Hitzeroth II, Rybicki EP (2016) Production of human papillomavirus pseudovirions in plants and their use in pseudovirion-based neutralisation assays in mammalian cells. *Scientific Reports* 6, 20431. https://doi.org/10.1038/srep20431

810 Leder C, Kleinschmidt JA, Wiethe C, Müller M (2001) Enhancement of capsid gene expression: Preparing
811 the human papillomavirus type 16 major structural gene L1 for DNA vaccination purposes. *Journal of*
812 *Virology* 75, 9201–9209. https://doi.org/10.1128/JVI.75.19.9201-9209.2001
813
814 Lenzi P, Scotti N, Alagna F, Tornesello ML, Pompa A, Vitale A, De Stradis A, Monti L, Grillo S, Buonaguro
815 FM, Maliga P, Cardi T (2008) Translational fusion of chloroplast-expressed human papillomavirus type 16
816 L1 capsid protein enhances antigen accumulation in transplastomic tobacco. *Transgenic Research* 17,
817 1091–1102. https://doi.org/10.1007/s11248-008-9186-3
818
819 Ljubojevic S, Skerlev M (2014) HPV-associated diseases. *Clinics in Dermatology* 32, 227-234.
820 https://doi.org/10.1016/j.clindermatol.2013.08.007
821
822 Mach H, Volkin DB, Troutman RD, Wang B, Luo Z, Jansen KU, Shi L (2006) Disassembly and reassembly of
823 yeast-derived recombinant human papillomavirus virus-like particles (HPV VLPs). *Journal of*
824 *Pharmaceutical Sciences* 95, 2195–2206. https://doi.org/10.1002/jps.20696
825
826 Maclean J, Koekemoer M, Olivier AJ, Stewart D, Hitzeroth II, Rademacher T, Fischer R, Williamson AL,
827 Rybicki EP (2007) Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants:
828 comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment
829 localization. *Journal of General Virology* 88, 1460–1469. https://doi.org/10.1099/vir.0.82718-0
830
831 Marsian J, Lomonossoff GP (2016) Molecular pharming – VLPs made in plants. *Current Opinion in*
832 *Biotechnology* 37, 201-206. https://doi.org/10.1016/j.copbio.2015.12.007
833
834 McCarthy MP, White WI, Palmer-Hill F, Koenig S, Suzich JA (1998) Quantitative disassembly and
835 reassembly of human papillomavirus type 11 virus like particles *in vitro*. *Journal of Virology* 72, 32–41.
836 https://doi.org/10.1128/JVI.72.1.32-41.1998
837
838 Michelfelder S, Parsons J, Bohlender LL, Hoernstein SNW, Niederkrüger H, Busch A, Krieghoff N, Koch J,
839 Fode B, Schaaf A, Frischmuth T, Pohl M, Zipfel PF, Reski R, Decker EL, Häffner K (2017) Moss-produced,
840 glycosylation-optimized human factor H for therapeutic application in complement disorders. *Journal of*
841 *the American Society of Nephrology* 28, 1462–1474. https://doi.org/10.1681/ASN.2015070745

842

843 Michelfelder S, Fischer F, Wäldin A, Hörle KV, Pohl M, Parsons J, Reski R, Decker EL, Zipfel PF, Skerka C,
844 Häffner K (2018) The MFHR1 fusion protein is a novel synthetic multitarget complement inhibitor with
845 therapeutic potential. *Journal of the American Society of Nephrology* 29, 1141–1153.
846 https://doi.org/10.1681/ASN.2017070738
847

848 Mori S, Ozaki S, Yasugi T, Yoshikawa H, Taketani Y, Kanda T (2006) Inhibitory cis-element-mediated
849 decay of human papillomavirus type 16 L1-transcript in undifferentiated cells. *Molecular and Cellular*
850 *Biochemistry* 288, 47–57. https://doi.org/10.1007/s11010-006-9117-7
851

852 Murén E, Nilsson A, Ulfstedt M, Johansson M, Ronne H (2009) Rescue and characterization of episomally
853 replicating DNA from the moss *Physcomitrella*. *Proceedings of the National Academy of Sciences USA*
854 106, 19444–19449
855

856 Muthamilselvan T, Khan MRI, Hwang I (2023) Assembly of human papillomavirus 16 L1 protein in
857 *Nicotiana benthamiana* chloroplasts into highly immunogenic Virus-Like Particles. *Journal of Plant*
858 *Biology* 66, 331–340. https://doi.org/10.1007/s12374-023-09393-6
859

860 Naupu PN, van Zyl AR, Rybicki EP, Hitzeroth II (2020) Immunogenicity of plant-produced human
861 papillomavirus (HPV) Virus-Like Particles (VLPs). *Vaccines* 8, 740.
862 https://doi.org/10.3390/VACCINES8040740
863

864 Nooraei S, Bahrulolum H, Hoseini ZS, Katalani C, Hajizade A, Easton AJ, Ahmadian G (2021) Virus-like
865 particles: preparation, immunogenicity and their roles as nanovaccines and drug nanocarriers. *Journal of*
866 *Nanobiotechnology* 19, 59. https://doi.org/10.1186/s12951-021-00806-7
867

868 Orellana-Escobedo L, Rosales-Mendoza S, Romero-Maldona A, Parsons J, Decker EL, Monreal-Escalante
869 E, Moreno-Fierros L, Reski R (2015) An Env-derived multi-epitope HIV chimeric protein produced in the
870 moss *Physcomitrella patens* is immunogenic in mice. *Plant Cell Reports* 34, 425-433.
871 https://doi.org/10.1007/s00299-014-1720-6
872

873 Park MA, Kim HJ, Kim HJ (2008) Optimum conditions for production and purification of human
874 papillomavirus type 16 L1 protein from *Saccharomyces cerevisiae*. *Protein Expression and Purification* 59,
875 175–181. https://doi.org/10.1016/j.pep.2008.01.021
876
877 Parsons J, Altmann F, Arrenberg CK, Koprivova A, Beike AK, Stemmer C, Gorr G, Reski R, Decker EL (2012)
878 Moss-based production of asialo-erythropoietin devoid of Lewis A and other plant-typical carbohydrate
879 determinants. *Plant Biotechnology Journal* 10, 851–861. https://doi.org/10.1111/J.1467-
880 7652.2012.00704.X
881
882 Pineo CB, Hitzeroth II, Rybicki EP (2013) Immunogenic assessment of plant-produced human
883 papillomavirus type 16 L1/L2 chimaeras. *Plant Biotechnology Journal* 11, 964–975.
884 https://doi.org/10.1111/pbi.12089
885
886 Rempfer C, Wiedemann G, Schween G, Kerres KL, Lucht JM, Horres R, Decker EL, Reski R (2022)
887 Autopolyploidization affects transcript patterns and gene targeting frequencies in Physcomitrella. *Plant*
888 *Cell Reports* 41, 153-173. https://doi.org/10.1007/s00299-021-02794-2
889
890 Reski R (1998) Development, genetics and molecular biology of mosses. *Botanica Acta* 111, 1–15.
891 https://doi.org/10.1111/j.1438-8677.1998.tb00670.x
892
893 Reski R (2018) Quantitative moss cell biology. *Current Opinion in Plant Biology* 46, 39–47.
894 https://doi.org/10.1016/J.PBI.2018.07.005
895
896 Reski R, Abel WO (1985) Induction of budding on chloronemata and caulonemata of the moss,
897 *Physcomitrella patens*, using isopentenyladenine. *Planta* 165, 354-358.
898 https://doi.org/1007/BF00392232
899
900 Reski R, Parsons J, Decker EL (2015) Moss-made pharmaceuticals: from bench to bedside. *Plant*
901 *Biotechnology Journal* 13, 1191–1198. https://doi.org/10.1111/pbi.12401
902
903 Reski R, Bae H, Simonsen HT (2018) *Physcomitrella patens*, a versatile synthetic biology chassis. *Plant*
904 *Cell Reports* 37, 1409–1417. https://doi.org/10.1007/S00299-018-2293-6

905
906 Roden RBS, Stern PL (2018) Opportunities and challenges for human papillomavirus vaccination in
907 cancer. *Nature Reviews Cancer* 18, 240-254. https://doi.org/10.1038/nrc.2018.13
908
909 Rommel O, Dillner J, Fligge C, Bergsdorf C, Wang X, Selinka HC, Sapp M (2005) Heparan sulfate
910 proteoglycans interact exclusively with conformationally intact HPV L1 assemblies: Basis for a virus-like
911 particle ELISA. *Journal of Medical Virology* 75, 114–121. https://doi.org/10.1002/jmv.20245
912
913 Rosales-Mendoza S, Orellana-Escobedo L, Romero-Maldonado A, Decker EL, Reski R (2014) The potential
914 of *Physcomitrella patens* as a platform for the production of plant-based vaccines. *Expert Reviews of*
915 *Vaccines* 13, 203-212. https://doi.org/10.1586/14760584.2014.872987
916
917 Rosales-Mendoza S, Salazar-Gonzalez JA, Decker EL, Reski R (2016) Implications of plant glycans in the
918 development of innovative vaccines. *Expert Reviews of Vaccines* 15, 915-925.
919 https://doi.org/10.1586/14760584.2016.1155987
920
921 Ruiz-Molina N, Parsons J, Müller M, Hoernstein SNW, Bohlender LL, Pumple S, Zipfel PF, Häffner K, Reski
922 R, Decker EL (2022a) A synthetic protein as efficient multitarget regulator against complement over-
923 activation. *Communications Biology* 5, 152. https://doi.org/10.1038/s42003-022-03094-5
924
925 Ruiz-Molina N, Parsons J, Schroeder S, Posten C, Reski R, Decker EL (2022b) Process engineering of
926 biopharmaceutical production in moss bioreactors *via* model-based description and evaluation of
927 phytohormone impact. *Frontiers in Bioengineering and Biotechnology* 10, 837965.
928 https://doi.org/10.3389/fbioe.2022.837965
929
930 Rybicki EP (2020) Plant molecular farming of virus-like nanoparticles as vaccines and reagents. *WIREs*
931 *Nanomedicine and Nanobiotechnology* 12, e1587. https://doi.org/10.1002/wnan.1587
932
933 Sapp M, Fligge C, Petzak I, Robin Harris J, Streeck RE (1998) Papillomavirus assembly requires
934 trimerization of the major capsid protein by disulfides between two highly conserved cysteines. *Journal*
935 *of Virology* 72, 6186 – 6189. https://doi.org/10.1128/jvi.72.7.6186-6189.1998
936

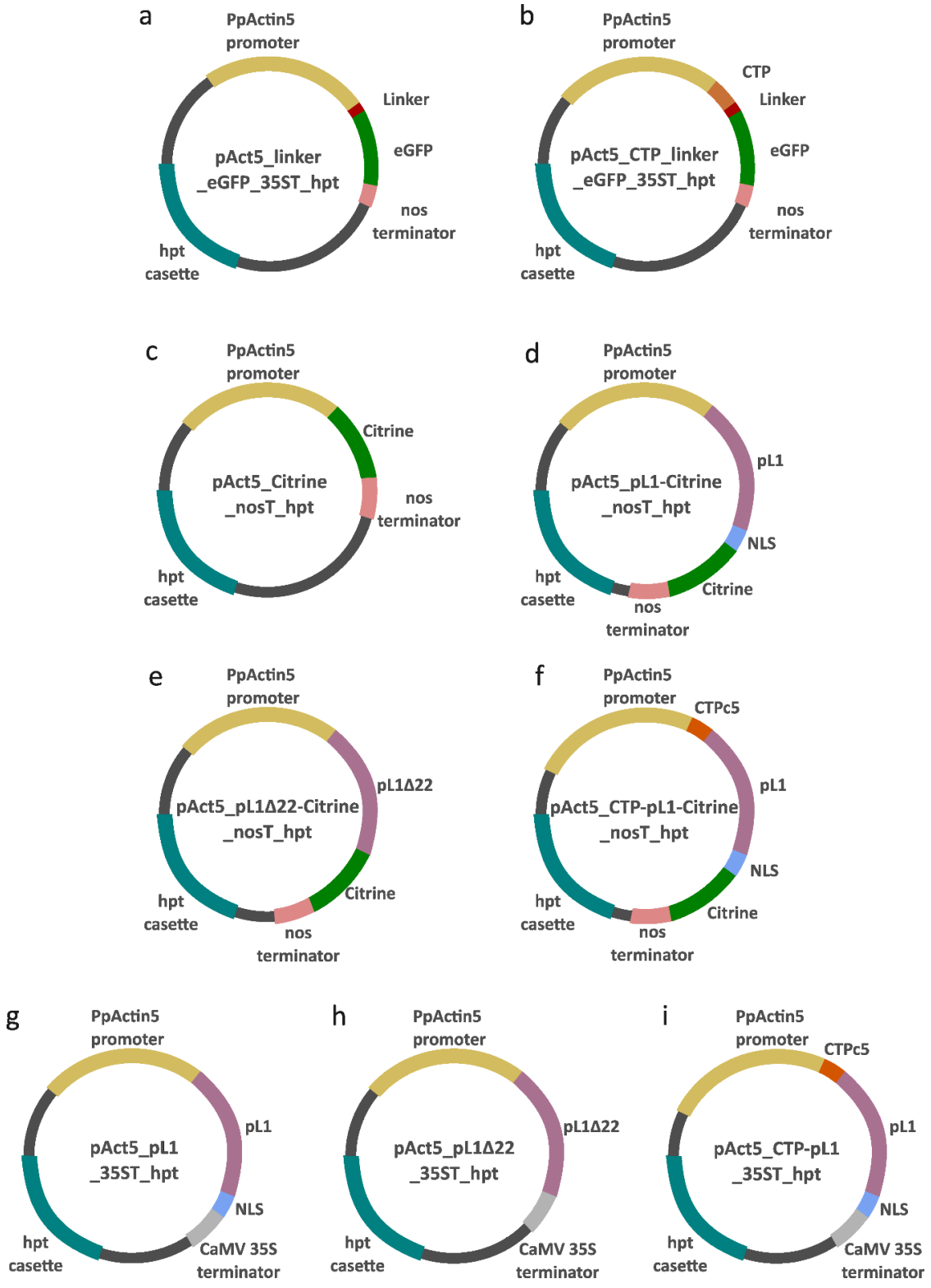
937 Schween G, Hohe A, Koprivova A, Reski R (2003) Effects of nutrients, cell density and culture techniques
938 on protoplast regeneration and early protonema development in a moss, *Physcomitrella patens*. *Journal*
939 *of Plant Physiology* 160, 209-212. https://doi.org/10.1078/0176-1617-00855
940
941 Scotti N, Rybicki EP (2013) Virus-like particles produced in plants as potential vaccines. *Expert Review of*
942 *Vaccines* 12, 211–224. https://doi.org/10.1586/erv.12.147
943
944 Shen JS, Busch A, Day TS, Meng XL, Yu CI, Dabrowska-Schlepp P, Fode B, Niederkrüger H, Forni S, Chen S,
945 Schiffmann R, Frischmuth T, Schaaf A (2016) Mannose receptor-mediated delivery of moss-made α-
946 galactosidase A efficiently corrects enzyme deficiency in Fabry mice. *Journal of Inherited Metabolic*
947 *Disease* 39, 293–303. https://doi.org/10.1007/S10545-015-9886-9
948
949 Shi L, Sanyal G, Ni A, Luo Z, Doshna S, Wang B, Graham TL, Wang N, Volkin DB (2005) Stabilization of
950 human papillomavirus virus-like particles by non-ionic surfactants. *Journal of Pharmaceutical Sciences*
951 94, 1538–1551. https://doi.org/10.1002/jps.20377
952
953 Shi L, Sings HL, Bryan JT, Wang B, Wang Y, Mach H, Kosinski M, Washabaugh MW, Sitrin R, Barr E (2007)
954 GARDASIL®: Prophylactic human papillomavirus vaccine development – From bench top to bed-side.
955 *Clinical Pharmacology & Therapeutics* 81, 259–264. https://doi.org/10.1038/SJ.CLPT.6100055
956
957 Šmídková M, Müller M, Thönes N, Piuko K, Angelisová P, Velemínský J, Angelis KJ (2010) Transient
958 expression of human papillomavirus type 16 virus-like particles in tobacco and tomato using a tobacco
959 rattle virus expression vector. *Biologia Plantarum* 54, 451–460. https://doi.org/10.1007/s10535-010-
960 0081-4
961
962 Studentsov YY, Schiffman M, Strickler HD, Ho GYF, Pang YYs, Schiller J, Herrero R, Burk RD (2002)
963 Enhanced enzyme-linked immunosorbent assay for detection of antibodies to virus-like particles of
964 human papillomavirus. *Journal of Clinical Microbiology* 40, 1755-1760.
965 https://doi.org/10.1128/JCM.40.5.1755-1760.2002
966

967 Strepp R, Scholz S, Kruse S, Speth V, Reski R (1998) Plant nuclear gene knockout reveals a role in plastid
968 division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proceedings of*
969 *the National Academy of Sciences USA* 95, 4368–4373. https://doi.org/10.1073/PNAS.95.8.4368
970
971 Top O, Parsons J, Bohlender LL, Michelfelder S, Kopp P, Busch-Steenberg C, Hoernstein SNW, Zipfel PF,
972 Häffner K, Reski R, Decker EL (2019) Recombinant production of MFHR1, a novel synthetic multitarget
973 complement inhibitor, in moss bioreactors. *Frontiers in Plant Science* 10, 260.
974 https://doi.org/10.3389/FPLS.2019.00260
975
976 Top O, Milferstaedt SWL, van Gessel N, Hoernstein SNW, Özdemir B, Decker EL, Reski R (2021)
977 Expression of a human cDNA in moss results in spliced mRNAs and fragmentary protein isoforms.
978 *Communications Biology* 4, 964. https://doi.org/10.1038/s42003-021-02486-3
979
980 U.S. Food and Drug Administration (2018) FDA approves expanded use of Gardasil 9 to include
981 individuals 27 through 45 years old | FDA. https://www.fda.gov/news-events/press-
982 announcements/fda-approves-expanded-use-gardasil-9-include-individuals-27-through-45-years-old
983
984 van Wijk KJ (2024) Intra-chloroplast proteases: A holistic network view of chloroplast proteolysis. *Plant*
985 *Cell* 36, 3116–3130. https://doi.org/10.1093/plcell/koae178
986
987 Varsani A, Williamson AL, Rose RC, Jaffer M, Rybicki EP (2003) Expression of Human papillomavirus type
988 16 major capsid protein in transgenic *Nicotiana tabacum* cv. Xanthi. *Archives of Virology* 148, 1771–
989 1786. https://doi.org/10.1007/s00705-003-0119-4
990
991 Weise A, Rodriguez-Franco M, Timm B, Hermann M, Link S, Jost W, Gorr G (2006) Use of *Physcomitrella*
992 *patens* actin 5' regions for high transgene expression: importance of 5' introns. *Applied Microbiology*
993 *and Biotechnology* 70: 337–345. https://doi.org/10.1007/s00253-005-0087-6
994
995 Wiedemann G, van Gessel N, Köchl F, Hunn L, Schulze K, Maloukh L, Nogué F, Decker EL, Hartung F, Reski
996 R (2018) RecQ helicases function in development, DNA repair, and gene targeting in *Physcomitrella*
997 *patens*. *Plant Cell* 30, 717–736. https://doi.org/10.1105/tpc.17.00632
998

999 World Health Organisation (2024a) Human papillomavirus and cancer. https://www.who.int/news-
1000 room/fact-sheets/detail/human-papilloma-virus-and-cancer, accessed 16/12/2024
1001
1002 World Health Organization (2024b) Considerations for human papillomavirus (HPV) vaccine product
1003 choice (Issue 2). https://iris.who.int/bitstream/handle/10665/376546/9789240089167-eng.pdf
1004
1005 Wu M, Huang H, Tang Y, Ren X, Jiang X, Tian M, Li W (2024) Unveiling the multifaceted realm of human
1006 papillomavirus: a comprehensive exploration of biology, interactions, and advances in cancer
1007 management. *Frontiers in Immunology* 15, 1430544. https://doi.org/10.3389/fimmu.2024.1430544
1008
1009 Zahin M, Joh J, Khanal S, Husk A, Mason H, Warzecha H, Ghim SJ, Miller DM, Matoba N, Jenson AB
1010 (2016) Scalable production of HPV16 L1 protein and VLPs from tobacco leaves. *PLOS ONE* 11, e 0160995.
1011 https://doi.org/10.1371/JOURNAL.PONE.0160995
1012
1013 Zhao X, Schwartz S (2008) Inhibition of HPV-16 L1 expression from L1 cDNAs correlates with the
1014 presence of hnRNP A1 binding sites in the L1 coding region. *Virus Genes* 36, 45–53.
1015 https://doi.org/10.1007/s11262-007-0174-0
1016
1017 zur Hausen H (2002) Papillomaviruses and cancer: From basic studies to clinical application. *Nature*
1018 *Reviews Cancer* 2, 342-350. https://doi.org/10.1038/nrc798
1019
1020 zur Hausen H, de Villiers EM, Gissmann L (1981) Papillomavirus infections and human genital cancer.
1021 *Gynecologic Oncology* 12, S124–S128. https://doi.org/10.1016/0090-8258(81)90067-6

1022 # Supplementary Information

1023 ## Supplementary Figure S1 – Vector maps



1024

1025 **List of all vectors used in this study.** Vectors **a** & **b** were used for the characterization of chloroplast transit peptides (CTPs).
1026 Vector **a** served as a control and was used as a cloning template for inserting different CTP candidates. Vectors **c**, **d**, **e** & **f** were
1027 used for the analysis of L1 subcellular targeting. Vector **c** served as a control and was used as a cloning template for inserting
1028 either pL1, pL1Δ22 or CTP-pL1 CDS (**d**, **e** & **f**, respectively). Vectors **g**, **h** & **i** were used for generation of stable transgenic
1029 Physcomitrella lines. The vectors contain either the full length pL1 including the native nuclear localization signal (NLS) (**g**), a
1030 truncated version lacking the NLS (**h**), or the full-length version including the native NLS and an N-terminal chloroplast transit
1031 peptide (**i**). All vectors contain the PpActin5 promoter (yellow), the CaMV 35S terminator (grey), and a hygromycin
1032 (hpt)(turquoise) resistance cassette for selection. Other elements are the GGGGGA linker (red), CTP (orange), NLS (blue) and eGFP
1033 or Citrine (green). Genetic elements are not depicted true to scale. Vector sizes: 7260 bp (**a**), 7461 bp (**b**), 7299 bp (**c**), 8802 bp
1034 (**d**), 8736 bp (**e**), 9021 bp (**f**), 7998 bp (**g**), 7932 bp (**h**), 8217 bp (**i**).

Supplementary Figure S2 - Generation of Physcomitrella-optimized (p)L1

HPV-16 hL1

ATGTCCCTGTGGCTGCCCAGCGAGGCCACCGTGTACCTGCCCCCCGTGCCCGTGAGCAAGGTGGTGAGCACCGATGAGTACGTGGCCCGGA
CCAACATCTACTACCACGCCGGCACCTCCAGACTGCTGGCCGTGGGCCACCCCTACTTCCCCATCAAGAAGCCCAACAACAACAAGATCCTGG
TGCCCAAGGTGAGCGGCCTGCAATACCGGGTGTTCAGAATCCACCTGCCCGACCCCAATAAGTTCGGCTTCCCCGACACCAGCTTCTACAACC
CCGACACCCAGAGACTGGTGTGGGCCTGCGTGGGCGTGGAGGTGGGCAGAGGCCAGCCTCTGGGCGTGGGCATCAGCGGCCACCCTCTGC
TGAACAAGCTGGACGACACCGAGAACGCCAGCGCCTACGCCGCCAACGCCGGCGTGGATAACAGAGAATGCATCAGCATGGACTACAAGCA
GACCCAGCTGTGCCTCATCGGCTGCAAGCCCCCCATCGGCGAGCACTGGGGCAAGGGCAGCCCCTGCACCAACGTGGCCGTGAATCCTGGC
GACTGTCCTCCCCTGGAACTCATCAACACCGTGATCCAGGACGGCGACATGGTGGACACCGGCTTCGGCGCCATGGACTTCACCACCCTCCAG
GCCAATAAGAGCGAGGTGCCCCTGGACATCTGCACCAGCATCTGCAAGTACCCCGACTACATCAAGATGGTGAGCGAGCCCTACGGCGATAG
CCTGTTCTTCTACCTGCGGCGGGAGCAGATGTTCGTGCGGCACCTGTTCAACAGAGCCGGCGCCGTGGGCGAGAACGTGCCCGACGACCTGT
ACATCAAGGGCAGCGGCAGCACCGCCAACCTGGCCAGCAGCAACTACTTCCCTACCCCCAGCGGCTCCATGGTGACCAGCGACGCCCAGATC
TTCAACAAGCCCTACTGGCTCCAGAGAGCCCAGGGCCACAACAATGGCATCTGCTGGGGCAACCAGCTGTTCGTGACCGTGGTGGATACCAC
CCGGAGCACCAACATGTCCCTGTGCGCCGCCATCAGCACCAGCGAGACCACCTACAAGAACACCAACTTCAAGGAGTACCTGAGGCACGGCG
AGGAGTACGACCTCCAGTTCATCTTCCAGCTGTGCAAGATCACCCTCACCGCCGACGTGATGACCTACATCCACAGCATGAACAGCACCATCC
TGGAGGACTGGAACTTCGGCCTGCAGCCCCCTCCTGGCGGCACCCTGGAGGACACCTACAGATTCGTGACCAGCCAGGCCATCGCATGCCAG
AAGCACACCCCTCCCGCCCCTAAGGAGGACCCCCTGAAGAAGTACACCTTCTGGGAGGTGAACCTGAAGGAGAAGTTCAGCGCCGACCTGG
ACCAGTTCCCTCTGGGCAGAAAGTTCCTGCTGCAAGCCGGCCTGAAGGCCAAGCCTAAGTTCACCCTGGGCAAGAGAAAGGCCACCCCCACC
ACAAGCAGCACCAGCACCACCGCCAAGCGGAAGAAGCGCAAGCTGTGA

HPV-16 pL1

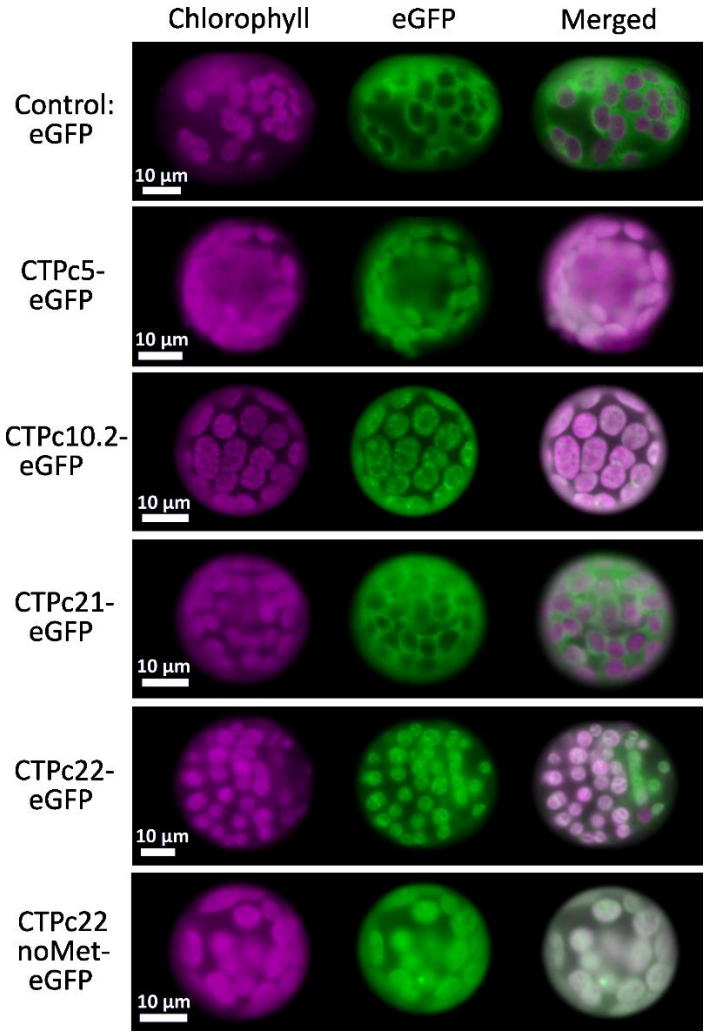
ATGTCCCTGTGGCTGCCCAGCGAGGCCACCGTGTACCTGCCCCCCGTGCCCGTGAGCAAAGTGGTGAGCACCGATGAGTACGTGGCCCGGAC
CAACATCTACTACCACGCCGGCACCTCCAGACTGCTGGCCGTGGGCCACCCCTACTTCCCCATCAAGAAGCCCAACAACAACAAGATCCTGGT
GCCCAAAGTGAGCGGCCTGCAGTACCGGGTGTTCAGAATCCACCTGCCCGACCCCAACAAGTTCGGCTTCCCCGACACCAGCTTCTACAACCC
CGACACCCAGAGACTGGTGTGGGCCTGCGTGGGCGTGGAAGTGGGCAGAGGCCAGCCTCTGGGCGTGGGCATCAGCGGCCACCCTCTGCT
GAACAAGCTGGACGACACCGAGAACGCCAGCGCCTACGCCGCCAACGCCGGCGTGGATAACAGAGAGTGCATCAGCATGGACTACAAGCA
GACCCAGCTGTGCCTCATCGGCTGCAAGCCCCCCATCGGCGAGCACTGGGGCAAGGGCAGCCCCTGCACCAACGTGGCCGTGAACCCTGGC
GACTGCCCTCCCCTGGAGCTCATCAACACCGTGATCCAGGACGGCGACATGGTGGACACCGGCTTCGGCGCCATGGACTTCACCACCCTCCA
GGCCAACAAGAGCGAAGTGCCCCTGGACATCTGCACCAGCATCTGCAAGTACCCCGACTACATCAAGATGGTGAGCGAGCCCTACGGCGATA
GCCTGTTCTTCTACCTGCGGCGCGAGCAAATGTTTGTGCGGCACCTGTTCAACAGAGCCGGCGCCGTGGGCGAGAACGTGCCCGACGACCTG
TACATCAAGGGCAGCGGCAGCACCGCCAACCTGGCCAGCAGCAACTACTTCCCTACCCCCAGCGGCTCCATGGTGACCAGCGACGCCCAGAT
CTTCAACAAGCCCTACTGGCTCCAGAGAGCCCAGGGCCACAACAACGGCATCTGCTGGGGCAACCAGCTGTTCGTGACCGTGGTGGATACCA
CCCGGAGCACCAACATGTCCCTGTGCGCCGCCATCAGCACCAGCGAGACCACCTACAAGAACACCAACTTCAAGGAGTACCTGAGGCACGGC
GAGGAGTACGACCTCCAGTTCATCTTCCAGCTGTGCAAGATCACCCTCACCGCCGACGTGATGACCTACATCCACAGCATGAACAGCACCATC
CTGGAGGACTGGAACTTCGGCCTGCAGCCCCCTCCTGGCGGCACCCTGGAGGACACCTACAGATTCGTGACCAGCCAGGCCATCGCATGCCA
GAAGCACACCCCTCCCGCCCCTAAGGAGGACCCCCTGAAGAAGTACACCTTCTGGGAAGTGAACCTGAAGGAGAAGTTCAGCGCCGACCTG
GACCAGTTCCCTCTGGGCAGAAAGTTCCTGCTGCAGGCCGGCCTGAAGGCCAAGCCTAAGTTCACCCTGGGCAAGAGAAAGGCCACCCCCAC
CACAAGCAGCACCAGCACCACCGCCAAGCGGAAGAAGCGCAAGCTGTGA

1073 **Human codon-optimized HPV-16 (h)L1 CDS and Physcomitrella-optimized HPV-16 (p)L1 CDS.** Grey marked nucleotides were
1074 changed in HPV-16 pL1 through codon usage optimization by the tool physCO (Top et al., 2021). Turquoise marked motifs refer
1075 to the Physcomitrella alternative splicing recognition sites. The green marked segment refers to a predicted miRNA binding site
1076 identified by the tool psRNATarget (Dai & Zhao, 2011; Dai et al., 2018). Light orange coloured segments represent the nuclear
1077 localization signal. Nucleotide changes are underlined in the HPV-16 pL1 sequence. (hL1, GenBank accession no. DQ067889,
1078 (Maclean et al., 2007))

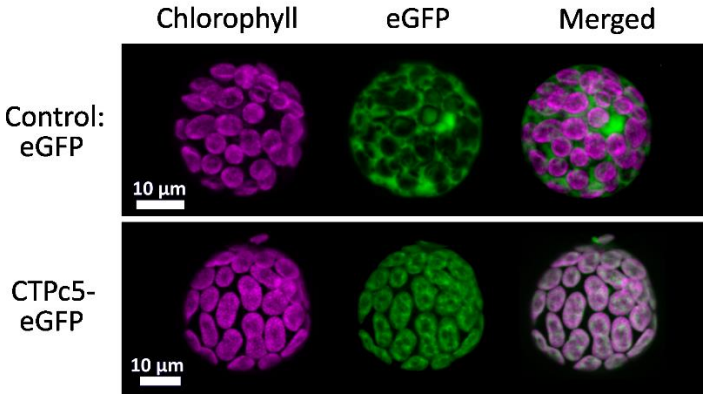
1079 Supplementary Figure S3 - Characterization of CTPs by analysis of subcellular eGFP

1080 localization

1081



1082



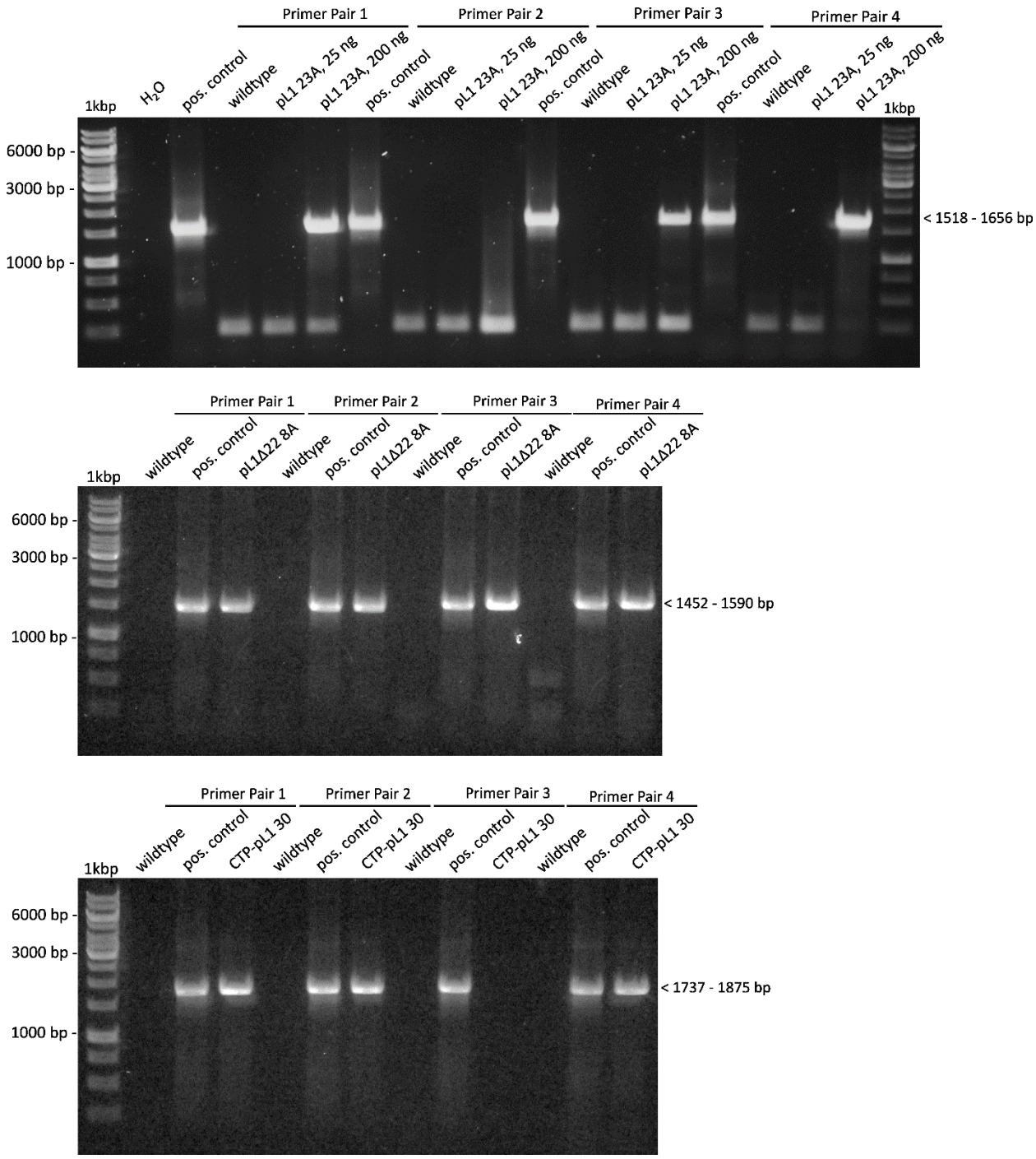
1083

1084

1085 **Microscopy pictures of Physcomitrella protoplasts transformed with CTP-eGFP constructs.** Samples were analysed using a
1086 rhodamine and a FITC filter to identify chloroplast (magenta) and eGFP localization, respectively, as well as a merge of the two
1087 signals. For sample CTPc21-eGFP no eGFP signal could be observed in chloroplasts. For samples CTPc5-eGFP, CTPc10-eGFP,
1088 CTPc22-eGFP and CTPc22noMet-eGFP green fluorescence signal is observed in chloroplasts. eGFP lacking a CTP localized to the
1089 cytoplasm and served as a negative control. The experiment was repeated two times. CTPc5 was examined a third time using
1090 confocal microscopy. Scale bars represent 10 µm.

1091 ## Supplementary Figure S4 – Analysis of CDS integrity in L1-producing lines

1092



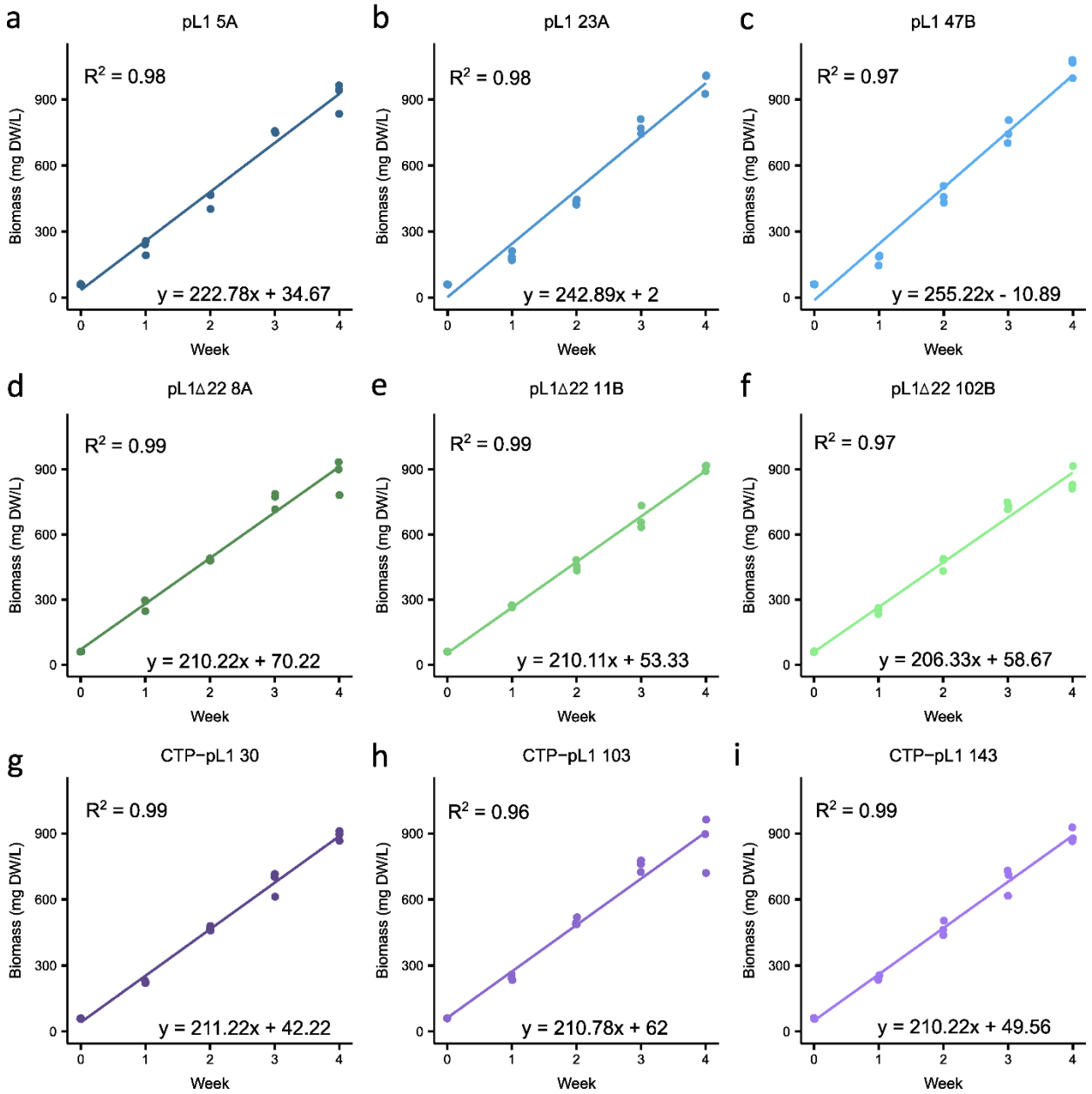
1093

1094

1095 **Analysis of rearrangements in the pL1 CDS using RT PCR.** Lines pL1 23A, pL1Δ22 8A and CTP-pL1 30 were checked for DNA or
1096 RNA rearrangements by performing PCR on cDNA using four different primer pairs. Primer pairs were selected to amplify the
1097 entire cDNA. The forward primer binds at the start codon and is the same in each primer pair (pL1_fwd_1 for lines pL1 23A &

1098 pL1Δ22 8A, CTP_pL1_fwd_1 for line CTP-pL1 30). The reverse primers are different in each pair and bind at the 3' end of the CDS
1099 or within the terminator (**Supplementary Table S1**). For line pL1 23A usage of 25 ng or 200 ng cDNA in PCR reactions was
1100 compared. Following experiments used 200 ng. Absence of distinct bands of a size smaller than the expected size is indicative for
1101 no larger rearrangements on both DNA and RNA level. No amplicon is visible for lines pL1 23A using Primer Pair 2 and CTP-pL1 30
1102 using Primer Pair 3. The constructs used for transformation of Physcomitrella protoplasts (**Supplementary Figure S1 g - i**) served
1103 as a positive control.

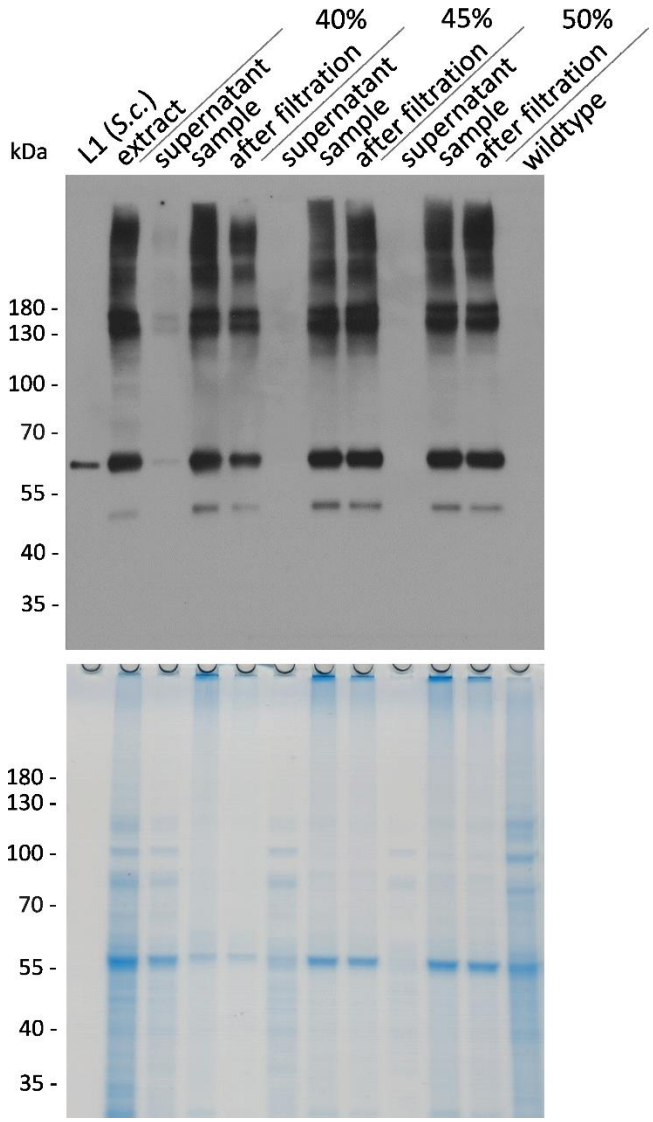
Supplementary Figure S5 – Comparison of biomass accumulation rates



Analysis of biomass accumulation rates among nine Physcomitrella lines of three different constructs. Biomass accumulation rates were observed over a course of 4 weeks for lines pL1 5A, pL1 23A, pL1 47B (blue); pL1Δ22 8A, pL1Δ22 11B, pL1Δ22 102B (green) and CTP-pL1 30, CTP-pL1 103, CTP-pL1 143 (purple). The data was plotted using linear regression analysis and all slopes were analysed using ANOVA and post hoc multiple comparisons test (TukeyHSD). Comparison of the slopes between the constructs of pL1, pL1Δ22 and CTP-pL1 showed pL1 lines accumulated biomass faster than pL1Δ22 ($p = <0.0001$) and CTP-pL1 ($p = 0.0001$) lines. Per week n = 3 biological samples were taken.

1113 ## Supplementary Figure S6 – Ammonium sulphate precipitation

1114

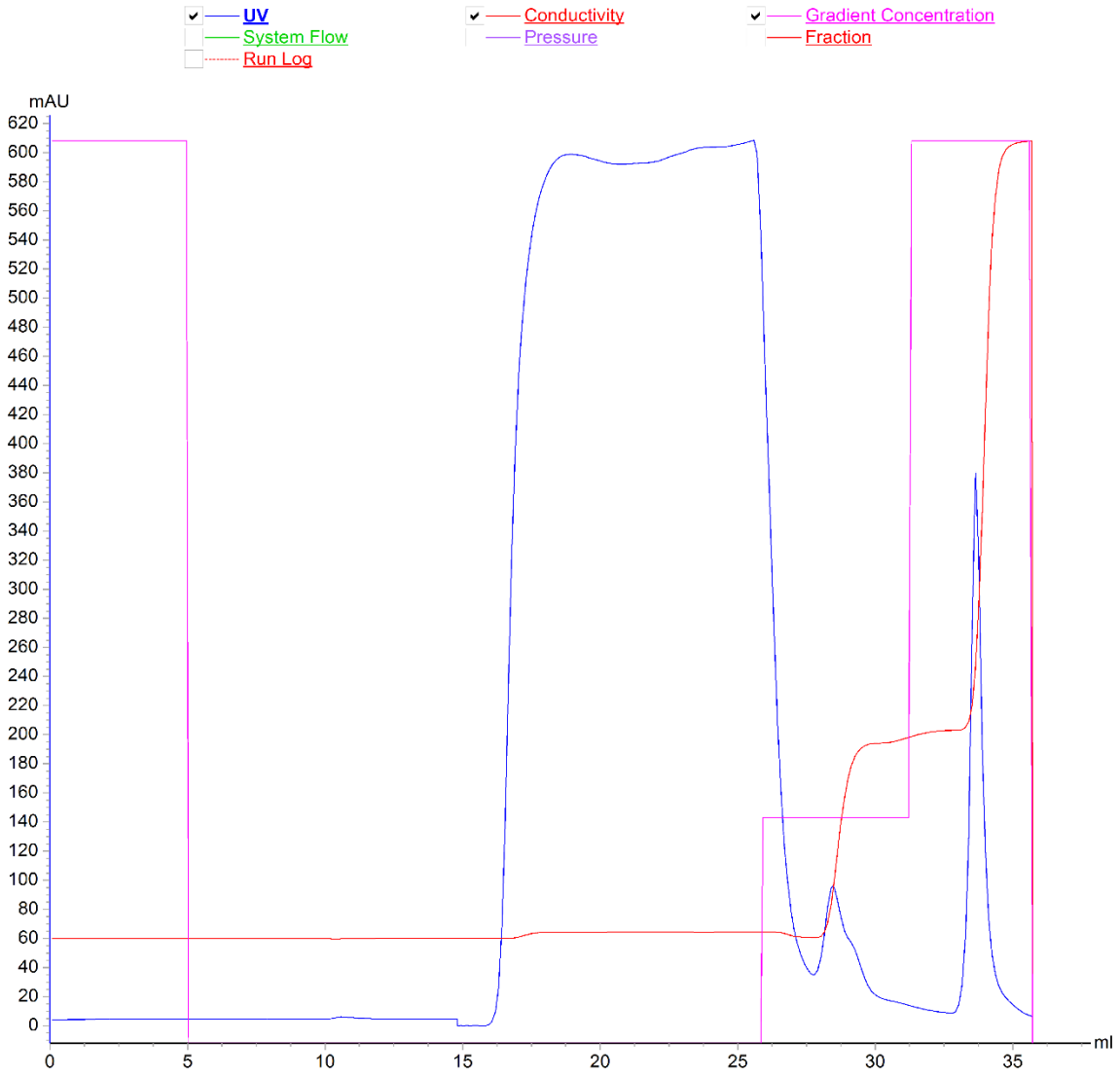


1115

1116

1117 **Ammonium sulphate precipitation of L1 from moss extracts.** Samples from the supernatant and the resuspended and dialysed
1118 pellet before and after filtration were analysed for the presence of L1 content and residual host proteins under reducing
1119 conditions *via* Western blot and Coomassie-stained gel, respectively. CTP-pL1 30 extracts were precipitated with either 40%, 45%,
1120 or 50% ammonium sulphate (AMS). After centrifugation, the supernatant and the precipitate were taken up and dialysed against
1121 cation exchange chromatography binding buffer to remove excess AMS. Non-dissolved protein was cleared from the sample using
1122 0.22 µm filters. 45% AMS was found to be sufficient for full precipitation of L1 from crude extracts. 30 ng *S. cerevisiae*-produced
1123 L1 served as a positive control and Physcomitrella wildtype as a negative control.

Supplementary Figure S7 – Chromatogram of Cation Exchange purification



Chromatogram of HPV-16 L1 purification from moss using cation exchange chromatography. The chromatogram shows the UV absorbance, indicative for protein concentration, in blue, gradient concentration in pink, and conductivity, indicative for salt concentration, in red. The y axis depicts milli-Absorbance Units and the x axis depicts the volume in mL. The large peak between 16 – 26 mL shows the sample application or flow through, the following small peak at 28 mL the wash fraction and the larger peak at 33 – 35 mL the elution fractions.

1132 ## Supplementary Table S1 – List of primers

1133 Capital letters refer to the part of the primer binding to the template whereas lower case letters represent overhangs added to
1134 the template via PCR. Abbreviations: fwd or fw (forward), rev or rv (reverse), CTP (chloroplast transit peptide), MMEJ (micro
1135 homology-mediated end joining), GFP (green fluorescent protein), OH (overhang (lower case letters)).

Primer name	Sequence (5'→3')	Application
AA2_3integFP (hpt3'seq)	ACTCGCCGATAGTGGAAACC	Amp of linker-eGFP vector part 1
PAN_PpActin-Pro_Rv	CTGCACAACAACCACCAATATTTAATTTC	Amp of linker-eGFP vector part 1
linker:EGFP_p1840_Fw	GGAGGTGGAGGTGGAGC	Amp of linker-eGFP vector part 2
HPTLUCfwd2	TTGCCCTCGGACGAGTGC	Amp of linker-eGFP vector part 2
CTP_Pp3c10_Fw	aattaaatattggtggttgttgtgcagATGGCGTTATTGGCGCGC	Amp of CTPc10
CTP_Pp3c10_Rv	gctccacctccacctccATTGGTGACCCTCAAAGGTCCAG	Amp of CTPc10
CTP_Pp3c5_Fw	aattaaatattggtggttgttgtgcagATGGCTACGATGAAGATTGCTGTGG	Amp of CTPc5
CTP_Pp3c5_Rv	gctccacctccacctccAGCTCGCACGGAGCTGTTC	Amp of CTPc5
CTP_Pp3c22_Fw	aattaaatattggtggttgttgtgcagATGCGGGCTCTCTGTGGAG	Amp of CTPc22 & CTPc22_noMet
CTP_Pp3c22_Rv	gctccacctccacctccCATAGCGACAATGCACCGTCTTC	Amp of CTPc22
CTP_Pp3c22_noMet_Rv	gctccacctccacctccAGCGACAATGCACCGTCTTCC	Amp of CTPc22_noMet
CTP_Pp3c21_Fw	aattaaatattggtggttgttgtgcagATGGCCGCCGTCGGAATG	Amp of CTPc21
CTP_Pp3c21_Rv	gctccacctccacctccGGCACGGACAGCGAACGC	Amp of CTPc21
Act5_seq	TCCACATGGCTACAGCTG	Colony PCR CTP-eGFP constructs
EGFP_NoLinker_Rev	CTTGTACAGCTCGTCCATG	Colony PCR CTP-eGFP constructs
XhoI-pL1-BamHI_fw	gtgcagctcgagATGTCCCTGTGGCTGCC	Amp of pL1 and pL1Δ22
XhoI-pL1-BamHI _Rv	tctagaggatccTCACAGCTTGCGCTTCTTCC	Amp of pL1
XhoI-pL1delta22-BamHI _Rv	tctagaggatcctcaGCCCAGGGTGAACTTAGGC	Amp of pL1Δ22
pL1_w/oATG_Fw	TCCCTGTGGCTGCCCAG	Amp of CTP-pL1 (no Start codon)
pL1_Rv_35St-OH	GACTGGTGATTTTTGCGGACTCTAG	Amp of pL1, Heterosplicing and MMEJ check of pL1 lines

CTPc5_Fw_Act5p-OH	tattggtggttgttgtgcagctcgagATGGCTACGATGAAGATTGCTGTG	Amp of CTPc5
CTPc5_Rv_pL1-OH	ctcgctgggcagccacagggaAGCTCGCACGGAGCTGTTC	Amp of CTPc5
Seq_hL1+HSP_Fw	CGTGATGACCTACATCCACAG	Colony PCR pL1 constructs
pL1_Rv_35St-OH	GACTGGTGATTTTTGCGGACTCTAG	Colony PCR pL1 constructs
pL1_fwd_1	ATGTCCCTGTGGCTGCCC	Heterosplicing and MMEJ check of pL1, pL1Δ22 lines (primer pair 1-4, forward)
pL1_rev_1	TCACAGCTTGCGCTTCTTCC	Heterosplicing and MMEJ check of pL1, CTP-pL1 lines (primer pair 1, reverse)
pL1d22_rev_1	TCAGCCCAGGGTGAACTTAGG	Heterosplicing and MMEJ check of pL1Δ22 lines (primer pair 1, reverse)
CTP-pL1_fwd_1	ATGGCTACGATGAAGATTGCTGTG	Heterosplicing and MMEJ check of CTP-pL1 lines (primer pair 1-4, forward)
pL1_Rv_35St-OH	GACTGGTGATTTTTGCGGACTCTAG	Heterosplicing and MMEJ check of pL1, pL1Δ22, CTP-pL1 lines (primer pair 2, reverse)
101_35S-T_rev	CCCTTATCTGGGAACTACTCACACATTATTCTGGA	Heterosplicing and MMEJ check of pL1, pL1Δ22, CTP-pL1 lines (primer pair 3, reverse)
3342_LdC-HATag-35ST-rev	GGGTTTCGCTCATGTGTTGAGCA	Heterosplicing and MMEJ check of pL1, pL1Δ22, CTP-pL1 lines (primer pair 4, reverse)
pL1-fwd	GCTCGAGATGTCCCTGTG	pL1(Δ22)-GFP fusion construct
pL1-rev	AGGATCCCAGCTTGCGCTTCTTC	(CTP-)pL1-GFP fusion construct
pL1d22-rev	AGGATCCGCCCAGGGTGAACTTAG	pL1Δ22-GFP fusion construct
CTPc5-pL1 fwd	GCTCGAGATGGCTACGATGA	CTP-pL1-GFP fusion construct

1136
1137

1138 Supplementary Table S2 – Predicted chloroplast transit peptides for six selected proteins
1139 in Physcomitrella (the predictions were performed with TargetP-2.0; Armenteros et al., 2019)
1140

Name: CTPc5

Protein: **Pp3c5_1370V3.1**

Predicted chloroplast transit peptide (74 aa):

MATMKIAVAAVAGTPAAAVGTRSAAATETARLCASSAQGLGLSTFVGLRAGSNASRGVTPRVRPLKAENSSVRA

Name: CTPc10.1

Protein: **Pp3c10_10490V3.1**

Predicted chloroplast transit peptide (63 aa):

MEVLPRAIATSHVVVSSSSAPKESSESGSKNNSQVSLSSVPFLRPASALAASIVSSATPAALA

Name: CTPc10.2

Protein: **Pp3c10_2940V3.1**

Predicted chloroplast transit peptide (70 aa):

MALLARGAVAAVAAVGFSNTTSKQGHQNGGSNGVLAFRGNAFCRAVAARTSSVGCQVRVKRGAGPLRVTN

Name: CTPc21

Protein: **Pp3c21_9980V3.1**

Predicted chloroplast transit peptide (43 aa):

MAAVGMTMSLKSSTALKGDFLGAALKQTVAAPKAANVAFAVRA

Name: CTPc22

Protein: **Pp3c22_5470V3.1**

Predicted chloroplast transit peptide (62 aa):

MRALCGVPSCASLSSHRLASPLLAYRACEFALGSSPGLRLEFVGRAVGLQSRSIGRRCIVAM

1141