Methods

An Agrobacterium-mediated stable transformation technique for the hornwort model Anthoceros agrestis

Eftychios Frangedakis¹, Manuel Waller²,³, Tomoaki Nishiyama⁴, Hirokazu Tsukaya⁵, Xia Xu⁶, Yuling Yue²,³, Michelle Tjahjadi⁶, Andika Gunadi⁶, Joyce Van Eck⁶,⁷, Fay-Wei Li⁶,⁸, Péter Szövényi²,³* and Keiko Sakakibara⁹**

¹Department of Plant Sciences, University of Cambridge, Cambridge, CB3 EA, UK

²Department of Systematic and Evolutionary Botany, University of Zurich, 8008, Switzerland

³Zurich-Basel Plant Science Center, Zurich, 8092, Switzerland

⁴Advanced Science Research Center, Kanazawa University, Ishikawa, 920-8640, Japan

⁵Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, 113-0033, Japan

⁶Boyce Thompson Institute, Ithaca, New York, 14853-1801, USA

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/NPH.17524

This article is protected by copyright. All rights reserved
7Plant Breeding and Genetics Section, Cornell University, Ithaca, New York, 14853-1801, USA
8Plant Biology Section, Cornell University, Ithaca, New York, 14853-1801, USA
9Department of Life Science, Rikkyo University, Tokyo, 171-8501, Japan

Received: 17 March 2021
Accepted: 20 May 2021

ORCID IDs

Eftychios Frangedakis: 0000-0002-3483-8464
Manuel Waller: 0000-0002-6060-0740
Tomoaki Nishiyama: 0000-0003-1279-7806
Hirokazu Tsukaya: 0000-0002-4430-4538
Andika Gunadi: 0000-0002-0097-6906
Joyce Van Eck: 0000-0002-8005-365X
Fay-Wei Li: 0000-0002-0076-0152
Péter Szövényi: 0000-0002-0324-4639
Keiko Sakakibara: 0000-0003-4420-9351

*corresponding author email address: peter.szoevenyi@uzh.ch, +41 634 84 40
**corresponding author email address: bara@rikkyo.ac.jp

This article is protected by copyright. All rights reserved
Summary

- Despite their key phylogenetic position and their unique biology, hornworts have been widely overlooked. Until recently there was no hornwort model species amenable to systematic experimental investigation. *Anthoceros agrestis* has been proposed as the model species to study hornwort biology.
- We have developed an *Agrobacterium*-mediated method for the stable transformation of *A. agrestis*, a hornwort model species for which a genetic manipulation technique was not yet available.
- High transformation efficiency was achieved by using thallus tissue grown under low-light conditions. We generated a total of 274 transgenic *A. agrestis* lines expressing the β-glucuronidase (GUS), cyan, green, and yellow fluorescent proteins under the control of the CaMV 35S promoter and several endogenous promoters. Nuclear and plasma membrane localization with multiple color fluorescent proteins was also confirmed.
- The transformation technique described here should pave the way for detailed molecular and genetic studies of hornwort biology, providing much needed insight into the molecular mechanisms underlying symbiosis, carbon-concentrating mechanism, RNA editing, and land plant evolution in general.

Keywords: *Anthoceros*, development, evolution, hornworts, transformation.

Introduction

The hornworts are one of the three lineages of bryophytes that diverged from the monophyletic group of liverworts and mosses over 460 million years ago (Morris *et al.*, 2018; One Thousand Plant Transcriptomes Initiative, 2019; Li *et al.*, 2020). While having only around 220 extant species (Söderström *et al.*, 2016), hornworts are key to address diverse questions about land plant evolution and terrestrialization. Hornworts display a unique combination of features (Frangedakis *et al.*, 2020) such as a sporophyte, that is produced by an indeterminate basal meristem, and has stomata similar
to mosses and vascular plants (Renzaglia et al., 2017). In addition, it is the only extant land plant lineage (together with a few Selaginella species (Liu et al., 2020)), that has a single (or just a few) algal-like chloroplast(s) per cell. The chloroplasts resemble those of algae in that they may contain pyrenoids, a carbon-concentrating structure that is shared with many algal lineages (Villarreal & Renner, 2012; Li et al., 2017). Hornwort plastids are also unique by exhibiting the highest RNA editing rates amongst land plants (Yoshinaga, 1996, 1997; Kugita, 2003; Small et al., 2019). Finally, hornworts are among the very few plant lineages that can establish symbiotic relationships with both endophytic cyanobacteria (Renzaglia et al., 2009) and various glomeromycotina and mucoromycotina fungal partners (mycorrhiza) (Desirò et al., 2013).

*Anthoceros agrestis* has been established as an experimental model system for hornworts and two isolates are currently available (Oxford and Bonn) (Szövényi et al., 2015). *A. agrestis*, like other bryophytes, has a haploid-dominant life cycle through which the haploid gametophyte phase alternates with the diploid sporophyte phase. The life cycle of *A. agrestis* starts with the germination of the haploid spores which develop into an irregularly shaped thallus (Fig. 1a-c). Sexual reproduction occurs through fusion of the egg (produced in archegonia) and the motile sperm (produced in antheridia). The resulting embryo develops within the gametophyte and gives rise to the sporophyte (Fig. 1d). *A. agrestis* can be easily grown under laboratory conditions and its haploid-dominant life cycle makes genetic analysis straightforward. The nuclear genome of *A. agrestis* was recently sequenced (Li et al., 2020; available at https://www.hornworts.uzh.ch/en/about.html) which is one of the smallest genomes amongst land plants (genome size: 116/123 Mb for the Bonn and Oxford isolates, respectively). The species is monoicous, with male and female reproductive organs produced by the same individual. The sexual life cycle of *A. agrestis* can be completed under laboratory conditions within approximately 2-3 months (Szövényi et al., 2015). Efficient genetic manipulation methods have been published for the mosses *Physcomitrium patens* (Rensing et al., 2020), *Ceratodon purpureus* (Zeidler et al., 1999; Trouiller et al., 2007; Finiuk et al., 2014), *Scopelophila cataractae* (Nomura et al., 2016), and the liverworts *Marchantia polymorpha* (Kohchi et al., 2021) and *Riccia fluitans* (Althoff & Zachgo, 2020). However, transformation of *A. agrestis* was not feasible until recently, posing a major obstacle to the analysis of gene function in hornworts, and more generally to land plant evo-devo studies.
Several approaches have been used for gene delivery in bryophytes, including polyethylene glycol (PEG)-mediated uptake of DNA by protoplasts (Schaefer et al., 1991), particle bombardment (Cho et al., 1999; Chiyoda et al., 2008) and *Agrobacterium tumefaciens* (revised scientific name *Rhizobium radiobacter* (Young et al., 2001)) mediated transformation (Ishizaki et al., 2008; Kubota et al., 2013; Althoff & Zachgo, 2020). *Agrobacterium*-mediated transformation is a commonly used method for various plant species (Gelvin, 2003), it is relatively simple and does not require specialised or expensive equipment. In addition, *Agrobacterium*-mediated transformation has several advantages over other transformation methods, such as the integration of a lower number of transgene copies into the plant genome and the ability to transfer relatively large DNA segments with intact transgene genome integration.

In this study we report the first successful stable genetic transformation method for hornworts. The method is based on *Agrobacterium*-mediated transformation of *A. agrestis* thallus. We also show the successful expression and targeting of four different fluorescent proteins in two different cellular compartments, the plasma membrane, and the nucleus. Finally, we characterise a number of native *A. agrestis* promoters for their potential to drive strong constitutive transgene expression that can be useful for future hornwort genetic studies.

**Materials and methods**

**Plant material and maintenance**

In this study we used the *Anthoceros agrestis* Oxford and Bonn isolates (Szövényi et al., 2015). *A. agrestis* thallus tissue was propagated on KNOP medium (0.25 g/L KH₂PO₄, 0.25 g/L KCl, 0.25 g/L MgSO₄•7H₂O, 1 g/L Ca(NO₃)₂•4H₂O and 12.5 mg/L FeSO₄•7H₂O). The medium was adjusted to pH 5.8 with KOH and solidified using 7.5 g/L Gelzan CM (SIGMA) in 92x16 mm petri dishes (SARSTEDT) with 25-30 mL per plate. Plants were routinely grown in a tissue culture room (21°C, 12 h of light and 12 h of dark, 3-5 or 35 μmol m⁻² s⁻¹ light intensity, Philips TL-D 58W (835)). To subculture the thallus tissue, a small part of it was cut using sterile disposable scalpels.
For sporophyte induction, *A. agrestis* Bonn gametophyte tissue was subcultured as described above and grown for 2-3 months. Fertilization was facilitated by adding 5 mL sterile water per petri dish when the first archegonia were observed. Several sporophytes appeared after two weeks, at which point excess moisture was removed from the petri dishes and the sporophytes were allowed to grow further and mature. Once the majority of sporophytes started to turn brown at their tips, the petri dishes were removed from the humidity controlled light chamber and put into a dryer environment (e.g., lab bench) to facilitate sporophyte maturation. Once all the sporophytes have turned completely brown and looked dry, they were collected with tweezers under a laminar hood and stored up to 8 months at 4°C (in a fridge) in 60x15 mm petri dishes (#82.1194.500, SARSTEDT). Sporophytes were used for downstream applications at the earliest after three months of storage at 4°C to allow for complete desiccation.

**Co-cultivation medium**

Co-cultivation medium was liquid KNOP supplemented with 2% sucrose (0.25 g/L KH$_2$PO$_4$, 0.25 g/L KCl, 0.25 g/L MgSO$_4$•7H$_2$O, 1 g/L Ca(NO$_3$)$_2$•4H$_2$O, 12.5 mg/L FeSO$_4$•7H$_2$O and 20 g/L sucrose, pH 5.8 adjusted with KOH).

**Tissue preparation for transformation**

Approximately 2 g of thallus tissue (15-20 petri dishes) was divided into 4 parts, and each part (approximately 0.5 g) was homogenised in 15 mL of sterile water using a homogeniser (#727407, IKA Ultra-Turrax T25 S7 Homogenizer) and corresponding dispensing tools (#10442743, IKA Dispersing Element), for 5 seconds, using the lowest speed of 8000 rpm. The homogenized tissue was washed with 50 mL of sterile water using a 100 µm cell strainer (#352360, CORNING), plated onto solid KNOP medium and placed at 21°C under 12 hours light and 12 hours dark at a light intensity of 3-5 µmol m$^{-2}$ s$^{-1}$. After 4 weeks the tissue was re-homogenised in 15-20 mL of sterile water and filtered using a 100 µm cell strainer. The re-homogenised tissue was transferred again
onto 4 plates with solid KNOP medium and was allowed to grow for 2 days at 21°C under continuous light (35 μmol m\(^{-2}\) s\(^{-1}\), PHILIPS, TL-D58W/835).

For sporeling-transformation, 20 mature and desiccated sporophytes were pooled in a 1.5 mL Eppendorf tube and the spores were surface sterilised as described in Sauret-Güeto et al., 2020. The spores were incubated in the sterilizing solution for 15 minutes and then centrifuged at 10000 x g for 2 minutes. The spore pellet was resuspended in 150 μL fresh sterilizing solution and plated onto solid KNOP medium with an additional 5 mL liquid KNOP medium added to the petri dish. The petri dishes were closed with micropore tape and kept at 21°C under 12 hours light and 12 hours dark at a light intensity of 35 μmol m\(^{-2}\) s\(^{-1}\). The first germinating spores were observed 5-8 days after plating.

Sporelings were let to grow for 7, 14 and 21 days (after the observation of the first spore germination). For Agrobacterium co-cultivation, the sporelings were scraped off from the surface of the solid KNOP medium with sterile scalpels and collected and washed using a 40 μm cell strainer (#431750, CORNING). Sporelings from a single petri dish (20 sporophytes) were transferred into a well of a 6-well plate with Agrobacterium and transformation buffer (for co-cultivation conditions see Agrobacterium culture preparation and co-cultivation conditions).

**Agrobacterium culture preparation**

One to three Agrobacterium colonies (AGL1 strain) were inoculated in 5 mL of LB medium supplemented with rifampicin 10 μg/mL (#R0146, Duchefa), carbenicillin 50 μg/mL (#C0109, MELFORD) and the plasmid-specific selection antibiotic spectinomycin 100 μg/mL (#SB0901, Bio Basic). The pre-culture was incubated at 28°C for 2 days with shaking at 120 rpm. OD600 was ~2.7 and was measured using an OD600 DiluPhotometer (IMPLEN).

**Co-cultivation conditions**

5 mL of 2 days Agrobacterium culture was centrifuged for 7 minutes at 2000 xg. The supernatant was discarded, and the pellet was resuspended in 5 mL liquid KNOP supplemented with 2% (w/v) sucrose (#S/8600/60, Fisher) and 100 μM 3',5'-dimethoxy-4'-hydroxyacetophenone (acetosyringone) (#115540050, Acros Organics, dissolved in dimethyl sulfoxide (DMSO) (#D8418, SIGMA)). The culture was incubated with shaking (120 rpm) at 28°C for 5 hours. The regenerating thallus tissue
was transferred (1 plate – 2 days after the second homogenisation) into a well of a 6-well plate with 5 mL of liquid KNOP medium supplemented with 2% (w/v) sucrose. 80 μL of *Agrobacterium* culture and acetosyringone at final concentration of 100 μM were added to the medium.

The tissue and *Agrobacterium* were co-cultivated using a 6-well plate (#140675, ThermoFisher) for 3 days with shaking at 110 rpm at 22°C on a shaker without any additional light supplemented (only ambient light from the room, 1-3 μmol m$^{-2}$ s$^{-1}$). After 3 days, the tissue was drained using a 100 μm cell strainer (#352360, CORNING) and moved onto solid KNOP plates (2 petri dishes from a single well) supplemented with 100 μg/mL cefotaxime (#BIC0111, Apollo Scientific) and 10 μg/mL Hygromycin (#10687010, Invitrogen). After 3-4 weeks, plants were transferred on fresh solid KNOP plates supplemented with 100 μg/mL cefotaxime and 10 μg/mL Hygromycin and grown at 22°C 12 hours light and 12 hours dark at light intensity of 35 μmol m$^{-2}$ s$^{-1}$ (PHILIPS, TL-D58W/835).

**GUS staining**

GUS staining was performed according to (Plackett *et al.*, 2014). GUS buffer stock solution was prepared by mixing 11.54 mL of 1M Na$_2$HPO$_4$, 8.46 mL of 1M NaH$_2$PO$_4$, 8 mL of 0.25M EDTA (pH7.0), 10μL of 20% Tween20 (v/v) and topped up to 200mL with dH$_2$O. 10mL of GUS staining solution was prepared by dissolving 10mg 5-bromo-4-chloro-3-indolyl-β-d-glucuronic acid (X-glcA) (#B72200-0.25, Melford) in 1mL DMSO and added to 9ml GUS buffer stock solution (final X-glcA concentration 1mg/mL). The tissue was first fixed in cold 90% (v/v) acetone for 10 minutes. The acetone was removed, and the tissue was rinsed with cold dH$_2$O. After that dH$_2$O was removed and GUS staining solution was added. The samples in the solution were placed under vacuum for 4 minutes twice to enable efficient infiltration to the tissues. Then the sample dish was wrapped in tinfoil and incubated at 37°C overnight. Prior to visualization, chlorophyll was removed (bleached) by a rising ethanol series (v/v): 15%, 30%, 50%, 70% and 100%.
Genomic DNA extraction

A modified CTAB protocol from (Porebski et al., 1997) was used for hornwort genomic DNA extraction. 0.5 g of tissue was harvested and frozen in liquid nitrogen. Tissue was ground into a fine powder using a chilled mortar and pestle and then added to 10 mL of DNA extraction buffer (100 mM Tris-HCl pH 8, 1.4 M NaCl, 20 mM EDTA pH 8, 2% (w/v) CTAB, 0.3% (v/v) β-mercaptoethanol and 100 mg of polyvinylpyrrolidone (PVP)/g of tissue) that had been prewarmed at 60°C, 100 μL of RNase A (100 mg/mL) was added and the solution was mixed well. The mix was incubated at 60°C for 30 minutes and then removed from heat and allowed to cool at room temperature for 4 minutes. 12 mL of chloroform:isoamyl alcohol (24:1) was added, mixed well by inversion and then centrifuged at 12000 g for 10 minutes at room temperature. The upper aqueous phase was transferred to a new 50 mL centrifugation tube and 10 mL of chloroform:isoamyl alcohol (24:1) was added, mixed well by inversion and then centrifuged at 6000 xg for 10 minutes at room temperature to remove any remaining PVP in the aqueous phase. The upper aqueous phase was transferred to a new 50 mL centrifugation tube and ½ volume of 5 M NaCl was added. 2 volumes of cold (-20°C) 95% (v/v) ethanol were also added and the contents of the tube were mixed well by inversion. The tube was spun at 20000 g for 6 minutes. The pellet was resuspended in 2 mL of TE buffer and the previous step was repeated. The pellet was washed with cold 70% (v/v) ethanol. The pellet was dried and dissolved in 500 μL of TE buffer and then stored at 4°C.

Promoter identification and isolation

We used RNA-seq data to find genes showing constantly high levels of expression under various developmental stages and experimental conditions ("constitutively expressed genes"). To do so, we estimated expression of genes under three developmental stages of the gametophyte and sporophyte phases and in symbiosis with cyanobacteria. We retrieved raw RNA-seq data for these experiments from (Li et al., 2020). We used trimmomatic to quality filter and trim the raw reads. Gene expression was estimated using Salmon (Patro et al., 2017) and expressed as normalized expression counts. We identified candidates by selecting those showing the highest average expression level and the least gene expression variability across all conditions investigated. We then manually selected a subset of genes taking account their genomic location, exact expression pattern, and the length and sequence composition of their putative promoter sequences. We also assessed
the suitability of our candidate promoters using the information available for *M. polymorpha* and *P. patens* (Supporting Information Table S1).

Putative promoter sequences were amplified from genomic DNA using the KOD Hot start polymerase (#71086-5, Merck Millipore) and cloned into pJET1.2 (#K1231, ThermoFisher) before Sanger sequencing. Loop assembly compatible DNA parts were generated according to (Sauret-Güeto *et al.*, 2020). Briefly, aliquots of the DNA parts (15 nM) and the pUAP4 acceptor vector (7.5 nM) were prepared. A type IIS assembly reaction was set up in a 0.2 mL tube (5 μL nuclease-free H2O, 1 μL 10× CutSmart buffer (#B7204S, NEB), 0.5 μL 1 mg/mL bovine serum albumin (#B9000S, NEB), 1 μL 10 mM Adenosine 5'-Triphosphate (ATP) (#P0756S, NEB), 0.25 μL 400 U/μL T4 DNA ligase (#M0202S, NEB), 0.25 μL 10 U/μL SapI (#R0569S, NEB), 1 μL of L0 DNA part, and 1 μL of pUAP4) to clone the amplified DNA part into pUAP4. Samples were incubated in a thermocycler using the following program. Assembly: 26 cycles of 37 °C for 3 min and 16 °C for 4 min. Termination and enzyme denaturation: 50 °C for 5 min and 80 °C for 10 min. 30 μL of chemically competent *E. coli* cells were transformed using 5 μL of the assembly reaction and then spread on LB agar plates containing 25 μg/mL chloramphenicol and 40 μg/mL X-gal. Construct sequences can be found in Supporting Information Table S2. List of primers can be found in Supporting Information Table S3. Promoter sequences, their genomic localization and RNA-seq coverage are also given in Notes S1.

**Construct generation**

Constructs were generated using the OpenPlant toolkit (Sauret-Gueto *et al.*, 2020). OpenPlant L0 parts used: OP-040 CTAG_linker-N7, OP-023 CDS12-eGFP, OP-020 CDS_hph, OP-027 CDS12_mTurquoise2, OP-029 CDS12_mVenus, OP-037, CTAG_Lti6b, OP-054 3TER_Nos-35S and OP-049 PROM_35S. **L1 construct generation**: L0 plasmids with the DNA parts to be assembled were prepared at a concentration of 15 nM, and the acceptor pCk vector at a concentration of 7.5 nM. The loop assembly Level 1 reaction master mix contained 2 μL nuclease-free H2O, 1 μL 10× CutSmart buffer (#B7204S, NEB), 0.5 μL 1 mg/mL bovine serum albumin (#B9000S, NEB), 1 μL 10 mM ATP (#P0756S, NEB), 0.25 μL 400 U/μL T4 DNA ligase (#M0202S, NEB), 0.25 μL 20 U/μL Bsal HF v2 (#R3733S, NEB). Cycling conditions were 26 cycles of 37 °C for 3 min and 16 °C for 4 min.
Termination and enzyme denaturation: 50 °C for 5 min, and 80 °C for 10 min. 30 μL of chemically competent *E. coli* cells were transformed using 5 μL of the Loop assembly reaction and spread on LB agar plates containing 50 μg/mL kanamycin and 40 μg/mL of X-Gal. **L2 construct generation:** The protocol is the same as the L1 BsaI assembly protocol with the exception of the master mix composition: 2 μL nuclease-free H₂O, 2 μL nuclease-free H₂O, 1 μL 10× CutSmart buffer (#B7204S, NEB), 0.5 μL 1 mg/mL bovine serum albumin (#B9000S, NEB), 1 μL 10 mM ATP (#P0756S, NEB), 0.25 μL 400 U/μL T4 DNA ligase (#M0202S, NEB), 0.25 μL 10 U/μL SapI (#R0569S, NEB) and spread on LB agar plates containing 100 μg/mL spectinomycin and 40 μg/mL of X-Gal. The donor plasmids were L1 constructs, and the acceptor plasmid was a pCsA vector (see Supporting Information Table S2 for full construct maps and sequences).

**Western blotting**

50 mg of *A. agrestis* thallus tissue was grown for 3 weeks on KNOP medium at 21 °C in continuous light (35 μmol m⁻² s⁻¹) and ground in liquid nitrogen. The tissue powder was resuspended in 500 μL 5× Laemmli loading buffer (0.2 M Tris-HCl pH 6.8, 5 % w/v SDS, 25 % v/v glycerol, 0.25 M DTT, 0.05 % w/v bromophenol blue, supplemented with Roche Complete protease inhibitor (# 11836170001, Roche), heated at 95 C° for 5 minutes and centrifuged at 10,000 g for 10 minutes. The supernatant was transferred to a new tube. Equal amounts of proteins were separated by denaturing electrophoresis in NuPAGE gel (#NP0322BOX, Invitrogen) and electro-transferred to nitrocellulose membranes using the iBlot2 Dry Blotting System (ThermoFisher). eGFP was immuno detected with anti-GFP antibody (1:4000 dilution) (JL-8, #632380, Takara) and anti-mouse-HRP (1:15000 dilution) (#A9044, Sigma) antibodies. Actin was immuno detected with anti-actin (plant) (1:1500 dilution) (#A0480, Sigma) and anti-mouse-HRP (1:15000 dilution) (#A9044, Sigma) antibodies, using the iBind™ Western Starter Kit (#SLF1000S, ThermoFisher). Western blots were visualised using the ECL™ Select Western Blotting Detection Reagent (#GERPN2235, GE) following the manufacturer's instructions. Images were acquired using a Syngene Gel Documentation system G:BOX F3.

**Sample preparation for Imaging**

A gene frame (#AB0576, ThermoFisher) was positioned on a glass slide and 30 μL of KNOP medium with 1.2% (w/v) Gelzan CM (#G1910, SIGMA) were placed within the gene frame. A thallus fragment was placed within the medium-filled gene frame together with 30 μL of milliQ water. The frame was then
sealed with a cover slip. Plants were imaged immediately using a Leica SP8X spectral fluorescent confocal microscope.

For the regeneration test experiment, five thallus fragments were placed into a KNOP medium-filled gene frame as described above (three slides and 15 plants in total). Images were acquired on a daily basis, for a total duration of a week, using a Leica SP8X spectral fluorescent confocal microscope and a 10× air objective (HC PL APO 10×/0.40 CS2).

**Imaging with Confocal Microscopy**

Images were acquired on a Leica SP8X spectral confocal microscope. Imaging was conducted using either a 10× air objective (HC PL APO 10×/0.40 CS2) or a 20× air objective (HC PL APO 20×/0.75 CS2). Excitation laser wavelength and captured emitted fluorescence wavelength window were as follows: for mTurquoise2 (442 nm, 460−485 nm), for eGFP (488 nm, 498−516 nm), for mVenus and eYFP (515 nm, 522−540 nm), and for chlorophyll autofluorescence (488 or 515, 670−700 nm). When observing lines expressing both eGFP and mTurquoise2, sequential scanning mode was used.

**Light microscopy**

Images were captured using a KEYENCE VHX-S550E microscope (VHX-J20T lens) or a Leica M205 FA Stereomicroscope (with GFP longpass (LP) filter).

**Sequencing transformant line genomes**

Transformant lines were grown on solid KNOP medium containing 10 μg/mL hygromycin and 100 μg/mL cefotaxime. Genomic DNA was extracted from 600 mg fresh tissue per line using either the DNeasy Plant Pro kit (#69204, Qiagen) (Cam-1 and Cam-2 lines) or the procedure from Li et al (2020) (BTI1-3 lines) to reach a total yield of at least 200 ng/line. Illumina libraries were prepared using the TruSeq DNA nano kit (#20015964, Illumina) and were sequenced on an Illumina Novaseq 6000 platform with an expected sequencing depth of 80-150x for all five lines in paired-end mode (read length: 151 bp).
After sequencing we quality filtered and trimmed reads using trimmomatic (command line: 
ALL_TruSeq-PE.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20
MINLEN:36) and assembled the reads with spades3.14.1 using the --isolate --cov-cutoff auto --only-assembler options as recommended (Nurk et al., 2013). To localize the insertion and its copy number, we used the insert sequence as a query in a BLASTN search (Altschul et al., 1990) against the database containing the assembly (Altschul et al., 1990) with an e-value threshold of $10^{-4}$. As evidence of genomic integration, we only accepted hits covering the full-length of the insert sequence with one or no mismatches. We manually inspected blast hits to eliminate false positives. Finally, we used the A. agrestis Oxford genome sequence (Li et al., 2020) to localize the hits on the genomic scaffolds.

**PCR confirmation of transformant insertion sites**

Based on the genome assemblies of the transformant lines, primers were designed to amplify regions of 0.6-1.3 kb spanning the 5’- and 3’- end of the T-DNA inserts and their respective adjacent genomic regions. Sequences were amplified from genomic DNA with Phusion High-Fidelity DNA Polymerase (#F-530S, ThermoFisher). For each amplified region, a second set of nested primers were designed to amplify a shorter amplicon, using 1:100 dilution of the previous PCR product as a template. Resulting nested PCR-products were either purified and Sanger sequenced (Eurofins) from both ends, or cloned into pJET1.2 (#K1231, ThermoFisher) before Sanger sequencing from both ends using pJET1.2 sequencing primers.

**Results**

We tested the potential of the Agrobacterium-mediated gene delivery method to recover stable A. agrestis transgenic lines. There are several critical factors that determine the efficiency of Agrobacterium-mediated gene delivery: (1) the selection of appropriate plant tissue and its preparation for infection, (2) the type and concentration of antibiotics applied to select for transgenic lines, (3) the choice of transformation vectors, (4) the choice of Agrobacterium strain, and (5) the optimal conditions for co-cultivation. Each of these factors is examined in this study.
Selection of tissue and optimal growth conditions

*Agrostis agrestis* gametophyte thallus was chosen as the appropriate tissue for transformation because it is easily accessible, has a remarkable regenerative capacity and is haploid. *Agrostis agrestis* thallus tissue cultures can easily be propagated and maintained by transfer of small thallus fragments (approximately 1 x 1 mm) onto fresh growth medium on a monthly basis (Supporting Information Fig. S1). *Agrostis* is similar to several other bryophytes in that an entire plant can regenerate from a small thallus fragment. This is in striking contrast with vascular plants, where usually the transition to an undifferentiated tissue state (callus), after treatment with extrinsic hormones such as auxin and cytokinin, is necessary for the regeneration of new plant tissue (Ikeuchi et al., 2013). The use of thallus has the additional advantage that the resulting transformants have a uniform genetic background.

We reasoned that similar to the liverwort *Marchantia polymorpha* (Kubota et al., 2013), fragmented tissue will be susceptible to *Agrobacterium* infection. In the case of *M. polymorpha*, the apical part of the thallus is removed to induce regeneration, followed by co-cultivation with *Agrobacterium* to generate transformed plants. *Agrostis* thallus regeneration is similarly induced by fragmentation and presumably removal of the apical parts of the thallus. Consequently, a transformation approach similar to the one used for *M. polymorpha* was utilized and adapted to *Agrostis*. However, unlike in *M. polymorpha*, the *Agrostis* notch area and apical cells are not easily distinguishable, thus determining which part of the thallus should be removed is not obvious (Frangedakis et al., 2020). Therefore, we tested whether homogenisation using dispensing tools is a suitable method for thallus tissue fragmentation. Different speed levels and duration of homogenisation were examined. We found that homogenisation of 0.5 g of thallus tissue in 15 mL of sterile water for 5 seconds (see Materials and methods) is sufficient to fragment the tissue and results in rapid tissue regeneration.

In addition, we found that the light intensity is a critical factor during the growth of the tissue used for transformation (Fig. 1e-g). Tissue that was grown under low light conditions, even though smaller in size, had a more regular and flattened shape and was optimal for transformation. (Fig. 1f-g and Supporting Information Fig. S1). When tissue was grown under high light intensity (above 40 μmol m⁻² s⁻¹), no transformants were obtained (for detailed explanation please see “Transformation efficiency and optimisation” section).
Besides fragmented thallus tissue, immature thalli grown from spores can also be used to achieve high numbers of stable transformants in *M. polymorpha* (Ishizaki et al., 2008). Therefore, we also tested the transformation of *A. agrestis* Bonn thalli grown from spores after 1, 2 and 3 weeks of germination. However, no stable transformants could be obtained using these tissues (see details in the materials and methods section).

**Selection of appropriate antibiotics**

A selectable marker gene, most commonly one that confers antibiotic-resistance, is necessary for efficient recovery of stable transgenic lines following co-cultivation with *Agrobacterium*. To identify antibiotics and their appropriate concentration with cytotoxic effect on *A. agrestis*, untransformed *A. agrestis* (Oxford and Bonn) thallus fragments were subjected to different concentrations of hygromycin and geneticin/G418 (an analogue of neomycin and kanamycin). The tested concentrations ranged from 0 μg/mL - 20 μg/mL for hygromycin and 0 μg/mL - 150 μg/mL for geneticin/G418. We found that a 3-week incubation period with 10 μg/mL hygromycin was sufficient to inhibit growth of untransformed thallus tissue for both the Oxford and Bonn isolates (Supporting Information Fig. S2 and S3), whereas thallus tissue was not susceptible to geneticin/G418 even when supplied in high concentration (150 μg/mL) (Supporting Information Fig. S4). Thus, we selected hygromycin as an appropriate selection agent for *A. agrestis* transformation.

**Preliminary tests with the GUS reporter**

Preliminary transformation experiments were performed using the Oxford isolate and the pCAMBIA1305.2 plasmid containing the *hygromycin B phosphotransferase* gene (*hph*, conferring hygromycin resistance) driven by the Cauliflower mosaic virus 35S (CaMV 35S, hereafter called 35S) promoter, terminated with a 35S polyadenylation signal, and a p-35S_s::GUSPlus (β-Glucuronidase) transcription unit. *GUSPlus* contains a catalase intron to ensure that the observed GUS expression is not due to the *Agrobacterium*.

Thallus grown under low light conditions (3-5 μmol m⁻² s⁻¹) (similar morphology with the tissue in (Fig.1f)) was homogenised, grown for one month under low light condition (3-5 μmol m⁻² s⁻¹) and
then homogenised for a second time (Fig. 1g-i). Two days after the second homogenisation, the regenerating thallus tissue was co-cultivated with the Agrobacterium AGL1 strain containing the pCAMBIA1305.2 plasmid, as well as Agrobacterium without a transformation vector as a negative control, in liquid KNOP media supplemented with sucrose. Media were also supplemented with 3',5'-dimethoxy-4'-hydroxyacetophenone (acetosyringone) since phenolic compounds such as acetosyringone have been shown to be important for the virulence genes activation (Stachel et al., 1985). Co-cultivation duration was 3 days at 22°C on a shaker without any light supplementation (only ambient light from the room). After co-cultivation, the tissue was spread on solid KNOP plates supplemented with cefotaxime and hygromycin. After 3-4 weeks the surviving, regenerating and putatively transformed tissue was transferred on fresh selective media (tissue morphology used for transformation is shown in Fig. 1g-i, transformation outline in Fig. 2 and Supporting Information Fig. S5 for detailed step-by-step protocol description). Emergence of rhizoids on surviving tissue fragments is a reliable indicator of successful transformation events (Fig. 3a,b). One to two months later successful transformants (plant fragments producing rhizoids) were visible. Finally, surviving plants were subjected to a third round of antibiotic selection to ensure false positives were eliminated. Thallus surviving selection on hygromycin (Fig. 3c) exhibited GUS expression (Fig. 3d, Supporting Information Fig. S6). No surviving plants were observed for those trials using the Agrobacterium without the transformation vector. These results indicated that A. agrestis Oxford thallus tissue is susceptible to Agrobacterium infection and that the 35S promoter driving the hph gene is sufficient for selection of transformants.

**Tests with eGFP as a reporter**

Subsequent experiments were carried out using the A. agrestis Oxford isolate and a construct containing the enhanced Green Fluorescent Protein (eGFP) reporter gene (Cormack et al., 1996). eGFP makes the identification of successful transformation events easier without the need of laborious GUS staining. For the construction of the eGFP transformation vector, we used the OpenPlant toolkit (Sauret-Gueto et al., 2020) which is based on the Loop assembly Type IIS cloning system (Pollak et al., 2019).

Loop Assembly is a Type IIS assembly system that alternates between two enzymes and, unlike other systems, requires only two sets of vectors (Fig. 3e). The restriction enzymes are Bsal which
generates four base 5’ overhangs and SapI which generates three base 5’ overhangs. Alternate use of BsaI and SapI and corresponding vector sets allows efficient and parallel assembly of large DNA circuits. In a single step reaction, the standardized DNA parts can be assembled into a Transcriptional Unit (TU) and TUs can then be combined into multi-transcriptional unit constructs. More specifically, level 2 parts can be digested with BsaI to combine 16 transcriptional units in a Level 3 part, which could be digested again with SapI to create a Level 4 part with 64 transcriptional units and so on.

All the DNA parts described here are generated following the common syntax (Patron et al., 2015) and are compatible with Type IIS cloning systems, such as GoldenGate and Loop assembly, facilitating the exchange of DNA parts between different laboratories. The transformation vector contained the hph gene driven by the 35S promoter and terminated with a 35S polyadenylation signal. It also contained a p-35S_s::eGFP-Lti6B transcription unit (same 35S promoter with the one driving GUSPlus in pCAMBIA1305.2 plasmid) terminated by the double nopaline synthase (Nos) - 35S polyadenylation signal (Sauret-Güeto et al., 2020) (Fig. 3f) which was fused to the Low Temperature Induced Protein 6B (Lti6B) signal for membrane localization from Arabidopsis thaliana (Arabidopsis) (Cutler et al., 2000) (see Supporting Information Notes S1 section “Map of the transformation vector used for the generation of the p-35S_s::eGFP-Lti6B expressing lines” and Supporting Information Table S2 for full map and sequence). The vector was transformed into A. agrestis using the method described above and eGFP was successfully expressed in A. agrestis with the expected localization in the plasma membrane (Fig. 3f,j). During the course of this study we generated 215 stable A. agrestis transgenic lines (in a total of 23 transformation experiments) expressing the p-35S_s::eGFP-Lti6B (Table I). There is variability in eGFP expression patterns between different transgenic lines (Supporting Information Fig. S7) presumably due to differences in the transgene copy number or genome location of transgene insertion. eGFP expression remains mosaic during the first, second and third round of selection. During the initial round of selection, the first transgene expression (based on fluorescence) is observed in regenerating thallus fragments that consist of only a few cells (~five). Thus, we speculate that the mosaic expression observed relates to the preferential expression of the promoter driving eGFP, in some parts of the thallus (see section “Comparison of the 35S and AaEF1a promoters”) rather than to chimeric plants that consist of transformed and untransformed sections/cells. A small fraction of hygromycin resistant lines (four
out of 157) do not show visible eGFP fluorescence which could be attributed to potential silencing events or truncation of the inserted T-DNA. These plants have been through at least five rounds of hygromycin antibiotic selection, so it is unlikely they are false positives. A total of 15 lines have been propagated vegetatively for more than 2.5 years without abolishing transgene expression (antibiotic resistance or eGFP).

Testing additional fluorescent proteins, Bonn isolate and transgene inheritance

Fluorescent proteins have been proven to be a powerful tool for plant cell biology studies, permitting temporal and spatial monitoring of gene expression patterns at a cellular and subcellular level (Berg et al., 2008). In order to expand the palette of fluorescent proteins that can be used in A. agrestis, we tested the expression of the monomeric Turquoise 2 fluorescent protein (mTurquoise2) (Kremers et al., 2006; Goedhart et al., 2012), the enhanced yellow fluorescent (eYFP) protein (Orm et al., 1996), and the mVenus fluorescent protein (Kremers et al., 2006). We used a construct similar to the one for the expression of eGFP protein, but with different subcellular localization signals. mTurquoise2 and mVenus were fused to the nuclear-localization peptide sequence of At4g19150/N7 (Cutler et al., 2000) with a linker to the amino (C)-terminus (Cutler et al., 2000), and eYFP was fused to the membrane-targeting myristoylation (myr) signal to the amino (N)-terminus (Resh, 1999) (see Supporting Information Table S2 for plasmid sequences). mTurquoise2 (Fig. 3g), eYFP (Fig. 3h) and mVenus (Fig. 3i) were successfully expressed in A. agrestis and were targeted to the predicted cellular compartments (for further information on the number of lines generated see Table I).

We then tested whether the protocol developed for the Oxford isolate can be used successfully for the Bonn isolate (Supporting Information Fig. S8). Four trials resulted in two successful transformants, which is considerably less than the average number of transformants obtained for the Oxford isolate (Fig. 3k).

We finally tested whether the transgene can be stably inherited through the sexual life cycle. Two eGFP expressing transgenic lines for both Oxford and Bonn isolates were brought to sexual reproduction. Sporophytes were produced and young gametophytes germinating from the spores (sporelings) were expressing eGFP indicating that the transgene and its expression was successfully passed on to the next generation (Fig. 3l,m).
Identification and selection of *A. agrestis* endogenous gene promoters

It is important to identify promoters that can be used to drive constitutive transgene expression (i.e. high-level expression across almost all tissues and development stages). Commonly used constitutive promoters in other bryophyte model species include the *M. polymorpha* ELONGATION FACTOR 1 ALPHA (*EF1a*) promoter (Althoff *et al.*, 2014), the rice *Actin1*, and the *M. polymorpha* ubiquitin-conjugating enzyme E2 promoter (Sauret-Güeto *et al.*, 2020). Using the genomic sequence (Li *et al.*, 2020) and RNA sequencing data, (Fig. 4a,b) we identified a series of candidate promoter regions as constitutive *A. agrestis* promoters. In particular we selected the promoter regions of the putative *A. agrestis* homologs of *EF1a*, *Ubiquitin*, *Actin*, and the Arabidopsis GAMMA TONOPLAST INTRINSIC PROTEIN (*Tip1;1*) genes (Fig. 4a,b, Supporting Information Table S1). We amplified a 1532 bp long stretch of the 5’ flanking region including the 5’UTR of the *EF1a*, a 933 bp segment for the *Ubiquitin* (*Ubi*), two fragments (1729 bp and 1516 bp) for the *Actin* (that correspond to two different predicted translational start sites), and a 1368 bp putative promoter for the *Tip1;1* gene (Sequence cloned, gene models, position of promoters and the corresponding RNAseq coverage tracks are shown in Supporting Information Fig. S9).

The candidate promoter regions were cloned (and if necessary domesticated in order to generate a Loop assembly cloning system compatible DNA part), fused with the eGFP or the mTurquoise2 reporter genes, and terminated with the double Nos - 35S terminator (Sauret-Güeto *et al.*, 2020). The *AaEF1a* (Fig. 5a) promoter region was sufficient to drive expression of eGFP throughout the thallus. Similarly, the *AaTip1;1* promoter region was sufficient to drive expression of eGFP and mTurquoise2 (Fig. 5b). However, only three independent lines were obtained for the *p-AaTip1;1::mTurquoise2-N7* construct (with one showing growth retardation probably due to the insertion site of the T-DNA) and one for the *p-AaTip1;1::eGFP* construct (Supporting Information Fig. S10). Thus, further characterisation of the *AaTip1;1* promoter is needed. The *AaUbi* (Fig. 5c) promoter gave less uniform expression patterns throughout the thallus, and the two *AaActin* promoters produced no detectable signal (Fig. 5d) (summary of the number of lines generated is shown in Table 1). Our data thus indicate that out of the five candidates, *AaEF1a* is the best promoter for driving relatively strong expression across cells of the gametophyte thallus. We generated a total of nine *p-AaEF1a::eGFP* lines, four of which are shown in Supporting Information Fig. S10. Out of the nine hygromycin resistant lines, two do not express eGFP which could be due to transgene silencing or truncation of the inserted T-DNA. In addition, we showed that the *AaEF1a* promoter can drive adequate *hph*
expression (Fig. 5e). Finally, we were also able to successfully express simultaneously three different transcription units, p-AaEF1a::mTurquoise2-N7, p-35S_s::eGFP-LTI6b and the p-35S::hph (Fig. 5f), which was the largest construct (approximately 7.4 kb) we successfully introduced into the A. agrestis genome. Three lines were obtained, one of which did not show eGFP expression (Table 1). The three lines have been propagated vegetatively for more than 1.5 years without abolishing transgene expression (antibiotic resistance or fluorescence protein).

**Comparison of the 35S and AaEF1a promoters**

Expression of eGFP driven by the CaMV 35S promoter seems to be weaker in newly grown parts of the thallus (Fig. 6a and Supporting Information Fig. S7). This is similar to the expression patterns of transgenes driven by the CaMV 35S promoter in M. polymorpha, which has a strong activity in all parts of the thallus except the notch area (Althoff et al., 2014). Expression of eGFP driven by the AaEF1a promoter seems to be stronger in the putatively younger parts of the thallus (Fig. 6b and Supporting Information Fig. S10). This is again similar to the expression patterns of transgenes driven by the EF1a promoter in M. polymorpha, showing a strong activity in all parts of the thallus particularly the notch area (Althoff et al., 2014). We also assessed the amount of eGFP protein accumulated in plants containing either the CaMV 35S or the AaEF1a promoter. Our western blot analysis showed that a similar amount of eGFP is accumulated in both types of plants (Fig. 6c). Therefore, our experiments imply that the CaMV 35S and the AaEF1a promoters have similar overall activity in the gametophyte thallus of A. agrestis. Nevertheless, the CaMV 35S drives the transgene weaker in newly grown parts of the thallus which needs to be taken into account when designing experiments.

**Transformation efficiency and optimisation**

In order to estimate the transformation efficiency of the protocol, we performed 10 independent trials using approximately as starting material 2 g of tissue per trial and the p-35S::hph - p-35S_s::eGFP-LTI6b construct. The number of successful transformation events per experiment varied from 3 to 23 (Table 2).
We then carried out further experiments to optimize transformation efficiency. We reasoned that tissue susceptibility to *Agrobacterium* infection may differ during different stages of regeneration after homogenisation, thereby affecting transformation efficiency. To estimate when plant regeneration is initiated, we set up a microscopy time course using homogenised thallus fragments. The first cell division was observed five days after homogenisation (Fig. 6d). Based on this result, we carried out an optimisation experiment starting co-cultivation at two, five, seven, 14 and 21 days after homogenisation. We found that the number of stably transformed lines decreased when using tissue that was recovered for more than five days after homogenisation (Fig. 6e). The highest number of transformants could be obtained when using tissue two days after homogenisation.

We also tested the effect of tissue growth (after the first homogenisation) under three different light intensities, 7, 35 and 80 μmol m⁻² s⁻¹ on transformation efficiency. Transformants were obtained only when using the lowest light intensity of 7 μmol m⁻² s⁻¹ (Fig. 6f).

Finally, we tested another *Agrobacterium* strain, the GV3101, for its ability to infect *A. agrestis* thallus. However, only up to two successful transformation events were obtained when the GV3101 strain harboring the *p-35S::hph - p-35S::eGFP-LT16b* construct was used (Fig. 6g).

**Verification of transgene incorporation into the *A. agrestis* genome**

To confirm the genomic integration of the transgene, we sequenced and assembled the genomes of five stable transformant lines (Fig. 7a-f). For all five lines, we found a single integration site, with one line showing a single full-length insertion of the T-DNA. The other four lines additionally showed one or multiple partial insertions in inverted and/or tandem directions (Supporting Information - see in Notes S1 section “Detailed description of insertion localization in the sequenced transformant lines”).

To confirm the transgene integration site in the five lines, fragments overlapping the 5’- and 3’- ends of the inserts and their adjacent genomic regions were amplified by nested PCR and Sanger-sequenced. The resulting sequences confirmed the integration sites identified by the genome assemblies (Fig. 7g and Supporting Information - see in Notes S1 section “Detailed description of insertion localization in the sequenced transformant lines”). We conclude that the transformation
method described here results in the stable integration of one or more targeted transcriptional units into the *A. agrestis* nuclear genome.

**Discussion**

The protocol described in this study successfully generated stable transformants in *A. agrestis* Oxford and Bonn isolates and may be applicable to other hornwort species. We generated a total of 274 stable lines in 38 transformation experiments. We showed that transgenic lines can be propagated for more than two years without abolishing transgene expression. Additionally, we confirmed that the transgene is integrated into the genome of *A. agrestis* and can be successfully inherited.

The genome sequencing of transgenic lines showed that the integration occurs in a single locus with one or more copies, which is similar to the reports in other organisms based on DNA gel blot analysis (Feldmann & David Marks, 1987; Ishizaki *et al.*, 2008; Plackett *et al.*, 2014). The utilization of recent high throughput sequencing technologies combined with the genome size of the *A. agrestis* allows precise determination of the transgene insertion site. Thus, it should be relatively simple to perform enhancer-trap or T-DNA based mutagenesis experiments in *A. agrestis*.

The light conditions under which the plant tissue was grown significantly affected the thallus morphology and were critical for successful transformation. It is likely that high light intensity triggers the accumulation of secondary metabolites (such as mucilage) and/or affects the composition of the cell wall, thereby significantly reducing transformation efficiency. Light intensity might affect the expression of genes such as the Cytochromes P450 (CYP) gene families for primary and secondary metabolism, an expansion of which family has been reported in the genome of *Anthoceros angustus* (Zhang *et al.*, 2020). Multiple photoreceptors are also present in *A. agrestis* genome (Li *et al.*, 2014, 2015a,b). Identifying which receptors determine the response to high light intensity could help to further improve the transformation efficiency.
We have tested different methods for tissue fragmentation before co-cultivation with *Agrobacterium*, such as vortexing, use of scalpels or razor blades as an alternative to tissue homogeniser. The use of scalpels or razor blades resulted in successful transformation events for both Oxford and Bonn strains, however this method is still under optimisation.

We are currently developing genome editing tools for *A. agrestis* using CRISPR/Cas9 (Jinek et al., 2012). More selection markers are needed, and preliminary data suggest that blasticidin (Tamura et al., 1995; Ishikawa et al., 2011) is a promising candidate. We are also testing whether inducible gene expression systems such as the glucocorticoid receptor (Schena et al., 1991) or estrogen receptor (Zuo et al., 2000) can be applied successfully in hornworts. Finally, we are testing alternative gene delivery methods for both *A. agrestis* Oxford and Bonn isolates, such as particle bombardment.

The development of a hornwort transformation method, in combination with the recently published genome, will greatly facilitate more comprehensive studies of the mechanisms underpinning land plant evolution. It can also help with engineering hornwort traits into plants with agronomic value. For example engineering pyrenoids in crops, has the potential to increase carbon fixation and therefore increase crop yield (Li et al., 2017).

**Acknowledgments**

We would like to thank Juan Carlos Villarreal for introducing us to the world of hornworts and always generously sharing his expertise. We are also thankful to Jim Haseloff for support during the latest stage of the project. Figure 1D picture credit: John Baker, Oxford University. Finally, constructive comments of three reviewers on an earlier version of the manuscript are greatly acknowledged.

This research was supported by the Japanese Society for the Promotion of Science (JSPS) Short Term Postdoctoral Fellowship grant no. PE14780 to E.F., the UZH Forschungskredit Candoc grant no. FK-19-089 to M.W., the MEXT and JSPS Grants-in-Aid for Scientific Research on Innovation Areas nos. 25113001 and 19H05672 to H.S., JSPS KAKENHI 26650143 and 18K06367 to K.S., 15H04413 to T.N., 19K22448 to T.N. and K.S., the Swiss National Science Foundation (grants 160004, 131726 and 184826) to P.S., The Deutsche Forschungsgemeinschaft (DFG – German Research Foundation) under the Priority Programme “MAdLand – Molecular Adaptation to Land:
Author Contributions

KS, EF, HT, MW and PS, conceived and designed the experiments. TN and MW identified gene promoter regions. EF and MW performed cloning. EF generated and characterized the transgenic lines and performed imaging. XX, MT and JVC generated the BTI lines and AG imaged the lines. MW, PS, F-WL and AG analysed the transcriptomic data. EF performed qPCR analysis. MW performed NGS of transgenic lines, PS and MW assembled the genomes. AG and MW confirmed the insert locations. YY provided technical assistance. KS, EF, TN, MW and PS wrote the article with contributions from all the authors.

Data availability

Accession numbers

Raw sequencing reads have been submitted to the NCBI SRA under the BioProject ID PRJNA683066 (SRR13209765-SRR13209769).

Protocols are also available on http://protocols.io/workspaces/hornworts and https://www.hornworts.uzh.ch. Plasmids will be available from Addgene.
References

Althoff F, Kopischke S, Zobell O, Ide K, Ishizaki K, Kohchi T, Zachgo S. 2014. Comparison of the MpEF1α and CaMV35 promoters for application in Marchantia polymorpha overexpression studies. Transgenic research 23: 235–244.

Althoff F, Zachgo S. 2020. Transformation of Riccia fluitans, an amphibious liverwort dynamically responding to environmental changes. International Journal of Molecular Sciences 21: 5410.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of molecular biology 215: 403–410.

Berg RH, Beachy RN. 2008. Fluorescent protein applications in plants. Methods in Cell Biology 85:153-177.

Chiyoda S, Ishizaki K, Kataoka H, Yamato KT, Kohchi T. 2008. Direct transformation of the liverwort Marchantia polymorpha L. by particle bombardment using immature thalli developing from spores. Plant cell reports 27: 1467–1473.

Cho SH, Chung YS, Cho SK, Rim YW, Shin JS. 1999. Particle bombardment mediated transformation and GFP expression in the moss Physcomitrella patens. Molecules and cells 9: 14–19.

Cormack BP, Valdivia RH, Falkow S. 1996. FACS-optimized mutants of the green fluorescent protein (GFP). Gene 173: 33–38.

Cutler SR, Ehrhardt DW, Griffitts JS, Somerville CR. 2000. Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. Proceedings of the National Academy of Sciences 97: 3718–3723.

Desirò A, Duckett JG, Pressel S, Villarreal JC, Bidartondo MI. 2013. Fungal symbioses in hornworts: a chequered history. Proceedings. Biological sciences / The Royal Society 280: 20130207.

Feldmann KA, David Marks M. 1987. Agrobacterium-mediated transformation of germinating seeds of Arabidopsis thaliana: A non-tissue culture approach. Molecular and General Genetics MGG 208:
Finiuk NS, Chaplya AY, Mitina NY, Boiko NM, Lobachevska OV, Miahkota OS, Lemets AI, Blium B, Zaichenko OS, Stoika RS. 2014. Genetic transformation of moss *Ceratodon purpureus* by means of polycationic carriers of DNA. *Cytol. Genet.* 48: 345–351.

Frangedakis E, Shimamura M, Villarreal JC, Li F-W, Tomaselli M, Waller M, Sakakibara K, Renzaglia KS, Szövényi P. 2020. The Hornworts: Morphology, evolution and development. *The New phytologist* 229:735-754.

Gelvin SB. 2003. *Agrobacterium*-mediated plant transformation: The biology behind the ‘Gene-Jockeying’ tool. *Microbiology and Molecular Biology Reviews* 67: 16–37.

Goedhart J, von Stetten D, Noirclerc-Savoye M, Lelimousin M, Joosen L, Hink MA, van Weeren L, Gadella TWJ, Royant A. 2012. Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nature Communications* 3: 751.

Ishikawa M, Murata T, Sato Y, Nishiyama T, Hiwatashi Y, Imai A., Kimura M, Sugimoto N, Akita A, Oguri Y, Friedman WE, Hasebe M, Kubo M. 2011. *Physcomitrella* cyclin-dependent kinase A links cell cycle reactivation to other cellular changes during reprogramming of leaf cells. *Plant Cell* 23: 2924–2938.

Ishizaki K, Chiyoda S, Yamato KT, Kohchi T. 2008. *Agrobacterium*-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. *Plant & cell physiology* 49: 1084–1091.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337: 816–821.

Kohchi T, Yamato KT, Ishizaki K, Yamaoka S, Nishihama R. 2021. Development and molecular genetics of *Marchantia polymorpha*. *Annual Review of Plant Biology* 72:1.

Kremers G-J, Goedhart J, van Munster EB, Gadella TWJ. 2006. Cyan and yellow super fluorescent proteins with improved brightness, Protein Folding, and FRET Förster Radius†,‡. *Biochemistry* 45: 6570–6580.
Kubota A, Ishizaki K, Hosaka M, Kohchi T. 2013. Efficient Agrobacterium-mediated transformation of the liverwort Marchantia polymorpha using regenerating thalli. Bioscience, biotechnology, and biochemistry 77: 167–172.

Kugita M. 2003. RNA editing in hornwort chloroplasts makes more than half the genes functional. Nucleic Acids Research 31: 2417–2423.

Li F-W, Melkonian M, Rothfels CJ, Villarreal JC, Stevenson DW, Graham SW, Wong GK-S, Pryer KM, Mathews S. 2015a. Phytochrome diversity in green plants and the origin of canonical plant phytochromes. Nature Communications 6: 7852.

Li F-W, Nishiyama T, Waller M, Frangedakis E, Keller J, Li Z, Fernandez-Pozo N, Barker MS, Bennett T, Blázquez MA, et al. 2020. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nature plants 6: 259–272.

Li F-W, Rothfels CJ, Melkonian M, Villarreal JC, Stevenson DW, Graham SW, -S. Wong GK, Mathews S, Pryer KM. 2015b. The origin and evolution of phototropins. Frontiers in Plant Science 6: 637.

Liu J-W, Li S-F, Wu C-T, Valdespino IA, Ho J-F, Wu Y-H, Chang H-M, Guu T-Y, Kao M-F, Chesson C, et al. 2020. Gigantic chloroplasts, including bizonoplasts, are common in shade-adapted species of the ancient vascular plant family Selaginellaceae. American journal of botany 107: 562–576.

Li F-W, Villarreal Aguilar JC, Szövényi P. 2017. Hornworts: An overlooked window into carbon-concentrating mechanisms. Trends in plant science 22: 275–277.

Li F-W, Villarreal JC, Kelly S, Rothfels CJ, Melkonian M, Frangedakis E, Ruhsam M, Sigel EM, Der JP, Pittermann J, et al. 2014. Horizontal transfer of an adaptive chimeric photoreceptor from bryophytes to ferns. Proceedings of the National Academy of Sciences 111: 6672–6677.

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ. 2018. The timescale of early land plant evolution. Proceedings of the National Academy of Sciences of the United States of America 115: E2274–E2283.

Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD,
Pyshkin A, Sirotkin A, Sirotkin Y, et al. 2013. Assembling single-cell genomes and minimetagenomes from chimeric MDA products. *Journal of computational biology: a journal of computational molecular cell biology* 20: 714–737.

One Thousand Plant Transcriptomes Initiative. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574: 679–685.

Orm M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. 1996. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* 273: 1392–1395.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods* 14: 417–419.

Plackett ARG, Huang L, Sanders HL, Langdale JA. 2014. High-efficiency stable transformation of the model fern species *Ceratopteris richardii* via microparticle bombardment. *Plant physiology* 165: 3–14.

Pollak B, Cerda A, Delmans M, Álamos S, Moyano T, West A, Gutiérrez RA, Patron NJ, Federici F, Haseloff J. 2019. Loop assembly: a simple and open system for recursive fabrication of DNA circuits. *New phytologist* 222: 628–640.

Porebski S, Grant Bailey L, Baum BR. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter* 15: 8–15.

Rensing SA, Goffinet B, Meyberg R, Wu S-Z, Bezanilla M. 2020. The moss *Physcomitrium (Physcomitrella)* patens: A model organism for non-seed plants. *The Plant Cell* 32: 1361-1376.

Renzaglia KS, Villarreal Aguilar JC, Piatkowski BT, Lucas JR, Merced A. 2017. Hornwort stomata: architecture and fate shared with 400-Million-year-old fossil plants without leaves. *Plant physiology* 174: 788–797.

Resh MD. 1999. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochimica et biophysica acta* 1451: 1–16.

Sauret-Güeto S, Frangedakis E, Silvestri L, Rebmann M, Tomaselli M, Markel K, Delmans M,
West A, Patron NJ, Haseloff J. 2020. Systematic tools for reprogramming plant gene expression in a simple model, *Marchantia polymorpha*. *ACS synthetic biology* 9: 864–882.

Schaefer D, Zryd J-P, Knight CD, Cove DJ. 1991. Stable transformation of the moss *Physcomitrella patens*. *Molecular and General Genetics MGG* 226: 418–424.

Schena M, Lloyd AM, Davis RW. 1991. A steroid-inducible gene expression system for plant cells. *Proceedings of the National Academy of Sciences* 88: 10421–10425.

Small ID, Schallenberg-Rüdinger M, Takenaka M, Mireau H, Ostersetzer-Biran O. 2019. Plant organellar RNA editing: what 30 years of research has revealed. *The Plant journal: for cell and molecular biology* 101:1040-1056.

Söderström L, Hagborg A, von Konrat M, Bartholomew-Began S, Bell D, Briscoe L, Brown E, Cargill DC, Costa DP, Crandall-Stotler BJ, et al. 2016. World checklist of hornworts and liverworts. *PhytoKeys* 59: 1–828.

Stachel SE, Messens E, Van Montagu M, Zambryski P. 1985. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature* 318: 624–629.

Szövényi P, Frangedakis E, Ricca M, Quandt D, Wicke S, Langdale JA. 2015. Establishment of *Anthoceros agrestis* as a model species for studying the biology of hornworts. *BMC plant biology* 15: 98.

Tamura K., Kimura M., Yamaguchi I. 1995. Blasticidin S deaminase gene (BSD): A new selection marker gene for transformation of *Arabidopsis thaliana* and *Nicotiana tabacum*. *Bioscience, Biotechnology, and Biochemistry* 59: 2336–2338.

Nomura T, Sakurai T, Osakabe Y, Osakabe K, Sakakibara H. 2016. Efficient and heritable targeted mutagenesis in mosses using the CRISPR/Cas9 system. *Plant and Cell Physiology* 57: 2600–2610.

Trouiller B, Charlot F, Choinard S, Schaefer DG, Nogué F. 2007. Comparison of gene targeting efficiencies in two mosses suggests that it is a conserved feature of bryophyte transformation.

This article is protected by copyright. All rights reserved
Villarreal JC, Renner SS. 2012. Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 18873–18878.

Zeidler M, Hartmann E, Hughes J. 1999. Transgene expression in the moss *Ceratodon purpureus*. *Journal of Plant Physiology* **154**: 641–650.

Yoshinaga K. 1996. Extensive RNA editing of U to C in addition to C to U substitution in the rbcL transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. *Nucleic Acids Research* **24**: 1008–1014.

Yoshinaga K. 1997. Extensive RNA editing and possible double-stranded structures determining editing sites in the atpB transcripts of hornwort chloroplasts. *Nucleic Acids Research* **25**: 4830–4834.

Young JM, Kuykendall LD, Martinez-Romero E, Kerr A, Sawada H. 2001. A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *International journal of systematic and evolutionary microbiology* **51**: 89–103.

Zhang J, Fu X-X, Li R-Q, Zhao X, Liu Y, Li M-H, Zwaenepoel A, Ma H, Goffinet B, Guan Y-L, et al. 2020. The hornwort genome and early land plant evolution. *Nature plants* **6**: 107–118.

Zuo J, Niu QW, Chua NH. 2000. Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *The Plant journal: for cell and molecular biology* **24**: 265–273.

Supporting Information:

Supporting Information Fig. S1 Effect of light on *A. agrestis* growth and tissue culture.

Supporting Information Fig. S2 Hygromycin sensitivity of *A. agrestis* gametophytes (Oxford).
Supporting Information Fig. S3 Hygromycin sensitivity of *A. agrestis* gametophytes (Bonn).
Supporting Information Fig. S4 G418 sensitivity of *A. agrestis* gametophytes (Oxford).
Supporting Information Fig. S5 *A. agrestis* Oxford transformation protocol.
Supporting Information Fig. S6 GUS activity detected in transformed *A. agrestis* plants.
Supporting Information Fig. S7 Examples of *A. agrestis* Oxford transgenic lines I.
Supporting Information Fig. S8 *A. agrestis* Bonn transformation protocol.
Supporting Information Fig. S9 Promoters and their localization on the *A. agrestis* genome.
Supporting Information Fig. S10 Examples of *A. agrestis* Oxford transgenic lines II.

Supplementary Notes S1 Native promoters, vector maps, and insertion localization.

Supplementary Table S1 Highly and/or constitutively expressed *A. agrestis* genes.
Supplementary Table S2 List of construct sequences.
Supplementary Table S3 List of primers used in this study.

**Figure 1** Morphological features of *Anthoceros agrestis* and effect of light on growth.

a) Light micrograph (LM) of a germinating spore. Scale bar: 50 μm. b) Surface view of the irregularly shaped thallus (gametophyte). Scale bar: 1 mm. c) LM showing cells of mature gametophyte tissue with single chloroplasts. Scale bar: 10 μm. d) *A. agrestis* Oxford gametophyte with sporophytes. Scale bar: 4 mm. e) The conditions for the preparation of tissue used for transformation are critical. Plants must be propagated in axenic culture by transferring small thallus fragments (typically 1 x 1 mm) onto plates with fresh growth medium using sterile scalpels and then grown under low light conditions (3-5 μmol m$^{-2}$ s$^{-1}$) for 4 weeks. f) *A. agrestis* Oxford thallus tissue grown for 4 weeks.
under low light intensity (3-5 μmol m⁻² s⁻¹). g) *A. agrestis* Oxford thallus tissue grown for 4 weeks under high light intensity (80 μmol m⁻² s⁻¹). Scale bars 1 mm. Tissue morphology after: h) the first homogenization prior to co-cultivation, i) one month after the first homogenization, j) after the second homogenisation prior to co-cultivation, Scale bars 1 mm.

**Figure 2:** Outline of *Anthoceros agrestis* transformation method.

1a-b: Tissue grown under low light is homogenised, transferred on growth medium, and placed again under low light conditions. 2a-c: After 4 weeks, the tissue is homogenised again and grown for two additional days (the purpose of the first homogenisation is tissue amplification). 3: The tissue is co-cultivated with *Agrobacterium* for three days (under low light) and then 4a-b: spread on appropriate antibiotic-containing growth medium. 5: After 3 weeks, the tissue is transferred again onto freshly prepared antibiotic-containing growth medium for a second round of selection. 6: After approximately 4-8 weeks, putative transformants are visible. 7: A final round of selection is recommended to eliminate false-positive transformants.

**Figure 3** Schematic representation of transformation constructs and transgenic *Anthoceros agrestis* expressing different fluorescent proteins.

a-b) The emergence of rhizoids (shown with yellow arrows) is a reliable indicator of successfully transformed plant fragments. In (A) the red arrowheads show false positives regenerating (green) tissue fragments that lack rhizoids. Scale bars: 2 mm. c) Example of transgenic *A. agrestis* plants (gametophyte thallus). Petri dish dimensions: 92 x16 mm. d) GUS activity detected as blue staining in thallus tissue fragments from a plant transformed with the pCAMBIA1305.2 plasmid. Scale bar: 200 μm. e) Loop assembly Type IIS cloning system outline: Level 0 DNA parts (for simplicity only promoter, coding sequence and terminator genetic DNA parts are indicated) are assembled in Level
1 transcription units (TUs) into one of the four pCk vector, depicted with numbered circles, by BsdI-mediated Type IIS assembly (sequential restriction enzyme digestion and ligation reactions). Level 1 TUs are assembled to Level 2 multi-TUs into one the four pCs vectors, depicted with lettered circles, by SapI-mediated Type IIS assembly. This workflow is then repeated for higher level assemblies. 

f-i top) Schematic representation of constructs for the expression of two transcription units (TU): one TU for the expression of the *hygromycin B phosphotransferase (hph)* gene under the control of the cauliflower mosaic virus (CaMV) 35S promoter and one TU for the expression of *p-35S_s::eGFP-LTI6b* (f), *p-35S_s::mTurquoise2-N7* (g), *p-35S_s::myr-eYFP* (h) and *p-35S::mVenus-N7* (I) TU. *hph*: hygromycin B phosphotransferase; 35S: CaMV 35S promoter; eGFP: enhanced green fluorescent protein; mTurquoise2: monomeric turquoise 2 fluorescent protein; eYFP: enhanced yellow fluorescent protein; LTI6b: Low Temperature Induced Protein 6B signal for membrane localization; N7: Arabidopsis At4g19150/N7 nuclear localization signal; myr: myristoylation signal for membrane localization; nosT: 3′ signal of *nopaline synthase*. f-i middle and bottom) Images of *A. agrestis* Oxford thallus tissue expressing different combinations of CaMV 35S promoter - fluorescent protein - localization signal. f) *p-35S_s::eGFP-LTI6b* for plasma membrane localization (white arrow). Scale bars: top: 50 μm, bottom: 20 μm, g) *p-35S_s::mTurquoise2-N7* for nuclear localization (white arrow). Scale bars: top: 20 μm, bottom: 10 μm, h) *p-35S_s::myr-eYFP* for plasma membrane localization (white arrow). Scale bars: top: 50 μm, bottom: 20 μm, l) and *p-35S::mVenus-N7* for nuclear localization (white arrow). Scale bars: top: 50 μm, bottom: 25 μm. The bottom image is a magnification of the image in the middle. Red, chlorophyll autofluorescence. j) Light micrograph image: surface view of thallus (gametophyte), similar to the area imaged in f (and also g-i). Scale bar: 100 μm. k) Images of *A. agrestis* Bonn gametophyte tissue expressing the *p-35S_s::eGFP-LTI6b* TU for eGFP plasma membrane localization. Scale bar: 50 μm. l) *A. agrestis* Bonn with mature sporophytes indicated with white arrows. Scale bars: 2 mm m) *A. agrestis* Bonn transgenic spores expressing the *p-35S_s::eGFP-LTI6b* TU (white arrow indicates rhizoid and yellow arrow indicates young thallus). Top left: Light microscopy of *A. agrestis* Bonn wild type germinating spore. Scale bar: 20 μm.

**Figure 4** Identification of constitutive promoters for *Anthoceros agrestis*. 

This article is protected by copyright. All rights reserved
a) Analysis of expression levels from RNAseq experiments on *A. agrestis* Oxford using datasets from Li et al. 2020. To generate this dataset, gametophytes were grown under varying carbon sources in the growth medium (indicated as C depleted, C enriched and C normal), as well as two different nitrogen depleted conditions (N- with cyanobacteria symbiosis and N- without cyanobacteria symbiosis). Error bars indicate standard error based on three independent experimental replicates. (Included for comparison: “CN-Max 1, 3 and 5”: highest expressing genes based on N conditions and “CN-Max 2 and 4”: highest expressing genes based on C normal conditions (for gene ID, see Supporting Information Table S1).

b) Analysis of expression levels from RNAseq experiments on *A. agrestis* Bonn using data sets from Li et al. 2020. Note: Normalized expression level of *A. agrestis* Bonn genes selected to represent strong and constitutive expression across various developmental stages of the gametophyte and the sporophyte phases (Included for comparison: “T-Max 1 to 5”: Highest expressing genes under all conditions (for gene ID, see Supporting Information Table S1). Error bars indicate standard error based on two independent experimental replicates.

**Figure 5** Identification of constitutive promoters for *Anthoceros agrestis*.

a-f top: Schematic representation of constructs for the expression of two-three transcription units (TU): one TU for the expression of the *hph* gene under the control of the CaMV 35S or the p-AaEF1a promoter and one TU for the expression of plasma membrane localized eGFP and/or nucleus localized mTurquoise2 under the control of different native *A. agrestis* promoters. a) p-AaEF1a, b) p-AaTip1;1, c) p-AaUbi and d) p-AaActin_1. e) p-AaEF1a driving *hph* and f) p-AaEF1a driving mTurquoise2-N7.

a-f middle and bottom: Images of *A. agrestis* Oxford gametophyte tissue expressing different combinations of *A. agrestis* native promoter - fluorescent protein - localization signal. a) p-AaEF1a::eGFP-LTI6b for plasma membrane localization. Scale bars: top: 50 μm, bottom: 10 μm, b) p-AaTip1;1::mTurquoise2-N7 for nuclear localization. Scale bars: top: 20 μm, bottom: 10 μm, c) p-AaUbi::eGFP-LTI6b for plasma membrane localization. Scale bars: top: 20 μm, bottom: 10 μm. The bottom images are a magnification of the images in the top. d) p-AaActin::eGFP-LTI6b. Scale bars: 50 μm and e) p-AaEF1a::hph - p-35S::eGFP-LTI6b. Scale bar: 50 μm. f) Image of *A. agrestis*
Oxford gametophyte tissue expressing both p-AaEF1a::mTurquoise2-N7 for nuclear localization and p-35S_s::eGFP-LTI6b for plasma membrane localization. Scale bar: 50 μm. Red, chlorophyll autofluorescence.

**Figure 6** Comparison of the CaMV 35S with the AaEF1a promoter and factors affecting the efficiency of Agrobacterium-mediated transformation of Anthoceros agrestis.

a) Expression of eGFP driven by the CaMV 35S promoter. Younger part of the thallus indicated with a white arrow. Scale bar: 50 μm. b) Expression of eGFP driven by the AaEF1a promoter. Younger part of the thallus is indicated with white arrows. Scale bar: 100 μm. c) Western blot analysis of eGFP accumulation in transgenic lines. Total cellular proteins were separated by polyacrylamide gel electrophoresis, blotted and probed with anti-GFP and anti-actin antibodies. Numbering above blot images correspond to the identifier of independent lines. Information on the genomic location and copy number of T-DNA insertion for lines cam1 and cam2 is provided in Fig. 7 and in the Supporting Information in Notes S1 section “Detailed description of insertion localization in the sequenced transformant lines”. Images of CaMV 35S 5, 7 and 9 lines can be found in Supporting Figure S7. Images of the four p-AaEF1a lines can be found in Supporting Information Fig. S10. Lower band (marked with red asterisk) in “CaMV 35S line 1” potentially corresponds to degraded eGFP protein.

d) Confocal microscopy images of fragmented A. agrestis thallus tissue taken on seven consecutive days after homogenisation. Five days after homogenisation plants start to regenerate. Yellow arrowhead indicates the fragmented thallus part. White arrows indicate new cell divisions. Scale bars: 50 μm. e) Number of transgenic lines obtained using A. agrestis thallus tissue two, five, seven, 14 and 21 days after homogenisation. Values of three independent experimental replicates are shown. f) Effect of light intensity used to grow plant tissue for transformation after the first homogenisation on the number of successful transformants. Co-cultivation started two days after the second homogenisation. Values of three independent experimental replicates are shown. Scale bars: 1 mm. g) Effect of Agrobacterium strain on the number of transformants obtained. Dots represent values of individual experimental replicates. Co-cultivation started two days after the second homogenisation. Values of three independent experimental replicates are shown.

**Figure 7** Stable incorporation of transgene into Anthoceros agrestis genome.
a-f Left: Light micrograph (LM) of transgenic thallus of *A. agrestis* plants. Scale bar: 500 μm. Right: Confocal fluorescent microscopy images of thallus expressing eGFP in the plasma membrane, driven by the CaMV 35S promoter. Scale bars: a-e 100 μm and f 50 μm. Bottom: Location of transgene insertion in the genome (see details in the Supporting Information in Notes S1 section “Detailed description of insertion localization in the sequenced transformant lines”). Black arrows indicate directionality of T-DNA insert. g) PCR analysis of genomic DNA from transgenic plants. L: fragment amplified from sequences spanning the 5’- end of the T-DNA inserts and their respective adjacent genomic regions. R: fragment amplified from sequences spanning the 3’- end of the T-DNA inserts and their respective adjacent genomic regions. Note: a, e and f LM images were acquired using a KEYENCE VHX-S550E microscope (VHX-J20T lens) and confocal fluorescent images with a Leica SP8X microscope, in Cambridge University. d-d LM and fluorescent images were acquired using a Leica M205 FA Stereomicroscope with GFP longpass (LP) filter, in BT Institute.
Table 1: List of *Anthoceros agrestis* lines generated.

| Construct                                      | Total number of lines | Number of lines exhibiting fluorescence /GUS staining | Number of lines not exhibiting fluorescence /GUS staining |
|------------------------------------------------|-----------------------|-------------------------------------------------------|----------------------------------------------------------|
| pCambia1305.1                                   | 10                    | 10                                                    | -                                                        |
| p-35S::hph -p-35S_s::eGFP-LTI6b                 | 215                   | 210                                                   | 5                                                        |
| p-35S::hph -p-35S_s::eGFP-LTI6b (Bonn)          | 2                     | 2                                                     | -                                                        |
| p-35S::hph -p-35S_s::mTurquoise2-N7             | 6                     | 5                                                     | 1                                                        |
| p-35S::hph -p-35S_s::myr-eYFP                   | 5                     | 5                                                     | -                                                        |
| p-35S::hph -p-35S::mVenus-N7                    | 3                     | 3                                                     | -                                                        |
| p-35S::hph -p-AaEF1a::eGFP-LTI6b                 | 8                     | 6                                                     | 2                                                        |
| p-35S::hph -p-AaUbi::eGFP-LTI6b                  | 3                     | 3                                                     | -                                                        |
| p-35S::hph -p-AaTip1;1::mTurquoise2-N7           | 3                     | 3                                                     | -                                                        |
Table 2. Number of \textit{p-35S::eGFP-Lti6B} transgenic lines of \textit{Anthoceros agrestis} obtained from 10 different experiments to estimate transformation efficiency.

| Experiment | Amount of tissue (g) used in first homogenisation | Number of transformants |
|------------|---------------------------------------------------|-------------------------|
| 1          | 2.7                                               | 21                      |
| 2          | 3.1                                               | 12                      |
| 3          | 2.9                                               | 23                      |
| 4          | 2.5                                               | 5                       |
|   |   |   |
|---|---|---|
|5  | 1.7 | 3  |
|6  | 2.7 | 11 |
|7  | 2.6 | 19 |
|8  | 1.9 | 4  |
|9  | 2.23| 12 |
|10 | 2.1 | 6  |
