Methods

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

Eftychios Frangedakis[1] (iD), Manuel Waller[2,3] (iD), Tomoaki Nishiyama[4] (iD), Hirokazu Tsukaya[5] (iD), Xia Xu[6], Yuling Yue[2,3], Michelle Tjahjadi[6], Andika Gunadi[6] (iD), Joyce Van Eck[6,7] (iD), Fay-Wei Li[6,8] (iD), Péter Szövényi[2,3] (iD) and Keiko Sakakibara[9] (iD)

[1]Department of Plant Sciences, University of Cambridge, Cambridge, CB3 EA, UK; [2]Department of Systematic and Evolutionary Botany, University of Zurich, Zurich 8008, Switzerland; [3]Zurich-Basel Plant Science Center, Zurich 8092, Switzerland; [4]Advanced Science Research Center, Kanazawa University, Ishikawa 920-8640, Japan; [5]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan; [6]Boyce Thompson Institute, Ithaca, NY 14853-1801, USA; [7]Plant Breeding and Genetics Section, Cornell University, Ithaca, NY 14853-1801, USA; [8]Plant Biology Section, Cornell University, Ithaca, NY 14853-1801, USA; [9]Department of Life Science, Rikkyo University, Tokyo 171-8501, Japan

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 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.

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). Although 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). *Anthoceros 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). *Anthoceros 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) and 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 *c.* 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) and *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), as it is relatively simple and does not require specialized 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 characterize 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 *A. agrestis* Oxford and Bonn isolates (Szövényi *et al.*, 2015). *Anthoceros agrestis* thallus tissue was propagated on KNOP medium ($0.25\,g\,l^{-1}$ KH_2PO_4, $0.25\,g\,l^{-1}$ KCl,

$0.25\,g\,l^{-1}$ $MgSO_4.7H_2O$, $1\,g\,l^{-1}$ $Ca(NO_3)_2.4H_2O$ and $12.5\,mg\,l^{-1}$ $FeSO_4.7H_2O$). The medium was adjusted to pH 5.8 with KOH and solidified using $7.5\,g\,l^{-1}$ Gelzan CM (G1910; Sigma) in 92×16 mm Petri dishes (82.1473.001; Sarstedt, Leicester, UK) with 25–30 ml per plate. Plants were routinely grown in a tissue culture room (21°C, 12 h : 12 h light : dark, 3–5 or 35 $\mu mol\,m^{-2}\,s^{-1}$ light intensity, Philips TL-D 58W (835)). To subculture the thallus tissue, a small part of it was cut using sterile disposable scalpels (0501; Swann Morton, Sheffield, UK) and placed on fresh media on a monthly basis (see Supporting Information Fig. S1).

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 antheridia were observed. Several sporophytes appeared after 2 wk, 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 had turned completely brown and looked dry, they were collected with tweezers under a laminar hood and stored for up to 8 months at 4°C (in a fridge) in 60×15 mm Petri dishes (82.1194.500; Sarstedt). Sporophytes were used for downstream applications at the earliest after 3 months of storage at 4°C to allow for complete desiccation.

Cocultivation medium

Cocultivation medium was liquid KNOP supplemented with 2% sucrose ($0.25\,g\,l^{-1}$ KH_2PO_4, $0.25\,g\,l^{-1}$ KCl, $0.25\,g\,l^{-1}$ $MgSO_4.7H_2O$, $1\,g\,l^{-1}$ $Ca(NO_3)_2.4H_2O$, $12.5\,mg\,l^{-1}$ $FeSO_4.7H_2O$ and $20\,g\,l^{-1}$ sucrose, pH 5.8 adjusted with KOH).

Tissue preparation for transformation

Approximately 2 g of thallus tissue (15–20 Petri dishes) was divided into four parts, and each part (*c.* 0.5 g) was homogenized in 15 ml sterile water using a homogenizer (727407; IKA Ultra-Turrax T25 S7 Homogenizer, Oxford, UK) and corresponding dispensing tools (10442743; IKA Dispersing Element), for 5 s, using the lowest speed of 8000 rpm. The homogenized tissue was washed with 50 ml sterile water using a 100 μm cell strainer (352360; Corning, Flintshire, UK), plated onto solid KNOP medium and placed at 21°C under 12 h : 12 h, light : dark at a light intensity of 3–5 $\mu mol\,m^{-2}\,s^{-1}$. After 4 wk the tissue was rehomogenized in 15–20 ml sterile water and filtered using a 100 μm cell strainer. The rehomogenized tissue was transferred again onto four plates with solid KNOP medium and was allowed to grow for 2 d at 21°C under continuous light (35 $\mu mol\,m^{-2}\,s^{-1}$; TL-D58W/835; Philips).

For sporeling-transformation, 20 mature and desiccated sporophytes were pooled in a 1.5 ml Eppendorf tube and the spores were surface sterilized as described by Sauret-Güeto *et al.* (2020). The spores were incubated in the sterilizing solution for 15 min
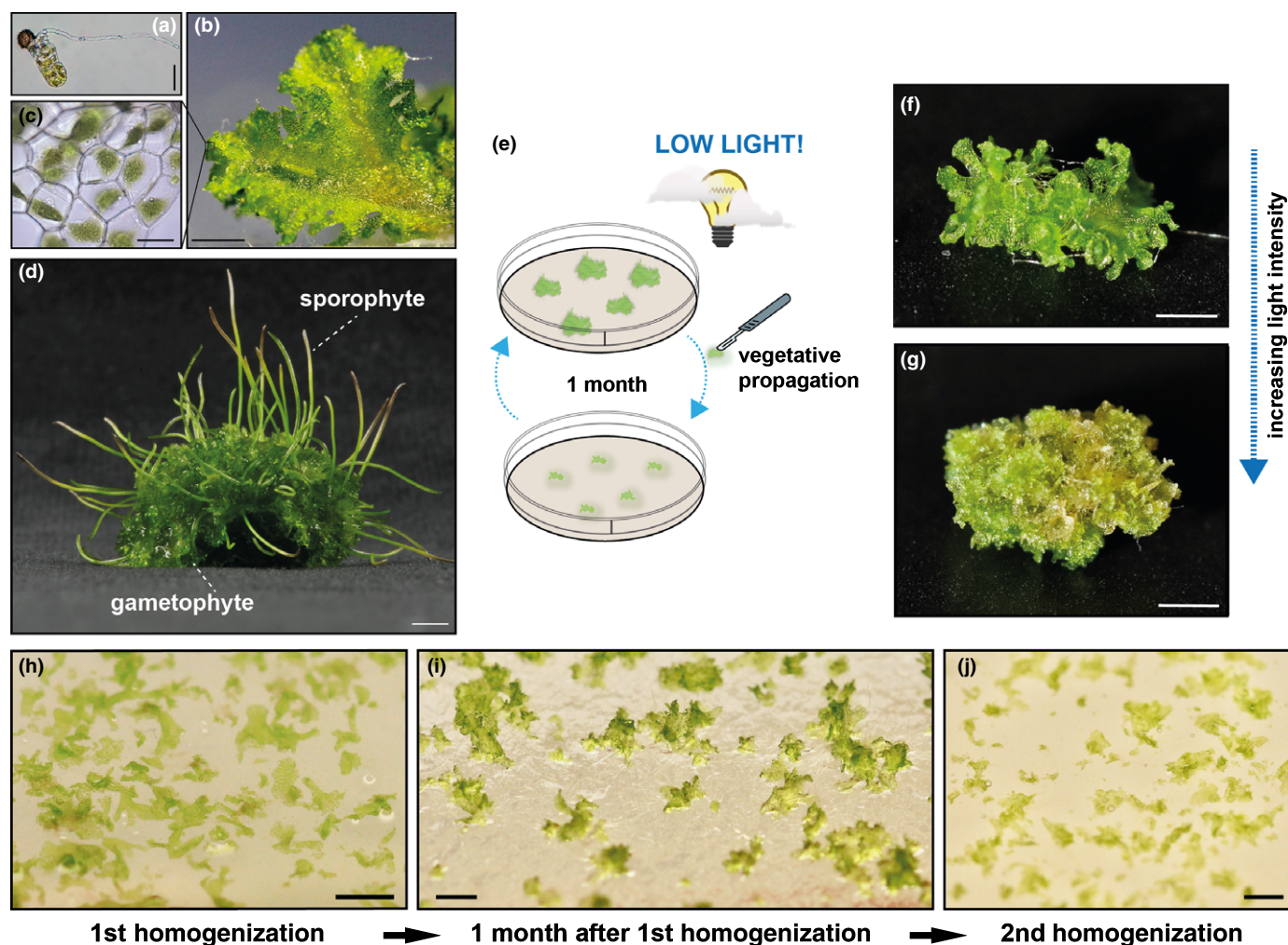
Fig. 1 Morphological features of *Anthoceros agrestis* and effect of light on growth. (a) Light micrograph (LM) of a germinating spore. Bar, 50 μm. (b) Surface view of the irregularly shaped thallus (gametophyte). Bar, 1 mm. (c) LM showing cells of mature gametophyte tissue with single chloroplasts. Bar, 10 μm. (d) *Anthoceros agrestis* Oxford gametophyte with sporophytes. 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 1×1 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 wk. (f) *Anthoceros agrestis* Oxford thallus tissue grown for 4 wk under low light intensity (3–5 μmol m^{-2} s^{-1}). (g) *Anthoceros agrestis* Oxford thallus tissue grown for 4 wk under high light intensity (80 μmol m^{-2} s^{-1}). Bars, 1 mm. Tissue morphology after: (h) the first homogenization before cocultivation, (i) 1 month after the first homogenization, (j) after the second homogenization before cocultivation. Bars, 1 mm.

and then centrifuged at 10 000 *g* for 2 min. 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 h : 12 h light : dark at a light intensity of 35 μmol m^{-2} s^{-1}. The first germinating spores were observed 5–8 d after plating. Sporelings were left to grow for 7, 14 and 21 d (after observation of the first spore germination). For *Agrobacterium* cocultivation, 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 six-well plate with *Agrobacterium* and transformation buffer (for cocultivation conditions see *Agrobacterium* culture preparation and cocultivation conditions).

Agrobacterium culture preparation

One to three *Agrobacterium* colonies (strain AGL1) were inoculated in 5 ml of LB medium supplemented with rifampicin 10 μg ml^{-1} (R0146; Duchefa, Haarlem, the Netherlands), carbenicillin 50 μg ml^{-1} (C0109; Melford, Chelsworth, UK) and the plasmid-specific selection antibiotic spectinomycin 100 μg ml (SB0901; Bio Basic, Toronto, Canada). The preculture was incubated at 28°C for 2 d with shaking at 120 rpm. OD$_{600}$ was *c.* 2.7 and was measured using an OD600 DiluPhotometer (Implen, München, Germany).

Cocultivation conditions

In total, 5 ml of a 2 d *Agrobacterium* culture was centrifuged for 7 min at 2000 *g*. The supernatant was discarded, and the pellet

was resuspended in 5 ml liquid KNOP supplemented with 2% (w/v) sucrose (S/8600/60; ThermoFisher, Loughborough, UK) 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 h. The regenerating thallus tissue was transferred (one plate – 2 d after the second homogenization) into a well of a six-well plate with 5 ml of liquid KNOP medium supplemented with 2% (w/v) sucrose. Then, 80 µl of *Agrobacterium* culture and acetosyringone at final concentration of 100 µM were added to the medium.

The tissue and *Agrobacterium* were cocultivated using a six-well plate (140675; ThermoFisher) for 3 d with shaking at 110 rpm at 22°C on a shaker without any additional supplementary light (only ambient light from the room, 1–3 µmol m^{-2} s^{-1}). After 3 d, the tissue was drained using a 100 µm cell strainer (352360; Corning) and moved onto solid KNOP plates (two Petri dishes from a single well) supplemented with 100 µg ml^{-1} cefotaxime (BIC0111; Apollo Scientific, Bredbury, UK) and 10 µg ml^{-1} hygromycin (10687010; Invitrogen). After 3–4 wk, plants were transferred to fresh solid KNOP plates supplemented with 100 µg ml^{-1} cefotaxime and 10 µg ml^{-1} hygromycin and grown at 22°C under 12 h : 12 h light : dark at a light intensity of 35 µmol m^{-2} s^{-1} (TL-D58W/835; Philips).

GUS staining

Glucuronidase (GUS) staining was performed according to Plackett *et al.* (2014). GUS buffer stock solution was prepared by mixing 11.54 ml of 1 M Na$_2$HPO$_4$, 8.46 ml of 1 M NaH$_2$PO$_4$, 8 ml of 0.25 M EDTA (pH 7.0), 10 µl of 20% Tween20 (v/v) and topped up to 200 ml with dH$_2$O. GUS staining solution (10 ml) was prepared by dissolving 10 mg 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-glcA) (B72200-0.25; Melford) in 1 ml DMSO and added to 9 ml GUS buffer stock solution (final X-glcA concentration 1 mg ml^{-1}). The tissue was first fixed in cold 90% (v/v) acetone for 10 min. 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 min twice to enable efficient infiltration to the tissues. Then the sample dish was wrapped in tinfoil and incubated at 37°C overnight. Before visualization, Chl 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 extraction of hornwort genomic DNA. Tissue (0.5 g) was harvested and frozen in liquid nitrogen. Tissue was ground to 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 polyvinylpyrrolidone (PVP)/g of tissue) that had been prewarmed at 60°C, 100 µl of RNAse A (100 mg ml^{-1}) was added and the solution was mixed well. The

mixture was incubated at 60°C for 30 min and then removed from heat and allowed to cool at room temperature for 4 min. Then, 12 ml of chloroform : isoamyl alcohol (24 : 1) was added, mixed well by inversion and then centrifuged at 12 000 *g* for 10 min 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 *g* for 10 min 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. Two 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 20 000 *g* for 6 min. 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-sequencing (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 TRIM-MOMATIC 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* (Table S1).

Putative promoter sequences were amplified from genomic DNA using KOD Hot start polymerase (71086-5; Merck Millipore, Darmstadt, Germany) 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 H$_2$O, 1 µl 10× CutSmart buffer (B7204S; NEB, Cambridge, UK), 0.5 µl 1 mg ml^{-1} BSA (B9000S; NEB), 1 µl 10 mM ATP (P0756S; NEB), 0.25 µl 400 U µl^{-1} T4 DNA ligase (M0202S; NEB), 0.25 µl 10 U µl^{-1} 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 programme: 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. Chemically

competent *Escherichia coli* cells (30 μl) were transformed using 5 μl of the assembly reaction and then spread on LB agar plates containing 25 μg ml^{-1} chloramphenicol and 40 μg ml^{-1} X-gal. Construct sequences are detailed in Table S2, and a list of primers in 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 were: 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. For 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 H$_2$O, 1 μl 10× CutSmart buffer (B7204S; NEB), 0.5 μl 1 mg ml^{-1} BSA (B9000S; NEB), 1 μl 10 mM ATP (P0756S; NEB), 0.25 μl 400 U μl^{-1} T4 DNA ligase (M0202S; NEB) and 0.25 μl 20 U μl^{-1} BsaI 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 was 50°C for 5 min and 80°C for 10 min. Chemically competent *E. coli* cells (30 μl) were transformed using 5 μl of the Loop assembly reaction and spread on LB agar plates containing 50 μg ml^{-1} kanamycin and 40 μg ml^{-1} X-Gal. For 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 H$_2$O, 2 μl nuclease-free H$_2$O, 1 μl 10× CutSmart buffer (B7204S; NEB), 0.5 μl 1 mg ml^{-1} BSA (B9000S; NEB), 1 μl 10 mM ATP (P0756S; NEB), 0.25 μl 400 U μl^{-1} T4 DNA ligase (M0202S; NEB) and 0.25 μl 10 U μl^{-1} SapI (R0569S; NEB), and spread on LB agar plates containing 100 μg ml^{-1} spectinomycin and 40 μg ml^{-1} X-Gal. The donor plasmids were L1 constructs, and the acceptor plasmid was a pCsA vector (see Table S2 for full construct maps and sequences).

Western blotting

In total, 50 mg of *A. agrestis* thallus tissue was grown for 3 wk on KNOP medium at 21°C under 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) sodium dodecyl sulfate, 25% (v/v) glycerol, 0.25 M dithiothreitol, 0.05% (w/v) bromophenol blue), supplemented with Roche cOmplete protease inhibitor (11836170001; Roche), heated at 95°C for 5 min and centrifuged at 10 000 **g** for 10 min. The supernatant was transferred to a new tube. Equal amounts of proteins were separated by denaturing electrophoresis in NuPAGE gel (NP0322BOX; Invitrogen) and electrotransferred to nitrocellulose membranes using the iBlot2 Dry Blotting System (ThermoFisher). Enhanced green fluorescent protein (eGFP) was immunodetected with anti-GFP

antibody (1 : 4000 dilution) (JL-8, 632380; Takara, Göteborg, Sweden) and anti-mouse-HRP (1 : 15 000 dilution) (A9044; Sigma) antibodies. Actin was immunodetected with antiactin (plant) (1 : 1500 dilution) (A0480; Sigma) and antimouse-HRP (1 : 15 000 dilution) (A9044; Sigma) antibodies, using the iBind™ Western Starter Kit (SLF1000S; ThermoFisher). Western blots were visualized using the ECL™ Select Western Blotting Detection Reagent (GERPN2235; GE) following the manufacturer's instructions. Images were acquired using a G: BOX F3 Syngene Gel Documentation system.

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) was placed within the gene frame. A thallus fragment was placed within the medium-filled gene frame together with 30 μl milliQ water. The frame was then sealed with a cover slip. Plants were imaged immediately using a Leica SP8X spectral fluorescence 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 1 wk, using a Leica SP8X spectral fluorescence 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 monomeric Turquoise 2 fluorescent protein (mTurquoise2) (442 nm, 460−485 nm), for eGFP (488 nm, 498−516 nm), for mVenus and enhanced yellow fluorescent protein (eYFP) (515 nm, 522−540 nm), and for Chl autofluorescence (488 or 515, 670−700 nm). When observing lines expressing both eGFP and mTurquoise2, the 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^{-1} hygromycin and 100 μg ml^{-1} 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 per line. Illumina libraries were prepared using the TruSeq DNA nano kit

(20015964; Illumina, Zurich, Switzerland) and were sequenced on an Illumina Novaseq 6000 platform with an expected sequencing depth of 80–150× 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′-ends 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 a 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: the selection of appropriate plant tissue and its preparation for infection; the type and concentration of antibiotics applied to select for transgenic lines; the choice of transformation vectors; the choice of *Agrobacterium* strain; and the optimal conditions for cocultivation. Each of these factors is examined in this study.

Selection of tissue and optimal growth conditions

Anthoceros agrestis gametophyte thallus was chosen as the appropriate tissue for transformation because it is easily accessible, has a remarkable regenerative capacity and is haploid. *Anthoceros agrestis* thallus tissue cultures can easily be propagated and maintained by transfer of small thallus fragments (*c.* 1 × 1 mm) onto fresh growth medium on a monthly basis (Fig. S1). *Anthoceros agrestis* is similar to several other bryophytes in that an entire plant can regenerate from a small thallus fragment. This is in

striking contrast to 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 *M. 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 cocultivation with *Agrobacterium* to generate transformed plants. *Anthoceros agrestis* 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 *A. agrestis*. However, unlike in *M. polymorpha*, the *A. agrestis* notch area and apical cells are not easily distinguishable, and thus determining which part of the thallus should be removed is not obvious (Frangedakis *et al.*, 2020). Therefore, we tested whether homogenization using dispensing tools is a suitable method for thallus tissue fragmentation. Different speeds and duration of homogenization were examined. We found that homogenization of 0.5 g of thallus tissue in 15 ml of sterile water for 5 s (see Materials and methods) is sufficient to fragment the tissue and results in rapid regeneration.

In addition, we found that the light intensity is a critical factor during 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 Fig. S1). When tissue was grown under high light intensity ($>40\,\mu\text{mol m}^{-2}\,\text{s}^{-1}$), no transformants were obtained (for detailed explanation see 'Transformation efficiency and optimization' 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 wk 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 cocultivation 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, respectively). The tested concentrations ranged from 0 to $20\,\mu\text{g ml}^{-1}$ for hygromycin and 0 to $150\,\mu\text{g ml}^{-1}$ for geneticin/G418. We found that a 3 wk incubation period with $10\,\mu\text{g ml}^{-1}$ hygromycin was sufficient to inhibit growth of untransformed thallus tissue for both the Oxford and

the Bonn isolates (Figs S2, S3), whereas thallus tissue was not susceptible to geneticin/G418 even when supplied in high concentration (150 μg ml) (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) 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 homogenized, grown for 1 month under low light conditions (3–5 μmol m^{-2} s^{-1}) and then homogenized a second time (Fig. 1g–i). Two days after the second homogenization, the regenerating thallus tissue was cocultivated with *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 acetosyringone because phenolic compounds
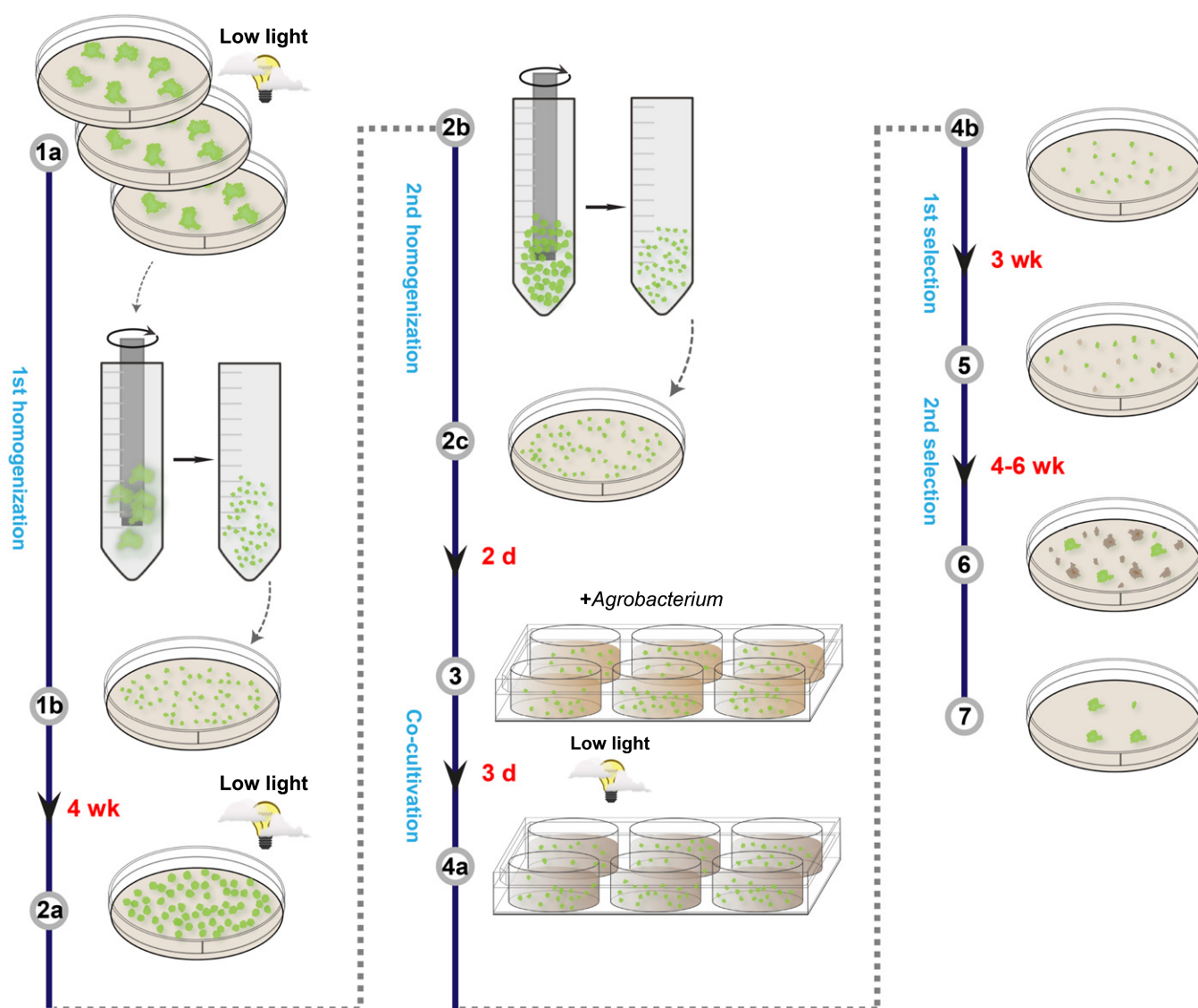
Fig. 2 Outline of *Anthoceros agrestis* transformation method. (1a–b), tissue grown under low light is homogenized, transferred to growth medium and placed again under low light conditions; (2a–c), after 4 wk, the tissue is homogenized again and grown for two additional days (the purpose of the first homogenization is tissue amplification); (3), the tissue is cocultivated with *Agrobacterium* for 3 d (under low light) and then (4a–b); spread on appropriate antibiotic-containing growth medium; (5), after 3 wk, the tissue is transferred again onto freshly prepared antibiotic-containing growth medium for a second round of selection; (6), after *c.* 4–8 wk, putative transformants are visible; (7), a final round of selection is recommended to eliminate false-positive transformants.

such as acetosyringone have been shown to be important for activation of *Agrobacterium* virulence genes (Stachel *et al.*, 1985). Cocultivation duration was 3 d at 22°C on a shaker without any light supplementation (only ambient light from the room). After cocultivation, the tissue was spread on solid KNOP plates supplemented with cefotaxime and hygromycin. After 3–4 wk 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; see 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 2 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 (Figs 3d, 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 CaMV 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 eGFP reporter gene (Cormack *et al.*, 1996). eGFP makes the identification of successful transformation events easier without the need of laborious GUS staining. For 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 BsaI, 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-TU constructs. More specifically, level 2 parts can be digested with BsaI to combine 16 TUs in a level 3 part, which could be digested again with SapI to create a level 4 part with 64 TUs 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 CaMV 35S promoter and terminated with a CaMV 35S polyadenylation signal. It also contained a *p-35S_s::eGFP-Lti6b* transcription unit (same CaMV 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 Notes S1 section 'Map of the transformation vector used for the generation of the *p-35S_s::eGFP-Lti6b* expressing lines' and 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 1). There was variability in eGFP expression patterns between different transgenic lines (Fig. S7) presumably due to differences in the transgene copy number or genome location of transgene insertion. eGFP expression remained mosaic during the first, second and third round of selection. During the initial round of selection, the first transgene expression (based on fluorescence) was observed in regenerating thallus fragments that consist of only a few cells (about 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 'Comparison of the CaMV 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. Fifteen lines have been propagated vegetatively for more than 2.5 yr without abolishing transgene expression (antibiotic resistance or eGFP).

Testing additional fluorescent proteins, Bonn isolate and transgene inheritance

Fluorescent proteins have been shown 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 & Beachy, 2008). To expand the palette of fluorescent proteins that can be used in *A. agrestis*, we tested the expression of the monomeric turquoise 2 (mTurquoise2) fluorescent protein (Kremers *et al.*, 2006; Goedhart *et al.*, 2012), the enhanced yellow fluorescent protein (eYFP) (Orm *et al.*, 1996) and the monomeric Venus (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 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 1).
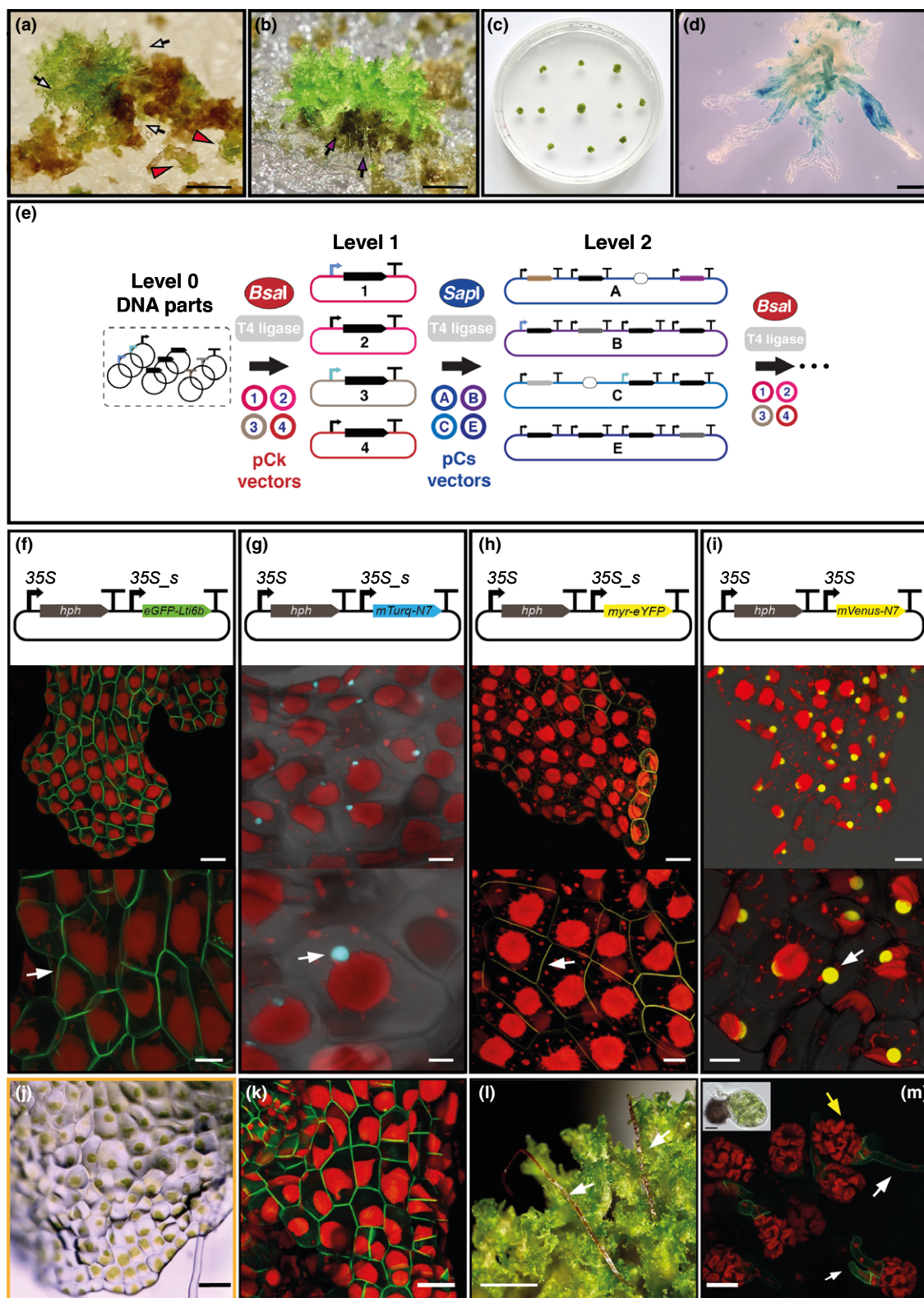
We then tested whether the protocol developed for the Oxford isolate can be used successfully for the Bonn isolate (Fig. S8). Four trials resulted in two successful transformants, which is considerably less than the average number of transformants obtained for the Oxford isolate (Fig. 3k).

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 gameto-phytes germinating from the spores (sporelings) were expressing

Fig. 3 Schematic representation of transformation constructs and transgenic *Anthoceros agrestis* expressing different fluorescent proteins. (a, b) The emergence of rhizoids (shown with (a) white and (b) purple 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. Bars, 2 mm. (c) Example of transgenic *A. agrestis* plants (gametophyte thallus). Petri dish dimensions: 92 × 16 mm. (d) GUS activity detected as blue staining in thallus tissue fragments from a plant transformed with the pCAMBIA1305.2 plasmid. 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 vectors, depicted with numbered circles, by BsaI-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 TUs: one TU for the expression of the *hygromycin B phosphotransferase* (*hph*) gene under 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). *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). Bars: (top) 50 µm; (bottom) 20 µm. (g) *p-35S_s::mTurquoise2-N7* for nuclear localization (white arrow). Bars: (top) 20 µm; (bottom) 10 µm. (h) *p-35S_s::myr-eYFP* for plasma membrane localization (white arrow). Bars: (top) 50 µm; (bottom) 20 µm. (i) *p-35S::mVenus-N7* for nuclear localization (white arrow). Bars: (top) 50 µm; (bottom) 25 µm. The bottom image is a magnification of the image in the middle. Red, Chl autofluorescence. (j) Light micrograph image: surface view of thallus (gametophyte), similar to the area imaged in (f) (and also g–i). Bar, 100 µm. (k) Images of *A. agrestis* Bonn gametophyte tissue expressing the *p-35S_s::eGFP-Lti6b* TU for eGFP plasma membrane localization. Bar, 50 µm. (l) *Anthoceros agrestis* Bonn with mature sporophytes indicated with white arrows. Bars, 2 mm. (m) *Anthoceros 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. Bar, 20 µm.

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-seq 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; Table S1). We amplified a 1532 bp long stretch of the 5′ flanking region including the 5′ untranslated region (UTR) of the *EF1a* gene, a 933 bp

Table 1 List of *Anthoceros agrestis* lines generated.

Construct	Total number of lines	Number of lines exhibiting fluorescence/GUS staining	Number of lines lacking 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	–
p-EF1a::hph- p-AaTip1;1::eGFP-Lti6b	1	1	–
p-EF1a::hph-p-35S_s::eGFP-Lti6b -p-EF1a: mTurquoise2-N7	3	2	1
p-EF1a::hph- p-35S::eGFP-Lti6b	3	3	–
p-35S::hph -p-AaActin_1::eGFP-Lti6b	3	–	3
p-35S::hph -p-AaActin_2::eGFP-Lti6b	9	–	9

See Supporting Information and Table S2 for promoter sequences and full plasmid maps.
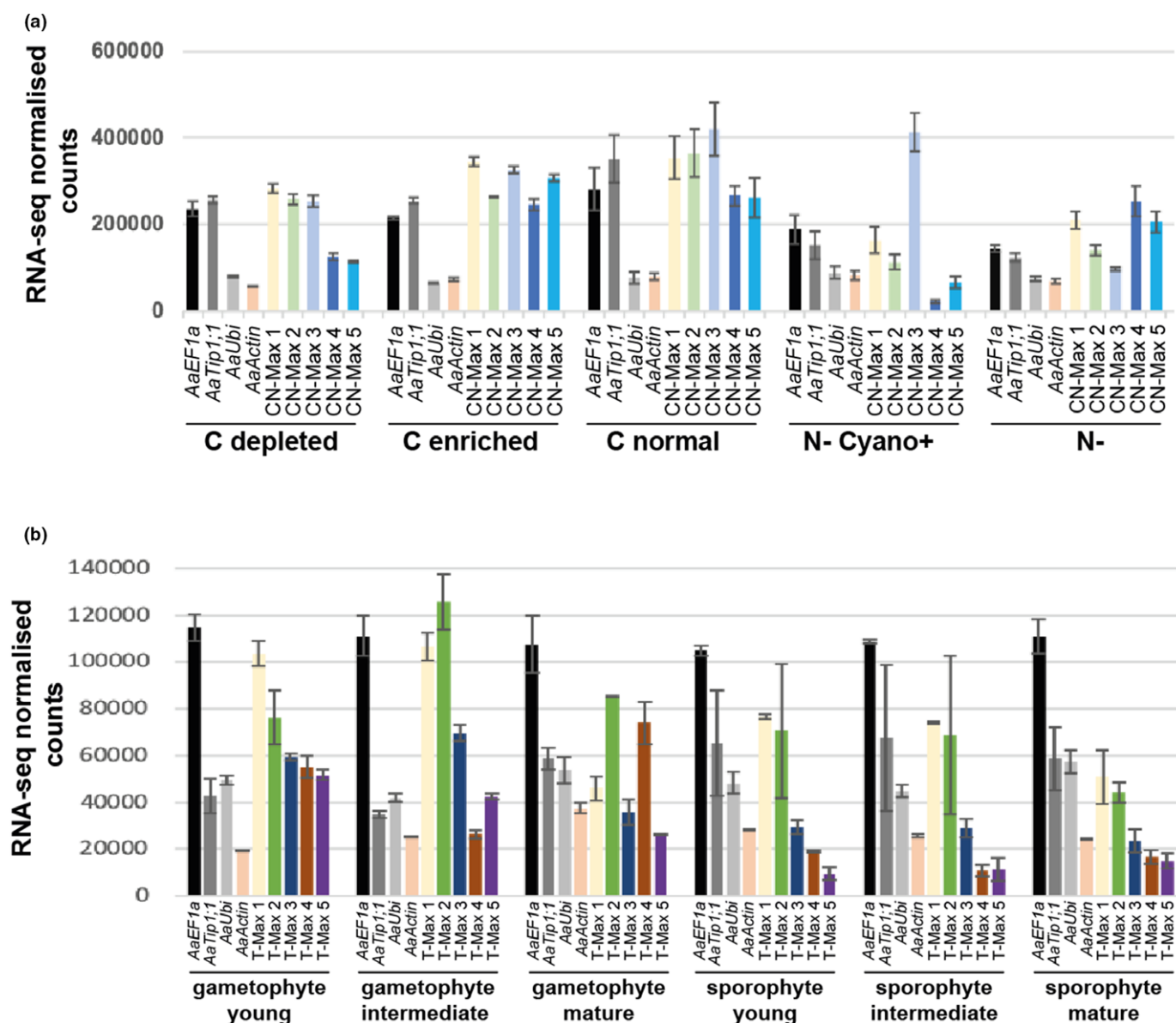
Fig. 4 Identification of constitutive promoters for *Anthoceros agrestis*. (a) Analysis of expression levels from RNA-seq experiments on *A. agrestis* Oxford using datasets from Li *et al.* (2020). To generate this dataset, gametophytes were grown under varying carbon sources in the growth medium (indicated 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 RNA-seq 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 Table S1).) Error bars indicate standard error based on two independent experimental replicates.

segment for the *Ubiquitin* (*Ubi*) gene, two fragments (1729 and 1516 bp) for the *Actin* gene (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 RNA-seq coverage tracks are shown in Fig. S9).

The candidate promoter regions were cloned (and if necessary domesticated 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
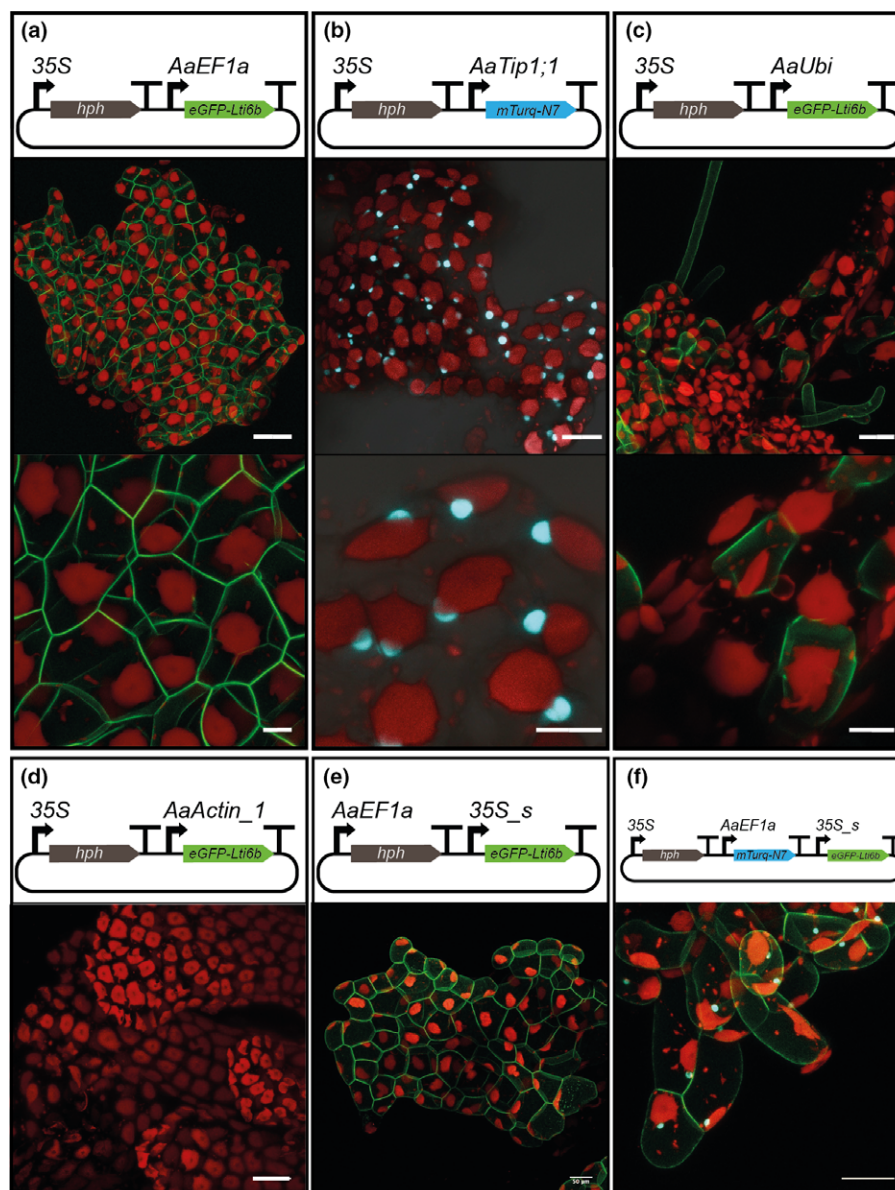
Fig. 5 Identification of constitutive promoters for *Anthoceros agrestis*. (a–f top) Schematic representation of constructs for the expression of two to three transcription units (TU): one TU for the expression of the *hph* gene under control of the CaMV 35S or the *p-AaEF1a* promoter and one TU for 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. Bars: (top) 50 μm; (bottom) 10 μm. (b) *p-AaTip1;1:mTurquoise2-N7* for nuclear localization. Bars; (top) 20 μm; (bottom) 10 μm. (c) *p-AaUbi::eGFP-Lti6b* for plasma membrane localization. Bars, top: 20 μm, bottom: 10 μm. The bottom images are a magnification of the images in the top. (d) *p-AaActin_1::eGFP-Lti6b*. Bars, 50 μm. (e) *p-AaEF1a::hph - p-35S_s::eGFP-Lti6b*. 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. Bar, 50 μm. Red, Chl autofluorescence.

(with one showing growth retardation probably due to the insertion site of the T-DNA) and one for the *p-AaTip1;1::eGFP* construct (Fig. S10). Thus, further characterization 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) (a 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 Fig. S10. 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
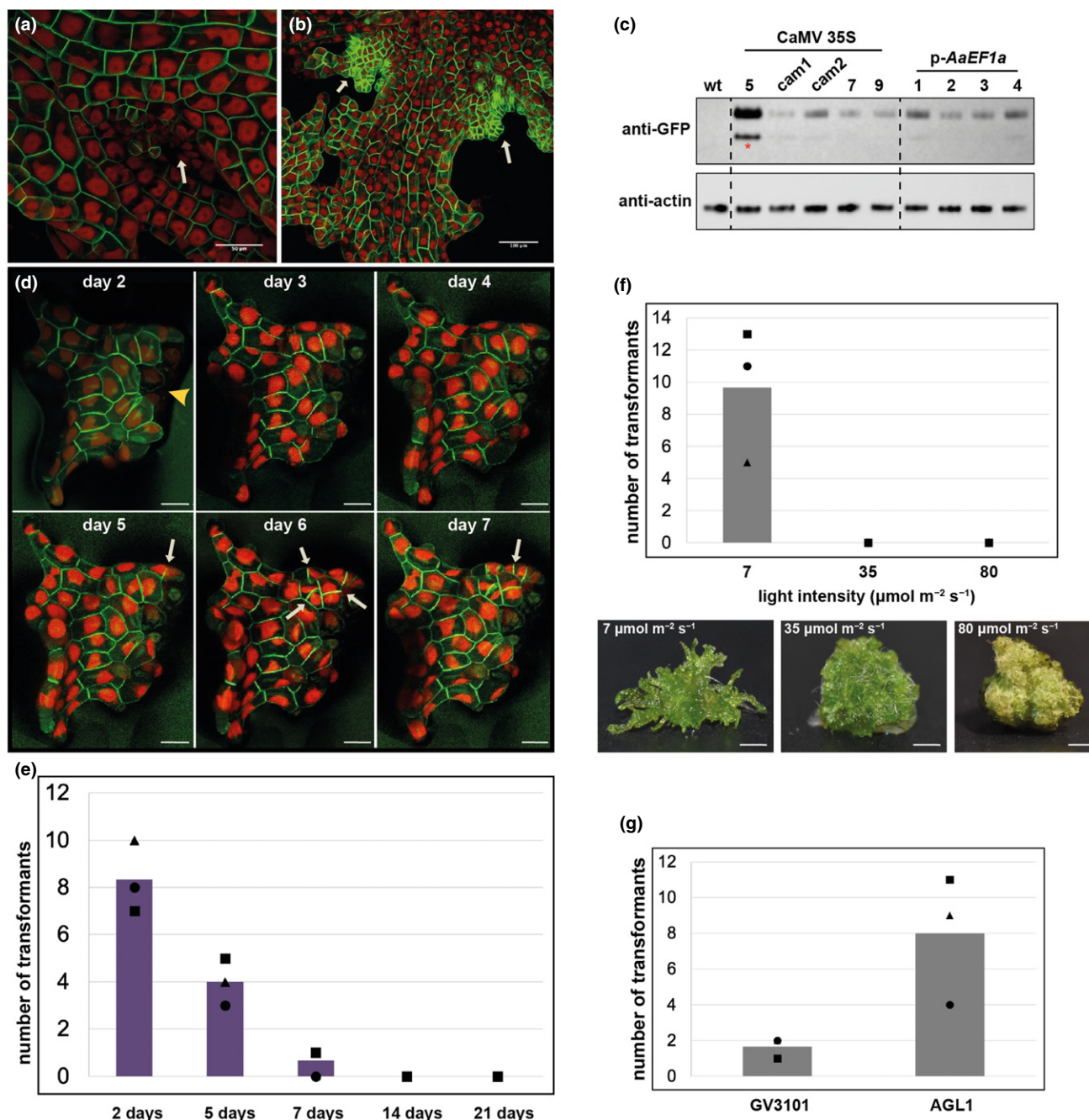
Fig. 6 Comparison of the CaMV 35S with the *AaEF1a* promoter and factors affecting the efficiency of *Agrobacterium*-mediated transformation of *Anthoceros agrestis*. (a) Expression of eGFP driven by the CaMV 35S promoter. Younger part of the thallus indicated with a white arrow. Bar, 50 μm. (b) Expression of eGFP driven by the *AaEF1a* promoter. Younger part of the thallus is indicated with white arrows. 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 corresponds 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 lines 5, 7 and 9 can be found in Fig. S7. Images of the four p-*AaEF1a* lines can be found in 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 homogenization. Five days after homogenization plants start to regenerate. Yellow arrowhead indicates the fragmented thallus part. White arrows indicate new cell divisions. Bars, 50 μm. (e) Number of transgenic lines obtained using *A. agrestis* thallus tissue 2, 5, 7, 14 and 21 d after homogenization. Values of three independent experimental replicates are shown. (f) Effect of light intensity used to grow plant tissue for transformation after the first homogenization on the number of successful transformants. Cocultivation started 2 d after the second homogenization. Values of three independent experimental replicates are shown. Bars, 1 mm. (g) Effect of *Agrobacterium* strain on the number of transformants obtained. Dots represent values of individual experimental replicates. Cocultivation started 2 d after the second homogenization. Values of three independent experimental replicates are shown.

successfully express simultaneously three different transcription units, *p-AaEF1a::mTurquoise2-N7*, *p-35S_s::eGFP-Lti6b* and *p-35S::hph* (Fig. 5f), which was the largest construct (*c.* 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 yr without abolishing transgene expression (antibiotic resistance or fluorescence protein).

Comparison of the CaMV 35S and *AaEF1a* promoters

Expression of eGFP driven by the CaMV 35S promoter seems to be weaker in newly grown parts of the thallus (Figs 6a, 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 (Figs 6b, 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, CaMV 35S drives the transgene more weakly in newly grown parts of the thallus, which needs to be taken into account when designing experiments.

Transformation efficiency and optimization

To estimate the transformation efficiency of the protocol, we performed 10 independent trials using as starting material *c.* 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 three to 23 (Table 2).

Table 2 Number of *p-35S::eGFP-Lti6b* transgenic lines of *Anthoceros agrestis* obtained from 10 different experiments to estimate transformation efficiency.

Experiment	Amount of tissue (g) used in first homogenization	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

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 homogenization, thereby affecting transformation efficiency. To estimate when plant regeneration is initiated, we set up a microscopy time course using homogenized thallus fragments. The first cell division was observed 5 d after homogenization (Fig. 6d). Based on this result, we carried out an optimization experiment starting cocultivation at 2, 5, 7, 14 and 21 d after homogenization. We found that the number of stably transformed lines decreased when using tissue that was recovered for more than 5 d after homogenization (Fig. 6e). The highest number of transformants could be obtained when using tissue 2 d after homogenization.

We also tested the effect of tissue growth (after the first homogenization) 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 (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-Lti6b* 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 (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 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 2 yr without abolishing transgene expression. Additionally, we confirmed that the transgene is integrated into the genome of *A. agrestis* and can be successfully inherited.
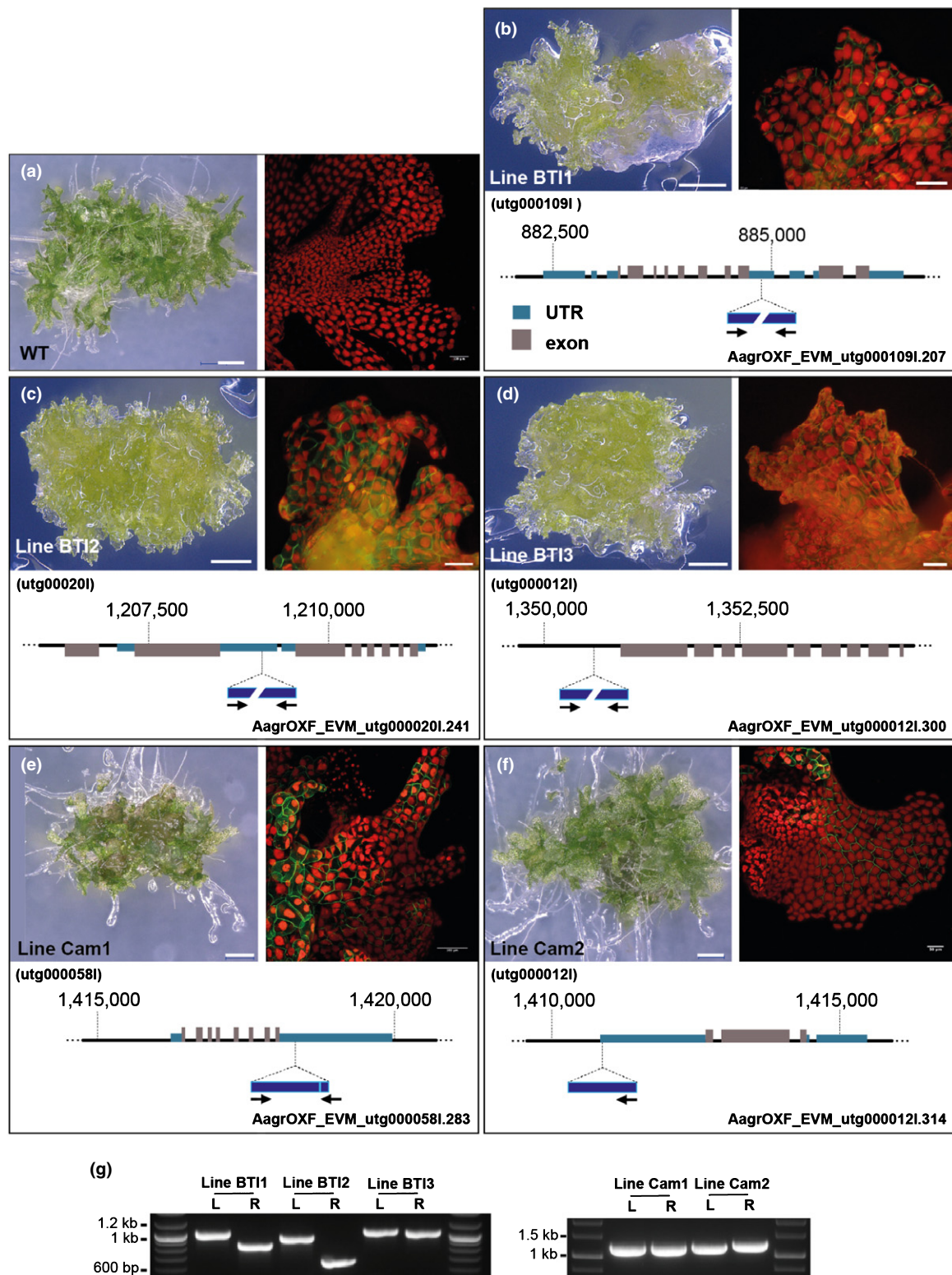
Fig. 7 Stable incorporation of transgene into *Anthoceros agrestis* genome. (a–f left) Light micrograph (LM) of transgenic thallus of *A. agrestis* plants. Bar, 500 μm. Right: confocal fluorescence microscopy images of thallus expressing eGFP in the plasma membrane, driven by the CaMV 35S promoter. Bars: (a–e) 100 μm; (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: LM images (a), (e) and (f) were acquired using a KEYENCE VHX-S550E microscope (VHX-J20T lens) and confocal fluorescence images with a Leica SP8X microscope, in Cambridge University. (d) LM and fluorescence images were acquired using a Leica M205 FA stereomicroscope with GFP longpass (LP) filter, in BT Institute.

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 & Marks, 1987; Ishizaki *et al.*, 2008; Plackett *et al.*, 2014). The utilization of recent high-throughput sequencing technologies combined with the genome size of *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 was 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 cytochrome P450 (CYP) gene family 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 the *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 cocultivation with *Agrobacterium*, such as vortexing, and use of scalpels or razor blades as an alternative to tissue homogenizers. The use of scalpels or razor blades resulted in successful transformation events for both Oxford and Bonn strains, but this method is still under optimization.

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).

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 analyzed the transcriptomic data. EF performed qPCR analysis. MW performed next-generation sequencing 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.

ORCID

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

Data availability

Accession numbers: raw sequencing reads have been submitted to the NCBI SRA under 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, USA* 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, Marks MD. 1987. *Agrobacterium*-mediated transformation of germinating seeds of *Arabidopsis thaliana*: A non-tissue culture approach. *Molecular and General Genetics MGG* 208: 1–9.

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. *Cytology and Genetics* 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.

Ikeuchi M, Sugimoto K, Iwase A. 2013. Plant callus: mechanisms of induction and repression. *Plant Cell* 25: 3159–3173.

Ishikawa M, Murata T, Sato Y, Nishiyama T, Hiwatashi Y, Imai A, Kimura M, Sugimoto N, Akita A, Oguri Y et al. 2011. *Physcomitrella* cyclin-dependent kinase A links cell cycle reactivation to other cellular changes during reprogramming of leaf cells. *The 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, Wong G-S, Mathews S, Pryer KM. 2015b. The origin and evolution of phototropins. *Frontiers in Plant Science* 6: 637.

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, USA* 111: 6672–6677.

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.

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, USA* 115: E2274–E2283.

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.

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 mini-metagenomes from chimeric MDA products. *Journal of Computational 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 K, Villarreal JC, Duff R. 2009. New insights into morphology, anatomy, and systematics of hornworts. In: Goffinet B, Shaw A, eds. *Bryophyte Biology*. Cambridge, UK: Cambridge University Press, 139–172.

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* **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, USA* **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* **101**: 1040–1056.

Söderström L, Hagborg A, von Konrat M, Bartholomew-Began S, Bell D, Briscoe L, Brown E, Cargill DC, da 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.

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. *Biotechnology Letters* **29**: 1591–1598.

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, USA* **109**: 18873–18878.

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, Martínez-Romero E, Kerr A, Sawada HA. 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.

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

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* **24**: 265–273.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Effect of light on *Anthoceros agrestis* growth and tissue culture.

Fig. S2 Hygromycin sensitivity of *Anthoceros agrestis* gametophytes (Oxford).

Fig. S3 Hygromycin sensitivity of *Anthoceros agrestis* gametophytes (Bonn).

Fig. S4 G418 sensitivity of *Anthoceros agrestis* gametophytes (Oxford).

Fig. S5 *Anthoceros agrestis* Oxford transformation protocol.

Fig. S6 GUS activity detected in transformed *Anthoceros agrestis* plants.

Fig. S7 Examples of *Anthoceros agrestis* Oxford transgenic lines I.

Fig. S8 *Anthoceros agrestis* Bonn transformation protocol.

Fig. S9 Promoters and their localization on the *Anthoceros agrestis* genome.

Fig. S10 Examples of *Anthoceros agrestis* Oxford transgenic lines II.

Notes S1 Native promoters, vector maps and insertion localization.

Table S1 Highly and/or constitutively expressed *Anthoceros agrestis* genes.

Table S2 List of construct sequences.

Table S3 List of primers used in this study.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.