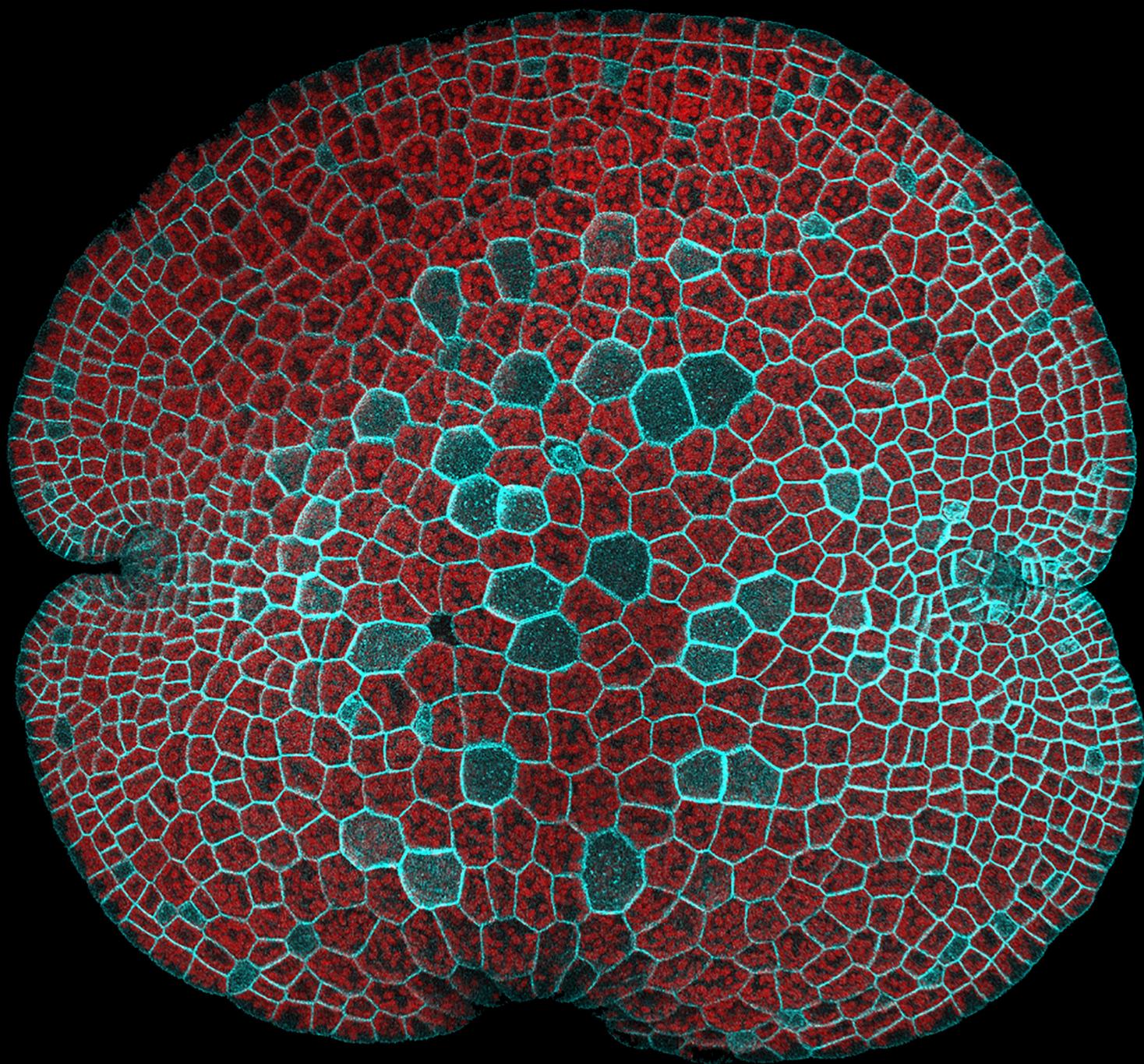


ACS Synthetic Biology

April 2020 • Volume 9, Issue 4

pubs.acs.org/acssynthbio



ACS Publications
Most Trusted. Most Cited. Most Read.

www.acs.org

Systematic Tools for Reprogramming Plant Gene Expression in a Simple Model, *Marchantia polymorpha*

Susanna Sauret-Güeto,[✉] Eftychios Frangedakis,[✉] Linda Silvestri, Marius Rebmann, Marta Tomaselli, Kasey Markel, Mihails Delmans, Anthony West, Nicola J. Patron, and Jim Haseloff*

Cite This: *ACS Synth. Biol.* 2020, 9, 864–882

Read Online

ACCESS |

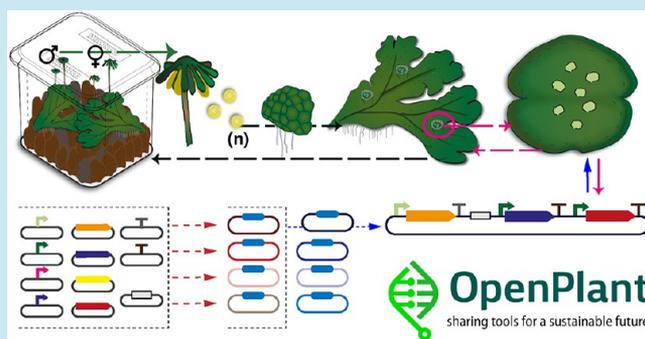
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: We present the OpenPlant toolkit, a set of interlinked resources and techniques to develop *Marchantia* as testbed for bioengineering in plants. *Marchantia* is a liverwort, a simple plant with an open form of development that allows direct visualization of gene expression and dynamics of cellular growth in living tissues. We describe new techniques for simple and efficient axenic propagation and maintenance of *Marchantia* lines with no requirement for glasshouse facilities. *Marchantia* plants spontaneously produce clonal propagules within a few weeks of regeneration, and lines can be amplified million-fold in a single generation by induction of the sexual phase of growth, crossing, and harvesting of progeny spores. The plant has a simple morphology and genome with reduced gene redundancy, and the dominant phase of its life cycle is haploid, making genetic analysis easier. We have built robust Loop assembly vector systems for nuclear and chloroplast transformation and genome editing. These have provided the basis for building and testing a modular library of standardized DNA elements with highly desirable properties. We have screened transcriptomic data to identify a range of candidate genes, extracted putative promoter sequences, and tested them *in vivo* to identify new constitutive promoter elements. The resources have been combined into a toolkit for plant bioengineering that is accessible for laboratories without access to traditional facilities for plant biology research. The toolkit is being made available under the terms of the OpenMTA and will facilitate the establishment of common standards and the use of this simple plant as testbed for synthetic biology.

KEYWORDS: *Marchantia*, plant, chloroplast, gene assembly, gene editing, toolkit, OpenPlant



Synthetic biology is a field that aims to bring formal engineering approaches to the design, construction, and analysis of biological systems.¹ Most recent advances in synthetic biology have used microbes as a platform. Transferring this approach to plants is challenging due to our limited understanding of plant genetic networks and increased complexity due to multicellularity.² However, plants harbor great potential for synthetic biology.³ Synthetic biology approaches can enable the production of compounds in specific organelles and cell types, useful for food, medicine, or industry. In plants, much energy production and biosynthetic capacity resides within chloroplasts. These photosynthetic organelles have a simple and prokaryote-like genome, with the ability to support high levels of transgene expression and metabolic activity. The proper harnessing of plants' capacity for bioproduction will require access to tools that allow multiscale manipulation of the nuclear and chloroplast genomes, building of synthetic metabolic pathways, and engineering of existing or new harvestable organs on the plant. A major benefit is that engineered plants can potentially

be grown at large scale, using globally accessible and low-cost agricultural production methods.

Much of our current knowledge of plant genetics and genome function comes from study of model angiosperm weeds and crops, such as *Arabidopsis*, maize, rice, tomato, and tobacco. Work with these plants is beset with one or more difficulties due to relatively long generation times, inefficient transformation, and slow regeneration, functional redundancy in diploid genomes and difficult access to early stages of growth. A simple prototype system is clearly needed. *Marchantia polymorpha* is the best-studied species of liverwort, an early divergent land plant lineage.⁴ *Marchantia* is proving to be a powerful and versatile experimental system for plant biology studies^{5–7} and an emerging platform for plant

Received: December 15, 2019

Published: March 12, 2020



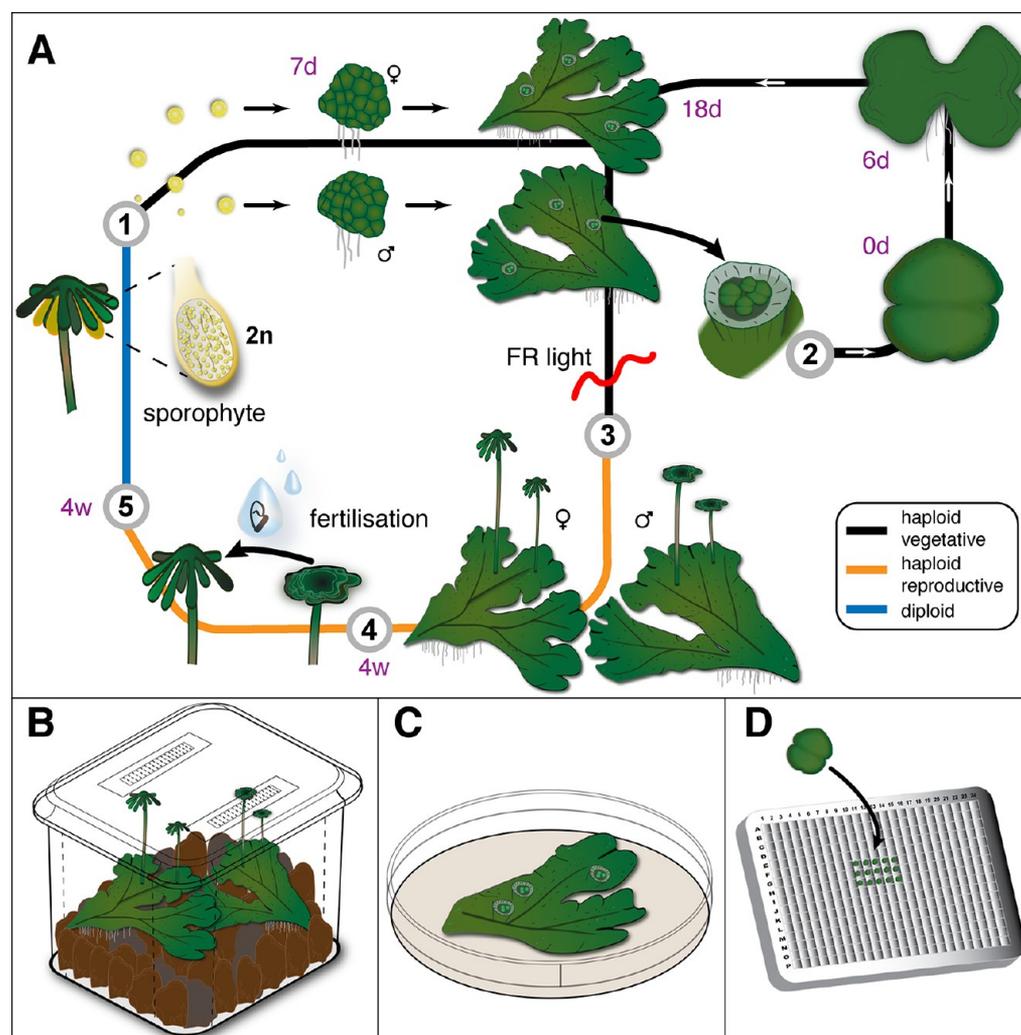


Figure 1. Use of the *Marchantia* life cycle for propagation in the laboratory. (A) *Marchantia* has two stages in its life cycle. As a bryophyte, the haploid (n) gametophyte form of the plant is dominant, larger, and longer-lived (stages 1–4) than the transient diploid ($2n$) sporophyte (stage 5). *Marchantia* is dioecious with separate male and female plants. The life cycle is shown starting with the germination of a mixture of male and female haploid spores (stage 1). These germinate to form male and female sporelings and plantlets after 7 days, which grow into similar-looking plants with a prostrate thallus. Asexual reproduction occurs via formation of clonal propagules called gemmae, which are produced spontaneously in cup-like structures on the dorsal side of both male and female thalli (stage 2). A gemma can be removed, germinated, and will develop into a plant with new gemmae cups in approximately 3 weeks. This provides a simple method for fast clonal propagation of lines. Vegetative tissues can be exposed to far red (FR, wavelength 730–740 nm) light in order to trigger the sexual phase of the life cycle (stage 3). Different reproductive organs develop on male and female plants, and mature within 4 weeks (w). Motile sperm cells are produced by male reproductive organs (antheridia) and can be transferred onto female reproductive organs (archegonia), where they can fertilize an egg cell (4). The resultant diploid zygote develops as a sporophyte (stage 5) and will produce a capsule (sporangia) with a large quantity of haploid spores (stage 1) after 4 weeks. (B) *Marchantia* can be propagated axenically throughout its life cycle for production of spores using Microboxes, with depth filtration for high-efficiency gas exchange and minimal water loss. (C) Petri dishes can be used for amplification of plantlets and gemmae propagation. (D) Gemmae can be cultured for microscopic observation in multiwell plates, allowing rapid screening of lines.

synthetic biology.⁸ *Marchantia* has a small size and grows fast and resiliently under laboratory conditions. In contrast to more recently diverged land plants, the dominant phase of its life cycle is haploid, with distinct male and female plants. It can be propagated both by spores derived from sexual reproduction and also through asexual clonal propagules called gemmae (Figure 1A). In just a week, both germinating spores and gemmae can give rise to tiny plantlets with a morphologically simple, sheet-like, and modular plant body called thallus. Thallus grows from apical stem cells located in invaginated notches at the thallus extremities.⁹ The early stages of development in *Marchantia* happen in an open fashion, allowing easy and direct observation of formative processes

during morphogenesis in a way amenable to rapid and high-throughput analysis. The morphological simplicity of *Marchantia* is matched by a small genome size (220 Mb) with low genetic redundancy.⁷ It shares most of the characterized gene regulatory components of more recently diverged land plants like angiosperms, but gene families are often sparsely populated, and haploidy makes genetic analysis easier.⁷ Altogether, *Marchantia* has the potential to be a unique multicellular testbed to study and engineer the links between genetic networks, cell processes, cell communication, cell fate decisions, and tissue-wide physical processes that drive morphogenesis.⁵

The implementation of synthetic biology approaches in plants calls for rapid and efficient methods for assembly of DNA circuits. Golden Gate Cloning methods (such as MoClo and GoldenBraid^{10,11}) are based on Type IIS restriction enzymes¹² and enable fast generation of complex DNA circuits from standardized basic DNA parts (e.g., promoters, CDS, and terminators). Golden Gate Cloning based toolkits have already been developed for other model plants.^{13,14} We have established a common syntax for modular DNA parts¹⁵ and developed a simplified recursive Type IIS assembly method, the Loop assembly.¹⁶

In this paper, we describe work that exploits these technical advances in DNA assembly and seek to improve and extend the set of compatible tools to develop an integrated system for reprogramming *Marchantia* and exploit its unique characteristics as a testbed for synthetic biology experiments in plants. (i) We adopted techniques to allow for axenic propagation of *Marchantia* throughout the life cycle, (ii) improved the throughput of *Agrobacterium*-mediated nuclear genome transformation of sporelings, and (iii) developed a simplified and more efficient technique for generation of homoplasmic plastid transformants. (iv) We swapped the plasmid backbone of Loop assembly vectors to provide higher yields of plasmid DNA, and (v) expanded the repertoire of vectors to allow chloroplast transformation and (vi) delivery of *Cas9* genes and guide sequences for efficient nuclear gene editing. (vii) We constructed a standardized and modular library of DNA parts for expression in *Marchantia*. These include antibiotic resistance genes for selection of successful transformants, signal peptides, and fluorescent proteins for multispectral imaging. (viii) We established higher throughput methods for screening transformed *Marchantia* plants using multiwell plates and fluorescence microscopy. Using this new suite of techniques, (ix) we surveyed the *Marchantia* genome for transcriptionally active genes, synthesized proximal promoter sequences, domesticated these as standardized DNA parts for Loop assembly, and constructed and tested reporter gene fusions *in vivo*. (x) We have created a novel set of compact gene elements with highly desirable properties, such as promoters that confer ubiquitous and specific patterns of gene expression. The increased scale of *Marchantia* experiments together with the need for standardized practices to support reproducibility, have demanded the establishment of research data management frameworks, and we have adopted the use of Benchling. Benchling provides a laboratory information management systems (LIMS), electronic notebook, molecular biology suite, and collaboration platform, and enables, for instance, a shared curated registry of annotated DNA parts and linked characterization data. (xi) To support access to the wider community, the OpenPlant toolkit sequences and metadata are shared through a Benchling public folder. (xii) We have also deposited plasmids in Addgene for global distribution under an OpenMTA license¹⁷ for public sharing and open access. The OpenMTA facilitates open exchange of biological materials and was designed to support openness, sharing, and innovation in global biotechnology. In contrast to the widely used UBMTA, the OpenMTA allows redistribution and commercial use of materials, while still acknowledging the creators and promoting safe practices. (xiii) We describe and distribute the relevant protocols at protocols.io/OpenPlant. The establishment of this set of open, interlinked resources and techniques for *Marchantia* is intended to encourage the establishment of a shared commons to facilitate scientific exchange and

innovation in advanced bioengineering with this new testbed for plant synthetic biology.

RESULTS AND DISCUSSION

Axenic Culture for Spore Production. Certain specialized features of the *Marchantia* life cycle provide substantial advantages to the experimental scientist. We have adopted techniques to allow axenic culture of *Marchantia* throughout its life cycle (Figure 1). The haploid gametophyte form of the plant is dominant, larger, and longer-lived than the transient diploid sporophyte stage (Figure 1A). The gametophyte takes the form of a flattened thallus with rhizoids emerging from the ventral surface, with photosynthetic chambers and other specialized cellular structures forming on the dorsal surface (Figure 1A-1).

Marchantia can reproduce both asexually and sexually. Asexual reproduction occurs via clonal propagules called gemmae, which are formed inside conical splash cups (Figure 1A-2). Each gemma is derived from a single epidermal cell at the base of a cup that expands and divides to form a group of cells carried on a short stalk. The cells continue to proliferate in regular fashion to form a bilobed gemma with a conserved morphology. This is eventually detached from the stalk and can be dispersed from the cup, typically by water splash. Gemmae can be used for simple vegetative propagation and amplification of plantlets during experiments, grow well in Petri dishes (Figure 1C), and can be stored in the cold or cryopreserved when long-term storage is necessary. We have simplified a previously published cryopreservation method to avoid the need for cryoplates, making it a more routine process¹⁸ (Figure S4 and Materials and Methods). Gemmae grow in exposed fashion, allowing easy access to morphogenetic processes, and we have adapted multiwell plates for culture and screening of lines under microscopic observation (Figures 1D and S16, and Materials and Methods).

Upon induction of formation of male and female reproductive organs (gametophores) and sexual reproduction, each cross can result in the formation of millions of single cell spores⁵ (Figure 1A-1). Efficient nuclear and chloroplast transformation methods both rely on the use of sterile sporelings (germinating spores) as a target tissue. In order to allow reliable production of large quantities of sterile *Marchantia* spores, we have adopted the use of Microbox micropropagation containers with a specially designed lid that allows high gas exchange and limited dehydration and blocks the entry of contaminants such as bacteria and fungi into the culture (Figures 1B and S1 and Materials and Methods). Plantlets or gemmae grown under axenic conditions (Figure 1C) can be used to inoculate the containers and grown in a growth room or controlled environment cabinet with limited need for watering or other maintenance. Supplementarily far red light triggers the sexual phase, with efficient formation of male and female gametophores after 4 weeks in Cam-1 and Cam-2 strains of *M. polymorpha* (Figure S1). In our experience, the widely used Tak-1 and Tak-2 isolates can be slower to respond to far-red light induction. The Cam-1 and Cam-2 isolates have not been extensively back-crossed yet, so a small degree of genetic variability will be introduced after cross fertilization, but no phenotypic variation is evident. Mature sperm can be harvested from antheridia (male sex organs), diluted, and transferred to archegonia (female sex organs) (Figures 1A-4 and S1). Yellow-colored sporangia containing huge numbers of spores, emerge after a further 4 weeks (Figure 1A-5), and can

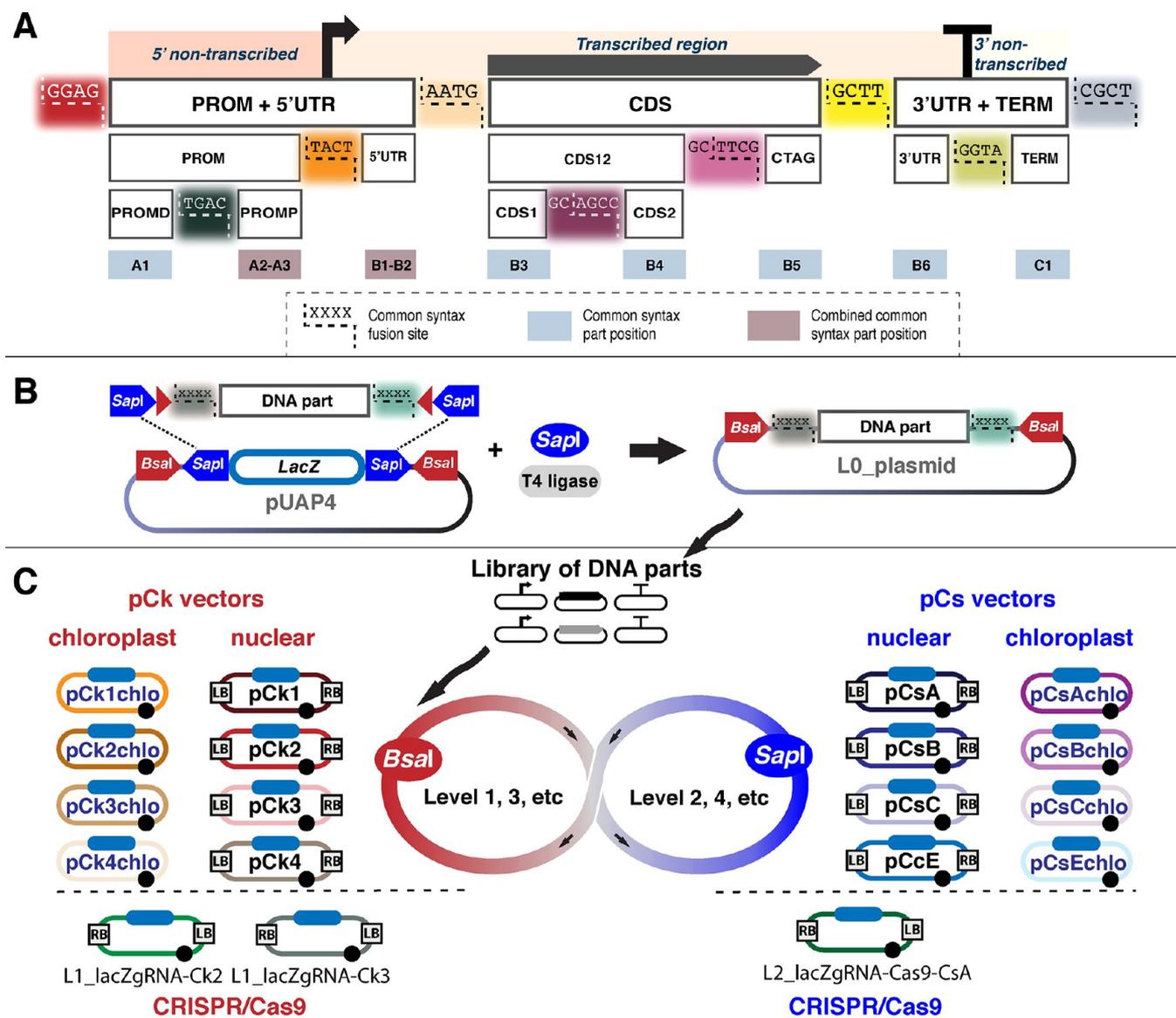


Figure 2. Key elements in the OpenPlant Loop assembly toolkit. (A) The Phytobrick common syntax defines 10 DNA part positions and 12 DNA part fusion sites. For *Marchantia*, we commonly use eight positions and nine fusion sites, by combining positions A2–A3 for proximal promoter (PROMP), and B1–B2 for 5' untranslated region (5'UTR). The other types of parts are A1 for distal promoter (PROMD), B3 for coding sequence with start codon and no stop codon (CDS1), B4 for coding sequence without start or stop codon (CDS2), B5 coding sequence without start codon and with stop codon (CTAG), B6 for 3' untranslated region (3'UTR), and C1 for transcription terminator (TERM). Parts can span multiple fusion sites, like A1–A3 for promoter (PROM), A1–B2 for promoter with 5' UTR (PROM5), B3–B4 for coding sequence with start codon and no stop codon for N-terminal fusion with CTAG (CDS12), B3–B6 for coding sequence with start and stop codons (CDS), or B6–C1 for 3' UTR with terminator (3'UTR). (B) Schematic representation of the pUAP4 vector with divergent *SapI* sites to accept L0 parts and convergent *BsaI* sites to assemble L0 parts into transcription units (L1). (C) Summary of the Loop acceptor vectors of the OpenPlant toolkit. For nuclear genome transformation: pCk (1,2,3,4) can be used for assembly of L0 parts into a Level 1 plasmid using *BsaI*, and pCs (A,B,C,E) can be used for assembly of up to four Level 1 plasmids into a Level 2 construct using *SapI*. The vectors pCkchlo (1,2,3,4) and pCsChlo (A,B,C,E) can be used for chloroplast applications. The vectors L1_lacZgRNA-Ck2, L1_lacZgRNA-Ck3, and L2_lacZgRNA-Cas9 are designed for CRISPR/Cas9 genome editing. LB and RB: left and right border repeats respectively from nopaline C58 T-DNA. Filled blue rounded rectangle: *lacZa* cassette for blue-white screening. Filled black circles: *pSa* origin or replication.

be stored in a cold, desiccated state (Figure S1 and Materials and Methods) and sterilized (Materials and Methods) prior to plating on suitable medium for germination.

Transformation of Sporelings. We have improved the throughput of *Agrobacterium*-mediated nuclear genome transformation of sporelings¹⁹ by scaling down previous protocols in order to use multiwell plates instead of flasks. This allows one to easily perform multiple transformations in one experiment, eliminating the transformation step as a bottleneck

for the production of lines (Figure S2 and Materials and Methods). Positive transformants are obtained in about 2 weeks (Figure S2). The OpenPlant toolkit contains three antibiotic resistance genes commonly used as selectable markers for *Marchantia*:²⁰ the hygromycin phosphotransferase gene (*hptII*), the modified acetolactate synthase gene (*mALS*), and the neomycin phosphotransferase II gene (*nptII*), that confer hygromycin, chlorsulfuron, and kanamycin resistance, respectively (see Figure 8).

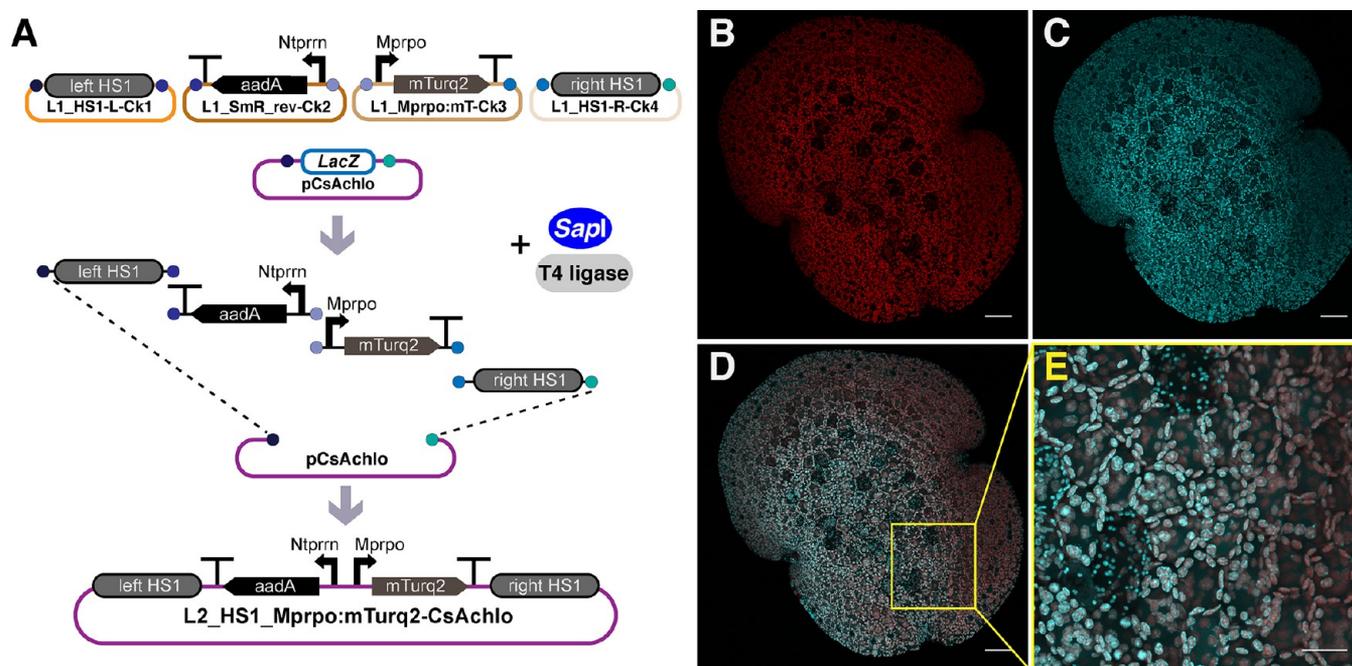


Figure 3. OpenPlant DNA tools for chloroplast genome modification. (A) Schematic representation of Level 2 Loop assembly of a device to express the chloroplast codon optimized mTurquoise2 fluorescent protein (mTurquoise2_chlo) under the control of the native Mprpo promoter using the homologous sequences left-HS1 and right-HS1 for integration in the chloroplast *trnG-trnfM* intergenic region. (B–E) Maximum intensity Z projection confocal images of *Marchantia* transplastomic gemmae transformed with the construct in A. (B) Chlorophyll autofluorescence channel, (C) mTurquoise2_chlo channel, and (D) merged channels. (E) Higher magnification images showing mTurquoise2_chlo accumulation inside the chloroplasts of all cells. Scale bar: 50 μm (B–D), 20 μm (E).

Chloroplast transformation of *Marchantia* sporelings is achieved through direct biolistic delivery of DNA-loaded microcarriers into the plastids. DNA constructs contain flanking homologous sequences and are integrated into a specific chloroplast genome location via homologous recombination.²¹ Previously published chloroplast transformation methods in *Marchantia* have required the bombardment of a large number of samples to generate a small number of transformed plants.²² We have modified the protocol by adopting nanoparticles called DNAdel as plasmid DNA carriers (Figure S3 and Materials and Methods). This reduces the time and labor required for loading of the plasmid DNA onto the microcarrier and increases efficiency and reproducibility. *Marchantia* sporelings regenerate after biolistic delivery without need for hormone treatment, and typically we can use the spores of a single sporangium to obtain about 10 transplastomic plants after 8 weeks of selection on antibiotic-containing growth media (Figure S3). PCR analysis of transformants in a typical experiment ($n = 30$) showed that greater than 90% of plants were homoplasmic (Figure S10). The OpenPlant kit includes the aminoglycoside adenyltransferase (*aadA*) resistance gene which confers spectinomycin resistance²² (see Figure 8).

DNA Part Syntax and Domestication. The OpenPlant consortium initiated the establishment of the common genetic syntax with wide support from the international plant science community,¹⁵ and we have adopted this convention for the composition (and exchange) of plant DNA parts (Phyto-bricks). The common syntax defines specific 5' end and 3' end fusion sites, of four nucleotides in length, for each type of standardized DNA part (called Level 0 parts, L0) (Figure 2A). In this way, it facilitates community-based efforts to build and exchange collections of L0 parts. Examples of different types of

L0 parts are promoter with 5' UTR (PROM5), coding sequence with start and stop codons (CDS), or 3' UTR with terminator (3'TERM). The 5' end and 3' fusion sites on each L0 part enables directional assembly of multiple L0 parts (Figure 2A, e.g., ggag-PROM5-aatg-CDS-gctt-3'TERM-cgct). We provide all L0 DNA parts cloned into a new universal acceptor plasmid called pUAP4 (Figure 2B) which is a derivative of the pUAP1, an OpenMTA vector used for cloning of L0 parts in the MoClo kit.¹³ pUAP4 is designed for Loop assembly, with *SapI* divergent sites flanking the L0 DNA part insertion site, in contrast to the *BbsI* sites in pUAP1 (Figure S5). Also, instead of the red fluorescence protein (RFP) as a marker for screening of colonies, the pUAP4 encodes LacZ activity, which allows for blue-white screening and is more visible by the naked eye than the RFP (Figure S5). The pUAP4 vector also has convergent *BsaI* sites to assemble multiple L0 parts into a transcription unit using *BsaI* (Figures 2C and S7). When a L0 part has internal *BsaI* or *SapI* sites, it is necessary to domesticate the DNA part (to remove restriction endonuclease recognition sites that would otherwise interfere with the cloning procedure), which can be done together with cloning the part into pUAP4 (Figure S6).

The OpenPlant toolkit is composed of a set of 50 L0 parts: 11 promoters and 5'UTR parts (PROM5 = PROM+5'UTR, PROM and 5'UTRs), 33 CDS-type parts (CDS, CDS1, CDS2, CDS12, and CTAG), and 6 3'UTR and terminators (3'TERM = 3'UTR+TERM) (Figure 8). L0 DNA parts were either amplified by PCR and cloned, or synthesized into the pUAP4 vector. All L0 parts have been verified by sequencing.

Recursive Assembly of Gene Parts. Pollak et al.¹⁶ previously described a method for gene assembly based on recursive use of two Type IIS restriction endonucleases, *SapI* and *BsaI*, called Loop assembly. Loop assembly allows rapid

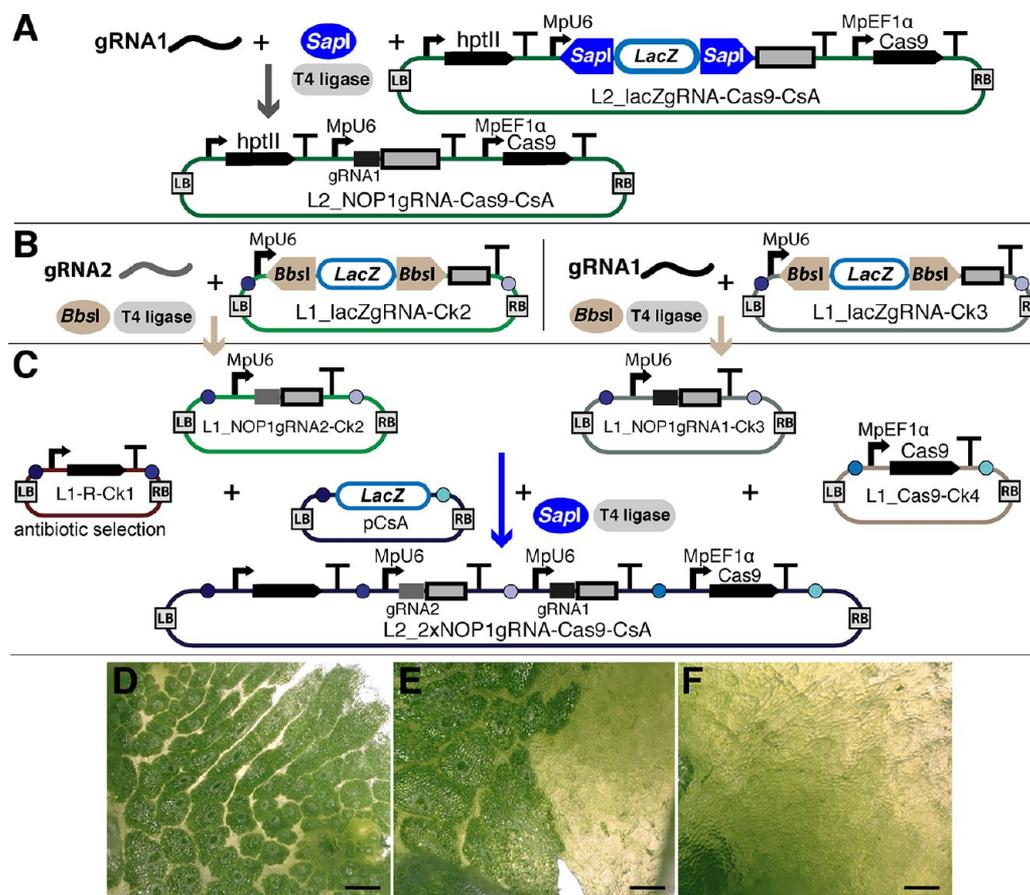


Figure 4. OpenPlant toolkit vectors for CRISPR/Cas9 genome editing. (A) Schematic representation of the plant transformation vector for simultaneous delivery of a single gRNA and Cas9 into *Marchantia* plants by *Agrobacterium*-mediated transformation. The gRNA sequence (in the example NOP1-gRNA1) was directly ligated into the L2_lacZgRNA-Cas9-CsA vector via *SapI* type IIS assembly. (B, C) Schematic representation of assembly of the plant transformation vector for delivery of two gRNAs and Cas9 into a *Marchantia* plant. (B) The two gRNA sequences were ligated into L1_lacZgRNA-Ck2 and L1_lacZgRNA-Ck3 vectors, respectively, via *BbsI* type IIS assembly. In the example, NOP1-gRNA2 into Ck2 and NOP1-gRNA1 into Ck3. (C) Then the two L1_gRNA transcription units were combined with an antibiotic resistance transcription unit and a MpEF1 α :Cas9 transcription unit via *SapI* Loop assembly. (D–F) The three phenotypic classes of regenerating *Marchantia* plants transformed with a L2 vector with NOP1gRNA and Cas9: wild-type (D), chimeric (E), and mutant (F). Scale bar: 400 μ m.

and efficient production of large DNA constructs for nuclear DNA transformation and is compatible with widely used standardized DNA parts such as Phytobricks. In a single-step reaction with *BsaI* and T4 ligase, L0 parts can be assembled into a transcription unit (Level 1, L1) (Figures 2C and S7). Then L1 constructs can be assembled together into DNA assemblies with up to four transcription units (Level 2, L2), in a single-step reaction with *SapI* and T4 ligase (Figure S8). The recursive nature of Loop assembly means that four L2 devices can be digested with *BsaI* and ligated to combine up to 16 transcription units in a Level 3 device, and so on (Figure 2C). Loop assembly is well suited to automation, and we developed methods for assembly using an acoustic-focusing non-contact liquid handling robot, which increases the speed and scale of assembly, while allowing reactions to be performed in sub-microliter volumes to reduce cost. We have scaled down both L1 and L2 assemblies to 500 nL using the Labcyte Echo acoustic focusing robot (Materials and Methods).

Modified Vector Systems. The manuscript describes an extension of the Loop vector system¹⁶ (Figure 2C). The plasmid backbones of the original Loop vectors were based on pGREEN II,^{23,24} which contains ColE1 and pSA origins of replication. We have replaced the plasmid backbone with that

from pCambia/pPZP.²⁵ The 7.5 kb pCambia backbone encodes a VS1 origin, the pVS1 RepA replication protein, and StaA protein that confer plasmid stability and higher yields in *E. coli* and stable maintenance in *Agrobacterium*. All of the modified Loop vectors used for *Agrobacterium* mediated transformation of the nuclear genome contain the right and left border repeats from nopaline CS8 T-DNA (RB and LB) derived from pCambia, and show similar efficiencies of transformation to other plasmids based on this well-tested backbone. The new odd (L1, L3, ...) acceptor vectors are collectively called pCk vectors (pCk1, pCk2, pCk3, and pCk4), and the new even (L2, L4, ...) acceptor vectors are collectively called pCs vectors (pCsA, pCsB, pCsC, and pCsE) (Figure S9A,B). We have not observed untoward toxic effects due to the increase in plasmid copy number, and the new plasmids are easier to harvest in useful quantities for routine cloning operations, without the need to increase the scale of host microbial cultures.

We have expanded the repertoire of vectors to include not only wide-host range plasmids for *Agrobacterium*-mediated nuclear transformation, but also vectors for chloroplast transformation via biolistic DNA delivery, and vectors for CRISPR-Cas9 mediated genome editing (Figure 2C).

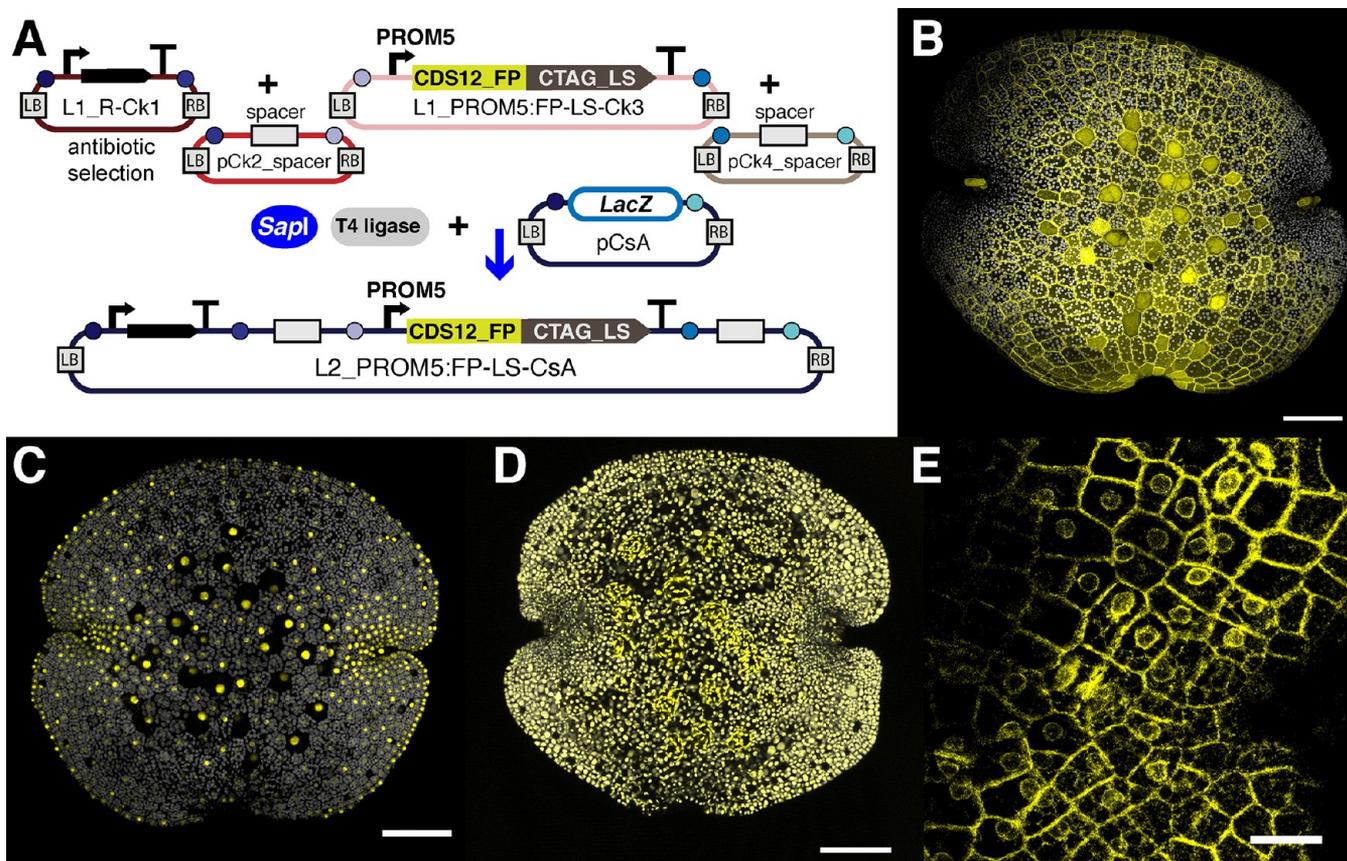


Figure 5. Testing different protein localization using the OpenPlant toolkit. (A) Schematic representation of the Level 2 Loop assembly of a construct for Agrobacterium-mediated nuclear transformation, built from Level 1 plasmids that contain an antibiotic resistance transcription unit in position 1 and gene for a fluorescent protein (CDS12_FP) fused to a localization signal (CTAG_LS) in position 3. Positions 2 and 4 contain spacer sequences. (B–E) Images of gemmae expressing different combinations of PROM5:FP-LS, with FP signal shown in yellow and chlorophyll autofluorescence in gray for all images. (B) p5-35S:mScarletI-LTI6b for plasma membrane localization (LTI6b is from the *Arabidopsis* gene AT3G05890). (C) p5-MpEF1 α :mTurquoise2-N7 for nuclei localization (N7 is from the *Arabidopsis* ankyrin-like protein, AT4G19150). (D) p5-MpEF1 α :MpSIG2-mVenus with chloroplast signal peptide from *Marchantia* SIG2 Mapoly0214s0004. (E) p5-35S:MpER-Targ-eGFP-HDEL with ER targeting peptide from a predicted *Marchantia* chitinase Mapoly0069s0092 (Figure S13). Images of gemmae at 0 d (B–D) or 4 d (E). Scale bar: 100 μ m (B–D), 20 μ m (E).

DNA Tools for Chloroplast Genome Modification. The Loop set of vectors for chloroplast transformation are collectively called pCkchlo (pCk1chlo, pCk2chlo, pCk3chlo, and pCk4chlo) and pCschlo (pCsAchlo, pCsBchlo, pCsCchlo, and pCsEchlo) (Figure S9C,D). These vectors are reduced derivatives of pCk and pCs, respectively, generated after removal of LB, RB, and elements necessary for stability in *Agrobacterium*. For most chloroplast applications, L2 assemblies are sufficient. A L2 device is integrated into the chloroplast genome via homologous recombination. For customized targeted integration (Figure 3), positions L1-Ck1 and L1-Ck4 contain the homologous sequences of the native chloroplast genome, positioned upstream and downstream to the integration site. Positions L1-Ck2 and L1-Ck3 contain the transcription units for integration into the chloroplast genome.

The OpenPlant kit contains left and right homologous sequences to allow integration in two intergenic regions (Figure S10). For integration in the *trnG-trnfM* region: the left homologous sequence (HS1-L, genome position: 40216–41843, Genbank accession no. MH635409) and right homologous sequence (HS1-R, genome position: 41847–43622, Genbank accession no. MH635409). For integration in the *rbcl-trnR* region: the left homologous sequence (HS2-L,

genome position: 56139–57662, Genbank accession no. MH635409) and the right homologous sequence (HS2-R, genome position: 57681–59304, Genbank accession no. MH635409). It should be noted that the homologous chloroplast sequences are derived from the *Marchantia* Cam-1/2 isolates, and the HS1 and HS2 sequences contain several nucleotide polymorphisms across their 1.5 kb lengths, compared to the Tak-1 isolate, for example.

The OpenPlant toolkit also contains the chloroplast codon optimized mTurquoise2 fluorescent protein coding sequence (mTurquoise2_chlo)²² (Figures 3 and S10). In addition to the previously used promoters from *Nicotiana tabacum*: the photosystem II protein D1 gene promoter (pNtpsba) and the plastid ribosomal RNA operon promoter (pNtprrrn),²² the kit also contains the *Marchantia* native chloroplast DNA-dependent RNA polymerase operon promoter (pMprpo) (genome position: 5591–5739, Genbank accession no. MH635409) (Figure 8). All parts have been functionally validated (Figure S10).

DNA Tools for Nuclear Genome Editing. CRISPR-Cas9 systems can be used to introduce double strand breaks at target DNA sequences *in vivo*.^{26,27} This allows the introduction of small deletion or insertion mutations via non-homologous end

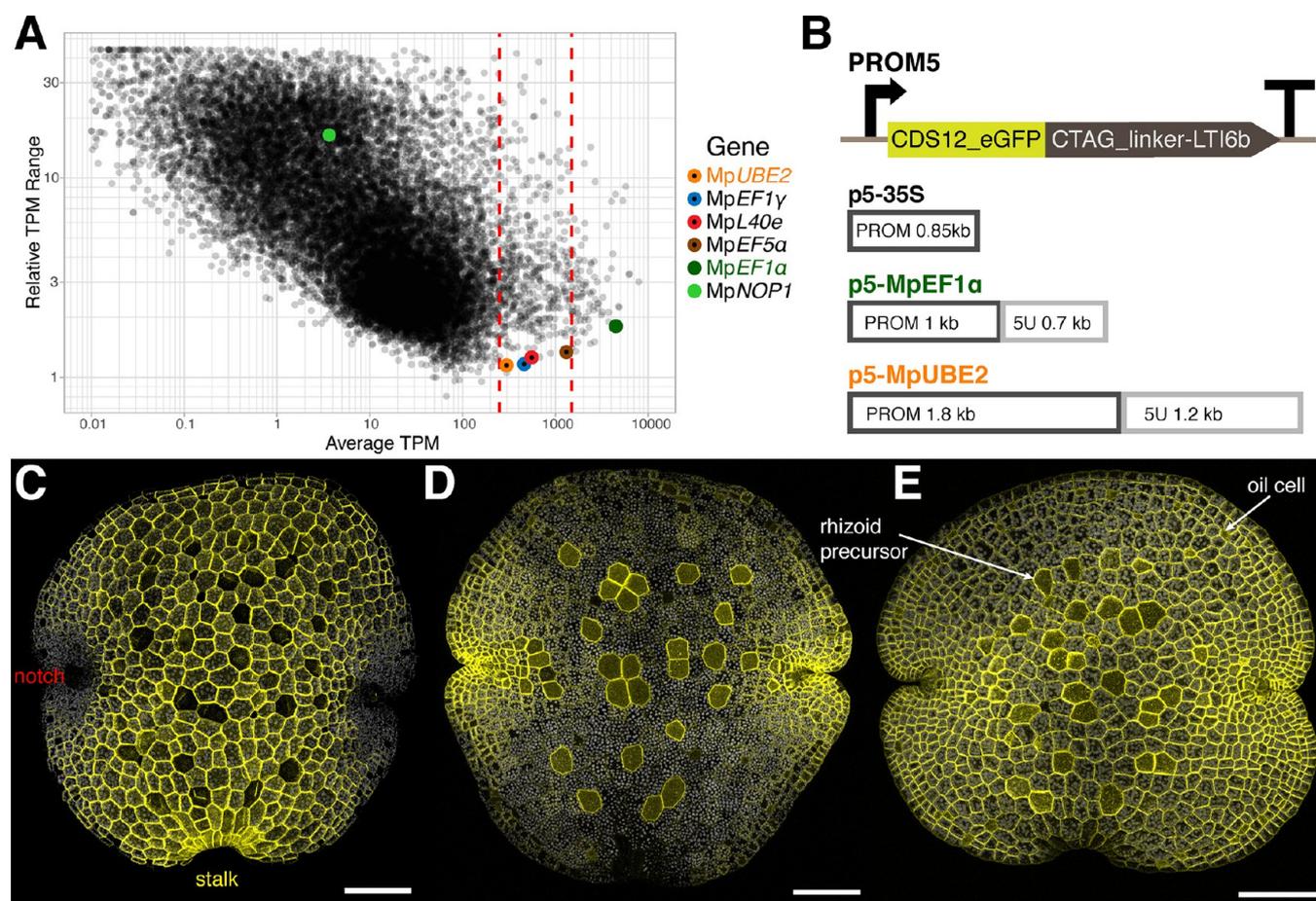


Figure 6. Identification of a ubiquitous and evenly expressed promoter for *Marchantia* gemmae. (A) Analysis of expression levels from publicly available RNAseq experiments for Mp genes. Variability of expression versus average expression levels for each gene (as relative TPM range $[(\max\text{TPM}-\min\text{TPM})/\text{avgTPM}]$ versus average TPM). Genes with an average TPM in between 250 and 1500 are shown bracketed by dashed red lines. Colored dots represent 4 housekeeping genes that were tested as a source of constitutive promoters (including MpUBE2). Dark and light green dots represent MpEF1 α and MpNOP1, with average TPM of ~ 4500 and ~ 25 , respectively. (B) Schema of transcription units for expression of plasma membrane localized eGFP. The eGFP-LTI6b gene was fused to the promoters: p5-35S (C), p5-MpEF1 α (D), and p5-MpUBE2 (E). (C–E) Images of gemmae just removed from the gemma cup (0 days old). Gemmae oriented with the two notch regions in the middle of the image and at the bottom, the region that was attached to the stalk in the gemma cup. There are two differentiated cell types on the surface of the gemmae that can be distinguished by their morphology: rhizoid precursors and oil cells (E). MpUBE2 is ubiquitously expressed at similar levels in all cells, while variable levels of expression were seen with the commonly used promoters 35S and MpEF1 α . (C–E) eGFP is shown in yellow and chlorophyll autofluorescence in gray. Scale bar: 100 μm .

joining reactions at specific chromosome locations. CRISPR-Cas9-based targeted mutagenesis has been demonstrated in *Marchantia*,²⁸ albeit with low frequency. This made the isolation of successful mutation events challenging. The efficiency of gene editing in plants has been improved by employing Cas9 genes with altered codon usage.²⁹ We chose a codon optimized Cas9 gene from the Puchta lab that was developed for use in *Arabidopsis thaliana*,³⁰ and this has proved to have higher genome editing efficiency than an earlier version.³¹

The OpenPlant kit contains a L2 vector (L2_lacZgRNA-Cas9-CsA) that allows *Agrobacterium*-mediated co-transformation of the Cas9 protein (expressed under the control of the p5-MpEF1 α promoter) and a guide RNA (gRNA) expression cassette (under the control of the *Marchantia* U6 promoter) (Figure 4A). To target a gene, the corresponding gRNA sequence is directly cloned into the L2 vector using SapI mediated Loop assembly (Figure S11A and Materials and Methods). Dual gRNA editing can be achieved by combining two L1 vectors with gRNA expression cassettes into an L2

vector (Figure 4B,C). The gRNA target sequences are cloned into the L1 vectors using BbsI mediated Loop assembly (Figures 4B and S11B and Material and Methods). Dual gRNA editing allows, for instance, the generation of knockout mutants with longer DNA deletions by targeting two nearby sites in the genome (Figure S12).

To test the efficacy of the new vectors (expressing either a single gRNA or a pair of gRNAs), we targeted the E3 ubiquitin ligase gene *Nopperabo1* (MpNOP1, Mapoly0008s0036) gene. Mutation of this gene results in loss of air chambers, an easily scored and non-lethal phenotype.³² The delivery of either single or double gRNAs resulted in around half of the primary transformants showing mutant phenotypes. We observed three categories of regenerating sporelings, wild-type, chimeric, and mutant plants respectively (Figure 4D–F). The percentage of mutant plants with no observable air chambers was between 16 and 23% of the total (Figures 4F and S12 and Table S2). The OpenPlant toolkit also contains L0 parts with the Cas9 protein as CDS and the MpU6 promoter,²⁸ allowing its users to create various devices for genome editing based on need (Figure 8).

DNA Tools for Nuclear Genome Transformation. The Loop pCk and pCs vectors allow assembly of constructs that combine different gene elements and can contain multiple transcription units. In each Loop reaction, four vectors of the same level need to be ligated together, and thus, if wanting to assemble less than four units, for the vector(s) without a unit, a spacer will be used instead (an arbitrary 200 base pair DNA sequence).¹⁶ Spacer sequences are provided in the form of four pCk and four pCs plasmids. An example of use is shown in Figure 5A with pCk2 and pCk4. Loop assembly plasmids contain four ordered sites for sequence insertion, and the LB is adjacent to the first position, used for insertion of pCk1 or pCsA. Ideally, nuclear transformation vectors are designed with genes for antibiotic resistance located adjacent to the LB (Figure 5A), as *Agrobacterium*-mediated T-DNA transfer is initiated from the RB,³³ and the presence of a selectable marker at the distal terminus can improve the rescue of full-length insertions after plant transformation.

Reporter genes are essential components for constructing genetic circuits in new systems. *Marchantia* is phylogenetically distant from other model plants, and thus well-tested components may need refactoring for use in this early divergent land plant. The OpenPlant toolkit contains L0 parts to assemble transcription units combining a wide range of multispectral fluorescent proteins (FPs) and targeting signals (Figures 5 and 8). We have systematically characterized and validated the performance of a series of modern fluorescent reporter genes. The kit contains genes for fluorescent proteins that are efficiently expressed and bright in *Marchantia*: eGFP,^{20,34} mVenus,^{16,35} mTurquoise2,^{16,36} and mScarlet-I.^{37,38} mScarlet is a monomeric bright RFP developed from a synthetic template.³⁷ Other monomeric RFPs, like tagRFP-T and mRuby3, have a residual tendency to dimerize, and carry additional problems due to incomplete or partial maturation.³⁷ The fluorescent protein coding sequences are supplied as standardized parts (CDS12 or CTAG) to allow construction of N- and C-terminal protein fusions. We supply versions with or without a linker (Ser-Gly-10xAla) for translational fusion with proteins of interest, such as proteins that confer specific subcellular localization, or functional peptides. The kit includes five parts for protein fusion and targeting to different cellular compartments (Figure 5B–D). LTI6b from *Arabidopsis*³⁹ can be used to localize a protein to the plasma membrane^{20,16} (Figure 5B). Peptide sequences N7 from *Arabidopsis*^{39,16} and the simian virus 40 large T-antigen nuclear localization signal (SV40-NLS)^{40,20} confer nuclear localization (Figure 5C). The N-terminal fusion of the MpSIG2 signal peptide allows targeting of a fusion partner into the chloroplast⁴¹ (Figure 5D).

The N-terminal signal peptide from *Arabidopsis* basic chitinase (AT3G12500), together with the C-terminal HDEL for protein retention in the ER, has been successfully used in different species of angiosperm plants for ER localization.⁴² Nevertheless, when tested, it did not produce the expected pattern in *Marchantia*. We found three predicted native chitinases in *Marchantia*, tested the leader peptide from one of them, Mapoly0069s0092, fused to eGFP-HDEL and could observe fluorescence in reticulate structures in the cytosol, around the nuclei and in the plasma membrane (Figure 5E). In the OpenPlant toolkit we provide three different color FPs fused to HDEL as well as eGFP with a GlyHis-tag before HDEL as described before.⁴³ The pumpkin 2S albumin leader peptide together with the ER retention signal HDEL has also

been successfully used in *Marchantia*²⁰ and is compared with chitinase signal peptide sequences (Figure S13).

Besides fluorescent proteins, other systems are widely used to report promoter activity like β -glucuronidase (GUS)⁴⁴ and firefly luciferase (LUC),⁴⁵ and the OpenPlant kit contains as L0 CDS parts both the GUS and LUC genes (Figure 8).

Synthetic Promoters for Nuclear Genome Transformation. Both for reference of constitutive expression and for extracting cell outlines, we need promoters expressed everywhere in the gemmae, at similar levels in all cell types, and with expression levels not too high, which could be toxic to the plant, or too low, which makes imaging more challenging.

Two constitutive promoters are commonly used in *Marchantia*, the cauliflower mosaic virus 35S (CaMV 35S) promoter and the endogenous *ELONGATION FACTOR 1 α* (MpEF1 α , Mapoly0024s0116) promoter.⁴⁶ There are two versions of the 35S promoter that are used in *Marchantia*: p5-35S^{16,47} and p5-35Sx2.²⁰ Both of them have a strong activity in all parts of the gemmae and thallus except the notch area, where its activity is weak, and the oil body cells, where it is not active⁴⁸ (Figures 5B and 6C). On the other hand, the MpEF1 α promoter is expressed in all parts of the thallus, but it is especially strong at the notch area⁴⁶ (Figures 5C and 6D). Also, as shown in published transcriptome studies (Table S3) MpEF1 α expression levels are very high throughout development (Figure 6A and Table S3). We observed some slower and deformed growth and higher frequency of silencing events in primary transformants when using p5-MpEF1 α to drive fluorescent protein expression. High levels of transcription driven by p5-MpEF1 α seems to be not well tolerated by plants. We exploited our pipeline for efficiently building constructs, transforming sporelings, and screening *Marchantia* gemmae, to test alternative promoters.

To identify new constitutive promoters without the limitations of the CaMV 35S and MpEF1 α promoters, we reanalyzed publicly available *Marchantia* transcriptome data and screened for genes with medium to high levels of expression and minimal variability (Figure 6A, Table S3). As reference genes we used the strongly expressed MpEF1 α gene and the lowly expressed MpNOpI gene.³² The average expression levels for MpEF1 α and MpNOpI are around 4500 and 25 TPM (transcripts per million) respectively (Figure 6A, Table S3). We selected candidate genes with an average expression between 250 and 1500 TPM and ranked them by their variability across all RNA seq experiments. We used the relative range as a measure of variability, calculated as (maximum TPM – minimum TPM)/average TPM. (Table S3). From the least variable genes, we selected four housekeeping genes with different average expression levels: the translation initiation factor 5a (MpELS α , Mapoly0045s0102, ~1300 TPM), the large subunit ribosomal protein (MpL40e, Mapoly0003s0257, ~560 TPM), the elongation factor 1 γ (MpEF1 γ , Mapoly0062s0115, ~460 TPM), and the ubiquitin-conjugating enzyme E2 (MpUBE2, Mapoly0106s0019, ~300 TPM) (Figure 6A and Table S2). MpELS α has been previously tested as a reference gene for qPCR.⁴⁹ We extracted putative promoter sequences for those candidate genes: MpELS α , MpL40e, and MpEF1 γ all as a PROM5 part of 1.8 kb upstream ATG, and MpUBE2 as a PROM 1.8 kb part plus a 5U 1.2 kb part upstream ATG (Figure 6B). We assembled Loop L2 constructs with those promoters driving mTurquoise2 localized to the nuclei and

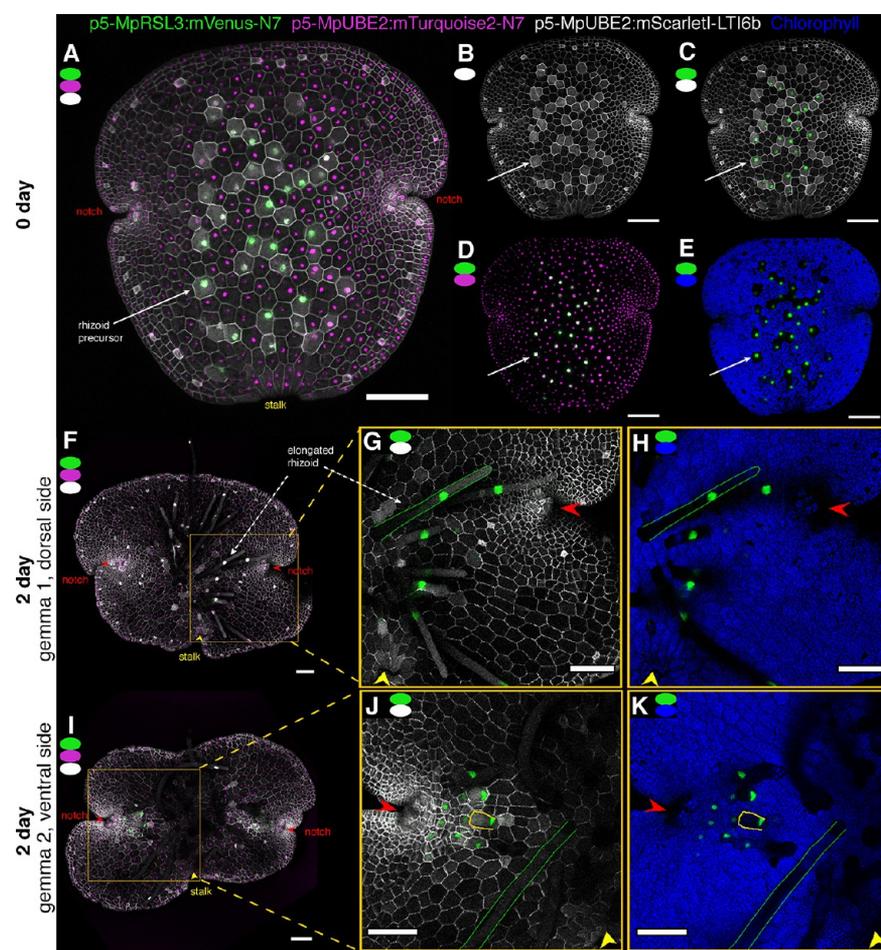


Figure 7. Construction of precise markers for visualizing cell differentiation using proximal promoter elements. The MpUBE2 and MprSL3 proximal promoters confer ubiquitous and rhizoid specific expression, respectively, and were used to build gene assemblies for multispectral imaging of live gemmae. Maximum intensity z-axis projection images of a gemma just removed from the gemmae cup (0 day). (A–E) Images from a line expressing a set of marker genes: p5-MprSL3:mVenus-N7 (nuclear-localized rhizoid cell marker, green), p5-MpUBE2:mTurquoise2-N7 (nuclear-localized ubiquitous expression, magenta) and p5-MpUBE2:mScarlet1-LTI6b (plasma membrane-localized ubiquitous expression, gray). (B–E) The relative levels and localization of gene expression can be dissected in separately combined image channels. (E) Chlorophyll autofluorescence, blue. (F–K) Images of whole 2 day old gemmae showing the dorsal (F) or the ventral side (I) (for details see Figure S14). Regions from the center of the gemmae to a meristem notch are shown enlarged (G, H, J, K). Rhizoid-specific expression (green) is shown with (gray) cell-outlines (G, J) and (blue) chlorophyll (H, K). MprSL3:mVenus-N7 marker is expressed in rhizoid precursor cells in the central part of the gemmae at day 0 (A–E), and in elongated rhizoids emerging from the central zone of the gemmae at day 2 (F–K). Hair-like rhizoids are outlined in green (G, H, J, K). Nuclei in elongated rhizoids move toward the tip of the cells (G, H). Expression of the MprSL3 marker is triggered in cells near the notch, only on the ventral side of gemmae (J, K). These subsequently differentiate into rhizoid hairs (Figure S15). An early, bulged rhizoid cell outlined in orange. Scale bar: 100 μm .

tested them *in vivo*. Localization of the fluorescent gene product to the nuclei allowed easier comparison of levels of gene expression in different size cell types. p5-MpEL5 α was weakly expressed in 0 day-old gemmae, only in the notch area and the rhizoid precursors, p5-MpL40e was expressed throughout the gemmae but with higher expression observed in nuclei in the region of the meristematic notch, and we could not detect expression driven by p5-MpEF1 γ . p5-MpUBE2 was expressed in all cell types and throughout gemmae development both when driving mTurquoise2 to the nuclei and eGFP to the plasma membrane (Figure 6E). Thus, we selected MpUBE2 promoter to drive constitutive expression of transgenes in early stages of gemmae development.

We then tested the new constitutive promoter p5-MpUBE2 in multispectral imaging, as a reference for expression and to ubiquitously label cell boundaries together with a marker for specific cell types. We made L2 devices containing tran-

scription units composed of the p5-MpUBE2 driving the expression of mTurquoise2 localized to the nuclei, the p5-MpUBE2 driving the expression of mScarlet1 localized to the plasma membrane, as well the MprSL3 proximal promoter (as a PROM5 1.8 kb part) driving the expression of mVenus targeted to the nuclei (Figure 7). MprSL3 (MpbHLH33, Mapoly0029s0135) is a transcription factor (TF) from the subfamily VIIIc of basic helix–loop–helix TFs, which in *Marchantia* has three members.⁵⁰ One of them, MprSL1 (MpbHLH14, Mapoly0039s0003) has been identified as a ROOTHAIR DEFECTIVE SIX-LIKE (RSL) class I gene, and it is required for the initiation and development of unicellular rhizoids, mucilage papillae, and gemmae.^{51–53} When imaging gemmae just displaced from the gemma cup (0 days-old gemma), MprSL3 promoter expression could be observed in the rhizoid precursor cells (Figure 7A–E). Both sides of a 0 days-old gemma have rhizoid precursor cells ready to elongate

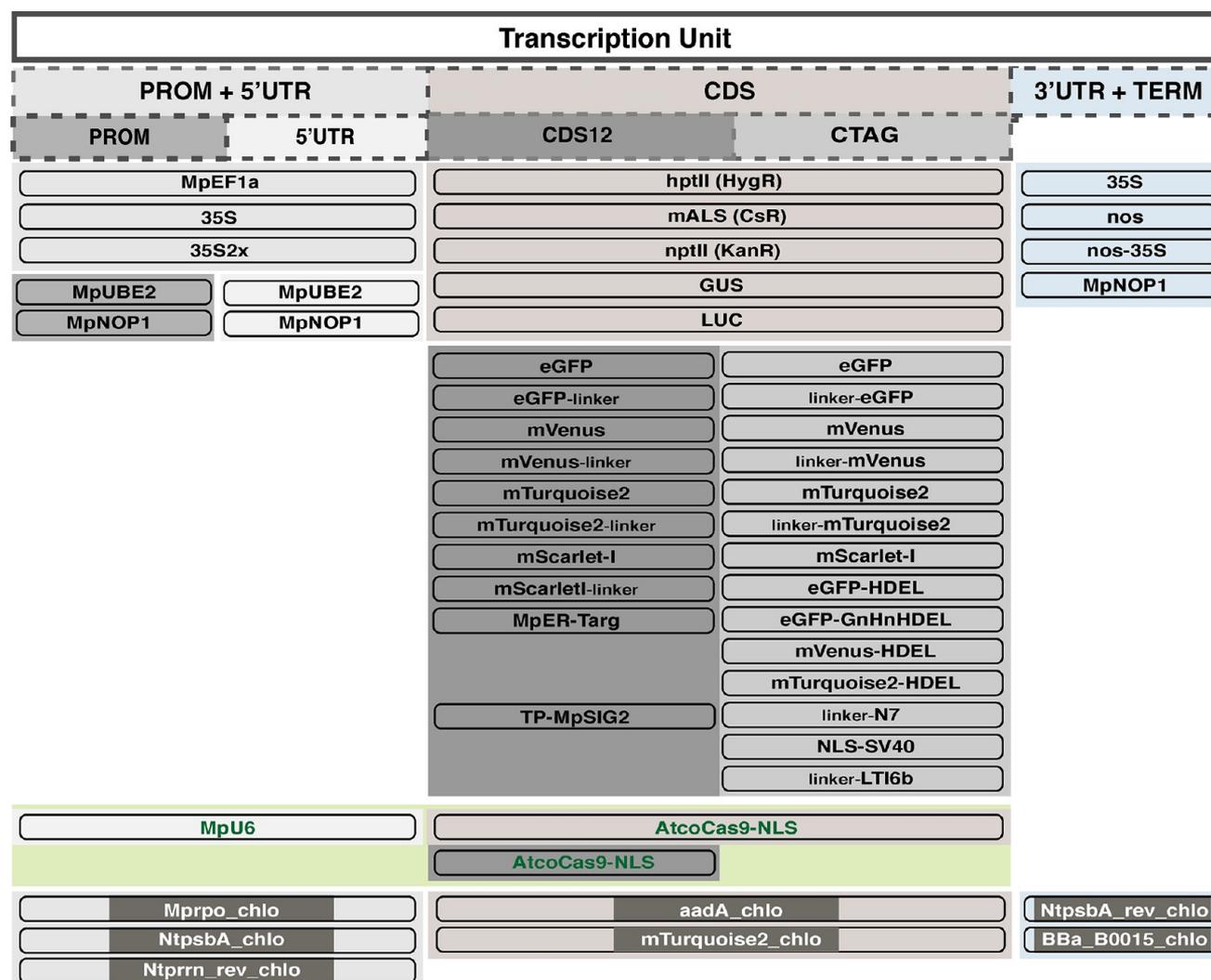


Figure 8. Level 0 DNA parts in the OpenPlant toolkit. The OpenPlant kit contains L0 parts to build transcription units, including: promoters with 5' UTR (PROM5), promoters (PROM) and 5' untranslated region (5UTR); coding sequences with start and stop codons (CDS), and for N-terminal (CTAG) and C-terminal (CDS12) protein fusions, and 3' UTR with terminator (3TERM). The OpenPlant toolkit contains 50 L0 parts, 43 for nuclear transformation, 3 for CRISPR-Cas9 genome editing (in green), and 7 for chloroplast transformation (name_chlo).

as rhizoid cells when landing. Once gemmae are displaced from the gemma cup,⁶ dormancy is released and dorsoventral polarity is specified with formation of air chambers on the dorsal side and formation of rhizoids on the ventral side (Figures 7 and S14). We tracked gemmae for 2 days, and the MprSL3 promoter kept expressing after the rhizoid precursors bulged and became elongated rhizoids with their nuclei moving toward the tip of the cells (Figures 7F–K and S15). On the dorsal side of a gemmae, the MprSL3 marker was not expressed in new cells and no new rhizoid precursor cells were specified (Figures 7F–H and S15). On the other hand, mTurquoise2 expressed under the control of the MpUBE2 promoter accumulated in all cells (Figures 7 and S15). To image the ventral side of gemmae, we grew them in inverted plates (Figure S14). On the ventral side, we observed that new cells near the notch started expressing the MprSL3 promoter and became rhizoid precursors which subsequently elongated (Figures 7I–K and S15). Thus, the MprSL3 proximal promoter can be used to study the dynamics of rhizoid precursor specification and elongation (position, number, spacing, timing), in combination with reference markers for

expression and ubiquitous labeling of cell boundaries under the control of the MpUBE2 promoter. Constructs like the one used here will advance the understanding of dynamic changes in expression patterns in *Marchantia*, as a first step to screen for regulatory sequences to engineer morphogenesis in this simple plant system.

CONCLUSIONS

The development of systematic approaches to the engineering of microbial genetic systems has facilitated substantial advances in science and biotechnology. There are many potential benefits in applying the same approach in plant systems. However, work with plants is beset with problems due to slow generation times, genetic redundancy, and difficulties in experimental manipulation and quantitative observation, compared to microbes. We have set out to establish a new technical platform for facile engineering of plants, to circumvent many of these difficulties.

We have adopted *Marchantia polymorpha* as a simple plant system that has an open form of development, which allows direct visualization of tagged gene expression and cellular

growth in living tissues. This liverwort plant system is fast and easy to culture and transform. Here, we describe methods to maintain the plant in axenic culture throughout its life cycle, using a combination of growth in agar-containing Petri dishes and on Jiffy-7 peat disks in Microboxes. We also describe methods for cryopreservation of *Marchantia gemmae*, to enable easier maintenance and storage of large numbers of plant lines. These streamlined methods allow routine propagation of *Marchantia* and provide a simple source of sterile propagules for manipulation or observation. These simple culture techniques allow laboratories to tackle advanced bioengineering projects in plants without the need for glasshouse facilities. We believe that the adoption of *Marchantia* as a prototype will widen access to synthetic biology experiments in plants.

Marchantia tissues show an extreme ability to regenerate without external cues. Regenerated plants spontaneously produce clonal propagules within a few weeks of an experiment. Patterns of gene regulation and cellular dynamics can be directly analyzed in these propagules. It has a simple cellular architecture and streamlined genome with highly reduced gene redundancy. The plant is haploid and has a genome of around 220 Mb and 19 000 encoded genes.⁷ Many families of regulatory genes are found with highly reduced numbers compared to higher plants, and whole tissue transcriptomic data is available for gene expression at different stages of growth. We have established a common genetic syntax for the composition (and exchange) of plant DNA parts (Phytobricks), recruiting wide support from the international plant science community.¹⁵ Further, MarpoDB was built as an open source database that handles the *Marchantia* genome as a collection of potential Phytobrick parts.⁵⁴ This is highly useful for automatic mining of new parts and managing part characterization.

We describe the modification of a published method for gene assembly.^{15,16} Loop assembly is based on recursive use of two Type IIS restriction endonucleases and allows rapid and efficient production of large DNA constructs, is compatible with widely used Level zero (L0) DNA parts such as Phytobricks, and can be easily automated. We have constructed an integrated set of plant transformation vectors for transformation and editing of the nuclear genome, and modification of the chloroplast genome. The new vectors are easier to handle in the laboratory and provide new features. The vectors maintain compatibility with modular DNA parts that follow the Golden Gate, MoClo, and Phytobrick syntax, and we have built and tested a wide range of new parts. These include a wide range of gene cassettes for multispectral fluorescent proteins, gene control, Cas9 and sgRNA expression, plastid gene regulation, and sequences for site-specific insertion into the *Marchantia* chloroplast genome by homologous recombination. The vectors and part collections are being made available through the Open Materials Transfer Agreement (OpenMTA¹⁷). Use of the OpenMTA places the materials into the public domain, allowing redistribution of materials and use for commercial projects.

Marchantia gemmae provide a new biological system with unique advantages for genetic reprogramming of plant growth. *Marchantia* plants produce gemmae spontaneously, as vegetative clonal propagules. New gemmae are formed around 2–3 weeks after germination of a vegetative parent gemma. New gemmae form inside conical splash cups and can be harvested prolifically. *Marchantia gemmae* have a stereotypical

and simple cellular architecture. They can be germinated and grown directly under microscopic observation. Gene expression can be visualized at the scale of individual cells. Dynamic cellular relationships within growing tissues can be directly visualized and quantified. Both short-range and organism-wide coordination of growth and differentiation can be perturbed and mapped in a way that is difficult or impossible in other plant systems due to tissue complexity, optical inaccessibility, genetic redundancy or all of the above. We have employed the new toolset described here to build and test new markers for visualizing cell dynamics.

An important aspect when engineering multicellular organisms, is the ability to target expression to specific cell types, at specific times in development, and at the desired expression levels. We have demonstrated the use of genome annotation and transcriptome data to design candidate promoter elements, and successfully characterized a new promoter MpUBE2, that drives ubiquitous gene expression in gemmae. Transcription factors initiate and maintain the transcriptional programs that specify the diversity of cell types in an organism and control growth. *Marchantia* has a low genetic redundancy, all main plant transcription factor families are comprised of only ~400 genes. We are conducting a systematic screen of TF promoter fusions in *Marchantia* using the tools for high-throughput transformation and fluorescence microscopy described here (Figures 1D and S13, and Materials and Methods). We believe that these new methods and the compound benefits of *Marchantia* as an experimental system will allow new approaches to important and fundamental questions about plant growth and engineering.

We have developed the OpenPlant toolkit as a set of interlinked resources and techniques, which have been designed to facilitate and standardize DNA assembly, and manipulation of both nuclear and chloroplast genomes, thus enabling the engineering of *Marchantia* development and metabolism. We hope that the tools developed in this work will encourage the adoption of *Marchantia* as a simple multicellular prototype for bioengineering, lower the logistical threshold for plant experiments and promote the recruitment of non-botanists to plant synthetic biology research, as well as encourage open science approaches to support standardization, reproducibility, and innovation.

■ MATERIALS AND METHODS

Plant Material and Growth Conditions. *Marchantia polymorpha* accessions Cam-1 (male) and Cam-2 (female) were used in this study.⁵⁴ Plants were grown and maintained on 0.5× Gamborg media (0.5× strength Gamborg B5 medium plus vitamins, Duchefa Biochemie G0210, pH 5.8) and 1.2% (w/v) agar (Melford capsules, A20021), under continuous light at 21 °C with light intensity of 150 μmol/m²/s.

Maintenance and Storage of Lines. To maintain a line, *Marchantia* plants can be grown under continuous light at 14 °C with light intensity below 100 μmol/m²/s. In this way plants can be kept for two-three months before transfer to new plates. Gemmae can be stored from 6 months to a year at 4 °C in the dark in 1.5 mL Eppendorf tubes half filled with 0.5× Gamborg media 1% (w/v) agar.

For longer term storage, lines can be cryopreserved at –80 °C for at least 1–2 years.¹⁸ We simplified the published cryopreservation protocol for *Marchantia polymorpha* gemmae.¹⁸ Gemmae were placed on preculture plates and incubated for 1–3 days under normal growth conditions. Preculture

plates consisted of 0.5× Gamborg media, 1.2% (w/v) agar, 0.3 M sucrose, 10 μM ABA. Media was autoclaved without ABA, and ABA added from a filter-sterilized stock solution just before pouring plates. For encapsulation of gemmae, small drops (approximately 40 μL) of alginate solution (0.5× Gamborg media, 3% sodium alginate) were placed on an empty 4.5 cm Petri dish to form beads. Gemmae were transferred from preculture plates to the Petri dish using forceps; up to five gemmae were placed inside each bead. A drop of CaCl₂ solution (0.5× Gamborg B5 media, 0.1 M CaCl₂) was added, and beads were allowed to solidify for 10 min. The plate was filled by serological pipet with dehydration buffer (0.5× Gamborg B5 media, 2 M glycerol, 1 M sucrose), submerging the beads. Beads were soaked for 30 min, and afterward the dehydration buffer was removed with a serological pipet without disrupting the beads. The lid was removed, and the beads were air-dried in the flow hood for at least 2 h. Using forceps or scalpels the beads were transferred into 1.5 mL Eppendorf tubes (max. 10 beads/tube). Tubes were flash frozen in liquid nitrogen for 2 min and stored at -80 °C. For thawing and recovery of lines, tubes were removed from the freezer and immediately placed in a 37 °C water bath or heat block for 2 min. 1.5 mL of thawing solution (0.5× Gamborg B5 media, 1 M sucrose) was added to each tube, and the tubes were incubated at room temperature for 10 min. Then the thawing solution was replaced with rinse solution (0.5× Gamborg B5 media, 0.03 M sucrose), and the tubes were incubated at room temperature for 10 min. Beads with gemmae were placed on 0.5× Gamborg B5 plates and incubated under normal growth conditions (Figure S4).

Marchantia Cultures for Spore Production. For spore production, 15 Jiffy-7 dehydrated peat disks were placed in a TP5000 + TPD5000 Microbox micropropagation containers with two #40 green filters (SacO₂, Belgium), 800 mL of water added, and the container autoclaved. After autoclaving, thallus fragments or gemmae were placed on each hydrated Jiffy-7 pellet, and another 200 mL of water added. Microboxes were placed at 21 °C, continuous light, 150 μE. After one month, Microboxes were transferred to long day conditions (16 h light/8 h dark), with light intensity of 150 μE supplemented with far-red light (peak emission around 730–740 nm, from Philips GreenPower LED HF far red, or Intelligent LED Solutions ILS-OW06-FRED-SD111 730 nm). After approximately 4 weeks mature male and female reproductive organs developed. For fertilization, one drop of water was transferred, with a disposable sterile plastic pipet, from the male reproductive organs on the top of the female reproductive organs. After one more month mature sporangia were visible and ready to be collected in large quantities and stored with desiccant at 4 °C for 1–2 months or at -80 °C for longer storage (Figure S1).

Sterilization of Marchantia Spores. Sterilization solution was prepared as previously described,⁷ by dissolving one Milton mini-sterilizing tablet (Milton Pharmaceutical UK, Cheltenham, UK) in 25 mL of sterile water, so the final concentration of the active ingredient, sodium dichloroisocyanurate (NaDCC) was 0.05% (w/v). Sporangia were transferred to a 1.5 mL Eppendorf tube. For Agrobacterium-mediated transformation with multiwell microplates, two sporangia were used per well-transformation (e.g., 12 sporangia for 6 transformations in a 6-well plate). Sterilization procedure was as follows. Sporangia were transferred to a 1.5 mL Eppendorf tube, and metal tweezers were used to crush them.

0.5 mL of sterilization solution was added to each tube, and the tubes were vortexed to get a finely pulped suspension. A 40 μM cell strainer was placed over a 50 mL falcon tube, and the 0.5 mL of crushed sporangia poured onto the filter. The Eppendorf tube was rinsed with 1 mL of fresh Milton solution and poured onto the filter. The filter was washed with 2.5 mL of fresh Milton solution. The filtrate containing the spores (total volume ~4 mL) was divided between 4 new 1.5 mL Eppendorf tubes. Spores were sterilized in the sterilization solution for 20–30 min. Tubes were then centrifuged at 16000g for 2 min, the supernatant discarded, and spores pooled into a single tube, resuspending the pellet in 25 μL for each planned transformation. Spores were plated on 0.5× Gamborg plates using 50 μL per plate (each plate provides enough material for 2 transformations). Plates were sealed with micro-pore tape and incubated (in an inverted orientation to prevent rhizoid growth into the agar media) for 5–7 days at 21 °C under continuous light.

Agrobacterium-Mediated Transformation in Multi-well Dishes. A modification of the published Agrobacterium-mediated sporeling transformation protocol¹⁹ was used. See Figure S2 and protocols.io/OpenPlant. For each construct of interest, a single colony of Agrobacterium, transformed with the plasmid of interest, was inoculated into 5 mL LB media plus the antibiotics required for selection and incubated at 28 °C with shaking at 150 rpm for 2 days. The Agrobacterium cultures were centrifuged for 10 min at 2000g, resuspended in 5 mL of 0.5× Gamborg B5, 1% (w/v) sucrose, and 100 μM acetosyringone (3',5'-dimethoxy-4'-hydroxyacetophenone, Sigma D134406), and incubated for 6 h at 28 °C at 150 rpm.

In each well of a 6-well plate, Marchantia 5–7 day old sporelings were co-cultivated with 100 μL of Agrobacterium culture in 4 mL of liquid 0.5× Gamborg B5 plus supplements (0.1% N-Z amino A (Sigma C7290), 0.03% L-glutamine, 2% (w/v) sucrose) 100 μM acetosyringone media. The 6-well plate was then placed on a shaker at 120 rpm for 2 days at 21 °C with continuous lighting (150 μmol/m²/s). For each well, the sporelings were washed with 50 mL of sterile water and plated on 0.5× Gamborg B5 media plus antibiotics plates. Plates contain 100 μg/mL cefotaxime and an antibiotic for selection of the transformant of interest; hygromycin 20 μg/mL (Invitrogen 10687010), chlorsulfuron (10 μM, Sigma 64902-72-3) or kanamycin (125 μg/mL, Sigma 25389-94-0).

Preparation of Seashell Nanoparticles for Chloroplast Transformation. Before bombardment, macro-carriers with the macro-carrier holders were placed in a glass Petri dish. Rupture disks and stopping screens were placed in another glass Petri dish. Both dishes were wrapped with aluminum foil and autoclaved. DNAdel nanoparticles (Seashell Technologies) were supplied as a 50 mg/mL suspension in binding buffer. To dissociate any aggregates prior to use, the suspension was agitated, sonicated briefly for 30 s using an ultrasonic water-bath sonicator, and vortexed for 5 s twice. 0.5 mg of nanoparticles per shot (transformation) were used. DNAdel nanoparticles were diluted into Binding Buffer (Seashell Technologies) to final concentration 30 mg/mL in a 1.5 mL Eppendorf tube. 1.5–2 μg of plasmid DNA per shot was added into the tube. Finally, an equal volume of Precipitation Buffer (Seashell Technologies) was added into the tube (volume = volume of DNAdel nanoparticles + volume of Binding Buffer + volume of plasmid DNA). The tube was vortexed and the mix was incubated at room temperature for 3 min. The tube was centrifuged at 8000g for 10 s, the

supernatant discarded, and the DNA-coated DNAdel nanoparticles were washed with 500 μL ice cold 100% EtOH. The tube was centrifuged at 8000g for 10 s again, the supernatant was discarded, and the nanoparticles were resuspended in 100% EtOH with a volume of n times 7 μL , n = number of shots planned. To resuspend, the nanoparticles were briefly sonicated using an ultrasonic water-bath sonicator, usually two rounds of 5 s sonication. 7 μL of DNA coated nanoparticles was pipetted into the center of each macro-carrier and left to dry. Biolistic bombardment was carried out with the Bio-Rad PDS-1000/He device according to the manufacturer's directions. The tissues used for bombardment were sporelings grown for 7 days on inverted 0.5 \times Gamborg media plates. After bombardment, the plates were cultured in a growth room for 2 days, and then transferred to 0.5 \times Gamborg B5 plates with spectinomycin (500 $\mu\text{g}/\text{mL}$). After 4–6 weeks successful transformants were visible (Figure S3).

Chloroplast Genome. Illumina short reads of Cam-1 nuclear, mitochondrial, and plastid DNA were sorted against a Tak-1 draft assembly generously provided by John Bowman, and reads that matched the chloroplast genome were assembled *de novo* using CLC Genomics workbench. *De novo* assembly generated 25 contigs spanning the entire plastome, which were assembled using the reference Tak draft assembly to generate the complete plastome. The assembly process was then repeated with Illumina reads from Cam-2, and the assembled sequences were identical. The Cam-1/2 sequence has been submitted to GenBank with accession number MH635409.

Genotyping. For genotyping,²² small pieces (3 \times 3 mm) of thalli from individual plants were placed in 1.5 mL Eppendorf tubes and crushed with an autoclaved micropestle in 100 μL genotyping buffer (100 mM Tris-HCl, 1 M KCl, 1 M KCl, and 10 mM EDTA, pH 9.5). The tubes were then placed at 80 $^\circ\text{C}$ for 5 min. 400 μL of sterile water was added to each tube. 5 μL aliquots of the extract were used as a template for PCR, and products were checked on a 1.5% (w/v) agarose gel.

L0 Reaction Protocol. To clone a L0 part into the pUAP4 vector, the following steps were followed. Primers were designed as described in Figure S5A. DNA parts were PCR amplified from the source DNA part (e.g., plasmid DNA or genomic DNA), using a high fidelity DNA polymerase. PCR products were run on a 1.5% (w/v) agarose gel, and the QIAquick Gel Extraction Kit was used to extract the band that corresponds to the size of the amplified DNA part. Aliquots of the DNA part were prepared at a concentration of 15 nM and of the pUAP4 acceptor vector at a concentration of 7.5 nM. A type IIS assembly reaction was set up into a 0.2 mL tube (5 μL nuclease-free H_2O , 1 μL 10 \times Tango Buffer (Thermo Fisher), 0.5 μL mg/mL bovine serum albumin (NEB), 0.25 μL T4 DNA Ligase at 5 U/ μL (Thermo Fisher), 1 μL 10 mM ATP (Sigma-Aldrich), 0.25 μL SapI (LguI) at 5 U/ μL (Thermo Fisher), 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 $^\circ\text{C}$ for 3 min and 16 $^\circ\text{C}$ for 4 min. Termination and enzyme denaturation: 50 $^\circ\text{C}$ for 5 min and 80 $^\circ\text{C}$ for 10 min. 20 μL of chemically competent *E. coli* cells (transformation efficiency of 1×10^7 transformants/ μg plasmid DNA) were transformed using 2 μL of the assembly reaction and then spread on LB agar plates containing 25 $\mu\text{g}/\text{mL}$ chloramphenicol and 40 $\mu\text{g}/\text{mL}$ X-gal. The presence of the correct insert was confirmed with Sanger sequencing using the

primers UAP_F and UAP_R and any additional DNA part specific primers (Figure S5).

L1 BsaI-Mediated Loop Assembly. L1 reactions were performed as previously described¹⁶ with minor modifications. Briefly, L0 plasmids with 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. Loop assembly Level 1 reaction master mix (MM) contained 3 μL nuclease-free H_2O , 1 μL 10 \times T4 DNA ligase buffer (New England Biolabs (NEB)), 0.5 μL 1 mg/mL bovine serum albumin (NEB), 0.25 μL T4 DNA ligase at 400 U/ μL (NEB), 0.25 μL BsaI at 10 U/ μL (NEB). Cycling conditions were 26 cycles of 37 $^\circ\text{C}$ for 3 min and 16 $^\circ\text{C}$ for 4 min. Termination and enzyme denaturation: 50 $^\circ\text{C}$ for 5 min, and 80 $^\circ\text{C}$ for 10 min. 20 μL of chemically competent *E. coli* cells were transformed using 2 μL of the Loop assembly reaction and plated on LB agar plates containing 50 $\mu\text{g}/\text{mL}$ kanamycin and 40 $\mu\text{g}/\text{mL}$ of X-Gal. The presence of the correct insert was confirmed with Sanger sequencing using primers pC_F and pC_R for pCk vectors and pC_F and pC_R2 for pCkchlo vectors (Figure S7).

L2 SapI-Mediated Loop Assembly. The protocol is the same as the L1 BsaI assembly protocol with the exception of MM composition: 2 μL nuclease-free H_2O , 1 μL 10 \times Tango Buffer (Thermo Fisher), 0.5 μL mg/mL bovine serum albumin (NEB), 0.25 μL T4 DNA ligase at 5 U/ μL (Thermo Fisher), 1 μL 10 mM ATP (Sigma-Aldrich), 0.25 μL SapI (LguI) at 5 U/ μL (Thermo Fisher) and plated on LB agar plates containing 100 $\mu\text{g}/\text{mL}$ spectinomycin and 40 $\mu\text{g}/\text{mL}$ of X-Gal. The donor plasmids were L1 constructs, and the acceptor plasmid was a pCs vector. The presence of the correct insert was confirmed with Sanger sequencing using the primers pC_F and pC_R for pCs vectors and pC_F and pC_R2 for pCschlo vectors.

Loop Automation Using Labcyte Echo 550. L1 and L2 assembly reactions were automated and miniaturized for assembly of 500 nL reactions using the Echo 550 acoustic-focusing non-contact liquid handling robot from Labcyte (San Jose, CA, USA). For L1 reactions an LDV source plate (Labcyte) with all necessary L0 plasmids was needed. For L2 plasmids another LDV plate with all L1 plasmids was provided. We used a PP source plate (Labcyte) for acceptor plasmids (AP) as well as master mix (MM) and water. Calibrations 384LDV_AQ_B2 and 384PP_AQ_BP2 were used, respectively. Both LDV and PP plates were sealed and stored at -20 $^\circ\text{C}$ for future use.

The assembly reaction conditions were identical to those described above for a final volume of 10 μL , but scaled down to a total volume of 500 nL. Plasmid DNAs in the source plates were 10 \times more concentrated than in the final reaction volume, and thus for each plasmid in a reaction/well the volume of plasmid to transfer to was 50 nL. For L1 reactions we used up to 6 L0 plasmids, the volume of plasmids was 300 nL, and the volume of L1-MM was 200 nL (with these components in each well: water: 100 nL, 10 \times T4 DNA ligase buffer: 50 nL, 1 mg/mL bovine serum albumin (NEB): 25 nL, T4 DNA ligase at 400 U/ μL (NEB): 12.5 nL, BsaI at 10 U/ μL (NEB): 12.5 nL). For L1 reactions assembling less than 6 plasmids, the volume was adjusted with water. For example, the volume of plasmid DNA in an L1 assembly with 4 plasmids (1 AP and 3 L0 plasmids) was 200 nL, and 200 nL L1-MM and 100 nL water were added to the final reaction. All L2 reactions involved assembly of 5 plasmids (1 AP and 4 L1 plasmids) with a total volume of 250 nL, added to 250 nL L2-MM (100 nL water, 10 \times Tango Buffer (Thermo Fisher): 50 nL, 1 mg/mL bovine

serum albumin (NEB): 25 nL, 10 mM ATP (Sigma-Aldrich): 50 nL, T4 DNA ligase at 5 U/ μ L (Thermo Fisher): 12.5 nL, *SapI* (*LguI*) at 5 U/ μ L (Thermo Fisher): 12.5 nL in each well. We first prepared the final volume of MM needed ((volume/well \times number of wells) plus dead volume of PP wells) in an Eppendorf tube, mixed by pipetting, and then transferred the MM to a PP-well. We created csv files in the Labcyte Cherry Pick software to add all the necessary information (source and destination wells for each sample and reaction, calibration, and volume).

Before running the Loop reaction, source plates were defrosted, spun down, and warmed at room temperature (\sim 22 $^{\circ}$ C). MM was prepared fresh, added to the PP plate, spun down, and kept at room temperature. Reactions were set up in 384 well PCR plates (4titude 4ti-0384), sealed with foil (4titude 4ti-0550), spun down, and incubated on a 384 well thermal cycling instrument using the conditions described above, with a lid temperature of 70 $^{\circ}$ C. Reactions were transformed into 2 μ L of Top10 chemically competent cells (transformation efficiency of 1×10^8 transformants/ μ g plasmid DNA). Heat shock was performed as follows: 2 μ L of cells were added to each well, the plate was incubated 15 min on ice, incubated at 42 $^{\circ}$ C for 40 s in a thermocycler, and cooled for 2 min on ice. 15 μ L of LB media was added to each well, the plate was sealed with a breath-easy sealing membrane (Z30059-1PAK) and incubated at 37 $^{\circ}$ C for 1 h. Then the whole volume in each well was plated in selective LB agar plates as described above.

Guide RNA Cloning for Gene-Editing Vectors. Two oligos for cloning of the gRNA target sequence were designed as described in Figure S11. 1 μ L of each oligo (100 μ M) were mixed with 8 μ L water into final volume of 10 μ L and then annealed in a thermocycler with the following parameters: 37 $^{\circ}$ C for 30 min, 95 $^{\circ}$ C for 5 min and then ramped down to 25 $^{\circ}$ C at 5 $^{\circ}$ C per min. After annealing, the gRNA target sequence was cloned into the L2_lacZgRNA-Cas9-CsA vector using L2 *SapI* mediated Loop assembly (Figure S11A). For cloning into the L1_lacZgRNA-Ck2 or L1_lacZgRNA-Ck3 vectors (Figure S11B), a *BbsI*-mediated type IIS assembly was used: 6 μ L nuclease-free H₂O, 1 μ L 10 \times T4 ligase buffer (NEB), 0.5 μ L mg/mL bovine serum albumin (NEB), 0.25 μ L T4 DNA ligase at 400 U/ μ L (NEB), 0.25 μ L *BbsI* (*BpiI*) at 5 U/ μ L (Thermo Fisher), 1 μ L vector, and 1 μ L annealed oligos.

Analysis of Public RNA-Sequencing Data. Fastq files for the following BioProject study accessions: PRJDB4420,⁵⁵ PRJDB5890,⁵⁶ PRJDB6579,⁵⁷ PRJDB7023,⁷ PRJNA218052,⁵⁸ PRJNA265205,⁵⁹ PRJNA350270,⁷ PRJNA397394,⁶⁰ PRJNA433456,⁶¹ and PRJNA251267⁷ were downloaded from the European Nucleotide Archive. FastQC was used to inspect read quality and TrimGalore to trim low-quality reads and remaining sequencing adapters. Reads were pseudo-aligned using kallisto⁶² to the Marchantia Genome⁷ version JGI3.1 (primary transcripts only, obtained from MarpolBase). Kallisto estimates abundance of each transcript in units of transcripts per million (TPM). We merged replicates by averaging and generated a gene-sample matrix of TPM values (Table S2). This matrix was used to calculate the average expression and relative range for each gene.

Screening of Transformants with Fluorescent Microscopy. Primary transformants growing on 0.5 \times Gamborg plates with antibiotic were screened using a stereo microscope with fluorescence (Leica M205 FA). Ten days after transformation, 20 plants were transferred onto new selection plates. We

performed a second screen for fluorescence at the stereo microscope, selecting around 10 plants with expected signals. In about 2 more weeks, the plants produced gemmae cups containing G1 (generation 1) gemma. For each construct, we selected at least three independent lines to image gemmae with laser scanning microscopy using either slides or 384 well plates.

The Leica M205FA was equipped with the following filters for observation of fluorescence: ET CFP (excitation filter ET436/20 nm, emission filter ET480/40 nm), ET GFP (ET470/40 nm, ET525/50 nm), ET YFP (ET500/20 nm, ET535/30 nm), ET Chlorophyll LP (ET480/40 nm, ET610 nm LP), and ET GFP LP500 (ET470/40 nm, ET500 nm LP), used respectively for observing mTurquoise, eGFP, mVenus, chlorophyll, and eGFP together with chlorophyll autofluorescence. We designed with Leica a new filter to allow for observation of RFPs like mScarlet-I without chlorophyll autofluorescence interference, the Plant RFP filter (excitation filter ET560/40 nm, emission filter ET594/10 nm).

Laser Scanning Confocal Microscopy. Preparation of Slides. A microscope slide was fitted with a Gene Frame (ThermoFisher AB0577, 65 μ L, 1.5 \times 1.6 cm) and a thin media pad made with 0.5 \times Gamborg media 1% agar. One to four Marchantia gemmae were placed on top of the pad using a small inoculation loop (ThermoFisher SL1S), a drop of sterile water placed on top of the gemmae, and a #0 coverslip placed on top of the Gene Frame.

Preparation of Multiwell Plates. We used a 384-well plate (Greiner bio-one 781186). 0.5 \times Gamborg media with 1.2% (w/v) agar was poured into the selected wells for imaging, making sure they were filled up to the top, and they did not contain any air bubbles inside. We filled extra sets of rows and columns of wells around our samples to reduce media evaporation. The media was allowed to solidify and leveled with a sterile scalpel if it bulged from the wells. Gene Frames (ThermoFisher AB0578, 125 μ L, 1.7 \times 2.8 cm) were placed to delimit the wells we were planning to image (Figure S16). One gemma was placed at the center of each filled well. The cover glass was cleaned with 70% EtOH and sprayed on one side with B-clean anti-fog spray. The anti-fog solution was let to evaporate and any residue removed with a lens tissue moving in one direction to avoid any visual artifacts. A coverslip was placed on top of the gene frame, making sure the anti-fog treated side was facing down, and the plate was ready for imaging. Gemmae could be imaged at several time points while growing inside the plate for up to 7 days.

Imaging of Plants in Petri Dishes. When imaging plants grown on media plates in order to follow growth processes (Figure 7, S14, S15), we let the plants grow under normal culture conditions, removed the lid for imaging, and when finished, replaced the lid and returned the plants to the growth chamber.

Imaging with Confocal Microscopy. Images were acquired on a Leica SP8X spectral confocal microscope upright system equipped with a 460–670 nm super continuum white light laser, 2 CW laser lines 405 nm, and 442 nm, and 5 Channel Spectral Scanhead (4 hybrid detectors and 1 PMT). For slides, imaging was conducted using either a 10 \times air objective (HC PL APO 10 \times /0.40 CS2), a 20 \times air objective (HC PL APO 20 \times /0.75 CS2) or a 40 \times water immersion objective (HC PL APO 40 \times /1.10 W CORR CS2). For plants on multiwell plates or on plates without lid, 10 \times or 20 \times objectives could be used. Imaging settings varied depending on number of channels. When observing more than one fluorescent protein, sequential

scanning mode was selected. 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 (515 nm, 522–540 nm), for mScarletI (569 nm, 585–610 nm), and for chlorophyll autofluorescence (488 or 515, 670–700 nm). Chlorophyll autofluorescence was imaged simultaneously with eGFP or mVenus.

Registries and Metadata. Benchling was used as DNA editor and registry of constructs. The Benchling Golden Gate assembly wizard allows for building of type IIS DNA constructs keeping the lineage information, list, and order of plasmids used to assemble the final one, like all the L0 plasmids used to make a L1 plasmid (<https://benchling.com/tutorials/11/golden-gate-assembly>). Curation of parts is facilitated by the auto-annotate sequence tool (<https://benchling.com/tutorials/34/auto-annotate>). The bioregistry (<https://benchling.com/tutorials/47/setting-up-your-bioregistry>) option allows establishing of registries for different type of parts with specific associated metadata (schema), which facilitates standardization (<https://help.benchling.com/articles/2725066-configure-your-registry>). All parts of a registry, with its unique identifier for each part, and all associated metadata can be exported in bulk as a csv file (<https://help.benchling.com/articles/2725575-export-registry-data-in-bulk>), and we use R studio and R packages tidy and ggplot2 to edit and further process them (<https://www.rstudio.com/>). The plasmids in the OpenPlant toolkit are shared through a public benchling folder (https://benchling.com/susana/f/_1ptxoMWP-openplant-kit/).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acssynbio.9b00511>. Detailed experimental protocols for DNA design, assembly, plant transformation, and microscopy can be found at <https://www.protocols.io/groups/openplant-project>.

Spreadsheet file with data used for reanalysis of all publicly available RNaseq experiments, including expression levels for all *Marchantia* genes and TPM stats (XLSX)

File with registry of OpenPlant kit plasmids, containing sequences and metadata (XLSX)

Figure S1, techniques for axenic propagation of *Marchantia*; Figure S2, agrobacterium-mediated transformation of *Marchantia* spores in multiwell plates; Figure S3, chloroplast transformation using *Marchantia* spores; Figure S4, cryopreservation of *Marchantia* gemmae; Figure S5, design of primers for cloning of Level 0 parts into pUAP4 vectors; Figure S6, domestication of new DNA parts; Figure S7, assembly of multiple Level 0 parts into a transcription unit; Figure S8, assembly of four transcription units into a Level 2 plasmid; Figure S9, schematic diagram of improved Loop vectors; Figure S10, testing of vectors and part DNAs for chloroplast transformation; Figure S11, design of primers for rapid cloning of gRNA sequences; Figure S12, Cas9-CRISPR mediated gene editing and generation of deletions; Figure S13, Alignment of ER signal peptide sequences; Figure S14, growth of gemmae on normal vs inverted plates; Figure S15, time course of MpRSL3 promoter

directed expression during early gemma development; Figure S16, gemmae culture for high magnification live microscopy; Table S1, analysis of the efficiency of CRISPR/Cas9 mediated editing of *MpNop1*; and Table S2, list of primers used in this study (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jim Haseloff – Department of Plant Sciences, University of Cambridge, CB2 3EA, U.K.; orcid.org/0000-0003-4793-8058; Email: jh295@cam.ac.uk

Authors

Susanna Sauret-Güeto – Department of Plant Sciences, University of Cambridge, CB2 3EA, U.K.; orcid.org/0000-0003-3792-8278

Eftychios Frangedakis – Department of Plant Sciences, University of Cambridge, CB2 3EA, U.K.; orcid.org/0000-0002-3483-8464

Linda Silvestri – Department of Plant Sciences and Department of Physiology, Development and Neuroscience, University of Cambridge, CB2 3EA, U.K.

Marius Rebmann – Department of Plant Sciences, University of Cambridge, CB2 3EA, U.K.; orcid.org/0000-0003-1462-0975

Marta Tomaselli – Department of Plant Sciences, University of Cambridge, CB2 3EA, U.K.; orcid.org/0000-0003-3470-2820

Kasey Markel – Department of Plant Sciences, University of Cambridge, CB2 3EA, U.K.; orcid.org/0000-0002-8285-3888

Mihails Delmans – Department of Plant Sciences, University of Cambridge, CB2 3EA, U.K.

Anthony West – Earlham Institute, Norwich NR4 7UZ, U.K.

Nicola J. Patron – Earlham Institute, Norwich NR4 7UZ, U.K.; orcid.org/0000-0002-8389-1851

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acssynbio.9b00511>

Author Contributions

⊗ S.S.-G. and E.F. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Suvi Honkanen for sharing constructs to test as constitutive promoters. We thank Connor Tansley for sharing his experience on cryopreservation protocol. We thank Emanuele Orsini and Michelle Lim for their contribution as summer students. We thank previous and present members of the lab working with *Marchantia*, Nuri Purswani, Christian Boehm, Owen Male, Liat Adler, Alicja Szałapak, Lukas Mueller, Alan Marron, Jenna Rever, and Bernardo Pollak; the latter especially for his doctoral thesis⁶³ that described a plate-based transformation method and a Nop1-targeted gRNA that provided a working design for this work. We thank the Cambridge *Marchantia* community, especially Chiara Airoidi, Giulia Arsuffi, and Philip Carella, for useful discussions and sharing of materials, and the ROC (Researchers with OpenPlant in Cambridge) for useful discussions on standardization and sharing tools. We thank Benchling for diligent support in implementing Registries. We

thank Plant Sciences Department facilities and administrative staff for their support. This work was funded as part of the BBSRC/EPSRC OpenPlant Synthetic Biology Research Centre Grant BB/L014130/1 to N.J.P. and J.H., BBSRC BB/F011458/1 for confocal microscopy to J.H., BBSRC Research Studentships for M.R. (1943399) and M.D. (1497818), and BBS/OS/GC/000013C Synthetic Biology Tools and Resources for genetic Improvement in the Developing World and BBS/E/T/000PR9815A Earlham BioFoundry, a BBSRC-funded National Capability, for N.J.P.

REFERENCES

- (1) Nemhauser, J. L., and Torii, K. U. (2016) Plant synthetic biology for molecular engineering of signalling and development. *Nat. Plants* 2, 16010.
- (2) Küken, A., and Nikoloski, Z. (2019) Computational approaches to design and test plant synthetic metabolic pathways. *Plant Physiol.* 179, 894–906.
- (3) Liu, W., and Stewart, C. N. (2015) Plant synthetic biology. *Trends Plant Sci.* 20, 309–317.
- (4) Puttick, M. N., Morris, J. L., Williams, T. A., Cox, C. J., Edwards, D., Kenrick, P., Pressel, S., Wellman, C. H., Schneider, H., Pisani, D., and Donoghue, P. C. J. (2018) The Interrelationships of land plants and the nature of the ancestral embryophyte. *Curr. Biol.* 28, 733–745.
- (5) Ishizaki, K., Nishihama, R., Yamato, K. T., and Kohchi, T. (2016) Molecular genetic tools and techniques for *Marchantia polymorpha* research. *Plant Cell Physiol.* 57, 262–270.
- (6) Shimamura, M. (2016) *Marchantia polymorpha*: Taxonomy, phylogeny and morphology of a model system. *Plant Cell Physiol.* 57, 230–256.
- (7) Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S. S., Althoff, F., Araki, T., Arteaga-Vazquez, M. A., Balasubramanian, S., Barry, K., Bauer, D., Boehm, C. R., Brigginshaw, L., Caballero-Perez, J., Catarino, B., Chen, F., Chiyoda, S., Chovatia, M., Davies, K. M., Delmans, M., Demura, T., Dierschke, T., Dolan, L., Dorantes-Acosta, A. E., Eklund, D. M., Florent, S. N., Flores-Sandoval, E., Fujiyama, A., Fukuzawa, H., Galik, B., Grimanelli, D., Grimwood, J., Grossniklaus, U., Hamada, T., Haseloff, J., Hetherington, A. J., Higo, A., Hirakawa, Y., Hundley, H. N., Ikeda, Y., Inoue, K., Inoue, S.-I., Ishida, S., Jia, Q., Kakita, M., Kanazawa, T., Kawai, Y., Kawashima, T., Kennedy, M., Kinose, K., Kinoshita, T., Kohara, Y., Koide, E., Komatsu, K., Kopsischke, S., Kubo, M., Kyojuka, J., Lagercrantz, U., Lin, S.-S., Lindquist, E., Lipzen, A. M., Lu, C.-W., De Luna, E., Martienssen, R. A., Minamino, N., Mizutani, M., Mizutani, M., Mochizuki, N., Monte, I., Mosher, R., Nagasaki, H., Nakagami, H., Naramoto, S., Nishitani, K., Ohtani, M., Okamoto, T., Okumura, M., Phillips, J., Pollak, B., Reinders, A., Rövekamp, M., Sano, R., Sawa, S., Schmid, M. W., Shirakawa, M., Solano, R., Spunde, A., Suetsugu, N., Sugano, S., Sugiyama, A., Sun, R., Suzuki, Y., Takenaka, M., Takezawa, D., Tomogane, H., Tsuzuki, M., Ueda, T., Umeda, M., Ward, J. M., Watanabe, Y., Yazaki, K., Yokoyama, R., Yoshitake, Y., Yotsui, I., Zachgo, S., and Schmutz, J. (2017) Insights into Land Plant Evolution Garnered from the *Marchantia polymorpha* Genome. *Cell* 171, 287–304.
- (8) Boehm, C. R., Pollak, B., Purswani, N., Patron, N., and Haseloff, J. (2017) Synthetic Botany. *Cold Spring Harbor Perspect. Biol.* 9, a023887.
- (9) Solly, J. E., Cunniffe, N. J., and Harrison, C. J. (2017) Regional growth rate differences specified by apical notch activities regulate liverwort thallus shape. *Curr. Biol.* 27, 16–26.
- (10) Sarrion-Perdigones, A., Falconi, E. E., Zandalinas, S. I., Juárez, P., Fernández-del-Carmen, A., Granell, A., and Orzaez, D. (2011) Golden Braid: an iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One* 6, No. e21622.
- (11) Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs. *PLoS One* 6, No. e16765.
- (12) Engler, C., Kandzia, R., and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One* 3, No. e3647.
- (13) Engler, C., Youles, M., Gruetzner, R., Ehner, T.-M., Werner, S., Jones, J. D. G., Patron, N. J., and Marillonnet, S. (2014) A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* 3, 839–843.
- (14) Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J. U., and Forner, J. (2013) Green Gate—a novel, versatile, and efficient cloning system for plant transgenesis. *PLoS One* 8, No. e83043.
- (15) Patron, N. J., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., Raitskin, O., Leveau, A., Farré, G., Rogers, C., Smith, A., Hibberd, J., Webb, A. A. R., Locke, J., Schornack, S., Ajioka, J., Baulcombe, D. C., Zipfel, C., Kamoun, S., Jones, J. D. G., Kuhn, H., Robatzek, S., Van Esse, H. P., Sanders, D., Oldroyd, G., Martin, C., Field, R., O'Connor, S., Fox, S., Wulff, B., Miller, B., Breakspear, A., Radhakrishnan, G., Delaux, P.-M., Loqué, D., Granell, A., Tissier, A., Shih, P., Brutnell, T. P., Quick, W. P., Rischer, H., Fraser, P. D., Aharoni, A., Raines, C., South, P. F., Ané, J.-M., Hamberger, B. R., Langdale, J., Stougaard, J., Bouwmeester, H., Udvardi, M., Murray, J. A. H., Ntougakis, V., Schäfer, P., Denby, K., Edwards, K. J., Osbourn, A., and Haseloff, J. (2015) Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytol.* 208, 13–19.
- (16) Pollak, B., Cerda, A., Delmans, M., Álamos, S., Moyano, T., West, A., Gutiérrez, R. A., Patron, N. J., Federici, F., and Haseloff, J. (2019) Loop assembly: a simple and open system for recursive fabrication of DNA circuits. *New Phytol.* 222, 628–640.
- (17) Kahl, L., Molloy, J., Patron, N., Matthewman, C., Haseloff, J., Grewal, D., Johnson, R., and Endy, D. (2018) Opening options for material transfer. *Nat. Biotechnol.* 36, 923–927.
- (18) Tanaka, D., Ishizaki, K., Kohchi, T., and Yamato, K. T. (2016) Cryopreservation of gemmae from the liverwort *Marchantia polymorpha* L. *Plant Cell Physiol.* 57, 300–306.
- (19) Ishizaki, K., Chiyoda, S., Yamato, K. T., and Kohchi, T. (2008) Agrobacterium-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. *Plant Cell Physiol.* 49, 1084–1091.
- (20) Ishizaki, K., Nishihama, R., Ueda, M., Inoue, K., Ishida, S., Nishimura, Y., Shikanai, T., and Kohchi, T. (2015) Development of gateway binary vector series with four different selection markers for the liverwort *Marchantia polymorpha*. *PLoS One* 10, No. e0138876.
- (21) Chiyoda, S., Yamato, K. T., and Kohchi, T. (2014) Plastid transformation of sporelings and suspension-cultured cells from the liverwort *Marchantia polymorpha* L. *Methods Mol. Biol.* 1132, 439–447.
- (22) Boehm, C. R., Ueda, M., Nishimura, Y., Shikanai, T., and Haseloff, J. (2016) A cyan fluorescent reporter expressed from the chloroplast genome of *Marchantia polymorpha*. *Plant Cell Physiol.* 57, 291–299.
- (23) Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S., and Mullineaux, P. M. (2000) pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant Mol. Biol.* 42, 819–832.
- (24) Watson, M. R., Lin, Y.-f., Hollwey, E., Dodds, R. E., Meyer, P., and McDowall, K. J. (2016) An improved binary vector and *Escherichia coli* strain for Agrobacterium tumefaciens-mediated plant transformation. *G3: Genes, Genomes, Genet.* 6, 2195–2201.
- (25) Hajdukiewicz, P., Svab, Z., and Maliga, P. (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. *Plant Mol. Biol.* 25, 989–994.
- (26) Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2579–86.
- (27) Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821.

- (28) Sugano, S. S., Shirakawa, M., Takagi, J., Matsuda, Y., Shimada, T., Hara-Nishimura, I., and Kohchi, T. (2014) CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L. *Plant Cell Physiol.* 55, 475–481.
- (29) Li, J.-F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G. M., and Sheen, J. (2013) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* 31, 688–691.
- (30) Schiml, S., Fauser, F., and Puchta, H. (2016) CRISPR/Cas-mediated site-specific mutagenesis in *Arabidopsis thaliana* using cas9 nucleases and paired nickases. *Methods Mol. Biol.* 1469, 111–122.
- (31) Sugano, S. S., Nishihama, R., Shirakawa, M., Takagi, J., Matsuda, Y., Ishida, S., Shimada, T., Hara-Nishimura, I., Osakabe, K., and Kohchi, T. (2018) Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in *Marchantia polymorpha*. *PLoS One* 13, No. e0205117.
- (32) Ishizaki, K., Mizutani, M., Shimamura, M., Masuda, A., Nishihama, R., and Kohchi, T. (2013) Essential role of the E3 ubiquitin ligase nopperab01 in schizogonous intercellular space formation in the liverwort *Marchantia polymorpha*. *Plant Cell* 25, 4075–4084.
- (33) Zupan, J. R., and Zambryski, P. (1995) Transfer of T-DNA from *Agrobacterium* to the Plant Cell. *Plant Physiol.* 107, 1041–1047.
- (34) Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173, 33–38.
- (35) Kremers, G.-J., Goedhart, J., van Munster, E. B., and Gadella, T. W. J., Jr. (2006) Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Förster radius. *Biochemistry* 45, 6570–6580.
- (36) Goedhart, J., von Stetten, D., Noirclerc-Savoye, M., Lelimosin, M., Joosen, L., Hink, M. A., van Weeren, L., Gadella, T. W. J., Jr, and Royant, A. (2012) Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat. Commun.* 3, 751.
- (37) Bindels, D. S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K. E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M. A., and Gadella, T. W. J., Jr. (2017) mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat. Methods* 14, 53–56.
- (38) Kato, H., Mutte, S. K., Suzuki, H., Crespo, I., Das, S., Radoeva, T., Fontana, M., Yoshitake, Y., Hainiwa, E., van den Berg, W., Lindhoud, S., Hohlbain, J., Borst, J. W., Roeland Boer, D., Nishihama, R., Kohchi, T., and Weijers, D. (2019) Design principles of a minimal auxin response system. *bioRxiv*. DOI: 10.1101/760876
- (39) Cutler, S. R., Ehrhardt, D. W., Griffiths, J. S., and Somerville, C. R. (2000) Random GFP::cDNA fusions enable visualization of subcellular structures in cells of *Arabidopsis* at a high frequency. *Proc. Natl. Acad. Sci. U. S. A.* 97, 3718–3723.
- (40) Kalderon, D., Roberts, B. L., Richardson, W. D., and Smith, A. E. (1984) A short amino acid sequence able to specify nuclear location. *Cell* 39, 499–509.
- (41) Kanazawa, T., Ishizaki, K., Kohchi, T., Hanaoka, M., and Tanaka, K. (2013) Characterization of four nuclear-encoded plastid RNA polymerase sigma factor genes in the liverwort *Marchantia polymorpha*: blue-light- and multiple stress-responsive SIG5 was acquired early in the emergence of terrestrial plants. *Plant Cell Physiol.* 54, 1736–1748.
- (42) Gnanasambandam, A., Polkinghorne, I. G., and Birch, R. G. (2007) Heterologous signals allow efficient targeting of a nuclear-encoded fusion protein to plastids and endoplasmic reticulum in diverse plant species. *Plant Biotechnol. J.* 5, 290–296.
- (43) Mitsuhashi, N., Shimada, T., Mano, S., Nishimura, M., and Hara-Nishimura, I. (2000) Characterization of organelles in the vacuolar-sorting pathway by visualization with GFP in tobacco BY-2 cells. *Plant Cell Physiol.* 41, 993–1001.
- (44) Jefferson, R. A., Kavanagh, T. A., and Bevan, M. W. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6, 3901–3907.
- (45) Howell, S. H., Ow, D. W., and Schneider, M. (1989) Use of the firefly luciferase gene as a reporter of gene expression in plants. *Plant Molecular Biol. Manual.*, 359–369.
- (46) Althoff, F., Kopischke, S., Zobell, O., Ide, K., Ishizaki, K., Kohchi, T., and Zachgo, S. (2014) Comparison of the MpEF1 α and CaMV35 promoters for application in *Marchantia polymorpha* overexpression studies. *Transgenic Res.* 23, 235–244.
- (47) Boissnard-Lorig, C., Colon-Carmona, A., Bauch, M., Hodge, S., Doerner, P., Bancharel, E., Dumas, C., Haseloff, J., and Berger, F. (2001) Dynamic Analyses of the Expression of the HISTONE::YFP Fusion Protein in *Arabidopsis* Show That Syncytial Endosperm Is Divided in Mitotic Domains. *Plant Cell* 13, 495–509.
- (48) Kanazawa, T., Era, A., Minamino, N., Shikano, Y., Fujimoto, M., Uemura, T., Nishihama, R., Yamato, K. T., Ishizaki, K., Nishiyama, T., Kohchi, T., Nakano, A., and Ueda, T. (2016) SNARE Molecules in *Marchantia polymorpha*: Unique and conserved features of the membrane fusion machinery. *Plant Cell Physiol.* 57, 307–324.
- (49) Saint-Marcoux, D., Proust, H., Dolan, L., and Langdale, J. A. (2015) Identification of reference genes for real-time quantitative PCR experiments in the liverwort *Marchantia polymorpha*. *PLoS One* 10, No. e0118678.
- (50) Catarino, B., Hetherington, A. J., Emms, D. M., Kelly, S., and Dolan, L. (2016) The Stepwise increase in the number of transcription factor families in the precambrian predated the diversification of plants on land. *Mol. Biol. Evol.* 33, 2815–2819.
- (51) Proust, H., Honkanen, S., Jones, V. A. S., Morieri, G., Prescott, H., Kelly, S., Ishizaki, K., Kohchi, T., and Dolan, L. (2016) RSL Class I genes controlled the development of epidermal structures in the common ancestor of land plants. *Curr. Biol.* 26, 93–99.
- (52) Bonnot, C., Hetherington, A. J., Champion, C., Breuninger, H., Kelly, S., and Dolan, L. (2019) Neofunctionalisation of basic helix-loop-helix proteins occurred when. *New Phytol.* 223, 993–1008.
- (53) Honkanen, S., Thamm, A., Arteaga-Vazquez, M. A., and Dolan, L. (2018) Negative regulation of conserved RSL class I bHLH transcription factors evolved independently among land plants. *eLife* 7, No. e38529.
- (54) Delmans, M., Pollak, B., and Haseloff, J. (2017) MarpoDB: An open registry for *Marchantia Polymorpha* genetic parts. *Plant Cell Physiol.* 58 (1), No. e5.
- (55) Higo, A., Niwa, M., Yamato, K. T., Yamada, L., Sawada, H., Sakamoto, T., Kurata, T., Shirakawa, M., Endo, M., Shigenobu, S., Yamaguchi, K., Ishizaki, K., Nishihama, R., Kohchi, T., and Araki, T. (2016) Transcriptional framework of male gametogenesis in the liverwort *Marchantia polymorpha* L. *Plant Cell Physiol.* 57, 325–338.
- (56) Kubo, H., Nozawa, S., Hiwatashi, T., Kondou, Y., Nakabayashi, R., Mori, T., Saito, K., Takanashi, K., Kohchi, T., and Ishizaki, K. (2018) Biosynthesis of riccionidins and marchantins is regulated by R2R3-MYB transcription factors in *Marchantia polymorpha*. *J. Plant Res.* 131, 849–864.
- (57) Yamaoka, S., Nishihama, R., Yoshitake, Y., Ishida, S., Inoue, K., Saito, M., Okahashi, K., Bao, H., Nishida, H., Yamaguchi, K., Shigenobu, S., Ishizaki, K., Yamato, K. T., and Kohchi, T. (2018) Generative cell specification requires transcription factors evolutionarily conserved in land plants. *Curr. Biol.* 28, 479–486.
- (58) Sharma, N., Bhalla, P. L., and Singh, M. B. (2013) Transcriptome-wide profiling and expression analysis of transcription factor families in a liverwort, *Marchantia polymorpha*. *BMC Genomics* 14, 915.
- (59) Frank, M. H., and Scanlon, M. J. (2015) Transcriptomic evidence for the evolution of shoot meristem function in sporophyte-dominant land plants through concerted selection of ancestral gametophytic and sporophytic genetic programs. *Mol. Biol. Evol.* 32, 355–367.
- (60) Mutte, S. K., Kato, H., Rothfels, C., Melkonian, M., Wong, G. K.-S., and Weijers, D. (2018) Origin and evolution of the nuclear auxin response system. *eLife* 7, No. e33399.
- (61) Flores-Sandoval, E., Eklund, D. M., Hong, S.-F., Alvarez, J. P., Fisher, T. J., Lampugnani, E. R., Golz, J. F., Vazquez-Lobo, A., Dierschke, T., Lin, S.-S., and Bowman, J. L. (2018) Class C ARFs

evolved before the origin of land plants and antagonize differentiation and developmental transitions in *Marchantia polymorpha*. *New Phytol.* 218 (4), 1612–1630.

(62) Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016) Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34, 525–527.

(63) Pollak Williamson, B. (2018) *Frameworks for reprogramming early diverging land plants*, doctoral thesis. DOI: [10.17863/CAM.20610](https://doi.org/10.17863/CAM.20610).

Supplementary-FigureS1-S16-TableS1-S2

Systematic tools for reprogramming plant gene expression in a simple model, *Marchantia polymorpha*.

Susanna Sauret-Güeto^{1,7}, Eftychios Frangedakis^{1,7}, Linda Silvestri^{1,2}, Marius Rebmann¹, Marta Tomaselli¹, Kasey Markel^{1,3}, Mihails Delmans^{1,4}, Anthony West^{5,6}, Nicola J. Patron⁵, and Jim Haseloff¹ *

¹ Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge. CB2 3EA. UK

² Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge. CB2 3EG. UK

³ Present address: Department of Plant Biology, University of California, Davis, US

⁴ Present address: Colorifix Ltd. Innovation Centre, Norwich Research Park, Norwich. NR4 7GJ. UK

⁵ Earlham Institute, Norwich, NR4 7UZ. UK

⁶ Present address: Isogenica Ltd., The Mansion, Chesterford Research Park, Little Chesterford, Essex. CB10 1XL. UK

⁷ equal contribution

*corresponding author email address: jh295@cam.ac.uk

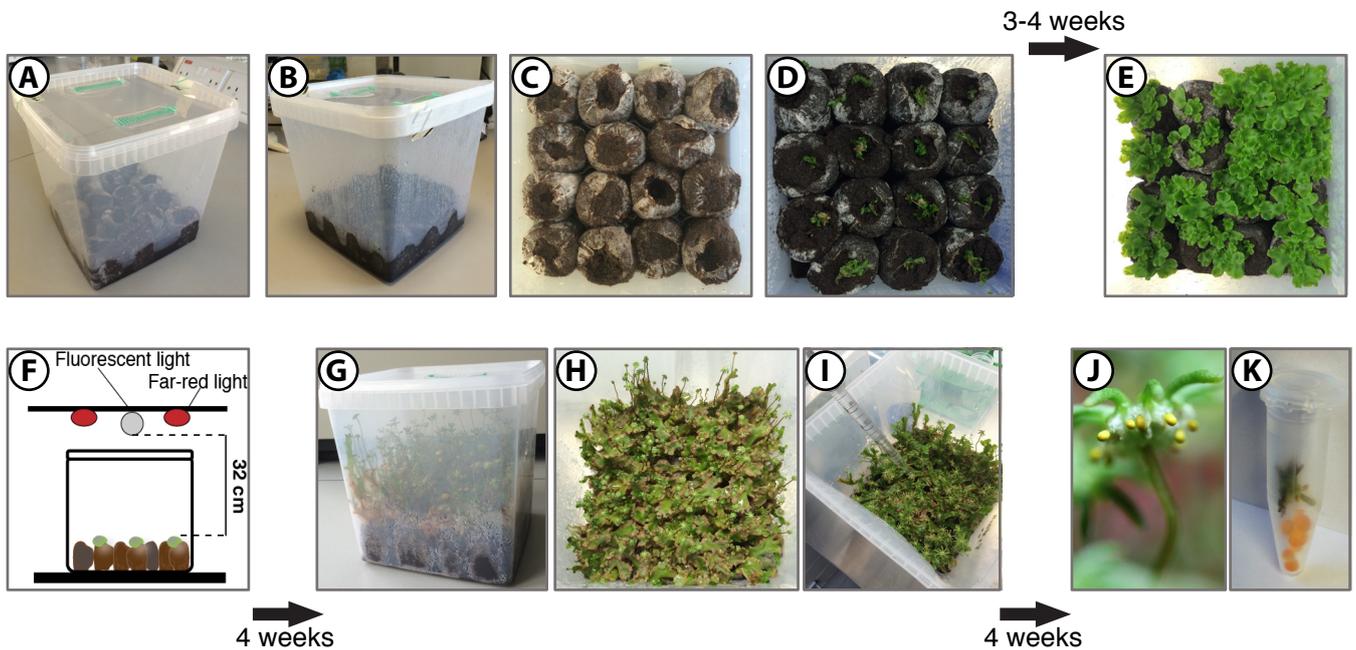


Figure S1. Techniques for axenic propagation of *Marchantia* spores. (A-B) 15 Jiffy-7 compressed peat disks were placed in a TP5000+TPD5000 Microbox with #40 green filter, 800 mL of water added, the lid was closed and the Microbox was autoclaved. (C-D) After autoclaving all work was performed in a laminar flow hood. One 10 mm x 10 mm thallus fragment or 3 gemmae were placed on each hydrated Jiffy-7 pellet using sterile tweezers. Another 150-200 mL of sterile water were added into the Microbox, the lid was sealed and the Microbox was placed at 21 °C under continuous light 150 μ E. (E-F) After one month, Microboxes were transferred to long day conditions (16 hours light / 8 hours dark) with light intensity of 150 μ E supplemented with far-red light (Philips GreenPower LED HF far red, peak emission around 730-740 nm, at a distance from plants of 32 cm). (G-H) In about 4 weeks, the plants produced mature male and female reproductive organs. (I) In a flow hood, the lid of the Microbox was opened, and using a pipette, 50-100 μ L of sterile water was added to the top of male reproductive organs. After a few seconds the sperm was visible as a cloudy exudate in the water drop. For fertilization, the drop was transferred, with a disposable sterile plastic pipette, to the top of a female reproductive organ. (J) After one more month, mature yellow colored sporangia were visible and ready for collection. In one Microbox, 60-70 fertile female gametophores usually developed, producing several hundred mature sporangia. (K) 1-2 mature sporangia were placed in a 1.5 mL centrifuge tube with 5 silica gel beads and stored at 4 °C if the spores were to be used within 1-2 months. For longer storage, sporangia were allowed to desiccate for a week at 4 °C and then stored at -80 °C.

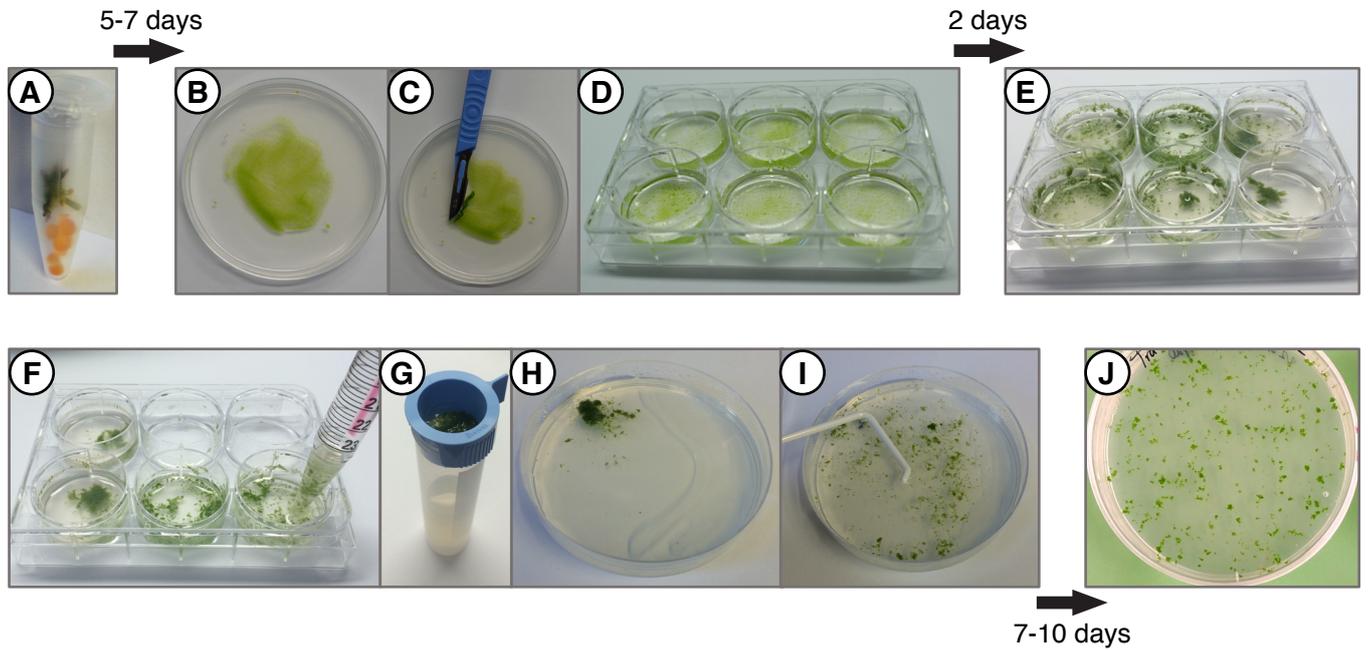


Figure S2. Agrobacterium-mediated transformation in multi-well plates. (A) Marchantia spores were sterilised and plated on 0.5x Gamborg B5 plates. (B-D) After 5-7 days, sporelings (germinated spores) were transferred into a 6-well plate containing 4 ml per well of 0.5x Gamborg B5 media plus supplements and 100 μ M acetosyringone, using a sterile scalpel. 100 μ L of an Agrobacterium culture transformed with a plasmid of interest was added to each well. (E) The plate was then placed on a shaker for 2 days at 21°C and continuous light. (F-G) After co-cultivation of sporelings and Agrobacterium, sporelings in each well were washed with 50 mL of sterile water on a cell strainer and (H-I) then plated on 0.5x Gamborg B5 plates with suitable antibiotics. (J) After 7-10 days, successful transformants were visible on the plate and could be screened with fluorescent microscopy.

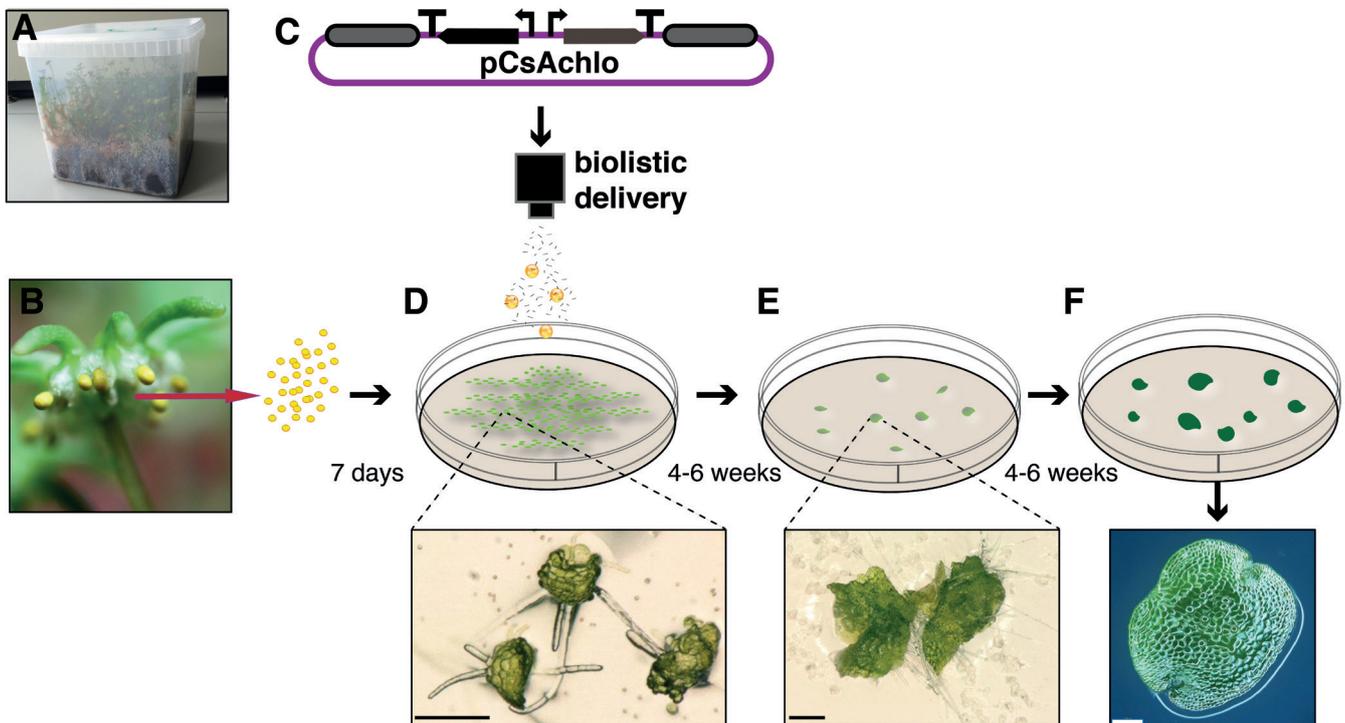


Figure S3. Outline of the chloroplast transformation method. (A) Microboxes were used to produce spores (Figure S1). (B) Spores were plated on 0.5x Gamborg B5 plates and incubated under continuous light for 7 days. (C-D) 7 day old sporelings were bombarded with DNAdelTM nanoparticles coated with the plasmid DNA of interest and placed in a growth room for 2 days. (E) They were then transferred to 0.5x Gamborg B5 plates with spectinomycin (500 $\mu\text{g}/\text{mL}$), and successful transformants were visible after 4-6 weeks. (F) Transformants were then transferred onto a fresh 0.5x Gamborg B5 plate containing spectinomycin. After 4-6 weeks in this second round of selection, plants produced gemmae cups, and gemmae could be tested for genotype and homoplasmy by PCR. Scale bar: (D-E) 200 μm , (F) 100 μm .

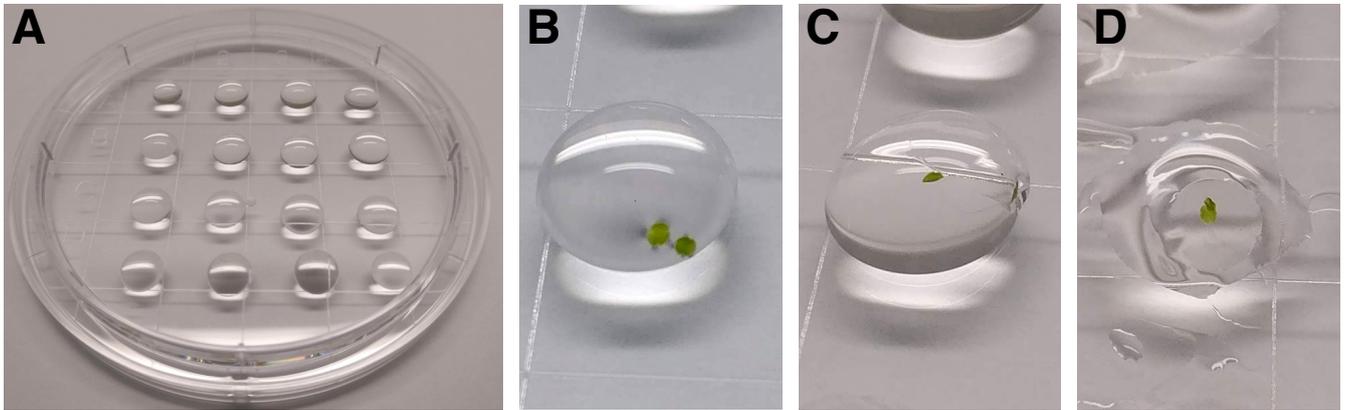


Figure S4. Cryopreservation of *Marchantia* gemmae. (A) Drops of alginate solution were prepared on a Petri dish, and (B) up to five *Marchantia* gemmae were transferred inside the alginate beads. (C) Beads polymerized upon addition of calcium chloride solution to the top of the bead, and gemmae were immobilized. (D) Gemmae were then soaked in dehydration solution and air-dried for at least 2 h. The dried beads with gemmae inside were placed in 1.5 mL Eppendorf tubes, frozen and kept at $-80\text{ }^{\circ}\text{C}$.

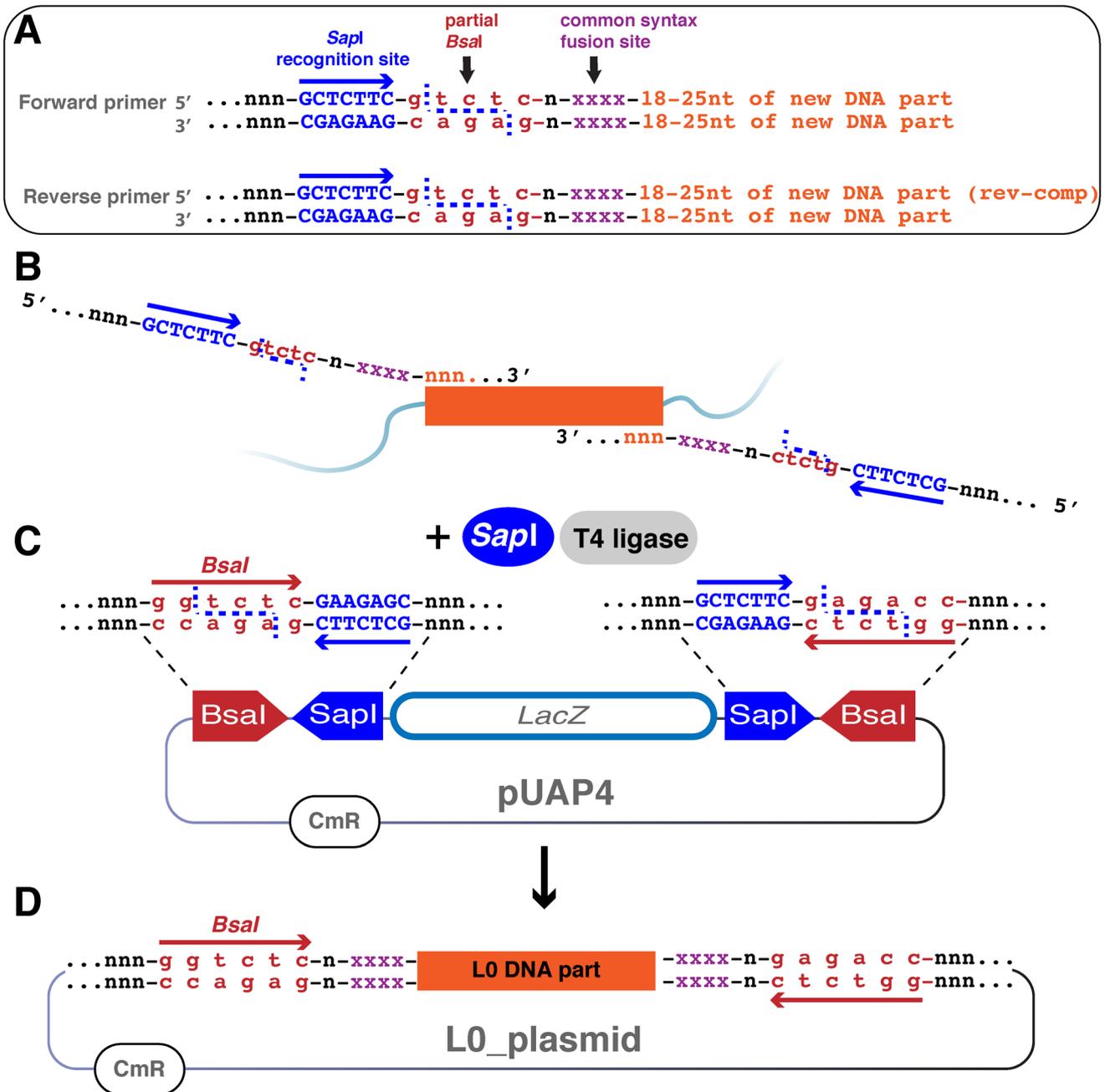


Figure S5. Design of primers for cloning of Level 0 parts into pUAP4 vectors. (A) To clone a standardized L0 part, 5' and 3' common syntax sites were fused to the DNA part by PCR, using specially designed primers. Primers included the sequences: 3 random bp (in black), the 7 bp *SapI* recognition site (in blue), 5 bp that correspond to a partial *Bsal* recognition site (in red), one random bp for spacing (in black), and the common syntax fusion site (in purple) followed by 18-25 bp complementary to the DNA part to be amplified (in orange). (B) DNA parts were amplified with a high fidelity DNA polymerase and (C) cloned into pUAP4 using *SapI* type IIS assembly to produce (D) the desired Level 0 plasmid. Once cloned into pUAP4, the full *Bsal* recognition site is reconstituted. *Bsal* type IIS cloning can be used to assemble multiple L0 parts into a transcription unit (L1). Blue arrows: *SapI* recognition site. Blue dashed lines: *SapI* cleavage site. Red arrows: *Bsal* recognition site. CmR: chloramphenicol bacterial resistance cassette. *LacZ*: lacZα cassette for blue-white screening of colonies (negative blue colonies contain undigested pUAP4, positive white colonies contain a L0 part inserted into pUAP4).

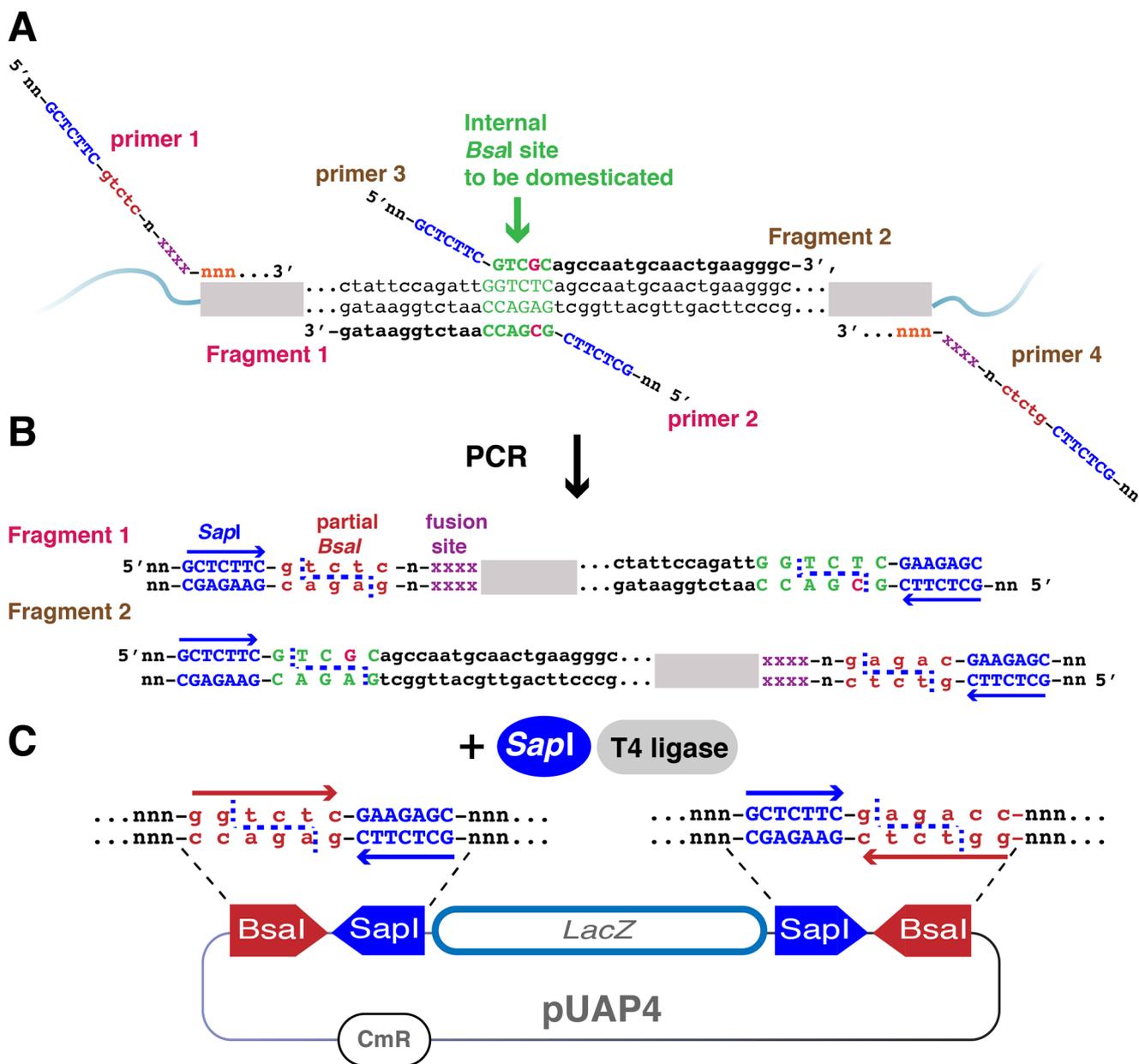


Figure S6. Guide to domestication of new DNA parts. Removal of internal *Bsal* and *SapI* sites is necessary when a DNA sequence of interest contains an internal *Bsal* or *SapI* recognition site. Overlapping PCR reactions can be used to mutate unwanted sites. Two PCR fragments are amplified upstream (fragment 1, with primers 1 and 2) and downstream (fragment 2, with primers 3 and 4) of the unwanted recognition site. Both primer 2 and 3 can be designed with a single nucleotide mismatch to alter the target sequence (taking care to not alter amino acid composition if the region is in the coding sequence). Primers also contain *SapI* recognition sites to allow the two fragments to be ligated together into pUAP4 using *SapI* type IIS assembly. Primer 1 and primer 4 are designed according to the guidelines for cloning of L0 parts into pUAP4 (see Figure S5). Blue arrows: *SapI* recognition site. Blue dashed lines: *SapI* cleavage site. Red arrows: *Bsal* recognition site. CmR: chloramphenicol bacterial resistance cassette. *LacZ*: lacZa cassette for blue-white screening of colonies.

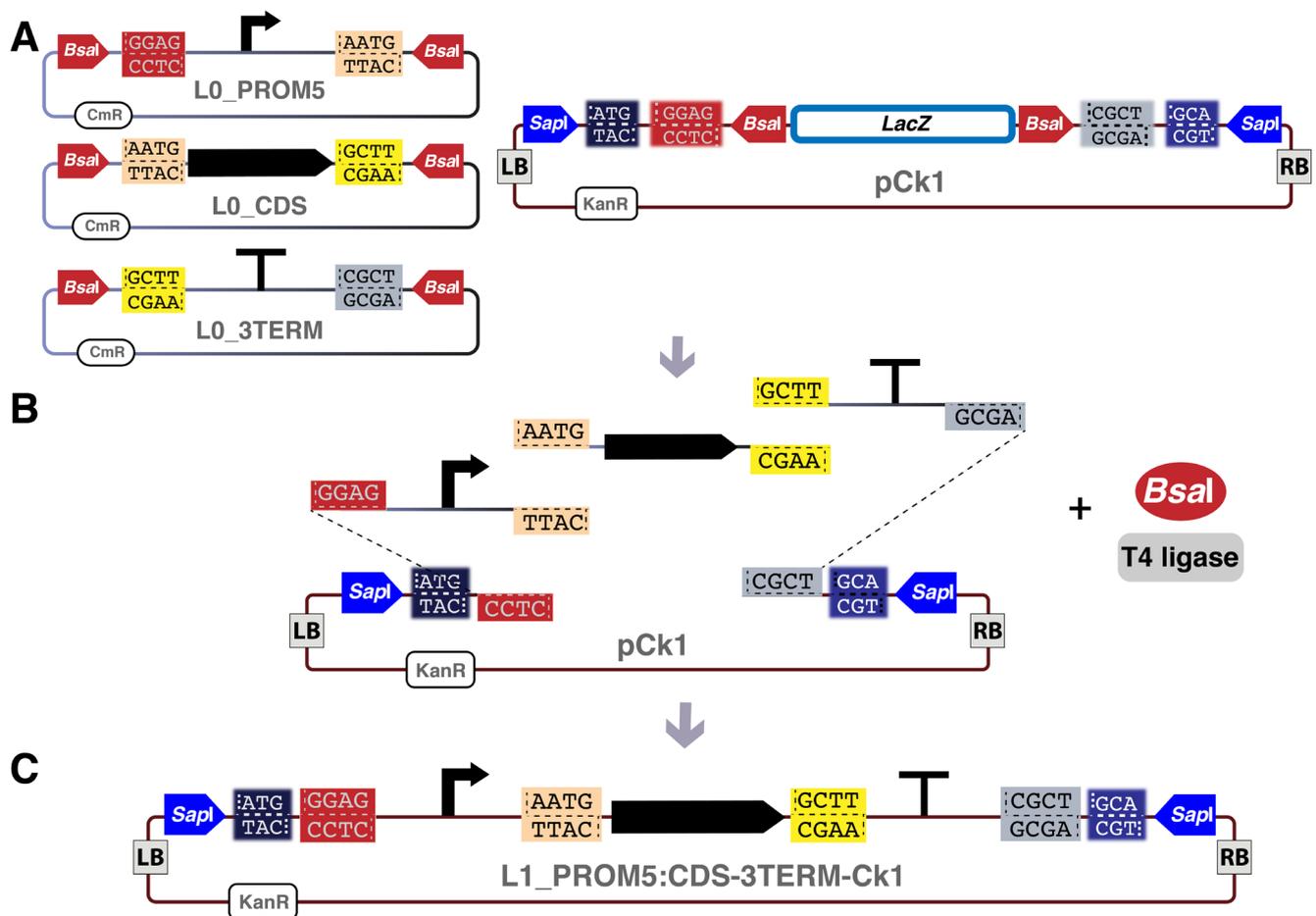


Figure S7. Assembly of multiple Level 0 parts into a transcription unit. (A-C) In a single-step digestion-ligation reaction, plasmids containing standard L0 parts (e.g. PROM5, CDS and 3TERM) are mixed with a pCk plasmid (e.g. pCk1), *Bsal* and T4 ligase enzymes. (A) L0 plasmids contain convergent *Bsal* sites and pCk plasmids contain divergent *Bsal* sites. (B) Digestion of a L0 plasmid with *Bsal* releases the part with 4 bp overhangs (e.g. GGAG) as sites for fusion according to the common syntax. Similarly, digestion of a pCk plasmid with *Bsal* releases the LacZ cassette and reveals the acceptor fusion sites as overhangs in the backbone (example CTCC). (C) L0 parts are ligated into the pCk vector with the fusion sites enabling directional assembly of DNA parts to build the desired L1 (in this case the L1_PROM5:CDS-3TERM-Ck1). Once the L1 has been assembled, it does not contain *Bsal* sites and thus cannot be further digested. After transformation into competent *E. coli* cells and plating on selection media with kanamycin, the colonies with undigested L0 plasmids will not be selected due to their different antibiotic resistance cassette (chloramphenicol) and colonies with undigested pCk plasmids will be blue in plates containing X-gal. Red arrows: *Bsal* restriction sites. Blue arrows: *SapI* restriction sites. CmR: chloramphenicol bacterial resistance cassette. KanR: kanamycin bacterial resistance cassette. LacZ: lacZa cassette for blue-white screening of colonies.

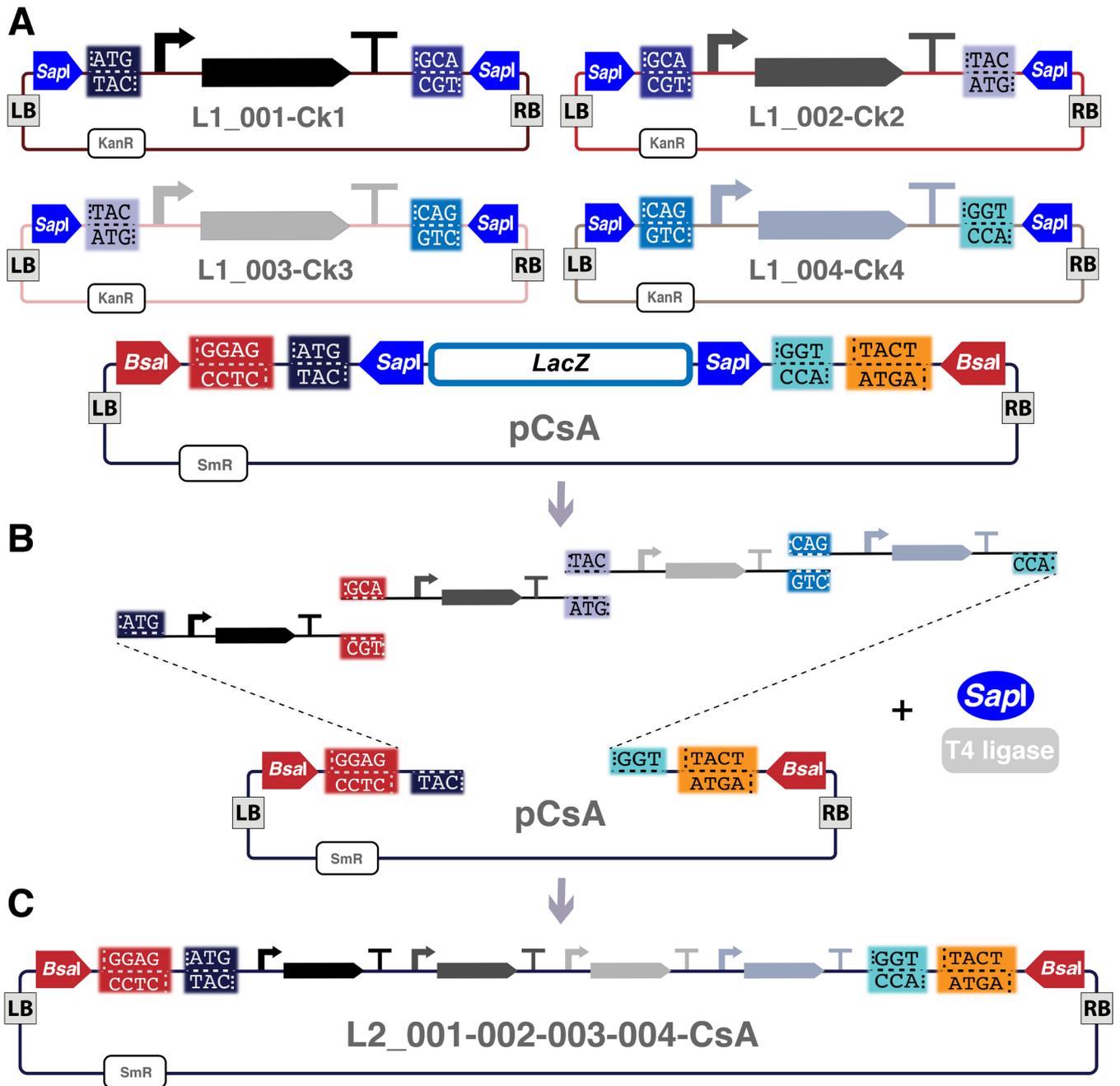


Figure S8. Loop assembly of four transcription units into a Level 2 plasmid. (A-C) In a one-step reaction, four transcription units assembled in pCk vectors (in this case L1_001-Ck1, L1_002-Ck2, L1_003-Ck3, L1_004-Ck4) were mixed with a pCs plasmid (in this case, pCsA), *SapI* and T4 ligase enzymes. (A) L1 plasmids contain convergent *SapI* sites and pCs plasmids contain divergent *SapI* sites. (B) Digestion of a L1 plasmid with *SapI* releases the L1 construct with its specific 3 bp Loop fusion sites as overhangs (e.g. ATG). Similarly, digestion of a pCs plasmid with *SapI* releases the LacZ cassette and reveals the acceptor 3 bp fusion sites as overhangs in the backbone (example CAT). (C) Released L1 fragments are ligated into the pCs vector, with the fusion sites enabling directional assembly of DNA parts and building of the desired L2 device (in this case the L2_001-002-003-004-CsA). Once the L2 has been assembled, it lacks *SapI* sites and thus cannot be further digested. The L2 reaction was transformed into competent *E. coli* cells and plated on selection media with spectinomycin. Undigested L1 plasmids do not carry the correct antibiotic resistance cassette and undigested pCs plasmids will be marked blue in plates containing X-gal. Red arrows: *Bsal* restriction sites. Blue arrows: *SapI* restriction sites. KanR: kanamycin bacterial resistance cassette. SmR: spectinomycin bacterial resistance cassette. LacZ: lacZ α cassette for blue-white screening of colonies.

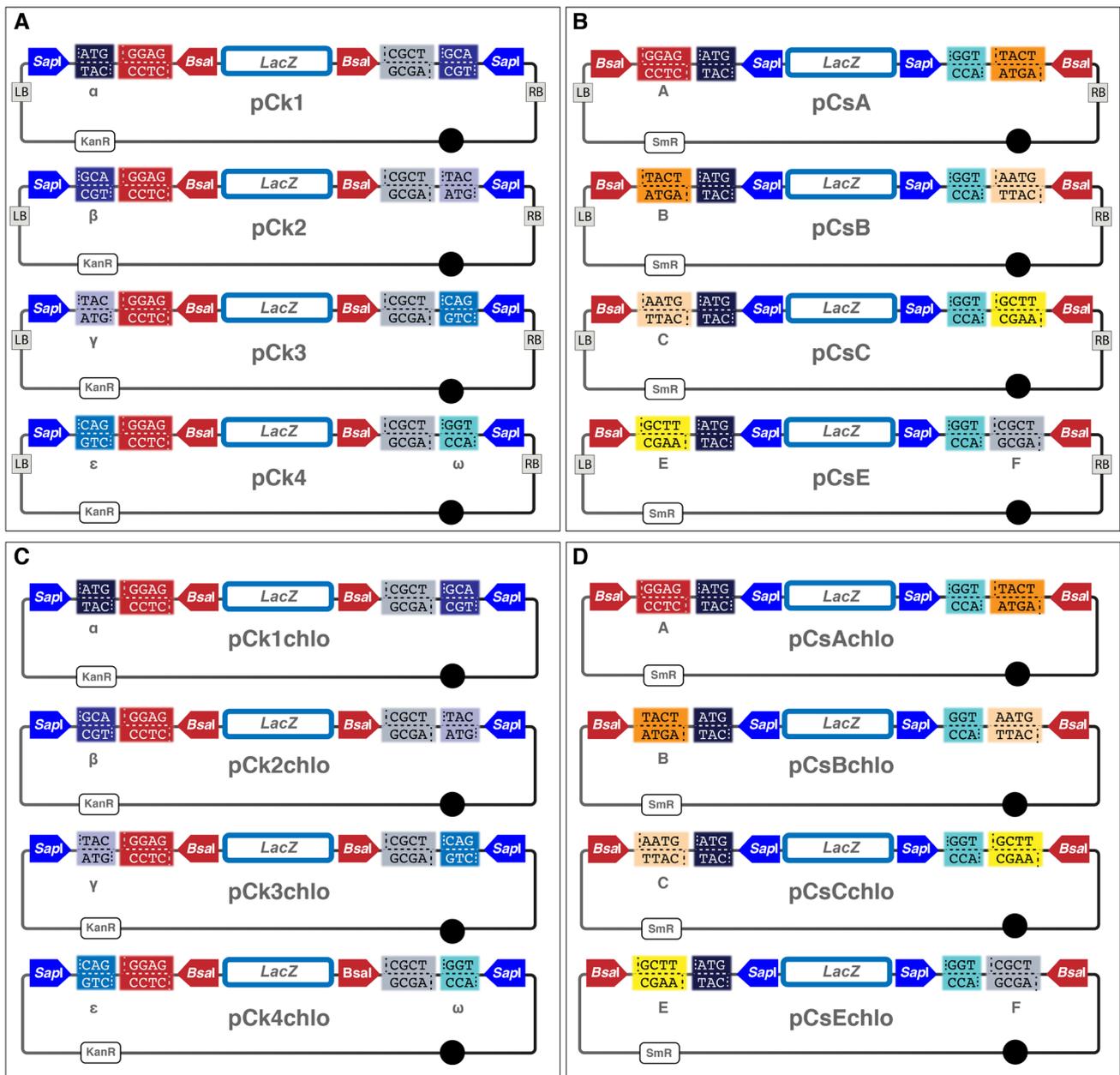


Figure S9. Improved Loop vectors for nuclear and chloroplast transformation. (A) pCk vectors (pCk1, pCk2, pCk3 and pCk4) all contain divergent *Bsal* restriction sites (red arrows) adjacent to specific Loop fusion sites [A (GGAG) and F (CGCT)] to allow assembly of different L0 parts into a L1 construct using a pCk vector and *Bsal*. They also contain convergent *SapI* restriction sites (blue arrows) next to specific Loop fusion sites to allow assembly of different L1 constructs into a L2 construct using a pCs vector and *SapI*. The fusion sites are: α (ATG), β (GCA), δ (TAC), ϵ (CAG) and ω (GCT), as pCk1 α - β , pCk2 β - δ , pCk3 δ - ϵ and pCk4 ϵ - ω . (B) pCs vectors (pCsA, pCsB, pCsC and pCsE) all contain divergent *SapI* restriction sites (blue arrows) adjacent to specific Loop fusion sites, α - ω , to assemble up to four L1 constructs into a L2 device using a pCk vector and *SapI*. They also contain convergent *Bsal* restriction sites (red arrows) next to specific Loop fusion sites to allow assembly of up to four L2 devices into a L3 device using a pCk vector and *Bsal*. The fusion sites are: A (GGAG), B (TACT), C (AATG), E (GCTT) and F (CGCT), which are present as: pCs1 A-B, pCs2 B-C, pCs3 C-E, pCs4 E-F. (C-D) Transformation vectors for chloroplast applications, pCkchlo vectors (C) and pCsChlo vectors (D), contain, respectively, same *Bsal*, *SapI* restriction sites, and fusion sites as the pCk and pCs sets, but they do not contain the left (LB) and right (RB) repeats from the nopaline C58 T-DNA for *Agrobacterium*-mediated nuclear transformation. KanR: kanamycin bacterial resistance cassette. SmR: spectinomycin bacterial resistance cassette. LacZ: lacZ α cassette for blue-white screening of colonies.

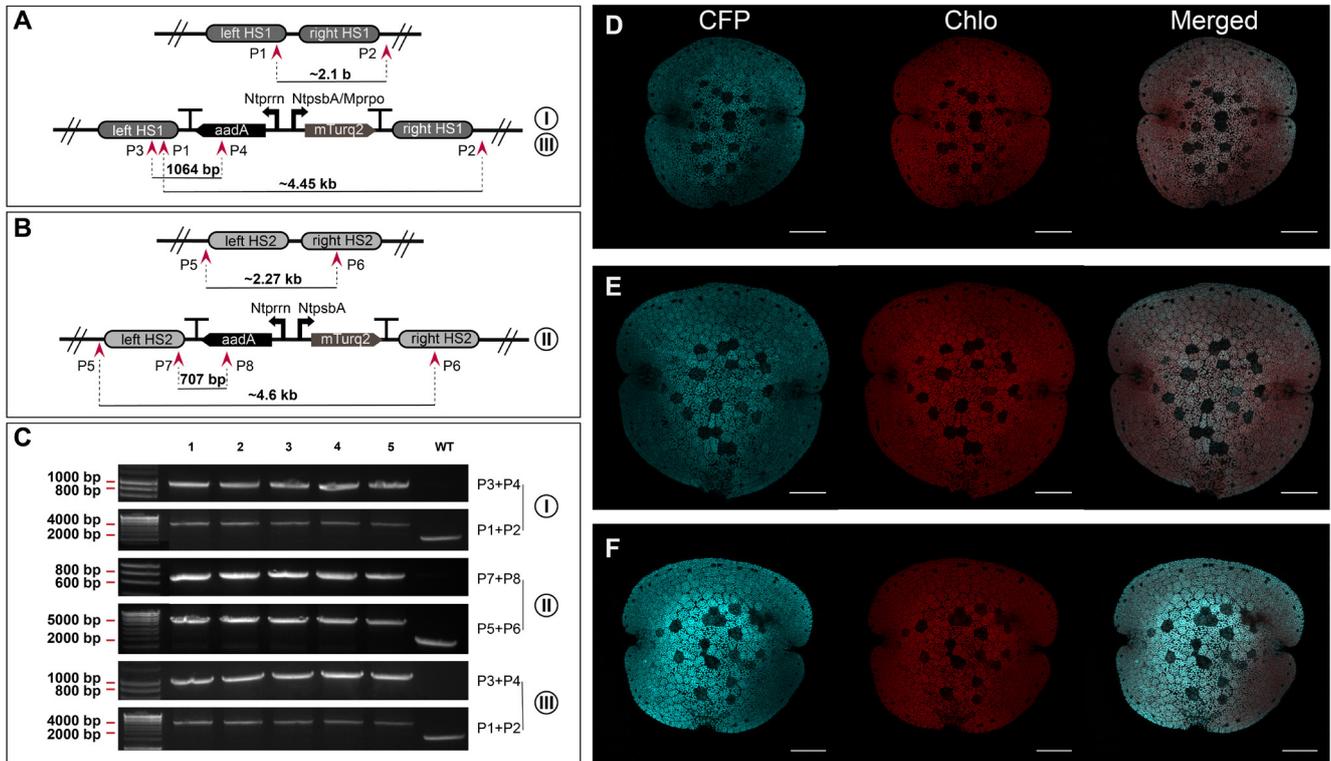


Figure S10. Validation of DNA parts and vectors for chloroplast transformation. (A) Schematic representation of the *trnG-trnM* target region (flanked by HS1-L and HS1-R) in the wild type chloroplast genome (top) and the same region after integration of the DNA construct I or III (bottom). Construct I: p5-NtspbA:mTurquoise2_chlo for integration into intergenic region 1. Construct III: p5-Mprpo:mTurquoise2_chlo for integration into intergenic region 1. (B) Schematic representation of the *rbcL-trnR* target region (flanked by HS2-L and HS2-R) in the wild type chloroplast genome (top) and the same region after integration of the DNA construct II (bottom). Construct II: p5-NtspbA:mTurquoise2_chlo for integration into intergenic region 2. (A-B) Red arrows indicate the position of the PCR primers used for the detection of wild type or homoplasmic transplastomic lines. Maps not in scale. (C) PCR analysis of genomic DNA isolated from wild type and transplastomic plants. Integrity of the reporter gene (top, primers P3+P4 or P7+P8) and homoplasmy (bottom, primers P1+P2 or P5+P6) were confirmed for transplastomic lines after 2 months of subculture under selective conditions. The primer pair used for each PCR are shown next to the gel images. (D-F): Images of 0 days transplastomic gemmae expressing (D) construct I, (E) II or (F) III. (D-F) Left: mTurquoise2cp, middle: Chlorophyll autofluorescence, right: merged images. Scale bar: 100 μ m.

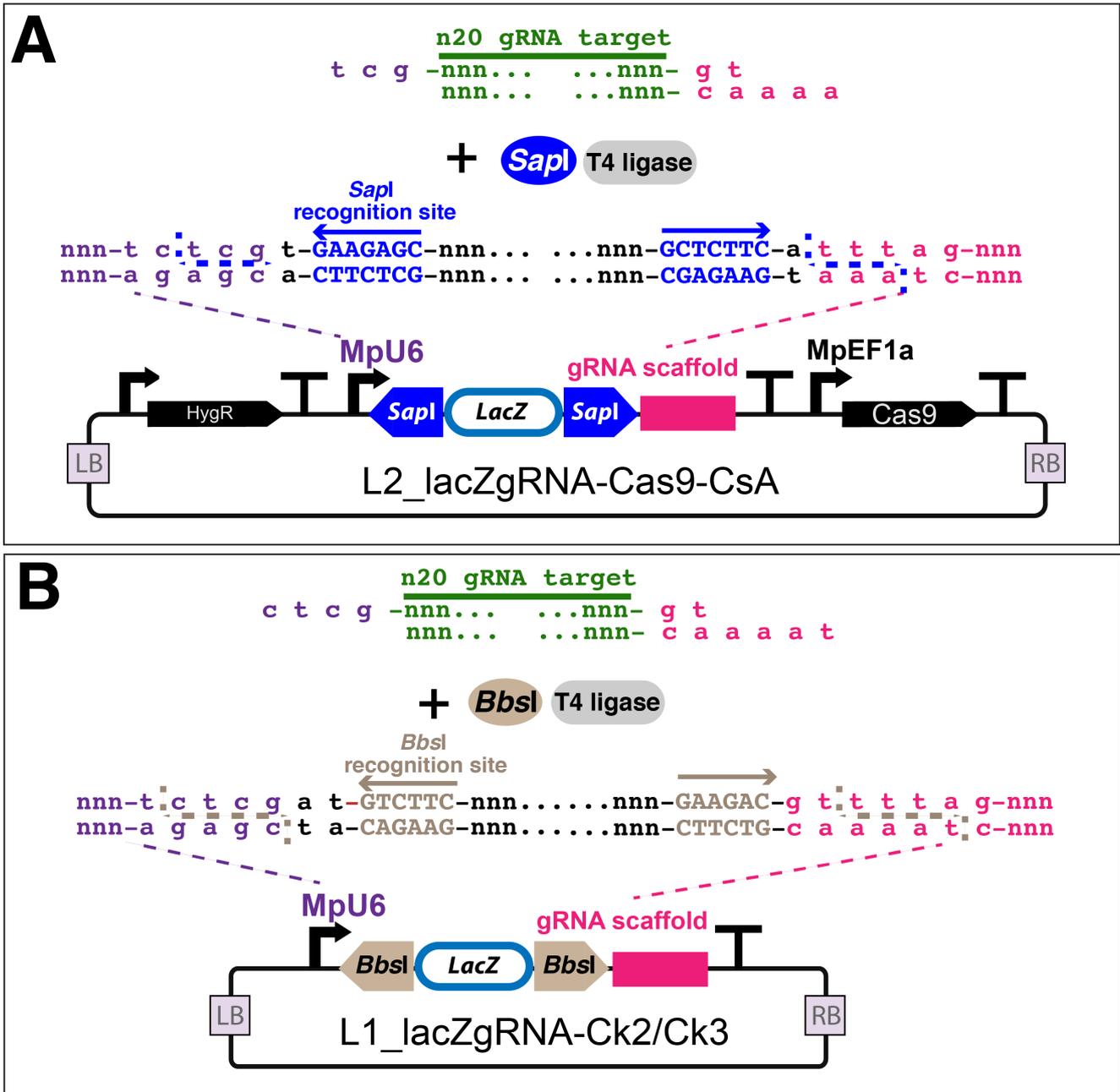


Figure S11. Design of oligonucleotide primers for rapid cloning of gRNAs. (A) Design of synthetic oligonucleotides for gRNA *SapI* mediated cloning into the L2_lacZgRNA-Cas9-CsA vector. The L2_lacZgRNA-Cas9-CsA *SapI* digested vector has AGC and TTT overhangs. Therefore, oligos for gRNA should be designed such that the forward strand has a 5' overhang of TCG and the reverse strand has a 5' overhang of gt-AAA (addition of "gt" nucleotides is necessary to reconstitute the full sequence of the gRNA scaffold in pink). Blue arrows: *SapI* recognition site. Blue dashed lines: *SapI* cleavage site. (B) Design of oligos for gRNA *BbsI* mediated cloning into L1_lacZgRNA-Ck2 or L1_lacZgRNA-Ck3 vectors. The L1_lacZgRNA-Ck2/3 *BbsI* digested vectors have GAGC and TTTA overhangs. Synthetic oligos for the gRNA should be designed so that the forward strand has a 5' overhang of CTCG and the reverse strand has a 5' overhang of gt-AAAT (addition of "gt" nucleotides is necessary to reconstitute the full sequence of the gRNA scaffold in pink). Light brown arrows: *BbsI* recognition site. Light brown dashed lines: *BbsI* cleavage site. (A-B) LacZ: lacZa cassette for blue-white screening of colonies.

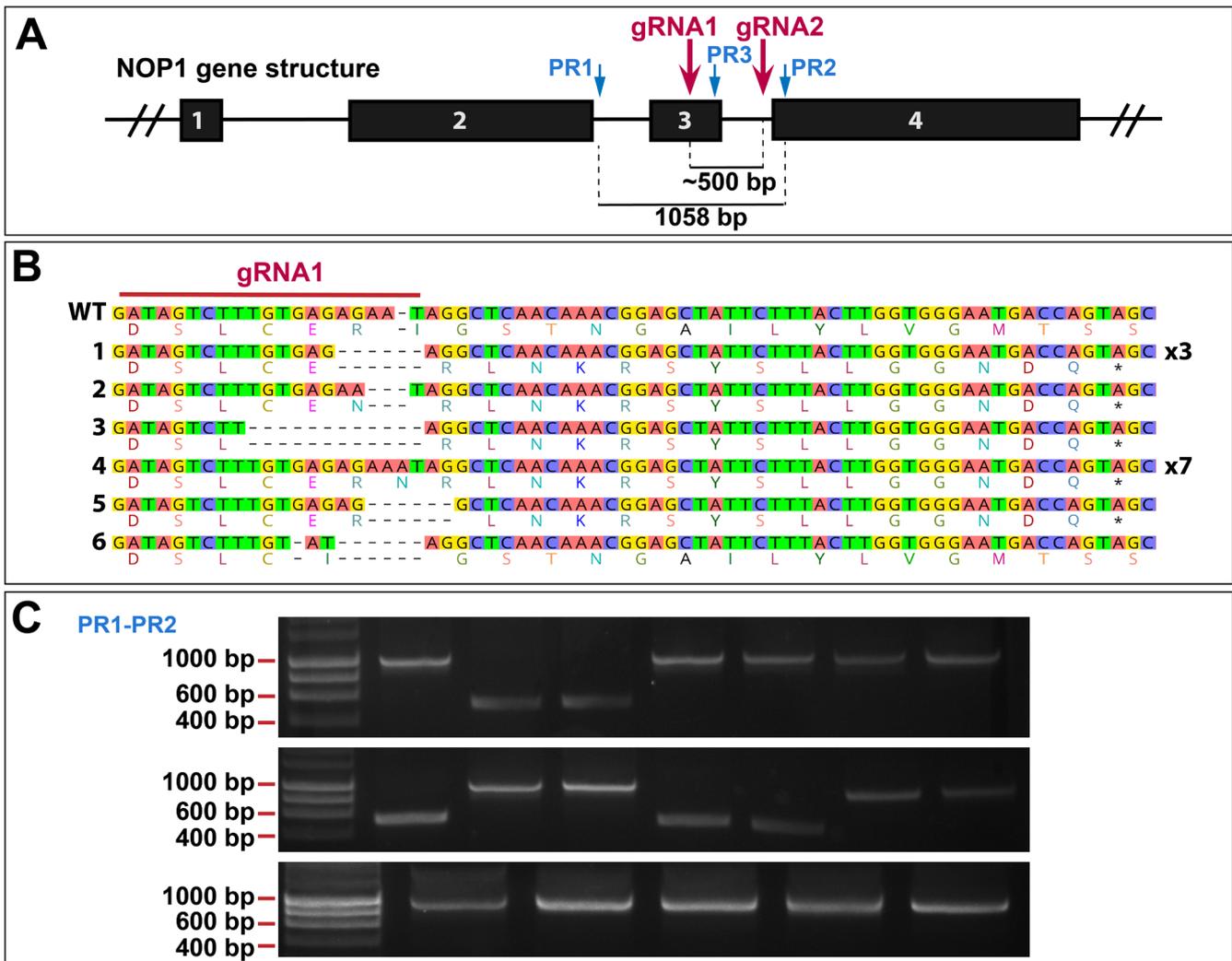


Figure S12. Cas9-CRISPR mediated gene editing and generation of deletions. (A) Schematic representation of the *MpnOP1* gene structure showing exons as black boxes. The gRNAs for CRISPR and primers for PCR-based genotyping are shown in red and blue, respectively. (B) Sequence analysis of *nop1* knockout mutant lines obtained with the L2 gRNA1 CRISPR vector (L2_NOP1gRNA-Cas9-CsA, Figure 4A). The wild type *Marchantia Cam-1* sequence is shown at the top, with the 20 bp gRNA1 target sequence highlighted with a red line. The amino acid sequence is depicted below the nucleotide sequence. Primers used: PR1-PR3. Of the 14 mutants analyzed, sequence deletions of identical to line 1 occurred 3 times and for line 4, 7 times. (C) Gel electrophoresis images of PCR-based genotyping of the mutants obtained with the L2 gRNA1-gRNA2 CRISPR vector (L2_2xNOP1gRNA-Cas9-CsA, Figure 4C). Primers used: PR1-PR2. Expected size of PCR products for wild type or mutants with small deletions (as in B) is approximately 1000 bp. Expected size of PCR products when large deletions occurred is approximately 500 bp. The large deletions were further confirmed by Sanger sequencing.

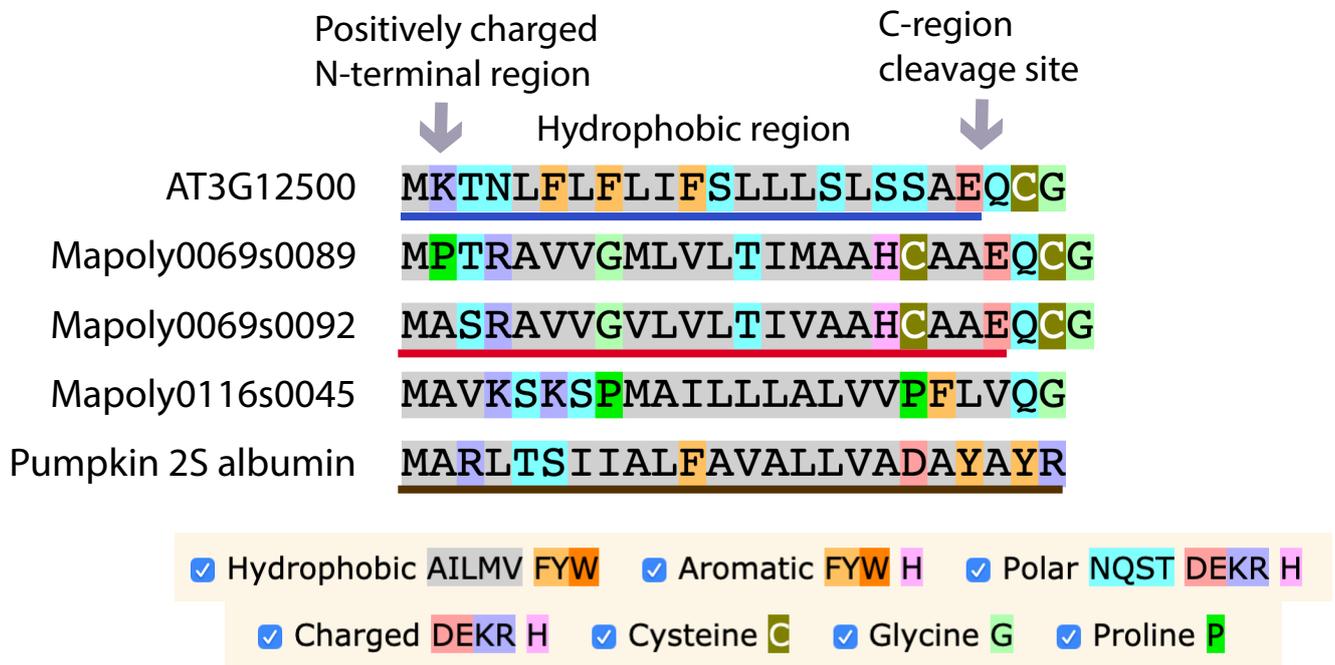


Figure S13. Alignment of ER signal peptides. ER signal peptides do not show conserved primary amino acid sequence, but typically include a positively charged N-terminal region, followed by a hydrophobic region, and C-region that defines the cleavage site ¹. The ER signal peptide from the Arabidopsis basic chitinase (AT3G12500) (underlined in blue) has been used in other plant species to confer ER localization. The ER signal peptide from pumpkin 2S albumin (underlined in brown) has been shown to work in Marchantia. In this work, we identified three predicted native chitinases in Marchantia (Mapoly0069s0089, Mapoly0069s0092, and Mapoly0116s0045), tested the leader peptide from one of them, Mapoly0069s0092 (underlined in red), and confirmed that it conferred ER localization in Marchantia. Protein alignment was performed using MSAReveal.org

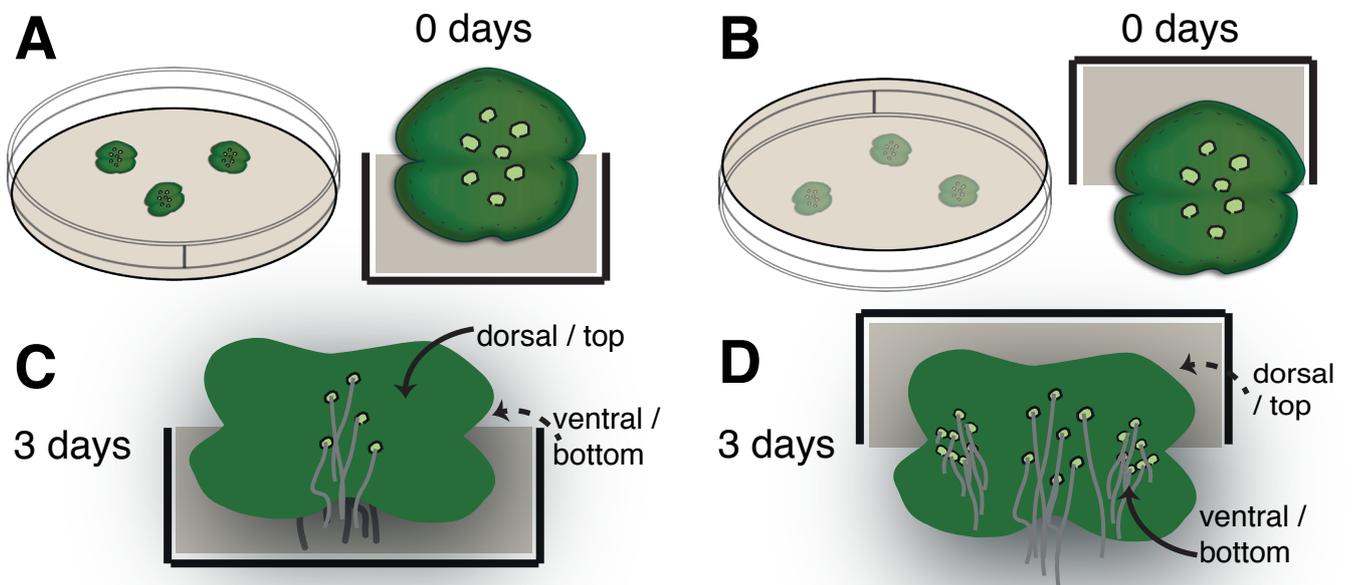


Figure S14. Schematic representation of gemmae growing in normal and inverted plates. The bottom side of gemmae acquires ventral identity both in normal and inverted plates. Gemmae removed the gemma cups (0 day) have rhizoid precursors present on both sides of the central part of the gemmae. Rhizoids are larger size cells that lack mature chloroplasts (light green cells in A,B). They elongate once gemmae are placed on media and dormancy is released (C,D). As gemmae grow, they acquire dorso-ventral polarity, with the dorsal (top) side developing air chambers and the ventral (bottom) side developing predominant rhizoids (A,C). If plates are inverted (B), the lower sides of the inverted gemmae form new rhizoids and adopt ventral identities, despite not touching the media (D).

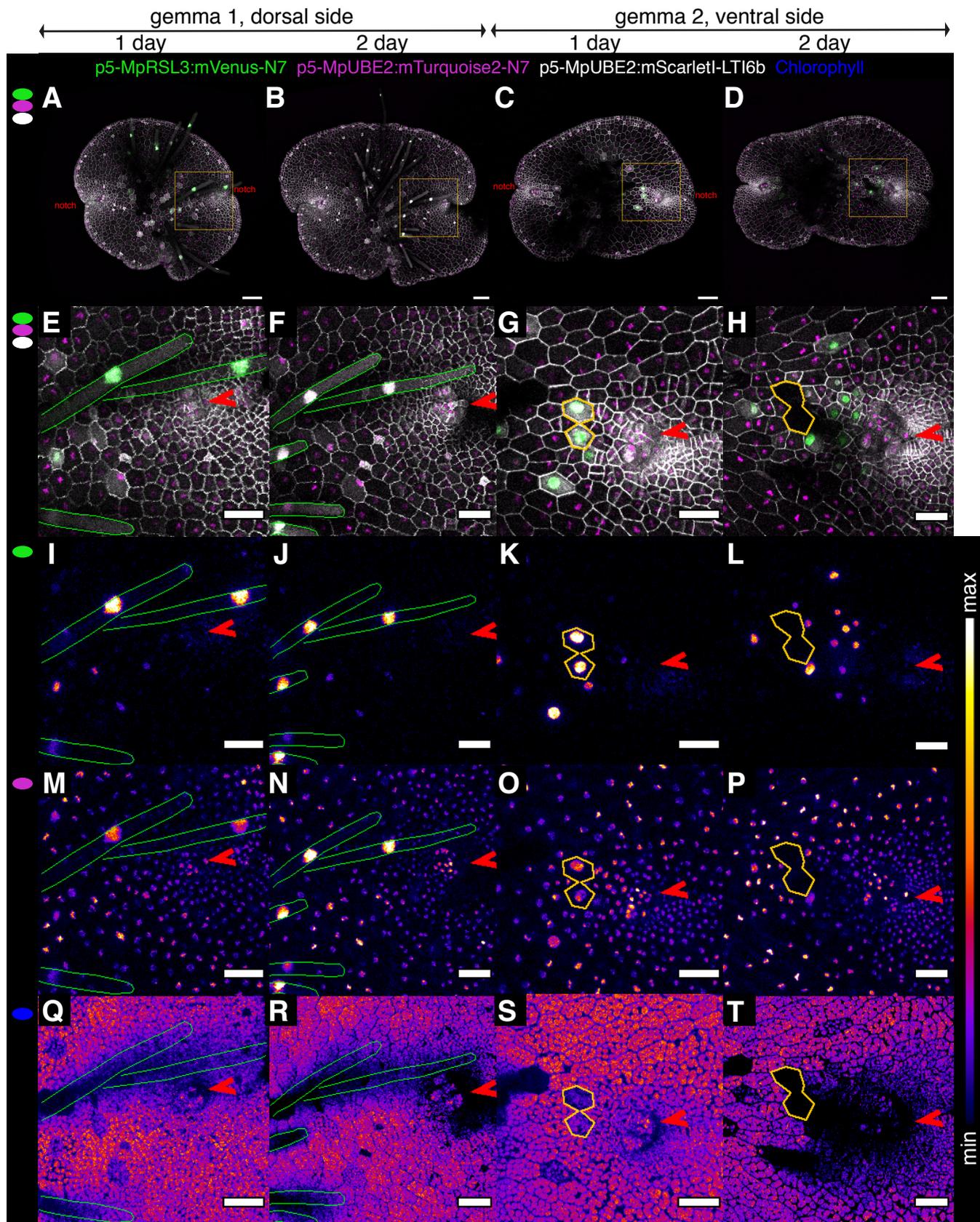


Figure S15. Time course of MpRSL3 expression in early gemma development. Images of the dorsal side of gemmae, or the ventral side of a gemmae. In order to image the ventral side of gemmae, we grew plants on agar in an inverted plate (see Figure S14). Line expressing: p5-MpRSL3:mVenus-N7 (rhizoid marker, in green), p5-MpUBE2:mTurquoise2-N7 (constitutive nuclear expression, in magenta) and p5-MpUBE2:mScarlet1-LTI6b (ubiquitous cell outline, in grey). Panels A-H show maximum intensity Z-axis projections gemmae surfaces with merged channels for mScarlet1 (grey), mTurquoise2 (magenta) and mVenus (green) channels. Image of whole gemmae (A-D) or a region (orange square in A-D and images E-T) from the center of the gemmae to one of the two notches (red arrow in E-T). At 1 and 2 days, the MpRSL3 marker can be detected in elongated rhizoids emerging from the central zone of the gemmae (A-D), like the ones delineated in green in the dorsal side images. In elongated rhizoids, the nuclei moves to the tip of the rhizoid cell (E,F). Moreover, only in the ventral side, the MpRSL3 marker is expressed in cells near the notch (G,H), which will become elongated rhizoids. Delineated in orange, a couple of cells expressing the MpRSL3 promoter at 1 day (G), which have already elongated at 2 days (with the nuclei above the Z-stack imaged and thus not in the image (H)). (I-L). Levels of accumulation of the mVenus marker shown in false-color. (M-P) Levels of accumulation of the mTurquoise2 marker shown in false-color. (Q-T) Levels of chlorophyll autofluorescence shown in false-color. Scale bar: 100 μm (A-D) or 50 μm (E-T).

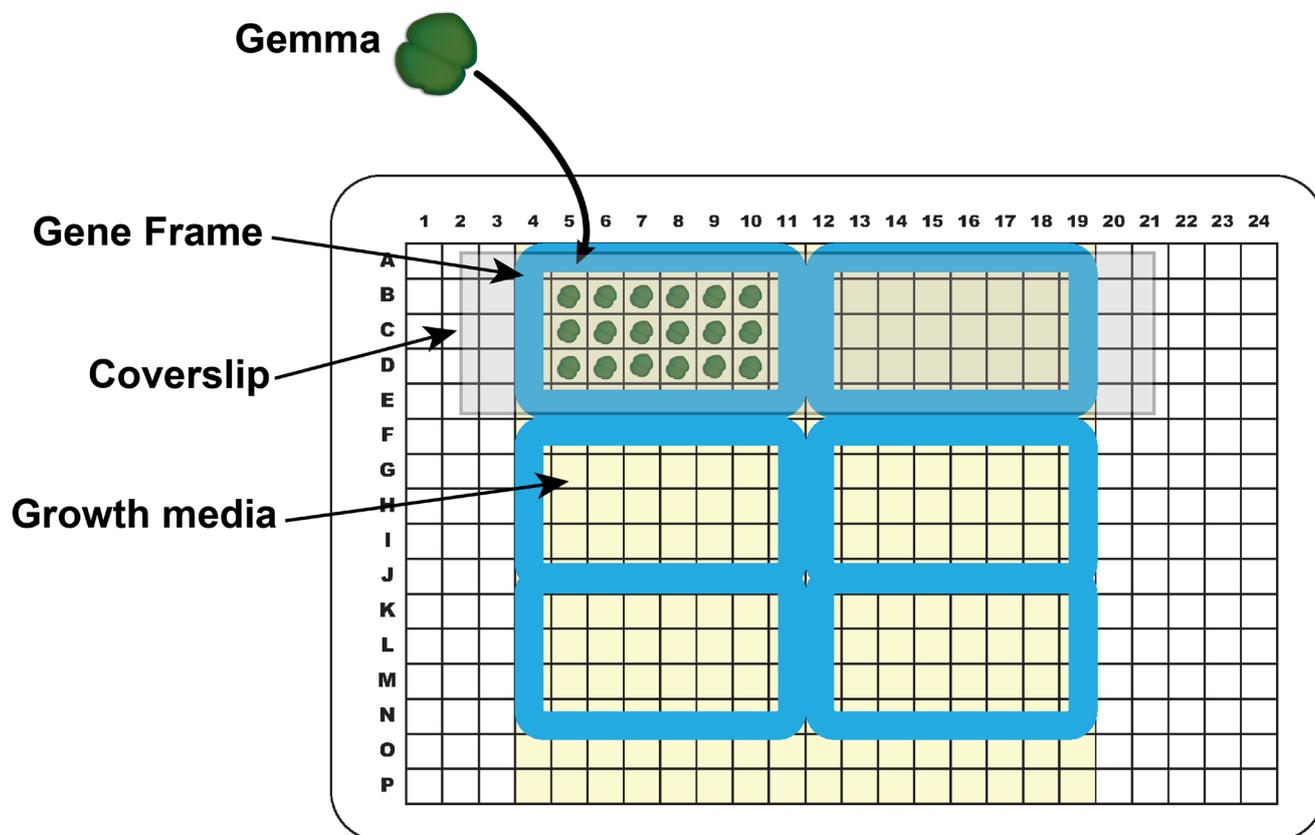


Figure S16. Culture of gemmae on multiwell plates for live imaging. Wells of a 384-well plate were filled with 0.5x Gamborg B5 media with 1.2% (w/v) agar. Single gemmae were placed at the center of each well. Gene frames and coverslips treated with anti-fog spray were used to cover and seal the wells. The setup allows automated imaging of samples.

Table S1. Quantification of CRISPR/Cas9 mediated editing of MpNOP1 using the OpenPlant toolkit CRISPR/Cas9 vectors

Number of plants screened (L2_NOP1gRNA-Cas9-CsA)				
	total	wild type	<i>nop1</i>	chimeric
experiment 1	121	58	28	35
	%	47.93	23.14	28.92
experiment 2	133	69	22	42
	%	51.87	16.54	31.57

Number of plants screened (L2_2xNOP1gRNA-Cas9-CsA)				
	total	wild type	<i>nop1</i>	chimeric
experiment 1	127	46	36	45
	%	38.01	29.75	37.19
experiment 2	125	61	25	39
	%	48.8	20	31.2

Table S2. List of primers used in this study

	For sequencing inserts in pUAP4 and pC vectors (pCk and pCs)
UAP_F	CTCGAGTGCCACCTGACGTCTAAGAAAC
UAP_R	CGAGGAAGCCTGCATAACGCGAAGTAATC
pC_F	GCAACGCTCTGTCATCGTTAC
pC_R	GTAACCTAGGACTTGTGCGACATGTC
pC_R2	CAATCTGCTCTGATGCCGCATAGTTAAG
	For genotyping chloroplast transformants with construct I and III (Figure S10)
P1_F	TTCGATTCCCCTACCCGC
P2_R	GCTGGAACAGCTATCTTAACTTTTATTGGTGG
P3_F	CTAGCTTCTCCTGAAGGTTGGTC
P4_R	GTTGTCCCGCATTGGTACAGC
	For genotyping chloroplast transformants with construct II (Figure S10)
P5_F	CGGAGACTAAAGCAGGTGTTGG
P6_R	ACAGATCCCATACTACCGCC
P7_F	CGGATCATGAAGTCAAGGGTTC
P8_R	GTTGTCCCGCATTGGTACAGC
	For cloning NOP1 gRNA1 and gRNA2 into L1-lacZgRNA-Ck2/3
gRNA1-BbsI-F	CTCG-ATAGTCTTTGTGAGAGAATgt
gRNA1-BbsI-R	TAAAc-ATTCTCTCACAAGACTAT
gRNA2-BbsI-F	CTCG-ATTAAGAGTGAAGTTGCTTgt (modified from ref 2)
gRNA2-BbsI-R	TAAAc-AAGCAACTTCCACTCTTAAT (modified from ref 2)
	For cloning NOP1 gRNA1 into L2-lacZgRNA-Cas9-CsA
gRNA1-SapI-F	TCG-ATAGTCTTTGTGAGAGAATgt
gRNA1-SapI-R	AAAc-ATTCTCTCACAAGACTAT
	For genotyping <i>nop1</i> CRISPR mutants
PR1_F	TGGGCCCATCTGTCACTAAG
PR2_R	CATGTCCACCAGTACGTG
PR3_R	GGTTGTATACGCCCGCTTTC

REFERENCES

- (1) Gnanasambandam, A.; Polkinghorne, I. G.; Birch, R. G. Heterologous signals allow efficient targeting of a nuclear-encoded fusion protein to plastids and endoplasmic reticulum in diverse plant species. *Plant Biotechnol. J.* **2007**, *5* (2), 290–296.
- (2) Sugano, S. S.; Nishihama, R.; Shirakawa, M.; Takagi, J.; Matsuda, Y.; Ishida, S.; Shimada, T.; Hara-Nishimura, I.; Osakabe, K.; Kohchi, T. Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in *Marchantia polymorpha*. *PLoS One* **2018**, *13* (10), e0205117.