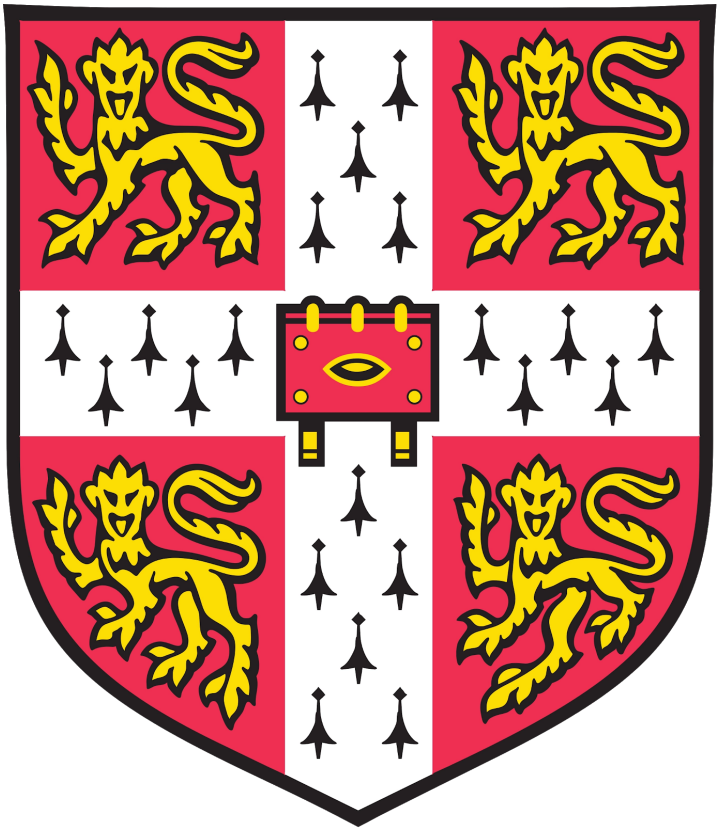
Improved plastid transformation in *Marchantia polymorpha*



Kasey Markel
Darwin College

Department of Plant Sciences
University of Cambridge

This dissertation is submitted for the degree of

Master of Philosophy

I would like to dedicate this thesis to my grandparents Bill and Jean Markel, for giving me every opportunity.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have never been submitted in whole or in part for consideration for any other degree at this or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

This dissertation contains fewer than 20,000 words exclusive of tables, footnotes, bibliography, and appendices.

Kasey Markel

August 2018

Acknowledgements

Throughout the last year, I have been fortunate to be surrounded by people who have helped to shape the course of this thesis, and, by extension, my personal development.

First and foremost, I would like to thank my supervisor, Dr. Jim Haseloff. Projects from his group have inspired me for years, and it has been an honor to transition from admiring the research group from afar to actively contributing from within the laboratory. His dedication to the research path continues to be an inspiration.

I would also like to thank Dr. Eftychios Frangedakis and Linda Silvestri, who together taught me all I know about *Marchantia* tissue culture and transformation protocols. Their patience, guidance, and troubleshooting assistance were invaluable throughout this process.

Furthermore, I would like to thank, in no particular order, the members of the Haseloff laboratory I've had the pleasure of working with throughout the last year. Dr. Susana Sauret-Gueto, Dr. Jenny Molloy, Dr. Lucas Mueller, Marta Tomaselli, Mihails Delmans, Dr. Emma Talbot, Liat Adler, Owen Male, Tom Bennett, Jet Mante, and Marius Rebmann. I would like to thank Imogen Black, Paula Cohen, and Jim Haseloff for comments which have greatly improved this manuscript.

Influence can of course last longer than a degree, so I would like to thank a few people who set me on the path of this project. Dr. Jeff Prince, for supervising so many projects during my undergraduate studies, Dr. Barbara Whitlock, for encouraging me to come to Cambridge, and Donna Widhalm for helping me find my passion for biology.

Finally, I would like to thank my family. Your love and support have kept me strong through many difficult times. I would not have been able to do this without you.

Abstract

Plastids are subcellular organelles which perform photosynthesis in plants and algae. They are descended from free-living photosynthetic bacteria which were engulfed by early eukaryotes ~1.5 billion years ago. Nearly all plastids retain a small genome separate from the nuclear genome of the plants in which they reside, which is termed the plastome (plastid genome). This genome encodes many of the proteins required for photosynthesis, as well as bacteria-like transcriptional and translational machinery. Biotechnological techniques for making precise modifications to the plastome have been available for three decades, nearly as long as the equivalent techniques for the plant nuclear genome. However, while nuclear genetic engineering quickly moved from a laboratory novelty to a tool used for the improvement of crop plants planted on over a billion acres, plastome engineering remains largely confined to research use. The primary reasons for this lack of application are the species restrictions and technical difficulty of the transformation process, which until recently was only possible in a few species.

In 2007, a protocol for plastid transformation was reported for *Marchantia polymorpha*, a thalloid liverwort classically used as a model species. *Marchantia* offers rapid generation time, small size, simple genetics, and asexual reproduction by means of gemmae, small disks of tissue which provide a powerful platform for live-tissue microscopy. In this thesis, tools for *Marchantia* plastid transformation are systematically improved by the generation of the first plastome sequence assembly for a widely used laboratory strain of *Marchantia*, optimisation of the transformation protocol itself, and a comparison of the *in vivo* activity of plastid regulatory elements through a fluorescent marker and quantitative microscopy. The highly conserved nature of the land plant plastome suggests the improvements to plastid transformation developed in *Marchantia* will translate to other species.

Table of Contents

List of figures

Figures in appendices

List of tables

Tables in Appendices

Abbreviations

kb - kilobase
DNA - deoxyribonucleic acid
RNA - ribonucleic acid
GFP - green fluorescent protein
PEP - plastid-encoded RNA polymerase
NEP - nucleus-encoded RNA polymerase
PPR - pentatricopeptide repeat protein
atpH - ATP synthase subunit C
LSC - Large single copy (plastome region)
SSC - Small single copy (plastome region)
IR - Inverted repeat (plastome region)
IEE - intercistronic expression element
PEG - polyethylene glycol
HR - homologous region
UTR - untranslated region
prrn - plastid rRNA operon promoter
TSP - total soluble protein
Rubisco - ribulose bisphosphate carboxylase oxygenase
rbcL - Rubisco large chain precursor
bp - base pairs
Nt - *Nicotiana tobacum* genetic parts
Mp - *Marchantia polymorpha* genetic parts
PCR - polymerase chain reaction
Tak - *Marchantia polymorpha* Takaragaike accession
Cam - *Marchantia polymorpha* Cambridge accession
Kit - *Marchantia polymorpha* Kitashirakawa accession
NEB - New England Biolabs
g - (Earth surface) gravity, unofficial unit of acceleration equal to 9.81 m s^{-2}
LAS AF - Leica Application Suite Advanced Fluorescence software
ROI - region of interest
CLAHE **-** contrast limited adaptive histogram equalisation algorithm
LSCM - laser scanning confocal microscope
Ti - tumor inducing
CI - confidence interval

Chapter 1: Introduction

The focus of this thesis is the development and improvement of techniques for directed modifications to the ~120 kb genome present within plant plastids, semi-autonomous intracellular organelles descended from free-living cyanobacteria. Plastids perform photosynthesis in plants and algae, and are therefore the foundation of nearly all terrestrial ecosystems. The plastome (plastid genome) encodes its own transcriptional and translational apparatus, as well as many of the genes essential for photosynthesis. The small size and bacteria-like properties of the plastome make it an interesting candidate for engineering, with potential applications in crop improvement and bioproduction. The development of improved plastid transformation techniques is here carried out in the model plant *Marchantia polymorpha*.

1.1 Plastids

Plastids are subcellular organelles mainly found in plant and algal lineages. They are descended from photosynthetic bacteria which were engulfed by early eukaryotes. They replicate independently and possess their own prokaryote-like genome, transcriptional, and translational machinery, each of which are retained in most plastid lineages but have been lost in certain clades.

1.1.1 Evolution

The presence of plastids is the defining ancestral characteristic of Archaeplastida, though some lineages have subsequently lost their plastids [1]. A substantial body of evidence points to monophyly of the extant Archaeplastida plastid lineages and thus a single ancient endosymbiotic event [2–4] but this remains in dispute [5,6]. Whether it was a single event or several, this ancient endosymbiosis occurred approximately 1.5 billion years ago [4]. This endosymbiosis event is called a primary endosymbiosis because a prokaryote was engulfed, generating double membraned plastid [3]. Secondary and higher-order endosymbioses involve eukaryotes being engulfed leaving three or more membranes structures, and have occurred repeatedly in certain algal lineages (reviewed in [7]).

The identity of the engulfed cyanobacteria in this endosymbiosis remains in dispute due to the erosion of relevant evidence over the extremely long timescales involved, but it is widely agreed to have been in the same lineage as the model species *Synechococcus elongatus [8].* The eukaryotic host in the ancient plastid endosymbiosis already contained mitochondria and a nuclear membrane as well as two flagella, the defining ancestral characteristic of the Bikonts [9]. Rhizaria and Excavata are additional Bikont lineages which later acquired plastids through separate endosymbiotic events. A modern primary endosymbiosis has been detected in the Rhizarian ameboid *Paulinella chromatophora* [10]. This endosymbiosis is more recent, between 90 and 140 million years ago, and has allowed for study of the early stages of endosymbiosis [11]. Among Excavata, many euglena lineages contain photosynthetic plastids. However, these were not obtained through a primary endosymbiosis, but rather through a secondary symbiosis wherein a nonphotosynthetic ancestral member of clade Excavata engulfed an ancient green alga, resulting in complex plastids with three membranes [12]. This surprising finding explains the presence of plastids Excavata and plant-like features in the classic model species *Euglena gracilis* [13]. A partial phylogeny of plastids demonstrating endosymbiotic events is shown in Figure 1.

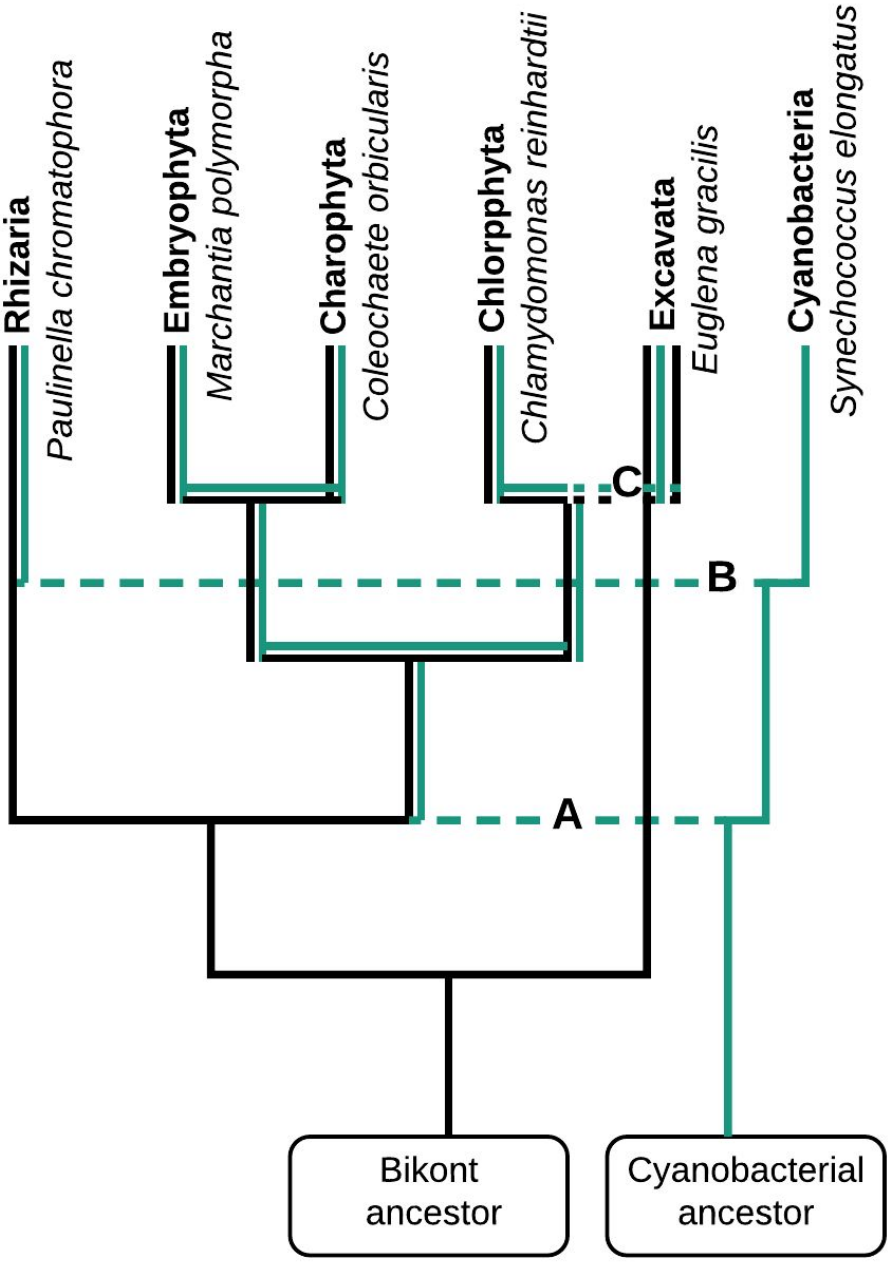


Figure 1: Partial phylogeny showing plastid endosymbiosis events. **A** indicates the ancient primary endosymbiosis ~1.5 billion years ago which defines the *Archaeplastida*. **B** indicates a more recent primary endosymbiosis in *Paulinella chromatophora*. **C** indicates a secondary endosymbiosis of chlorophyte algae into euglena. Branch lengths are not proportional.

Three major algal lineages contain plastids nearly without exception, chlorophytes, charophytes, and embryophytes. Chlorophytes are a clade of green algae containing unicellular and multicellular species including the model species *Chlamydomonas reinhardtii,* which has been instrumental in understanding plastid biology (reviewed in [14]). Charophytes are freshwater algae often studied in an evolutionary context as the closest relatives to land plants. This clade includes the model species *Coleochaete orbicularis*. Embryophytes are the land plants, and include *Marchantia polymorpha*, the model species used in this thesis.

1.1.2 Plastome architecture

The plastomes of photosynthetic land plants are 100 - 200 kb, whereas closely related free-living cyanobacteria, such as *Synechococcus elongatus* or *Synechocystis sp.* PCC6808, have genomes of 2-5 Megabase pairs [15,16]. As plastids have been "domesticated" from free-living bacteria to intracellular organelles, the plastome has shrunk considerably, mostly as a result of the migration of genes to the nuclear genome [17].

The plastome is well conserved among green algae and land plants [18]. In most instances the plastome maps as a single circular chromosome [19], though in some cases additional smaller circular DNA molecules are present [20]. Furthermore, some evidence suggests that plastome structure is linear *in vivo*, at least for some plastome copies [21]. From the several thousand genes originally present in the endosymbiont, land plant plastomes have been reduced to 100-130 genes [19], and smaller plastomes have been found among non-photosynthetic plants. The plastome of the parasitic plant *Epifagus virginiana* has shrunk to 72 kb and contains only 42 genes [22]. *Polytomella*, a clade of nonphotosynthetic algae, retain plastids as a compartment but have entirely lost the plastome [23]. Plastid genes are mostly organised into prokaryote-like operons [24].

The plastome is classically comprised of a tetrapartite structure of two inverted repeats separating two single copy regions. The inverted repeats range between 10 and 25 kilobases and code primarily for rRNA [25]. These repeats are separated by the large and small single copy regions (LSC and SSC, respectively) [25]. Notable exceptions to this genome organisation include legumes, many of

which have only a single copy of the inverted repeat [26], some *Coniferophyta* which have lost a single copy of the IR but generated short (200-500 bp) inverted repeats [27], and the genus *Erodium,* which lost one copy of the IR but subsequently regained it in some lineages [28].

1.1.3 Transcription

Plastid transcription occurs through two independent and overlapping systems, based on the eukaryotic nuclear-encoded polymerase (NEP) complex and the prokaryote-like plastid-encoded polymerase (PEP) complex [29]. These two polymerase complexes are recruited by different classes of plastid promoters, yet many genes can be transcribed by both [30]. The genes coding for NEP subunits are transcribed in the nucleus and translated in cytoplasm before being imported into the chloroplast as proteins and self-assembling into the functional complex. In a mutant barley (*Hordeum vulgare*) line deficient in plastid ribosomes and thus in PEP, many plastid genes ordinarily transcribed by PEP retained expression due to compensatory transcription by NEP [30]. The same study used differential RNAseq to distinguish between primary and processed plastid transcripts, revealing 244 transcriptional start sites, more than double the 113 genes in the plastome. This indicates many genes have multiple transcriptional start sites, presumably each with a separate promoter. These results validate an earlier observation that the transcriptional start site for the *atpB-E* operon is shifted further upstream under PEP-deficient conditions, a finding explained by compensatory transcription by NEP from a different promoter [31].

The promoter sequences recognised by NEP and PEP reflect their separate origins. NEP resembles the RNA polymerase present in plant mitochondria, and binds to promoters with the sequence motif 5'-YRTA-3' near the site of transcription initiation [32]. PEP resembles the eubacterial σ^{70} enzyme complex, and binds to promoters containing -10 (TATAAT) and -35 (TGACA) regions [33].

1.1.4 RNA Processing

Much like the transcription apparatus, mRNA processing and stability in plastids show characteristics of both classically prokaryotic and eukaryotic mechanisms [34]. One of the early surprises in plastid research was the prevalence of introns,

which are extremely rare in prokaryotic genomes [35]. The first plastid introns were discovered in *Chlamydomonas reinhardtii* [36], and land plant plastomes are now known to possess over 20 group I and group II introns [34]. Both of these intron groups are mobile genetic elements with autocatalytic splicing ability, meaning they can spontaneously splice out of mRNA before translation. This allows them to propagate with minimal interference to the host genome, facilitating their spread [37]. Group I introns require only guanosine triphosphate and magnesium in solution to complete splicing, whereas group II introns use the hydroxyl from an internal adenine, and thus are fully autocatalytic under appropriate salt and pH conditions [38]. However, under physiological conditions within the plastid, kinetic traps prevent efficient folding of the autocatalytic ribozyme, necessitating intron-encoded maturases as a cofactor for splicing [39]. The mobile nature of group II introns is dependent on a multifunctional intron-encoded reverse transcriptase/maturase which allows for RNA lariats spliced out of the host mRNA to re-integrate at similar sites in the genome [40]. In land plant plastomes, most open reading frames within introns have degenerated, reducing intron mobility through elimination of reverse transcriptase function and necessitating nuclear-encoded splicing cofactors for intron excision [41].

Distinct sets of introns exist in chlorophyta and streptophyta plastomes [25], consistent with the idea that the ancestral plastome had few introns [42]. A single group I intron in the leucine UAA- anticodon tRNA gene has been found to be conserved in extant cyanobacteria and several algal plastomes, as well as *Nicotiana tabacum* [43]. This intron is also present in the *Marchantia* plastome. This intron has been identified as an essential ncRNA for autotrophic growth in *Synechococcus elongatus* [44], demonstrating that in this case the autocatalytic splicing is intact and the tRNA gene is required for plastid translation.

The endosymbiotic theory for the origin of mitochondria and plastids was first suggested by Konstantin Mereschkowski in 1905, and more fully developed and argued in 1967 by Lynn Margulis, at which time the bacterial nature of the plastid ribosome was among the strongest evidence available for the theory [45]. In the years since, the endosymbiotic theory has accumulated overwhelming evidence, and is widely accepted (reviewed in [46]). More recently, additional prokaryote-like aspects of plastid transcription have been uncovered, including polycistronic

transcription and translation [47]. The presence of polyribosomes on plastid mRNA shows polycistronic translation can occur, but in many cases RNA processing into monocistronic transcripts has been shown to increase translation efficiency [48] or be required for translation [49]. Fei Zhou *et al.* recently discovered several RNA sequences necessary and sufficient for post-transcriptional processing into monocistronic transcripts, dubbed intercistronic expression elements (IEEs). These sequences allow multiple transgenes to be efficiently expressed from synthetic operons [50].

Post transcriptional single-base RNA editing is widespread in plant organelles. *Arabidopsis thaliana* has 43 edit sites in the plastid transcriptome, *Anthoceros angustus* has 942 [51]. RNA editing has been discovered in the plastomes of most plants (reviewed in [52]), but editing has not been reported in *Marchantia* despite extensive searches [53]. Nuclear encoded pentatricopeptide repeat (PPR) proteins bind to plastid mRNA in a sequence specific manner and recruit as-yet-unidentified RNA editing enzymes, which deaminate cytosine (C) to uridine (U) [54]. The reverse U to C editing has also been reported at lower frequency, and never among flowering plants [55]. The lack of RNA editing makes *Marchantia* more tractable as an engineering chassis.

1.1.5 Regulation

In modern plants, the majority of plastidial proteins are encoded by nuclear genes and imported to the plastid after translation [56]. The interconnected nature of protein pathways necessitates tight coordination of expression between the nuclear genome and plastome, which is accomplished through anterograde (nucleus to plastid) and retrograde (plastid to nucleus) signaling [57]. Much remains to be elucidated about the details of these signalling pathways, but it is instructive to discuss PPR proteins as an example of anterograde signalling.

PPR proteins are a large family of RNA binding proteins that has radiated within land plants - the family has single digit gene numbers in Animalia, 12 in *Chlamydomonas reinhardtii,* 103 in *Physcomitrella patens*, and 400-600 in angiosperms [58]. Most are associated with plastids or mitochondria. The sequence-specific RNA binding capacity of these proteins makes them an

excellent candidate for RNA degradation and translational regulation, a function that has been confirmed for many specific cases [59]. For example, PPR10 has been found to be necessary for chloroplast development. Its role is controlling translation of ATP synthase subunit C (atpH) transcripts by defining and stabilising the termini, protecting them from exonucleases [60]. This allows for nuclear regulation of plastid protein expression, primarily at the post-transcriptional level [57]. Coordination of protein expression between these two genomes is essential for maintaining homeostasis under changing environmental conditions.

Nucleus-encoded factors and environmental conditions cause proplastids to differentiate into several distinct types dedicated to particular biochemical processes (reviewed in [61]). Fully differentiated types include chloroplasts which carry out photosynthesis, etioplasts which are present in light-starved tissue and seeds, amyloplasts which store starch in tubers, and chromoplasts which are responsible for colouration in many fruits [62]. Particular plastid types are often localised within certain plant tissues [63], though some tissues contain multiple plastid types, and plastids can transition from one differentiated form to another [64].

1.2 *Marchantia polymorpha*

Most research in plant engineering has focused on angiosperms due to their agricultural relevance and overall familiarity. However, even the simplest angiosperm models pose challenges, such as several month generation time, highly redundant genetics, and adult size >40 cm tall, making highly parallel tissue culture and prototyping difficult. To address these challenges, many research laboratories are turning to simpler model species such as bryophytes. [65–67].

Bryophytes comprise three phyla of nonvascular plants united by the presence of an unbranched sporophyte with a single sporangium and a haploid-dominant life cycle: the mosses (Bryophyta), hornworts (Anthocerotophyta), and liverworts (Marchantiophyta). These lineages represent extant relatives of the earliest land plants, and have a stronger morphological resemblance to the early plant fossils

than angiosperms [68]. Figure 2 shows the location of bryophyte lineages within the broader context of green algal and land plant clades.

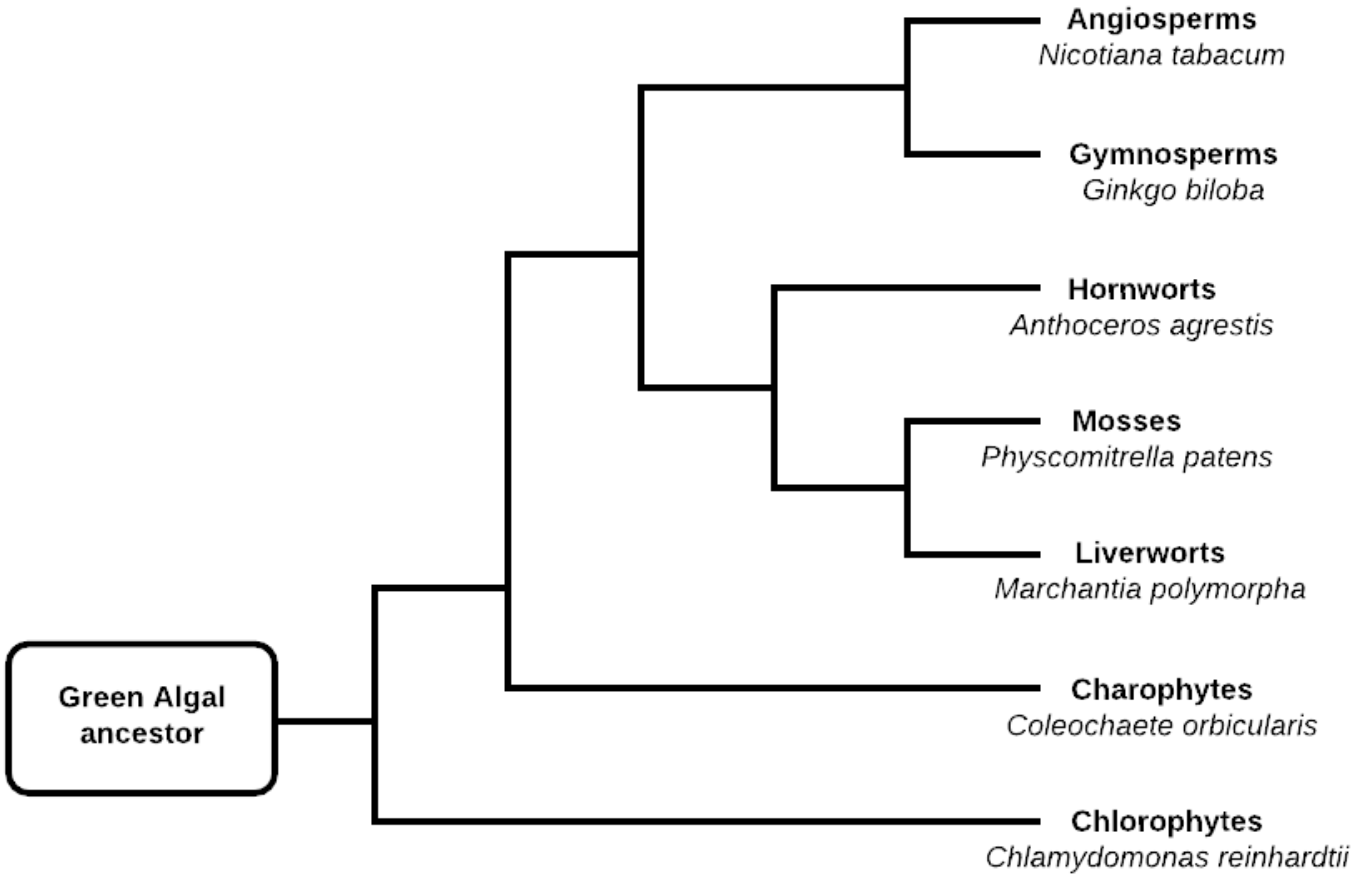


Figure 2: Partial phylogeny of green algae. Relationships as reported by Puttick *et al*. [69]. The position of hornworts is least certain, but is strongly supported by their model. Branch lengths are not proportional.

The evolutionary relationships among bryophytes have been repeatedly revised throughout the last two decades (reviewed in [70]). The precise topology is not central to this thesis, but this author is convinced by the systematic analysis published by Puttick *et al*. in March 2018 [69]. In their model, all bryophytes share a common ancestor which has already diverged from the tracheophyte lineage leading to gymnosperms and angiosperms. In this model, all three groups are equally basal with respect to land plant evolution. This model contrasts with the widely-held though contentious claim that liverworts are the most basal lineage,

with hornworts and mosses diverging later (as argued in [71,72]). Regardless of the exact topology, bryophytes represent early-diverging lineages in the evolution of land plants.

Figure 3: *Marchantia polymorpha*, image courtesy Jim Haseloff. Note the circular gemma cups projecting above the thallus. Within these are gemma, discoidal clonal reproductive propagules.

Marchantia polymorpha is a common and widely distributed thalloid liverwort that has been studied for nearly 200 years (eg [73–75]). It is a classical example of the liverworts, a basal group in the evolution of land plants. Figure 3 shows mature thallus tissue and gemma cups of *Marchantia* grown in laboratory conditions. The gemma cups are approximately 2 mm in diameter.

1.2.1 Use as a model system

Advantages as a model species

The extensive literature from classical experiments generated interest in using *Marchantia* as a modern model species for molecular biology. The 2 cm ● 2 cm ● 1 cm mature gametophyte size and 1-3 month generation time of these plants make them amenable to highly parallel tissue culture. Asexual reproduction through gemmae propagation provides a simple way to multiply and preserve a particular genotype, and can be accomplished in standard Petri dishes. Gemmae are small disk-shaped propagules whose growth can be visualised under confocal microscopy for several days, enabling studies of plant developmental biology. They are produced in a cup-shaped structure which develops out of the epithelium of adult thallus tissue. The induction of sexual organs is reliably obtained through a 16:8 light:dark photoperiod and exposure to far-red light [76], which can be achieved in axenic conditions in pots just 15 cm high. A single sexual cross generates millions of spores, allowing for rapid generation of phenotypic diversity and amplification of biological material. Despite these physiological, taxonomic, and morphological differences between *Marchantia* and flowering plants, the plastomes are similar. The majority of plastome genes are shared between most

photosynthetic plants, and the intergenic regions are sufficiently similar that homologous regions cloned from one species have successfully been used to transform plastomes of other species [77].

Tools available

Marchantia polymorpha was reported to be the first land plant with a published plastome [78], just weeks ahead of *Nicotiana tobacum* [79]. However, that sequence is now known to belong to *M. paleacea*, a morphologically and genetically similar plant. More recently, the nuclear genome has been published [80]. *Agrobacterium*-mediated transformation of the nuclear genome is routine [65], and faster and more efficient protocols have recently become available [81]. Nuclear transformation has also been demonstrated through biolistic delivery [82]. Plastid transformation has been established for suspension-cultured cells [83] and sporelings [84].

Marchantia tissue culture is over 100 years old [85], and has been optimised throughout the last 80 years [86,87]. *Marchantia* research has also benefited from the development of standardised conditions for liquid suspension cultures [87] as an alternative to growth on solid media. The combination of these two, in addition to *Marchantia*'s small size, allow for high-throughput parallel growth in axenic conditions. Vegetative propagation through thallus fragments or gemmae is possible with plants as small as $1 \cdot 1 \cdot 0.5$ cm, and is easily accomplished in standard 100 mm Petri dishes on simple nutrient media solidified with agar. Long term cryopreservation techniques originally developed by Takeuchi *et al.* [88] and improved and expanded by Tanaka *et al.* [89] allow for long-term maintenance and sharing of mutant plants.

1.3 Plastome engineering

Genetic manipulations of plants has accelerated scientific understanding of their basic biology and has facilitated the development of improved crop cultivars. The majority of genetic manipulations have been performed on the nuclear genome, which is more experimentally facile to modify. Plastome modification was first demonstrated in *Chlamydomonas reinhardtii* and *Nicotiana tabacum* three decades ago, and within the last decade plastid transformation protocols have

become available for a diverse group of plants, including crop species. Plastome engineering offers several benefits compared to traditional nuclear genetic engineering (discussed in section 1.3.2), and has been used on its own or in combination with nuclear genetic engineering for a variety of applications (discussed in section 1.3.5).

1.3.1 Basic principle

Plastid transformation consists of four fundamental steps: transformation vector assembly, DNA delivery, transgene integration, and selection to homoplasmy. The defining characteristic of plastid transformation vectors is the presence of two DNA regions which are homologous to the plastome of the plant to be transformed. Genes of interest are cloned between these two homologous regions, and can be integrated at the corresponding locus in the plastome. Vector design is discussed in section 1.3.3. DNA delivery is accomplished via a variety of techniques discussed in section 1.3.4, and in principle only requires that a single copy of the transformation vector reach the stroma of a single plastid within a plant cell. Transgene integration is mediated by homologous recombination, and hence occurs in a particular genetic locus determined by the sequence of the transformation vector. Most plastid transformation vectors include the selectable marker *aadA* between the two targeting homologous regions. The antibiotic spectinomycin can then be used to eliminate untransformed cells and confer a selective advantage to those plastids which have been transformed. Over a period of months, untransformed plastomes will be eliminated by selection and segregation until only the transformed plastomes remain, a state termed homoplasmy. *Marchantia* selection conditions are explored and discussed in section 4.5.

1.3.2 Comparison to nuclear genome engineering

The most widely used DNA delivery method for plant nuclear genome transformation is *Agrobacterium* species such as *A. tumefaciens, A. rhizogenes,* and *A. vitis.* [90,91]. This method is based on the tumor inducing (Ti) plasmid, which has a natural ability to integrate into the nuclear genome of plants. The gene of interest is cloned into the Ti plasmid along with a selectable marker, and this plasmid is transformed into a chosen strain of *Agrobacterium*. Plants are then

infected with *Agrobacterium*, allowing for the Ti-mediated transformation to take place (reviewed in [92]). This technique is versatile, rapid (1-3 hours of labor per construct), reliable, and low cost. However, transgene integration occurs in unpredictable locations and variable copy numbers. Biolistic transformation of the plant nuclear genome is also well established [93], but the problem of complex and unpredictable genomic integration remains [94]. The same obstacle affects more niche transformation methods such as electroporation, vacuum infiltration, silicon carbide fibres, laser microbeams, microinjection, and electrophoresis [95]. This random integration leads to positional effects, gene silencing, and large variance in expression for the same transgene construct. By contrast, plastome transformation is mediated by homologous recombination, so integration copy number can be controlled (usually one or two) and positional effects can be used as a tool for modulation of expression levels rather than causing uncontrolled variance in transformant phenotype [96].

The plastome is present in 100 - 50,000 copies per cell [97], so plastome transgenes have a much higher maximum transcription level than nuclear genome transgenes, despite lower copy number per genome. Nuclear expression of foreign proteins rarely exceeds 1% total soluble protein (TSP) [98,99], whereas plastome expression has repeatedly exceeded 5% TSP [100,101], and at the extreme has exceeded 70% TSP [102]. These proteins are also intrinsically localised within a biochemical powerhouse with access to a pool of biosynthetic precursor molecules, making plastome transformation particularly attractive for metabolic engineering [103].

In most crop species, the plastome inheritance is uniparental, usually through the maternal lineage [104,105], though this trend is not universal [106,107]. In species with uniparental maternal inheritance, accidental parental inheritance has been reported but only at frequencies less than one in ten thousand under strict laboratory screening conditions where paternally inherited plastomes would a substantial selective advantage [108]. Hence, transgenes in the plastome are less likely to escape to wild relatives of crop species than those in the nuclear genome [109]. This partially mitigates a major criticism of the agricultural use of transgenic plants. However, stacking this containment method with methods of containment may still be desirable for some transgenic traits likely to confer a selective

advantage, such as herbicide resistance, drought tolerance, or improvements to photosynthesis [110].

1.3.3 Vector design principles

In the three decades since the first successful transformation of land plant plastomes, substantial progress has been made in optimising the design of vectors for foreign gene expression. The basic structure of the vector is two homologous regions which flank the genes to be inserted. Homologous regions are usually 1-3 kb, and must match the sequence of the plastome to be transformed [96]. This design is to facilitate the integration of transgenes through homologous recombination. In most circumstances, the adenyltransferase gene *aadA* is included between the homologous regions, which allows for selection with spectinomycin and streptomycin, the most widely used antibiotics for plastome engineering [111]. Any other genes to be inserted, such as fluorescent markers, are adjacent, and often oriented in the opposite direction to *aadA.* The vector backbone contains a bacterial origin of replication for easy amplification, usually in *E. coli*. Most published vectors also include a bacteria-optimised antibiotic resistance cassette in the backbone for ease of plasmid maintenance in bacteria. Some studies have included a unique restriction enzyme site on one or both sides of the backbone to allow for easy linearisation, including previous plastome transformation experiments in *Marchantia polymorpha* [112]. However, optimisation studies have not shown an improvement in transformation efficiency from linearisation [113].

Several additional selection systems have been implemented, including use of neomycin phosphotransferase (*NPTII*) in conjunction with kanamycin [114] and a mutant anthranilate synthase α-subunit gene (*ASA2*) conferring insensitivity to synthetic indole analogues [115]. A variety of promoters and 5’ and 3’ untranslated regions (UTRs) have been tested for expression levels (reviewed in [116]). The most commonly used promoter is from the plastid rRNA operon (*prrn*). A substantial amount of the reported variation in protein accumulation levels can be explained by the identity of the UTRs, which confer stability and protection from exonucleases by RNA hairpin structures in addition to providing binding sites for stabilising and translation-enhancing PPR proteins [117]. The ribulose

bisphosphate carboxylase oxygenase (Rubisco) large chain precursor (*rbcL*) 3' UTR has been frequently employed to stabilise foreign transcripts, and has been shown to generate mRNA accumulation levels ~50% higher than the *psbA* 3' UTR, the most commonly used alternative [118].

1.3.4 DNA Delivery Techniques

Plastome transformation was first achieved in *Chlamydomonas* by JE Boynton *et al*. using biolistic delivery in 1988 [119]. This was quickly followed by successful transformation of *Nicotiana tabacum,* which remains the most widely used land plant model species in transplastomic studies [120,121]. Since then, several new techniques have been developed [122].

Biolistics, also known as particle bombardment or the "Gene Gun" approach, remains the most widely used technique [123]. This method, once optimised, has an efficiency of 1-10 transformants per Petri dish, and has proven to function in the largest variety of species, including crops such as rice [124], tomato [125], potato [126], eggplant [127], soybean [128], and wheat [129]. Recently an efficient transformation protocol was reported for the model species *Arabidopsis thaliana [130]*. The technique has the additional benefit of requiring few specialised skills and having a low marginal cost once the required equipment purchases have been made. However, the purchase of a biolistic device such as the PDS 1000/He represents a substantial capital investment.

Alternative methods of plastid transformation require removal of cell walls and the generation of 'naked cells' called protoplasts through enzymatic digestion. Protoplasts are contained only by a plasma membrane, and thus are extremely fragile to physical agitation and osmotic stress. For some model species, protocols have been developed allowing for the regeneration of plants from protoplasts, but for many plants this process is laborious or has not been developed (reviewed in [131]).

Polyethylene Glycol (PEG)-mediated transformation of protoplast plastids was first demonstrated by Timothy Golds *et al.* [132] and independently by Carmel O'Neill *et al*.[133] in 1993. It is routine in tobacco (*Nicotiana tabacum)* [134], and has been

demonstrated in lettuce (*Lactuca sativa*) [135], *Physcomitrella patens [136]*, and cauliflower (*Brassica oleracea*) [137]. PEG-mediated transformation offers the benefit of substantially simpler equipment requirements at the cost of longer regeneration time and a need for more advanced tissue culture to recover plants from protoplasts.

Transient expression of foreign genes has been accomplished through a highly specialised form of microinjection [138], but this technique has not been widely adopted due to costly equipment requirements and the specialised technical expertise needed for subcellular microinjection. Agitation in the presence of glass beads and DNA has been demonstrated for cell wall-deficient mutants of *Chlamydomonas [139]*, but the low efficiency and incompatibility with cell walls has limited the application of this technique.

1.3.5 Applications of transplastomic plants

High expression levels, controlled insertion location, and genetic containment have made the plastome a promising target for modification of agronomic traits, large-scale protein production, and metabolic engineering. The two most widespread traits in genetically modified crop plants are glyphosate resistance, which allows for in-crop use of a potent herbicide, and Bt Cry toxins, which are a family of highly selective protein insecticides from *Bacillus thuringiensis* (reviewed in [140]). To date, all commercially available genetically modified crops are nuclear-genome modified, but glyphosate resistance [141] and Bt toxin [142] have also been expressed from the plastome. Both of these traits warrant containment efforts because they confer a selective advantage to recipient plants and should not be allowed to escape into wild plants. Bt expression also benefited substantially from the high expression levels possible through plastome engineering - the plants generated contained Cry2Aa2 levels lethal even to Bt-resistant insects with toxin tolerance more than 20,000-fold higher than wild type [142]. The high expression levels made possible by plastid expression should help delay the development of resistance, lengthening the useful life of these agronomic tools [143].

Plants have often been touted as potential bioreactors for large-scale protein production, but low expression levels have constrained application (reviewed in [144]). They offer the potential for massive scaling, folding and post-translational modification conditions much better than microbes, and an existing production infrastructure. As such, plants have always been a promising candidate for edible vaccine production offering scale, low cost, and decentralised distribution of vaccines (reviewed in [145]). Plastid expression of the *Vibrio cholerae* toxin B subunit as a vaccine antigens reached over 4% TSP in tobacco, a high-yield crop suitable for large-scale cultivation [146]. Recently, a plastid-expressed vaccine for *Bacillus anthracis* (anthrax) toxin provided immunoprotection in a mouse study [147].

Metabolic engineering is the usage of biotechnology to manipulate the flux and accumulation of metabolites, usually through modulating the expression levels of relevant enzymes in synthesis pathways [148]. Recent major successes include the complete transfer of the opioid synthesis pathway from opium poppy (*Papaver somniferum)* to both yeast (*Saccharomyces cerevisiae*) [149] and *Escherichia coli [150]*, and the generation of large amounts of β-carotene in Golden Rice [151]. Successful plastome engineering projects have included β-carotene biofortification [103] and expression of astaxanthin, a high-value nutraceutical [152]. Plants are a highly desirable production chassis for these products due to the cheap inputs and simple scaling, and plastome engineering allowed for higher level production - in the case of astaxanthin, approximately 25 times higher than the best performing nuclear metabolic engineering biosynthesis plants [153].

1.3.6 Basic research

Elimination of a gene product and phenotypic study of the resulting organism is among the most common study designs in molecular biology. The most direct method to ‘knockout’ a gene is to disrupt the coding sequence. Plastid transformation generating knockout mutations is often used to study photosynthesis and other processes dependant on plastid-encoded gene products [154]. For example, the plastid gene *clpP1*, encoding the ClpP1 protease subunit, has been shown to be essential for shoot development [155]. This finding led to the discovery that ClpP1 regulates isoprenoid levels in a plastid-specific manner

[156]. Other plastid-encoded proteins have been shown to be essential for leaf development [157], or normal morphology [158].

1.4 Project Aims

1.4.1 *Marchantia polymorpha* Cam plastome sequencing

The Cam strain of *Marchantia polymorpha* is used for the majority of plant engineering projects in the Haseloff laboratory. However, no plastome sequence assembly exists. As a result, previous experiments with plastid transformation have used homologous regions clones from a different strain. The use of homologous regions not entirely homologous to the plastome to be transformed is problematic for two reasons: lowered efficiency and introduction of mutations. Lowered efficiency is expected because the homologous recombination process is known to be highly sensitive to even small regions of non-homology, which reduce the probability of recombination [159]. Introduction of mutations is especially troublesome due to the gene-dense nature of the plastome. The homologous regions used by Boehm *et al*. [112] contain a deletion which could cause a frameshift mutation in the nearby gene *psbC*. The plants were never sequenced, leaving open the possibility that this mutation was accidentally induced. The generation of a Cam plastome sequence will render the use of genetic parts cloned from different strains unnecessary. The Cam-1/2 plastome sequence has been submitted to Genbank for public use (accession MH635409).

1.4.2 Optimisation of plastid transformation

Previous projects involving plastid transformation in *Marchantia* have required the bombardment of a large number of samples to generate a small number of transformed plants. A more efficient transformation protocol would enable the screening of larger scale libraries such as those which can be assembled through combinatorial Type IIS assembly. An improved protocol for the generation and selection of transplastomic *Marchantia* will be developed through optimisation of microcarrier choice, bombardment distance, and selection conditions.

1.4.3 Generation and characterisation of plastid genetic parts

One integration locus has been demonstrated in transplastomic *Marchantia*, but more are required for super-transformation experiments of complex metabolic engineering. The requisite genetic parts for construction of transformation vectors targeted to different loci in the plastome will be developed. A fluorescent marker has been expressed previously from the *Marchantia* plastome, but only through a single construct. Here a series of transformation constructs containing different regulatory elements will be assembled, transformed into *Marchantia*, and compared with quantitative confocal microscopy.

Chapter 2: Materials and Methods

2.1 Genome Annotation

Access to a high-quality next generation sequencing Illumina dataset was generously provided by Bernardo Pollack. All data analysis was performed with CLC Genomics Workbench version 11.0. The original dataset included 72,519,716 100 bp single-end sequences from nuclear, mitochondrial, and plastid DNA. Reads were mapped to a reference genome provided by John Bowman, (Tak-1, personal communication) with strict parameters (required match length 75%, mismatch penalty = 2, gap penalty = 5). Nonspecific matches were mapped randomly to allow for good coverage of the inverted repeats. This resulted in a subset of 100,183,190 sequences, spanning the reference assembly at an average read depth of 83,480. Duplicate reads were eliminated, resulting in 9,432 unique sequences spanning the plastome. This subset was assembled in two ways, first by *de novo* assembly, then by reference-based assembly with the Tak-1 assembly. *De novo* assembly produced 25 contigs, which were aligned against the reference-based assembly to check for biases from reference use.
Reference-based assembly was used to assemble the contigs and the resulting assembly was validated by Sanger sequencing, and comparison to the newly published Kit-2 assembly (Accession NC_037507) as described in section 3.5. Annotation cleaning for NCBI submission standardisation was performed in Geneious version 11, and the annotated assembly was designated Cam-1/2 and submitted to Genbank (Accession MH635409).

2.2 *Escherichia coli* methods

E. coli strain TOP10 (Genotype F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG) was used for all experiments.

2.2.1 Culturing Conditions

Medium

LB media (Sigma-Aldrich) was made into solution as according to manufacturer's instructions. Solid media Petri dishes were made by adding 1.2% (w/v) agar (Melford Biolaboratories Ltd).

Antibiotics

The following were used for selection and plasmid maintenance.

- 25 μg mL^{-1} chloramphenicol (Duchefa Biochemie)
- 50 μg mL^{-1} kanamycin (Sigma-Aldrich)
- 100 μg mL^{-1} spectinomycin (Sigma-Aldrich)

2.2.2 Preparation of chemically competent cells

SOB Medium

- 0.5% (w/v) yeast extract
- 2% (w/v) tryptone
- 10 mM NaCl
- 2.5 mM KCl
- 20 mM $MgSO_4$
- Add 1.2% (w/v) agar for solid media Petri dishes

CCMB80 Buffer [160]

- 10 mM KOAc (potassium acetate)
- 80 mM $CaCl_2$
- 20 mM $MnCl_2$
- 10% (v/v) glycerol
- Made fresh day of use

Bacterial seed stock preparation

1. TOP10 cells were streaked onto a SOB Petri dish and grown overnight
2. A single colony was used to inoculate 2 mL SOB liquid media and incubated at 23°C 130 RPM overnight.
3. 500 μL cells and 500 μL 50% (v/v) glycerol were mixed in pre-chilled 1.5 mL microcentrifuge tubes, which were stored at -80°C until needed

Competent Cell Preparation

1. 1 mL of seed stock was used to inoculate 10 mL LB in a 50 mL centrifuge tube and incubated overnight at 37°C 130 RPM
2. 5 mL of the overnight culture was used to inoculate 500 mL SOB and grown at 20°C 200 RPM

3. Cells were grown until they reached OD_{600nm}= 0.2, split into 50 mL aliquots, and centrifuged at 4°C 1500 g for 10 minutes
4. -80°C freezer racks were filled with 1.5 mL microcentrifuge tubes and chilled in a -20°C freezer
5. Keeping pellets on ice, 3 mL ice cold CCMB80 buffer was added per 50 mL pellet and gently resuspended
6. Cells were incubated on ice a further 30 minutes, then combined into one 50 mL centrifuge tube and centrifuged at 4°C 15000 g for 10 minutes, then resuspended in 20 mL CCMB80 and incubated on ice a further 10 minutes
7. Bacteria were aliquoted into prechilled 1.5 mL microcentrifuge tubes on ice and flash frozen with liquid nitrogen, then stored at -80°C until needed

2.2.3 Transformation of chemically competent cells

Previously prepared TOP10 competent cells were removed from -80°C freezer and thawed on ice, then aliquoted into 1.5 mL microcentrifuge tubes, 25 µL per tube. 3 µL plasmid DNA were mixed into bacteria and allowed to incubate on ice for 5 minutes. Transformation was accomplished by heat shock at 42° C for 30-60 seconds, followed by 2 minutes incubation on ice. 500 µL LB broth was added and the bacteria were allowed to grow for 1-3 hours at 37° C in an incubator shaking at 130 RPM. 100 µL of liquid bacteria was then spread onto agar Petri dishes containing the appropriate antibiotic to select for the plasmid of interest and grown overnight at 37° C.

2.2.4 Plasmid DNA isolation

Plasmids were purified from bacteria using the QIAprep spin miniprep kit according to manufacturer's instructions. Protocol was partially automated through the use of a Qiacube robot, an automated benchtop platform for molecular biology reactions (both products from Qiagen, USA).

2.2.5 Sanger Sequencing

Sanger sequencing was carried out by Genewiz UK. DNA samples were prepared in water and the primer of interest was added. Plasmid DNA concentration was 30-50 ng μL^{-1}, PCR product concentration was determined by the formula below.

$$\frac{length\ (bp)}{750} = concentration\ needed\ (\frac{ng}{\mu L})$$

In all cases, a single primer at concentration 1.7 µM was included, and 15 µL of sample was sent for analysis.

2.3 DNA assembly

Plasmids in this work were assembled with Loop Assembly, a recently developed Type IIS DNA assembly standard developed by Bernardo Pollack and Fernan Federici [161]. Loop Assembly is discussed further in section 5.1.2.

2.3.1 Overview and credit

Loop assembly alternates between two sets of enzymes, buffers, and acceptor plasmids. These levels are designated as even and odd. In general, level 0 parts contain parts smaller than a transcriptional unit, level 1 parts contain a single transcriptional unit, and level 2 parts contain four transcriptional units.

2.3.2 Buffers and thermocycling conditions

Even Level Master Mix (1 reaction)

- 3 µL nuclease-free H_2O
- 1 µL 10x T4 DNA Ligase buffer (New England Biolabs (NEB))
- 0.5 µL 1 mg mL^{-1} Bovine Serum Albumin (NEB)
- 0.25 µL T4 DNA Ligase at 400 U μL^{-1} (NEB)
- 0.25 µL BsaI at 10 U μL^{-1} (NEB)

Odd Level Master Mix (1 reaction)

- 2 µL nuclease-free H_20
- 1 µL 10x Tango Buffer (Thermo Fisher)
- 0.5 µL mg mL^{-1} Bovine Serum Albumin (NEB)
- 0.25 µL T4 DNA Ligase at 5 U μL^{-1} (Thermo Fisher)
- 1 µL 10 mM ATP (Sigma Aldrich)
- 0.25 µL LguI/SapI at 5 U μL^{-1} (Thermo Fisher)

Loop thermocycling program

- Assembly
    - 37°C for 3 minutes
    - 16°C for 4 minutes

- Termination & enzyme denaturation
    - 50°C for 5 minutes
    - 80°C for 10 minutes

Procedure

1. DNA parts were prepared at a concentration of 15 nM. This was calculated from the more frequently used unit of ng/μL through the formula

$$\frac{length\ (bp)}{100} = concentration\ needed\ (\frac{ng}{\mu L})$$

which was used as an approximation of the more accurate formula

$$15 \bullet 10^{-9}\ \frac{mol}{L} \bullet \frac{(607.4N + 157.9)\ g}{mol \bullet bp} \bullet N\ bp\ \bullet 1000 = \ concentration\ (\frac{ng}{\mu L})$$

where N is the length of the DNA part of interest. Prepare a universal receiver plasmid at concentration 7.5 nM.

2. Master mix was prepared, 5 μL per reaction
3. 1 μL of each DNA part was added to receiver plasmid in a thin wall PCR tube
4. Nuclease free H_2O to 5 μL was added
5. 5 μL even level master mix was added
6. Samples were thermocycled according to the Loop thermocycling program (shown above)

2.3.3 Loop plasmids used

All primers, level 0, and level 1 parts used in this research are available in a communal registry of parts at the Haseloff lab, University of Cambridge. They have been assembled only with parts available under the Open Materials Transfer Agreement [162]. Table 1 summarises level 2 plasmids used in this work.

Table 1: Level 2 plasmids successfully transformed into plants and discussed in this thesis. 1: Homologous Region (HR) contains significant mismatches with the Cam plastome because it was assembled before a high-quality plastome assembly was available. May have introduced mutations into the homologous region upon transformation. 2: Not Loop-compatible, construct used in [112] and was the only validated fluorescent protein synthetic plastid DNA construct in *Marchantia* at the beginning of this project. Assembled by Christian Boehm. 3: Assembled by Eftychios Frangedakis. 4: Nuclear transformation plasmid. 5: Successfully transformed recently, data too preliminary to be included here. 35S promoter sequence taken from [163].

Name	Position 1	Position 2	Position 3	Position 4	Notes
pCS Clo*B	*tRNA-G* HR[1]	*Ntprrn::aadA::NtpsbA_term*	*NtpsbA::mturq2cp::BBa_B0012::rps16_term*	*tRNA-M* HR	2
L2_352	*tRNA-G* HR	*Ntprrn::aadA::NtpsbA_term*	*MppsbA::mTurq2cp::BBa_B0012::rps16_term*	*tRNA-M* HR	3
L2_354	*tRNA-G* HR	*Ntprrn::aadA::NtpsbA_term*	*Mpprrn::mTurq2cp::BBa_B0012::rps16_term*	*tRNA-M* HR	3
L2_355	*tRNA-G* HR	*Ntprrn::aadA::NtpsbA_term*	*NtpsbA::mturq2cp::BBa_B0012::rps16_term*	*tRNA-M* HR	3
L2_360	RNA pol HR	*Ntprrn::aadA::NtpsbA_term*	*Nt PsbA::mturq2cp - BBa_B0012::rps16_term*	*atpH* HR	5
L2_361	RNA pol HR	*Ntprrn::aadA::NtpsbA_term*	*Mp PsbA::mturq2cp - BBa_B0012::rps16_term*	*atpH* HR	5
L2_103	spacer	*35S::mturq2-N7 - 35S term*	*35S::mVenus::N7::35S_term*	*EF1α:HygR1-35S term*	4

2.4 Molecular biology methods

2.4.1 Primer design

Primers were designed *in silico* with Benchling [164], using both manual design and the primer designer tool. Multiple Primer Analyzer (Thermo Fisher [165], based on the algorithm from [166]) was used to check for primer dimerisation.

Primers for generation of standardised DNA parts included SapI restriction site and the four base pair overhangs in the 5’ overhang. The general design was

$$5'\underbrace{NNNN}_{\text{binding overhang}}\underbrace{GGTCTC}_{\text{SapI}}\underbrace{N}_{\text{spacer}}\underbrace{NNNN}_{\text{standardised overhang}}\underbrace{18-30\ bp}_{\text{homologous to part cloned}}3'$$

Special primer tails were used to generate fragments capable of Loop Type IIS assembly, with a combination of restriction enzyme sites designed by Eftychios Frangedakis. This allowed for insertion into the universal receiver plasmid pUAP4 (provided by Eftychios Frangedakis) with shorter primers, and caused integration direction to be random.

$$5'\underbrace{NN}_{\text{binding overhang}}\underbrace{GCTCTTC}_{\text{SapI}}\underbrace{GTCTC}_{\text{partial BsaI}}\underbrace{N}_{\text{spacer}}\underbrace{NNNN}_{\text{standardised overhang}}\underbrace{18-30\ bp}_{\text{homologous to part cloned}}3'$$

2.4.2 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was performed using Phusion High-Fidelity DNA polymerase (Thermo Fisher). Master mix buffer was prepared on ice.

Phusion buffer - 50 µL reactions

- Nuclease-free H_20 to 50 µL
- 10 µL 5x Phusion HF buffer
- 1 µL 10 mM dNTPs
- 2.5 µL 10 µM forward primer
- 2.5 µL 10 µM reverse primer
- 1 µL template DNA
- 0.5 µL Phusion DNA polymerase

Once primers and DNA were added to all reaction tubes (200 µL thin-wall PCR tubes), samples were moved to thermocycler for reaction.

Thermocycling conditions

- Initial denaturation
    - 98°C for 30 seconds
- Elongation - repeat 35 times
    - 98°C for 10 seconds
    - 52-62°C (based on predicted melting temperature) for 20 seconds
    - 72°C for 30 seconds per kb

- Final extension
    - 72°C for 5 minutes
    - Hold at 4°C

2.4.3 Gel electrophoresis

1-4% (w/v) agarose was prepared by adding the appropriate amount of agarose into Tris acetate buffer, pH EDTA buffer and bringing the solution to a boil. This solution was held at 55°C until needed, at which point it was poured into molds and mixed with 1x SYBR Safe (Thermo Fisher) to create gels. Depending on application, gels consisted of 8-32 wells, with fewer wells used when DNA was to be extracted and purified. Gels were run at 90 Volts for 30-90 minutes and imaged with blue light or UV excitation.

2.4.4 Gel DNA purification

Was performed with the MinElute kit (Qiagen) according to manufacturer recommendations.

2.5 Plant transformation methods

2.5.1 *Agrobacterium*-mediated transformation

Materials

- Wild type *M. polymorpha* Cam1 spore heads bearing sporangia
- 1.5 mL microcentrifuge tubes
- Milton solution (sodium troclosene)- dissolve 1 tablet in 25 mL H_2O
- 40 µm Falcon filter
- 50 mL centrifuge tubes
- Sterile hood
- Benchtop centrifuge
- Agar Petri dishes with 0.5x Gamborg growth medium [167]
- Sterile cellophane disks
- Aluminium foil

Antibiotics

- Rifampicin 10 µg mL^{-1}
- Carbenicillin 50 µg mL^{-1}
- Tetracycline 5 µg mL^{-1}

- Spectinomycin 100 µg mL^{-1}
- Cefotaxime 100 µg mL^{-1}
- Hygromycin 25 µg mL^{-1}

Sporeling preparation

1. 1 sporangia per Petri dish desired was placed into a 1.5 mL microcentrifuge tube and crushed with forceps
2. 500 µL Milton solution was added
3. Pour through filter into 50 mL tube, washing once
4. Spores were incubated for 10 minutes in Milton
5. Spores were transferred from 50 mL tube to 2 microcentrifuge tubes
6. Spores were centrifuged at 16,200 g for 2 minutes, pour off supernatant
7. Pellets were resuspended in 30 µL Milton solution per Petri dish desired
8. Spores were spread onto Petri dishes and grown at 21-23°C constant light at 150-200 µE for 5-8 days until visibly green

***Agrobacterium* growth**

1. 50 µL highly concentrated *Agrobacterium tumefaciens* strain GV2260 cells was removed from -80°C freezer and thawed in pre-chilled 2 cm electro cuvette
2. 100 ng plasmid DNA was added
3. Bacteria were electroporated at 12.5 kV cm^{-1} for 1 second and 500 µL LB both was immediately added.
4. Bacteria were incubated at 28°C 120 rpm for 2 hours
5. 100 µL transformed bacteria were spread on a LB agar Petri dish containing rifampicin, carbenicillin, tetracycline, and spectinomycin and incubated at 30°C for 2 days
6. Use a single colony to inoculate 5 mL LB media containing rifampicin, carbenicillin, tetracycline, and spectinomycin in a 50 mL centrifuge tube covered in foil to ensure darkness. Incubate at 28°C 150 RPM for 2 days.

Liquid *Marchantia* Growth Media

- 0.5x Gamborg B5 with vitamins
- 0.1% (w/v) N-Z amine A (Sigma Aldrich)
- 0.03% (w/v) L-Glutamine

- pH 5.8

Agrobacterium induction and co-cultivation

1. Agrobacterium cultures were centrifuged for 15 minutes at 3000 g, the supernatant discarded, and resuspended in 5 mL liquid *Marchantia growth media*
2. Acetosyringone was added to a final concentration of 100 µM
3. Samples were incubated 6 hours at 28°C 150 RPM in darkness
4. Sporelings were removed from solid media and suspended in liquid *Marchantia* growth media
5. 4 mL of sporelings + liquid media were pipetted into each well of 6 well culture Petri dishes, 100 µM acetosyringone was added, 180 µL *Agrobacterium* culture was added
6. Sporelings and bacteria were co-cultivated for 2 days
7. Sporelings were transferred onto a 40 µm cell strainer, then washed with 25 mL

2.5.2 Biolistic transformation

2.5.2.1 Preparation of microcarriers

BioRad Tungsten / gold microparticles - Adapted from Chiyoda *et al.* 2007 [83]

1. 20 mg microparticles were added - either 0.7 µm tungsten or 0.6 µm gold - to 1 mL EtOH in a 1.5 mL microfuge tube
2. Particles were sonicated 2 minutes, centrifuged the least amount possible to achieve pellet, supernatant removed. Wash step repeated 3 times
3. 1 mL sterile water was added, sonicated for 2 minutes, centrifuged, supernatant removed, repeated 2 times
4. Pellet was resuspended by vortexing in 1 mL sterile H_2O. 50 µL aliquots were stored in sterile 1.5 mL tubes, vortexing between each to ensure equal distribution of particles and stored at -20°C until needed
5. On the day of bombardment, particles were thawed, 50 µL for 10 shots
6. 230 µL H_2O was added, vortexed to mix
7. 25 µg plasmid DNA was added, vortexed to mix
8. 250 µL 2.5 M $CaCL_2$ was added, vortexed to mix

9. 50 µL 1 M spermidine was added, vortexed to mix
10. Microcarriers were incubated for 10 minutes on ice, vortexing 10 seconds every minute
11. Microcarriers were centrifuged at 2000 g for 3 seconds, supernatant removed
12. 500 µL ice-cold absolute EtOH was added, vortexed briefly, centrifuged at 2000 g for 3 seconds, repeated 3 times
13. Final pellet was resuspended in 60 µL absolute EtOH

Seashell nanoparticles

DNAdel 550 nm (Seashell technology) nanoparticles were loaded with DNA according to manufacturer recommendations. Particles are sold with proprietary buffers 'binding buffer' and 'precipitation buffer'.

1. DNAdel particles are supplied in solution at 50 mg mL^{-1}, 0.5 mg are used per bombardment
2. DNAdel nanoparticles were diluted to 30 mg mL^{-1} with binding buffer and sonicated
3. 5 µg DNA per mg particles was added and vortexed to mix
4. A volume of precipitation buffer equal to the total volume of DNAdel particles and DNA was added, vortexed and incubated at 22°C for three minutes
5. Samples were centrifuged at 8000 g for 10 seconds, the supernatant discarded, and 500 µL 4°C EtOH added.
6. Samples were centrifuged again, supernatant discarded, and resuspended in 7 µL per bombardment

Sporeling preparation

Sporeling preparation was the same as for *Agrobacterium*-mediated transformations (see 2.5.1), except sterile cellophane disks were placed on the surface of 0.5x Gamborg Petri dishes during the sodium troclosene sterilisation period. These disks facilitate sporeling transfer following bombardment.

2.5.2.2 Bombardment

Biolistic bombardment was carried out with the Bio Rad PDS 1000/He device according to manufacturer's directions. Rupture pressures of 900 and 1100 pounds per square inch (PSI) were used at a range of distances as described in

chapter 4. Bombarded plants were transferred onto 0.5x Gamborg Petri dishes containing Spectinomycin 50 µg mL^{-1} for selection 72 hours after bombardment.

2.6 Microscopy

2.6.1 Image Acquisition

Stereo microscopy

Stereomicroscopy was performed with a Leica M205 FA fluorescence microscope with default configurations. Planapo 1.6x objective and 10x ocular lenses were used with the variable magnification internal lens. Table 2 lists the excitation and emission wavebands of fluorescence filters used. Transmitted light and LED side lighting were used for finding areas of interest.

Table 2: Fluorescence filters used in stereomicroscopy. All filters produced by Leica. CFP, GFP, and YFP are intended for use with cyan, green, and yellow fluorescent proteins, respectively.

Filtercube name	Excitation Band	Emission Band
CFP	426-446 nm	460-500 nm
GFP	450-490 nm	500-550 nm
Chlorophyll land plant	460-500 nm	610+ nm Long Pass
YFP	490-510 nm	520-550 nm

Antibiotic tolerance assays

Growth and antibiotic tolerance was assessed with a Keyence VHX-5000 microscope using the 100-500x objective and ring illumination. Focal plane stacks were generated with the extended depth of focus “Depth-up” feature using standard settings.

Transient Nuclear Expression Assays

48 hours after bombardment, Petri dishes of cells were directly imaged using stereomicroscopy. Each entire dish was scanned at moderate magnification (50-70x) using the YFP channel for bright dots, candidate fluorescent nuclei. Candidates were then viewed under the CFP filter before being counted. Plants that had not been bombarded and plants that had been bombarded with microcarriers not coated with DNA were imaged first, and fluorescent nuclei were never observed in these control cases.

Confocal microscopy

Confocal microscopy was performed on a Leica SP5 laser scanning confocal microscope (LSCM) using Leica Application Suite Advanced Fluorescence (LAS AF) software version 2.7.3.9723. Table 3 lists objectives used. Table 4 lists excitation and collection wavelengths used.

Table 3: Objectives used for confocal microscopy.

Name	Magnification	Optical Specifications	Immersion	Application
PL FLUOTAR	2.5x	0.07, ∞	Air	Screening
HC PL Apo	10x	0.40 CS ∞/0.17	Air	Screening, analysis of microcarriers (see section 4.2)
HC Pl Apo	20x	0.7 ∞/0.17	Air	Nuclear fluorescence quantification, all images except plastid expression quantification
PL Apo	63x	1.2 ∞ 0.14/0.18	Water	Plastid expression quantification

Table 4: LSCM excitation and collection information. Excitation illumination generated by Argon laser and UV laser diode.

Fluorophore	Excitation wavelength	Collection band
mTurquoise2	458	462-479
SYBR Safe	405	520-540 nm
mVenus	514	520-560 nm
Chlorophyll	458 or 514	680-720 nm

For quantitative experiments, Z stack range was limited to the epidermal cell layer, 40-80 µm deep. This was to minimise the effects of absorption and scattering of the excitation laser and the fluorescent signal. Images not for quantification were generally acquired through the entirety of the useful signal, usually 80-150 µm.

2.6.2 Image Processing

Maximum Projection

Maximum projection images were generated from confocal Z stacks using LAS AF with s scaling factor of 1 and threshold of 10.

Macrocarrier loading

These images were collected using a fluorescent channel and a transmitted laser channel which records light that passes through the sample to a detector in the normal location of the brightfield illumination source. Transmitted light images were inverted to generate bright spots at the location of microcarriers where light was occluded, and the colocalisation was assessed by eye during microscopy and post-imaging using Fiji version 2.0.0-rc-64/1.51S [168].

Thresholding

Various local thresholding algorithms were implemented using Fiji, as described in section 4.4.2 and 5.3.1. After the implementation of a thresholding algorithm, Lookup Table (LUT) was inverted, and the analyze particles function was implemented. For nuclear fluorescence, particle size was limited to 30-200 μm^2, circularity 0.2-1.0. For plastid fluorescence, particle size was limited to 10-50 μm^2. In all cases, particles on edge were excluded.

Region of interest measurement

Regions of interest (ROIs) were measured with the native Fiji function. Area, minimum intensity, maximum intensity, and mean intensity values were captured in all relevant color channels.

Figure composition

Figure arrangement was performed in Fiji, Google slides, and Inkscape version 0.92.1, XQuartz version 2.7.11 running on macOS 10.13.4. All images were maintained either uncompressed or in lossless compression, either lif, png, or tiff file formats.

2.7 Statistics

Unless stated otherwise, statistical analysis was performed with Rstudio version R version 3.5.0 (2018-04-23) and ggplot2 version 3.0.0. Example instances of commands used are shown in Appendix A. Where stated, statistics were generated with Graphpad Instat Version 3.10

Chapter 3: Plastid Genome Annotation

3.1 History and existing assemblies

There is a history of confusion with the *Marchantia polymorpha* plastome, because the original sequence published in 1986 actually belonged to a related species, *Marchantia paleacea* [169]. Table 5 is provided as a guide to plastome assemblies discussed in this chapter.

Table 5: Summary table of plastome assemblies discussed in this chapter. NC_001319 was originally misannotated as *polymorpha*, leading to confusion. Length can be used to confirm assembly identity.

Assembly Name	**Description**	**Length**	**Genbank Accession**
M. paleacea	Is actually from *M. paleacea*	121024	NC_001319 **and** X04465
Cam-1/2	Assembled in this project	120314	MH635409
Kit-2 *subsp. ruderalis*	Assembly uploaded April 24, 2018, by John Bowman *et al.*	120304	NC_037507
Tak-1	Sequence identical to Kit-2, unannotated	120304	Personal correspondence with John Bowman

At the beginning of this project only the *M. paleacea* plastome, at the time mislabelled as *M. polymorpha*, was publicly available. However, a large *M. polymorpha* Cam total DNA NGS dataset which contained ~10% plastid reads was available from a previous project. Comparison of a preliminary Cam plastome assembly by Bernardo Pollack showed substantial differences to the *M. paleacea* genome, which led to doubt on the species identity and eventually to the discovery of the *M. paleacea* genome's mistaken identity, which had been mentioned in the literature but remained uncorrected in Genbank until recently [169]. The *M. paleacea* genome will remain useful for comparative genetics, but is

not appropriate for designing homologous regions for *M. polymorpha*, necessitating generation of a new plastome assembly.

3.2 Cam-1/2 Plastome Assembly

Our laboratory uses two accessions of *Marchantia polymorpha* for our research, Takaragaike (Tak), collected in a park in Kyoto, Japan, and Cambridge (Cam), collected in Cambridge, UK. The sexes are denoted as a number appended to the accession name, with 1 signifying male plants and 2 signifying female plants (this nomenclature is widely accepted in the *Marchantia* community). The maternal inheritance of the plastome means it is plausible that there are no sequence differences between the sexes, but to test this hypothesis an assembly was made of Cam1 and Cam2 DNA short reads separately, and the final assemblies were compared to each other and found to be identical in sequence.

The assembly process is described in detail in section 2.1. In brief, Illumina short reads of Cam-1 nuclear, mitochondrial, and plastid DNA were mapped against a Tak draft assembly generously provided by John Bowman, and the reads that matched were assembled *de novo* using CLC Genomics workbench. *De novo* assembly generated 25 contigs spanning the entire plastome, which were assembled using the reference to generate the complete plastome. The resulting assembly was compared to an assembly produced by Jim Haseloff, which differed in 4 locations. These minor differences were resolved by manual inspection of the Illumina dataset and the resulting assembly was designated Cam-1/2 Plastome. The assembly process was then repeated with Illumina reads from Cam-2, and the assembled sequences were identical. The Cam-1/2 plastome differs more from the *M. paleacea* plastome, an unsurprising find since these sequences are derived from different species (see Table 6).

Table 6: Comparison of the Cam-1/2 sequence and NC001319, the *M. paleacea* plastome. They align well, but have significant amounts of small-scale variation. The longest string of mismatches is 7 bp, the longest gap is 24 bp. Mismatches and gaps are here grouped by length.

Cam-1/2 vs *M. paleacea*	**Length=1**	**Length=2**	**Length=3**	**Length=4**	**Length=5**	**Length>5**
Mismatches	2306	105	18	7	1	3
Gaps	191	43	22	4	9	40

3.3 Validation

The Cam-1/2 plastome assembly was validated in two ways. It was first compared to Sanger sequencing data covering ~10% of the plastome, then later compared to the newly published Kit-2 plastome assembly [80]. In both cases, validation supports a highly accurate assembly process.

3.3.1 Sanger Sequencing

The primary thrust of this project is plastid transformation, which requires 1.5 kb homologous regions for construct assembly to facilitate integration by homologous integration (see section 1.3.3). Figure 4 shows the location of insertion loci selected.

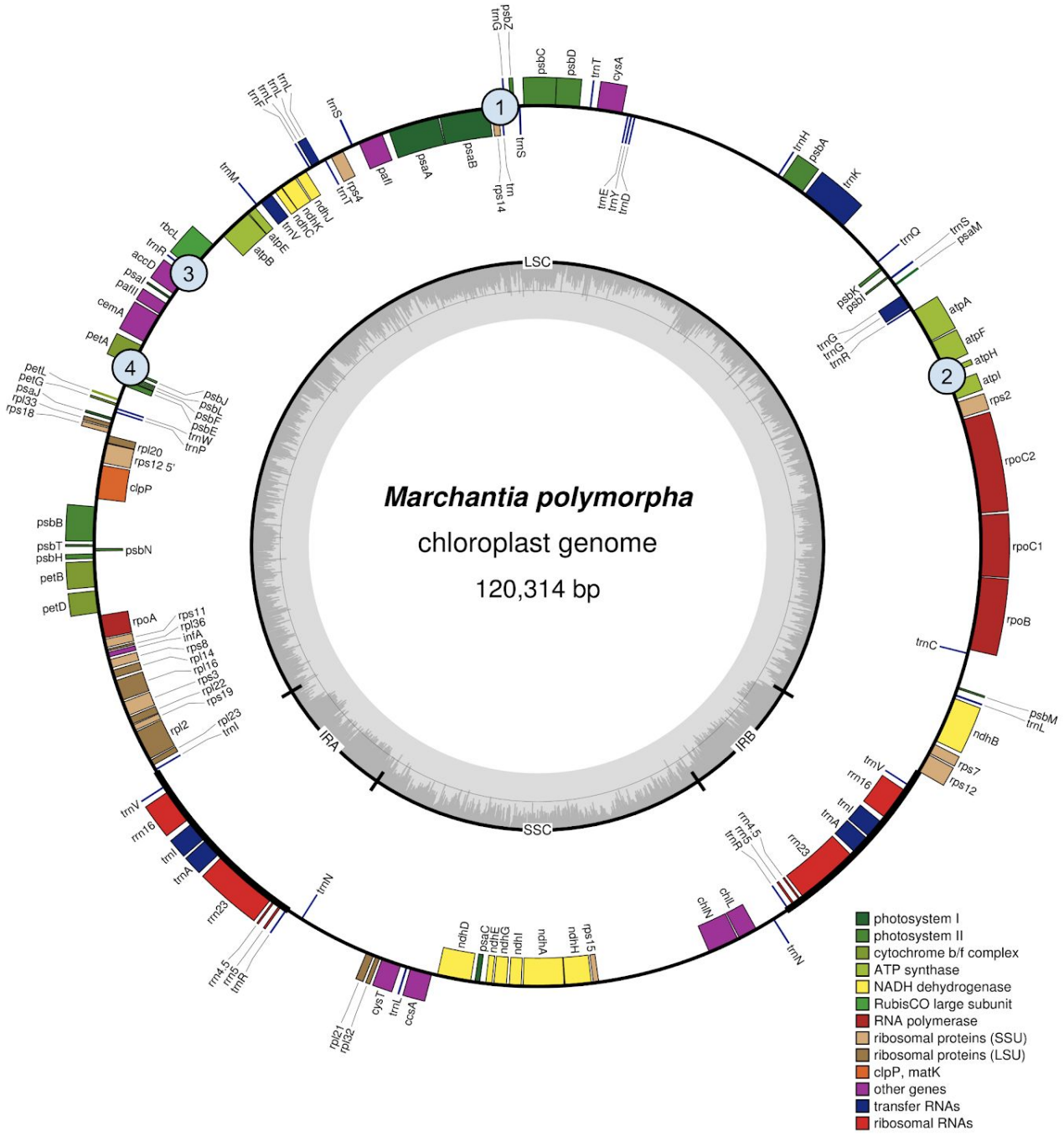


Figure 4: Map of the *Marchantia polymorpha* plastome, Cam-1/2 assembly. Numbers on the assembly indicate the transgene insertion locations used in this work. The inner circle shows a GC map, and shows the tetrapartite structure of LSC, SSC, and inverted repeats (IRA and IRB). 1: tRNA-G - tRNA-M intergenic region, the insertion location used in [112]. 2: RNA Polymerase Operon 3' end intergenic region, successfully transformed in this work. 3: *rbcL-tRNA-R* intergenic region, *Marchantia* constructs prepared, demonstrated in *N. tabacum* [170]. 4: *petA-psbJ* intergenic region, *Marchantia* constructs prepared, demonstrated in *N tabacum* [171]. Produced from the Cam-1/2 assembly with OrganellarGenomeDRAW [172].

In each case, two homologous regions were cloned from the wild type plastome and analysed with Sanger sequencing. In addition to the previously used *Marchantia* insertion locus in the intergenic region between tRNA-G and tRNA-M, three additional insertion loci were selected based on locations that had resulted in high expression levels in other species (see Figure 4 legend). The Sanger sequencing data which totalled just over 12 kb in 4 locations throughout the plastome also served as a method for assembly validation. No differences were observed between the Cam-1/2 plastome assembly and the Sanger data, whereas the data showed 16 bp of disagreement with the Kit-2 assembly, split over 8 loci and 303 loci of disagreement with the *M. paleacea* genome.

3.3.2 Comparison to Kit-2 Assembly

After the Cam-1/2 plastome was assembled, a new *M. polymorpha* genome was published as the Kitashirakawa-2 strain of subspecies *ruderalis* ([80], plastome assembly available 10 April 2018). This assembly is similar to the Cam-1/2 assembly. A complete list of differences is provided in Table 7. The similarity between these two makes any large-scale assembly errors unlikely.

Table 7: Complete list of the 16 differences between Cam-1/2 plastome and Kit-2. SNP: single nucleotide polymorphism. RLP: repeat length polymorphism. Location numbers taken from Kit-2.

Location	Difference type	Cam-1/2 sequence	Kit-2 Sequence
17737	SNP	T	A
22552	RLP	AT •23	AT • 24
23968	SNP	T	C
26870	SNP	C	T
30893	SNP	A	T
40980	SNP	A	T
40982	SNP	G	A
40983	Gap	T	_

41009	RLP	TTGTTTTTTTAC • 3	TTGTTTTTTTAC • 2
49005	Mismatch	TT	AA
55636	RLP	Polymorphism - AT • 12 or 13	AT • 13
80466	SNP	C	T
90906	SNP	C	A
94602	Gap	T	_
106453	SNP	G	T
111912	SNP	A	C

3.4 Discussion

I have generated the first full assembly of *Marchantia polymorpha* strain Cam, and demonstrated that the plastome is identical for Cam1 and Cam2. This plastome has been annotated and submitted to Genbank to serve as a resource for future engineering projects. In particular, this assembly allows for the accurate design of homologous regions for transgene insertion. In addition to the single previously available insertion locus, genetic parts for three more insertion loci have been generated, and more could easily be generated through cloning or synthesis aided by an accurate plastome assembly.

Chapter 4 - Optimisation of transformation

4.1 Introduction

The systematic optimisation of the physical and biological parameters included testing choice and preparation protocol of microcarrier, firing distance, tissue type bombarded, and recovery/selection conditions. Transient nuclear transformation with genes encoding fluorescent proteins was developed as a rapid assay of DNA delivery efficiency.

The Bio-Rad PDS 1000/He device (diagrammed in Figure 5) was used for biolistic transformation. It functions by accelerating microcarriers into the sample tissue. A shock wave of Helium is generated when the rupture disk is broken, which accelerates the macrocarrier into the stopping mesh. Microcarriers initially placed on the bottom of the macrocarrier continue downwards towards the sample at high velocity. The distance between the stopping mesh and the sample can be varied by selecting which shelf the sample is placed on (see section 4.3).

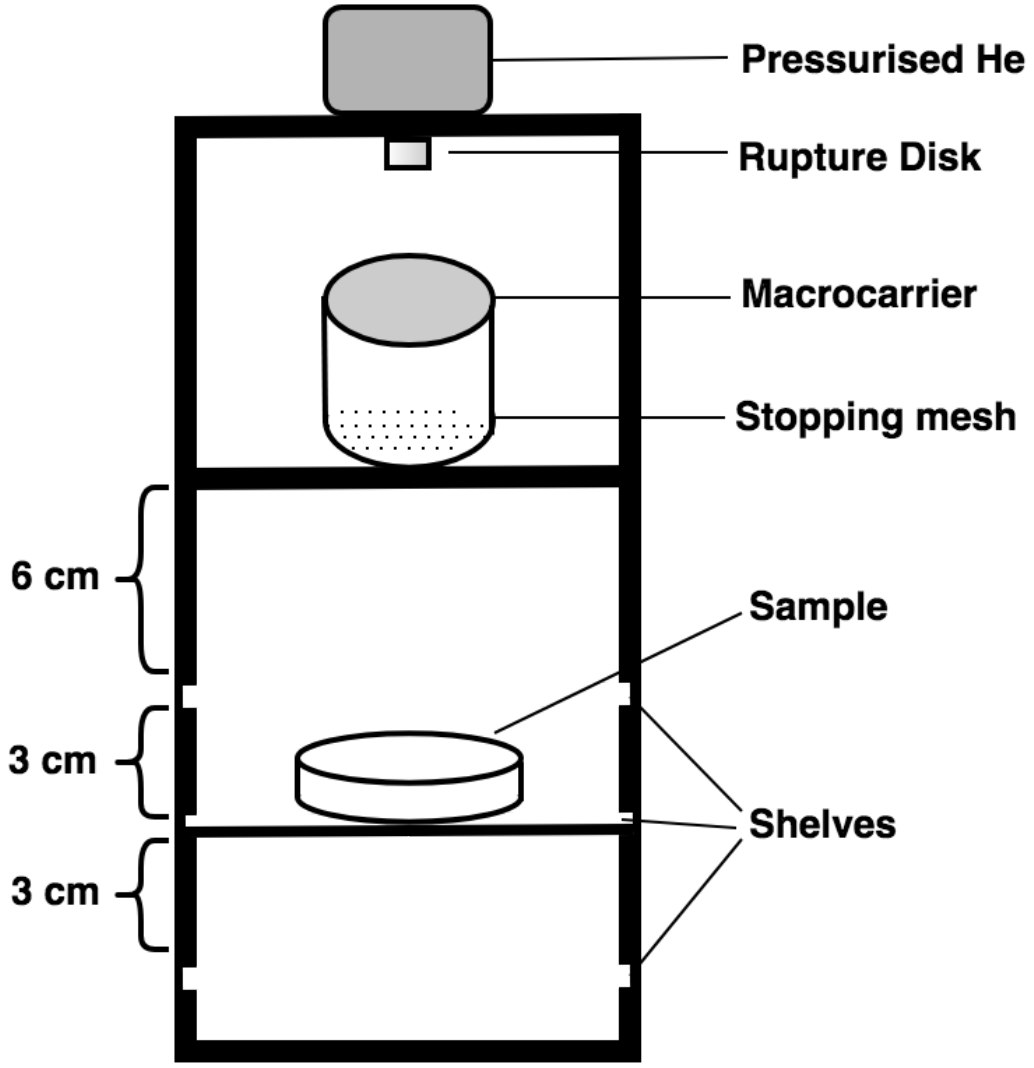


Figure 5: Schematic diagram of PDS 1000/He biolistic device. The three shelves indicated correspond to the short, medium, and long conditions discussed in section 4.3.

4.2 Microcarrier selection

4.2.1 Overview and experimental design

Microcarriers comprise the core of the DNA delivery mechanism, and therefore macrocarrier choice and DNA loading protocol is essential to successful transformation. Gold and tungsten particles 0.6 - 3.0 μm in diameter are generally used for transformation [90]. For plastid transformation, particles 0.6 - 1.0 μm are most frequently used [173].

To visually assess efficiency of DNA binding, DNA was mixed with a fluorescent stain, bound to microcarriers, then imaged with confocal microscopy. The images

were composed of two channels, transmitted light to localise the microcarriers and SYBR Safe fluorescence to localise DNA. The colocalisation of the two was then assessed, as demonstrated in Figure 6. To test whether the stain altered DNA binding, the opposite order of the staining and binding procedures was also tested, which did not affect the results. Additional controls are shown in appendix B, Figure 33.

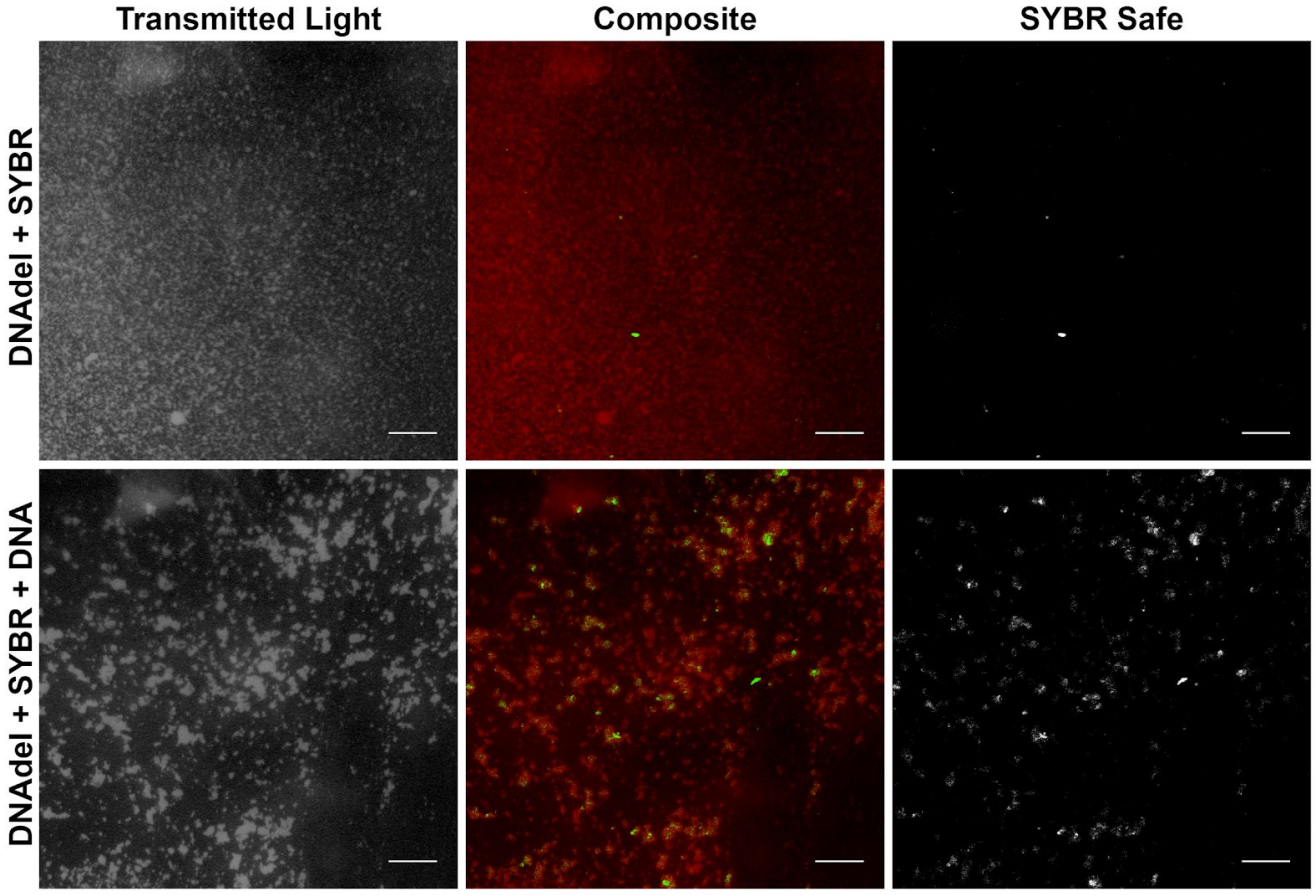


Figure 6: Example of confocal microscopy used to assess DNA binding onto microcarriers. Top row shows minimal SYBR safe fluorescence in the absence of DNA. Bottom row shows colocalisation of SYBR fluorescence and DNAdel beads. Left column shows transmitted light channel with lookup table inverted so particles show up as white spots, right column shows SYBR safe channel. Central column is a composite with transmitted light pseudocoloured red and SYBR safe pseudocoloured green. Scale = 40 µm

4.2.2 Results

Three microcarriers and a wide selection of binding protocols were tested in a rapid-prototyping regime, and the results are summarised in Table 8. The microcarriers tested were Biorad 0.6 µm gold, Bio-Rad 0.7 µm Tungsten, and Seashell DNAdel 0.55 µm gold with proprietary functionalisation. For Bio-Rad microcarriers, two binding protocol variables were tested where discrepancies

existed in published protocols, spermidine concentration and binding reagent order. DNAdel 550 nm nanoparticles (SeaShell technology) are frequently used for algal plastid transformation [174,175], are here tested for land plant plastid transformation.

Table 8: Summary of findings across different microcarriers and binding protocols. After review of protocols, the most variable aspect was the concentrations of spermidine used and the order of adding DNA, spermidine, and $CaCl_2$. Single, double, and triple asterisks represent minor, moderate, and high levels of fluorescence colocalisation, as an indicator of DNA binding efficiency. Figure 6 demonstrates the degree of colocalization awarded three asterisks, double and single asterisks show approximately three and ten times less colocalisation, respectively.

Microcarrier type	**Binding protocol**	**DNA Colocalisation**
0.6 µm gold	As described (see section 2.5.2.1)	**
0.6 µm gold	Standard order, spermidine concentration 0.1 M	*
0.6 µm gold	Standard order, spermidine concentration 1 mM	Negligible
0.6 µm gold	Binding order DNA, spermidine, $CaCl_2$	Negligible
0.6 µm gold	Binding order DNA, spermidine, $CaCl_2$, spermidine concentration 0.1 M	Negligible
0.6 µm gold	Binding order DNA, spermidine, $CaCl_2$, spermidine concentration 1 mM	Negligible
0.7 µm Tungsten	As described	*
0.7 µm Tungsten	Standard order, spermidine concentration 0.1 M	*
0.7 µm Tungsten	Standard order, spermidine concentration 1 mM	Negligible
0.7 µm Tungsten	Binding order DNA, spermidine, $CaCl_2$	Negligible
0.7 µm Tungsten	Binding order DNA, spermidine, $CaCl_2$, spermidine concentration 0.1 M	Negligible
0.7 µm Tungsten	Binding order DNA, spermidine, $CaCl_2$,	Negligible

	spermidine concentration 1 mM	
Seashell 550 nm (0.55 µm)	As described (Manufacturer recommendations)	***

The protocol in the laboratory, provided by Christian Boehm, used 1 mM spermidine, whereas most published protocols use 1 M or 0.1 M. These higher concentrations of spermidine were substantially improved DNA binding on Bio Rad microcarriers. However, even the best protocols for Bio-Rad particles were substantially outperformed by DNAdel particles, which also have a simpler preparation protocol (see section 2.5.2.1). To the author's knowledge, this is the first use of DNAdel particles for *Marchantia* plastid transformation. As can be seen from Table 8, these particles bound DNA much more efficiently, and were thereafter used.

4.3 Nuclear transformation as a rapid assay

4.3.1 Rationale

Transient nuclear expression of foreign DNA has been demonstrated in many species of plants, and requires only that foreign DNA of organism-appropriate regulatory elements and codon usage be delivered to the cytoplasm [176]. Transient expression occurs on the timescale of hours to days, rather than the weeks to months required to achieve homoplasmic plastid transformation.Transient nuclear transformation is here developed as a rapid assay for DNA delivery. Plasmid L2_103 (see Table 1) encodes two fluorescent proteins, mTurquoise2 and mVenus, both driven by the *CaMV 35S* promoter. Both proteins are localised in the nucleus by N7 targeting domains, resulting in clear punctate fluorescence expression upon successful DNA delivery.

4.3.2 Results

Transient expression was visible in stereo microscopy by the bright dots visible with both the eYFP and eCFP filter sets (see Table 2 for filter details), as shown in Figure 7. In most cases, single nuclei were fluorescent, but occasionally two or more adjacent nuclei showed fluorescence. Given the low frequency of

transiently-expressing cells compared to the total number of cells, the most likely explanation is the division of a single transformed cell. As such, these instances were counted as a single transformation event for DNA delivery quantification purposes.

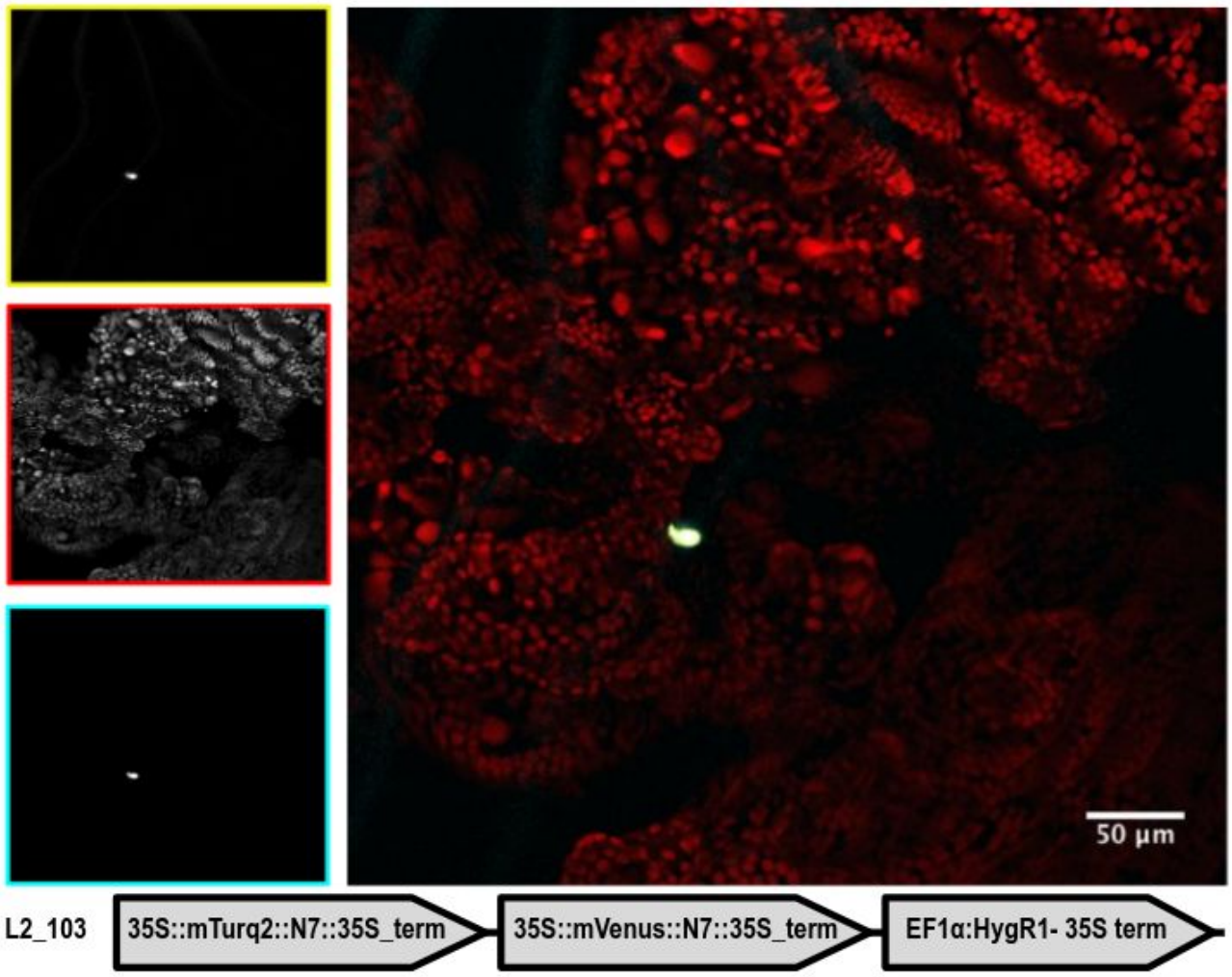


Figure 7: Transient expression of L2_103 visualised with stereomicroscopy for DNA delivery assays. Yellow box shows YFP channel, red box shows Chloroplast Land Plant channel, cyan box shows CFP channel. Bright, punctate fluorescence co-localising in the CFP and YFP channels was clearly seen and used as an assay of DNA delivery. Scale bar corresponds to 50 µm.

Transient nuclear transformation allowed for rapid feedback on DNA loading and microparticle delivery, shortening the testing cycle from the 6-12 weeks required for plastid transformation and selection to roughly 72 hours, allowing for a rapid prototyping process. To the author's knowledge, transient nuclear transformation

is reported here for the first time for two tissue types, 30 day old sporelings and 0-3 day old gemmae (see Appendix C, Table 10).

Firing distance optimisation

The PDS 1000/He has several adjustable parameters, and the parameter the manufacturer recommends as the most important to optimise is the distance from the topping disk to the tissue, the flight distance for the microcarrier. As the particles descend from the stopping mesh, their distribution of velocities and distance from the center increase, resulting in different patterns of microcarrier penetration and DNA delivery. DNA delivery was assessed through transient nuclear fluorescent protein expression, and the results are shown in Table 9.

Table 9: Transient nuclear expression for several firing distances. The distance indicated is between the macrocarrier stopping mesh and the plant tissue, and is varied by placing the plant tissue on one of three shelves within the PDS-1000/He device (see Figure 5). Numbers reported are transiently transformed cell per 10 cm petri dish coated in sporelings. This optimisation shows 9 cm is the optimal target distance 5-8 day sporelings, though this may vary by tissue type.

	6 cm - Top shelf	**9 cm - Middle shelf**	**12 cm - Bottom shelf**
Replicate 1	27	>50	3
Replicate 2	25	31	4
Average	26	>40	4

4.4 Stable Biolistic nuclear transformation

The relatively high efficiency of transient expression of the nucleus-expressed construct L2_103 suggested it would be possible to generate stable nuclear transformants with biolistic transformation. Transformation was attempted through both methods in parallel, and the resulting plants are shown in Figure 8. Biolistic transformation of the *Marchantial* nuclear genome was first reported by Chiyoda *et al.* in 2008 [82]. The timeline and efficiency of both methods is comparable, but biolistic nuclear transformation has the advantage of being entirely axenic and potentially functional outside the species restrictions dictated by *Agrobacterium* host range, which may be approaching the limit to its biotechnological expansion [177].

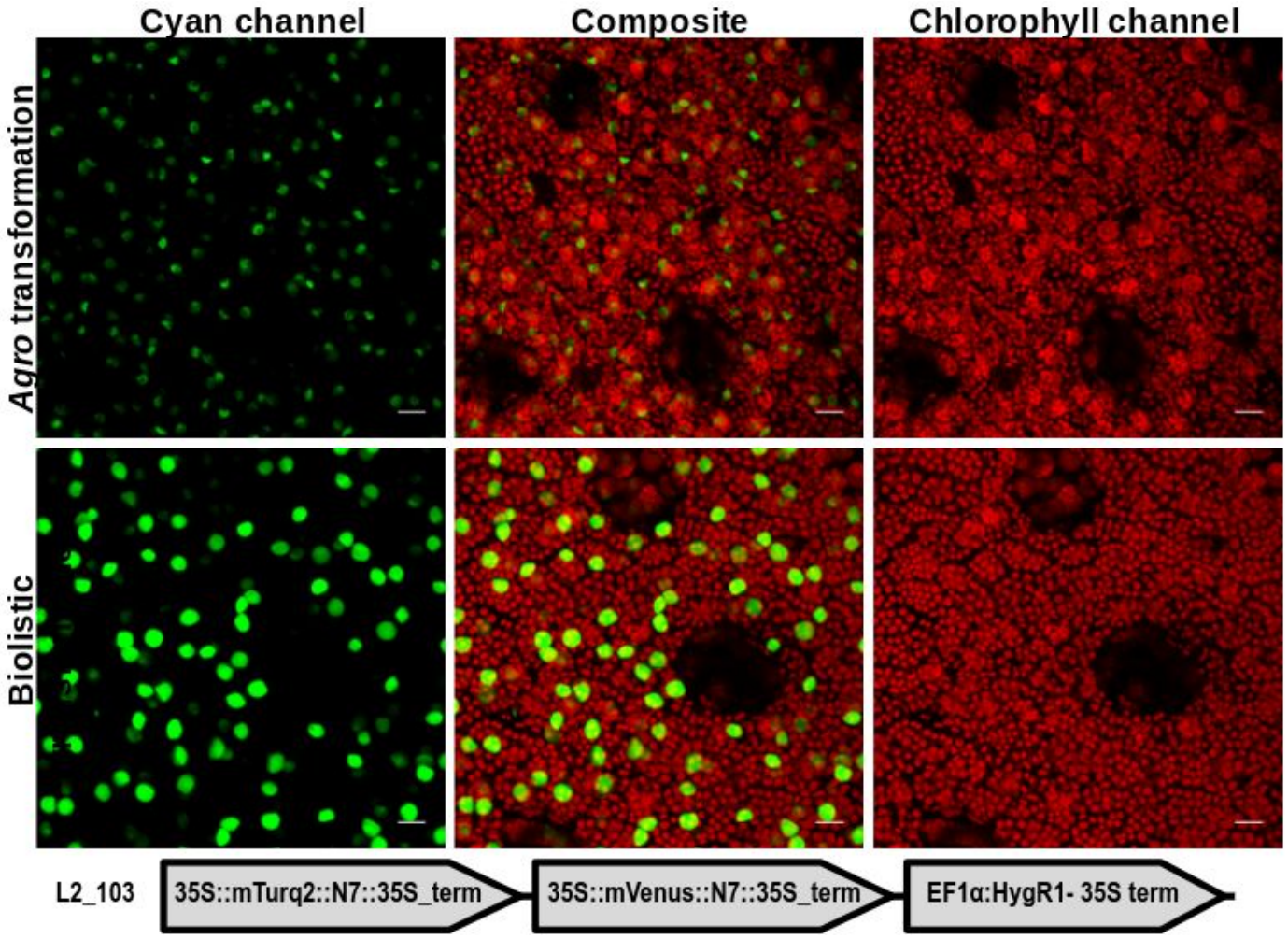


Figure 8: Biolistic nuclear transformation. Top row shows thallus tissue of plants stably transformed with L2_103 via *Agrobacterium*-mediated transformation. Bottom row shows thallus tissue of plants stably transformed with L2_103 via biolistic transformation. Hygromycin 25 µg mL^{-1} to establish stable transformants. Excitation and gain settings were kept constant between images. Scale bar corresponds to 20 µm.

Successful application of two distinct methods of DNA delivery for nuclear transformation invited a comparison of relative levels of fluorescence intensity achieved by these two methods. In order to do so, micrographs of plants transformed with each technique were generated, and a segmentation pipeline for their analysis was developed.

4.4.1 Strategy

Fluorescence intensity is measured through a three step process: microscopy, segmentation, and digital measurement. Segmentation is the process of grouping local regions of an image into discrete sets which can be digitally measured [178].

This can be achieved manually (for example, by circling nuclei by hand in a micrograph, then measuring their characteristics) or automatically. Automated segmentation was chosen to allow for larger sample size. The central challenge of automated segmentation is design of an image analysis pipeline that results in segmented regions that correspond appropriately to the biological features to be measured.

Automated segmentation is achieved by first duplicating the channel to be segmented, in this case the nuclear fluorescence channel, creating a ‘mask’ channel that can be manipulated without altering the original data. A series of mathematical transformations and algorithms are then executed on the mask channel, ending with a thresholding algorithm. Thresholding converts an image from a high bit depth (in this case 8-bit, for a total of 256 possible pixel intensity values) to a binary image (1 bit, so all pixels are either on or off). A simple example thresholding algorithm would be to convert all pixels with intensity higher than a particular ‘threshold’ value to 1, and all pixels with intensity lower than the threshold to 0. A variety of thresholding algorithms have been developed, many of which set different threshold values for different image regions according to the composition of local features. Which thresholding algorithm performs best, as well as which pre-processing is necessary, varies by image composition and properties [179]. In section 4.4.2, processing steps and thresholding algorithm are selected empirically by comparing the output images on the basis of segmentation quality. High quality segmentation is achieved when regions containing the object of interest have a pixel value of one and all other regions have a pixel value of zero.

This binary image is then scanned for contiguous regions of value one that match the spatial criteria for the biological object of interest. These regions of interest (ROIs) are then identified in Fiji by encirclement with a yellow line. Regions of interest can then be mapped back to the original eight bit image, whose values remain undistorted by the algorithms and binarisation. The properties of the original image within the boundaries of these ROIs can then be measured automatically. The average fluorescence intensity of more than one thousand nuclei was measured in this section, and fluorescence intensity of more than ten thousand plastids was measured in section 3.5. These sample sizes are

impractical for manual measurement, but can be analysed in under an hour with an automated segmentation pipeline once optimisation is complete.

The nuclear segmentation pipeline developed here consists of duplicating the nuclear fluorescence channel, applying a local contrast enhancement algorithm, applying a thresholding algorithm to generate a binary image, and ROI identification using the “Analyze particles” feature in Fiji. ROIs are then mapped back to the raw image nuclear channel for nuclear fluorescence intensity measurement. The workflow is diagrammed in Figure 9.

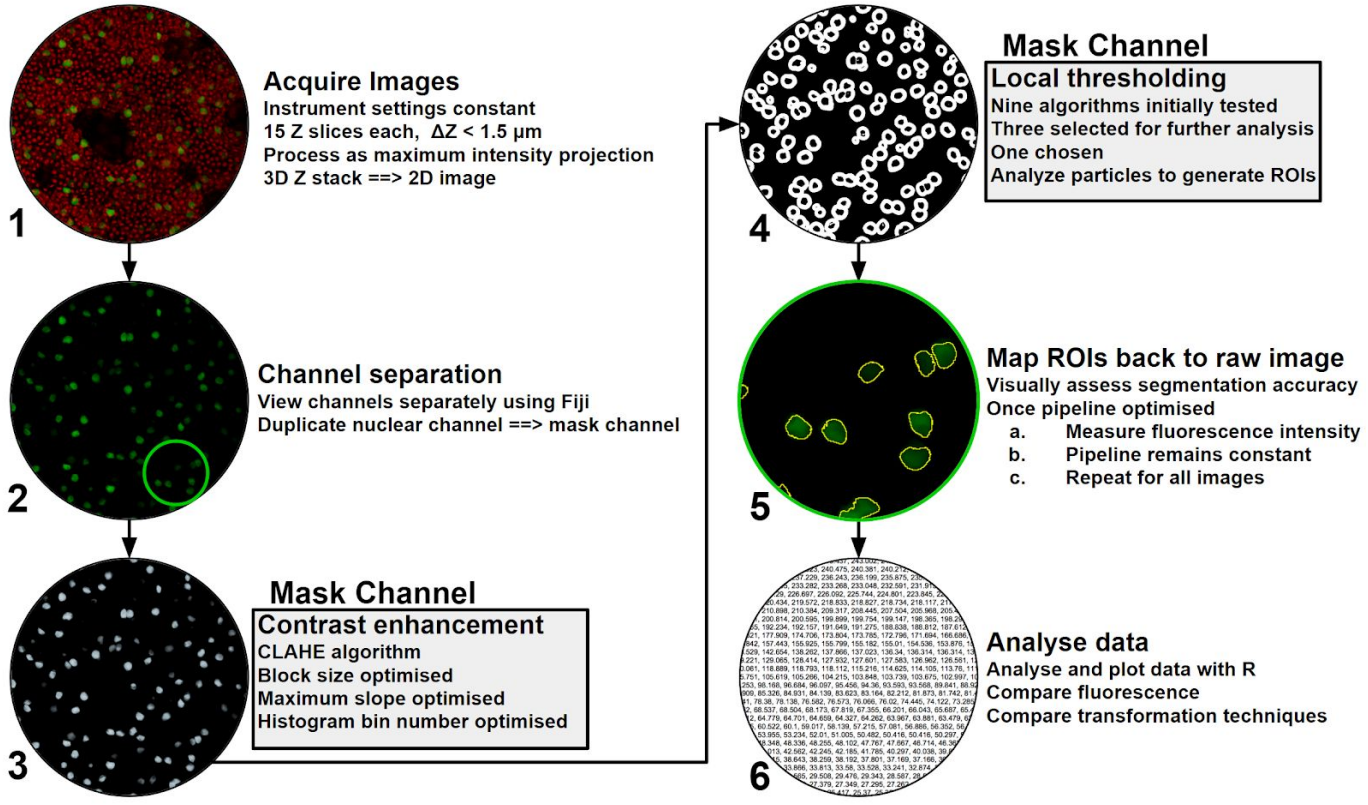


Figure 9: Schematic showing nuclear fluorescence analysis pipeline. Green circle in step 2 corresponds to the high magnification image in step 5. Step 5 is shown at higher magnification for illustrative purposes only to show yellow lines demarcating ROIs.

A segmentation-based workflow was selected to characterise the fluorescent behaviour of the transgenic mutants generated in this project. Once optimised, an automated segmentation pipeline is used to measure the fluorescence intensity of a large number of fluorescently labeled nuclei to compare the expression levels of the same fluorescent marker transformed through two methods:

Agrobacterium-mediated and biolistic DNA delivery. A similar method is used to assess plastid fluorescence in section 5.3.

4.4.2 Image pipeline optimisation

The pipeline is first optimised using a single maximum projection test image of thallus tissue stably transformed with L2_103 through biolistics. The test image is a two channel maximum intensity projection of 15 Z slices acquired on a Leica SP5. The image is 1024 by 1024 pixels at 8 bits per channel, and channels are pseudocoloured green and red for nuclear CFP and chlorophyll autofluorescence, respectively. A total of 27 such projections were acquired and are analysed in section 4.4.3.

Local contrast enhancement has been shown to improve segmentation accuracy [180], and was selected as a pre-processing step (step 3 in Figure 9). The Enhance Local Contrast function in Fiji was used for this purpose, based on the contrast limited adaptive histogram equalisation (CLAHE) algorithm [181]. The algorithm examines small spatial regions (blocks) of the image in series, and enhances contrast by transforming histogram shape, redistributing pixel intensity across the full range 0-255. There are three parameters in the program, block size, maximum slope, and histogram bin number. Block size is the size of the local region for which the histogram is equalised. Chloroplasts are 10-20 pixels across in the test image, and nuclei are 25-45 pixels. Candidate values selected were 10, 25, and 50 pixels. Maximum slope is the contrast stretch the algorithm is allowed, value of 1 results in unchanged images, high values cause substantial increases in contrast. Candidate values of 2, 3, and 4 were selected based on the examples included with the Fiji documentation. Histogram bin number is the number of input value bins used in the algorithm, where two bins would treat the image as binary and 256 bins applies no transformation.

Block size and maximum slope were first varied. Several values each were used and all permutations tested, as shown in Figure 10. At this stage, histogram bin number was held constant at 256, which applies no binning to the 8-bit data. From visual inspection of the output images, maximum slope = 3, block size = 25 was selected as a good compromise between high levels of contrast enhancement and

minimal distortion of the image. Next, histogram bin number was varied widely from 8 bins to 256 bins (see Appendix B, Figure 36). Low values for histogram bin number result in higher levels of contrast in the output image, but also cause distortion. As such, histogram bin number of 256 was selected and used for further applications.

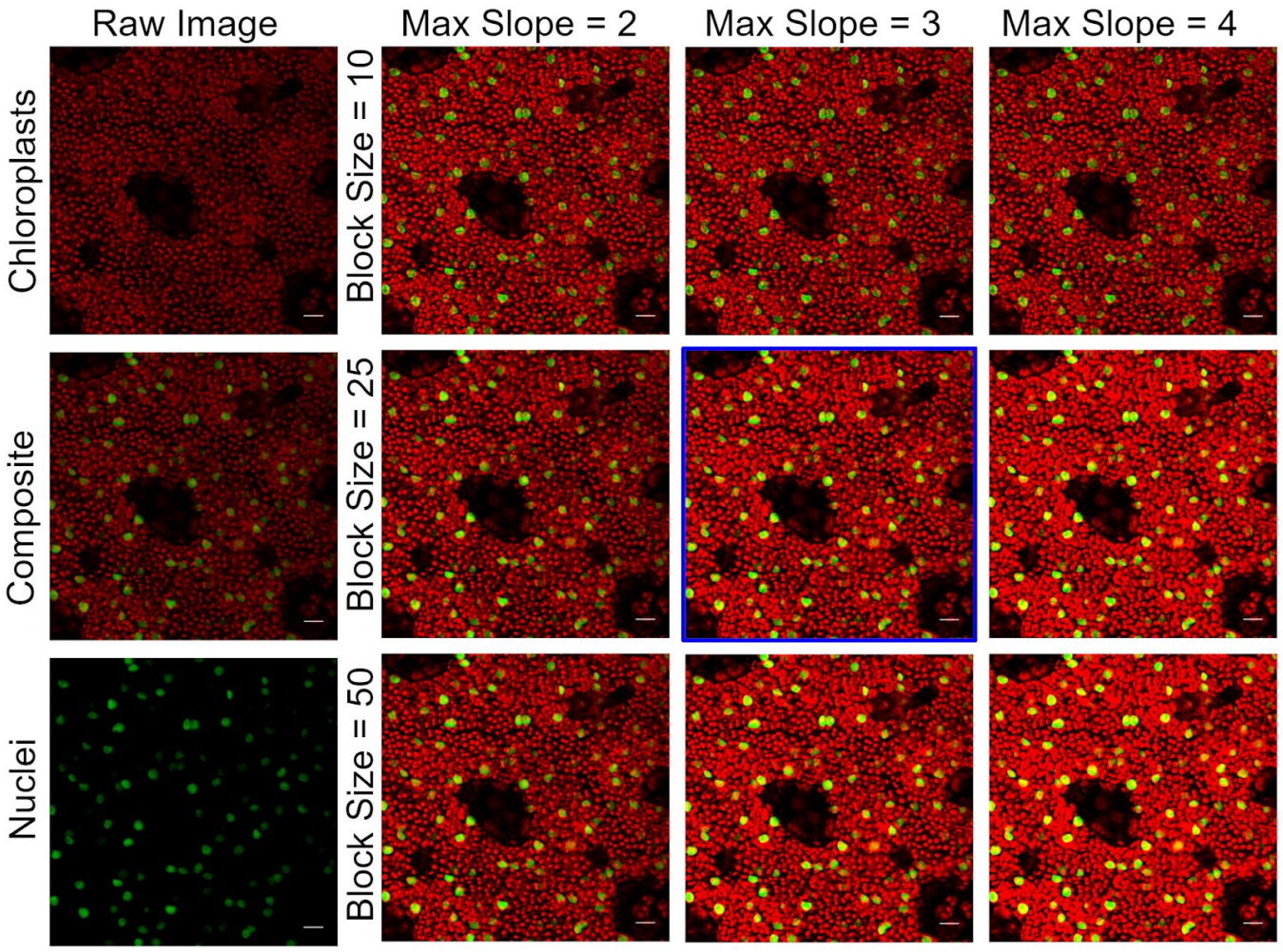


Figure 10: Slope and block size optimisation of CLAHE algorithm. Leftmost column shows raw image. Remaining columns show images after application of CLAHE algorithm with parameter permutations of maximum slope of 2, 3, and 4, and block sizes of 10, 25, and 50. Blue outline indicates parameters selected. Scale bar corresponds to 20 μm.

After CLAHE processing, a local thresholding algorithm was employed to identify nuclei (step 4 in Figure 9). A two-tiered breadth-first search was employed, where many algorithms were initially screened, three candidates were selected for further analysis, and one was chosen for the final pipeline. All local thresholding algorithms available in Fiji were examined, and the resulting thresholded images are shown in Figure 11.

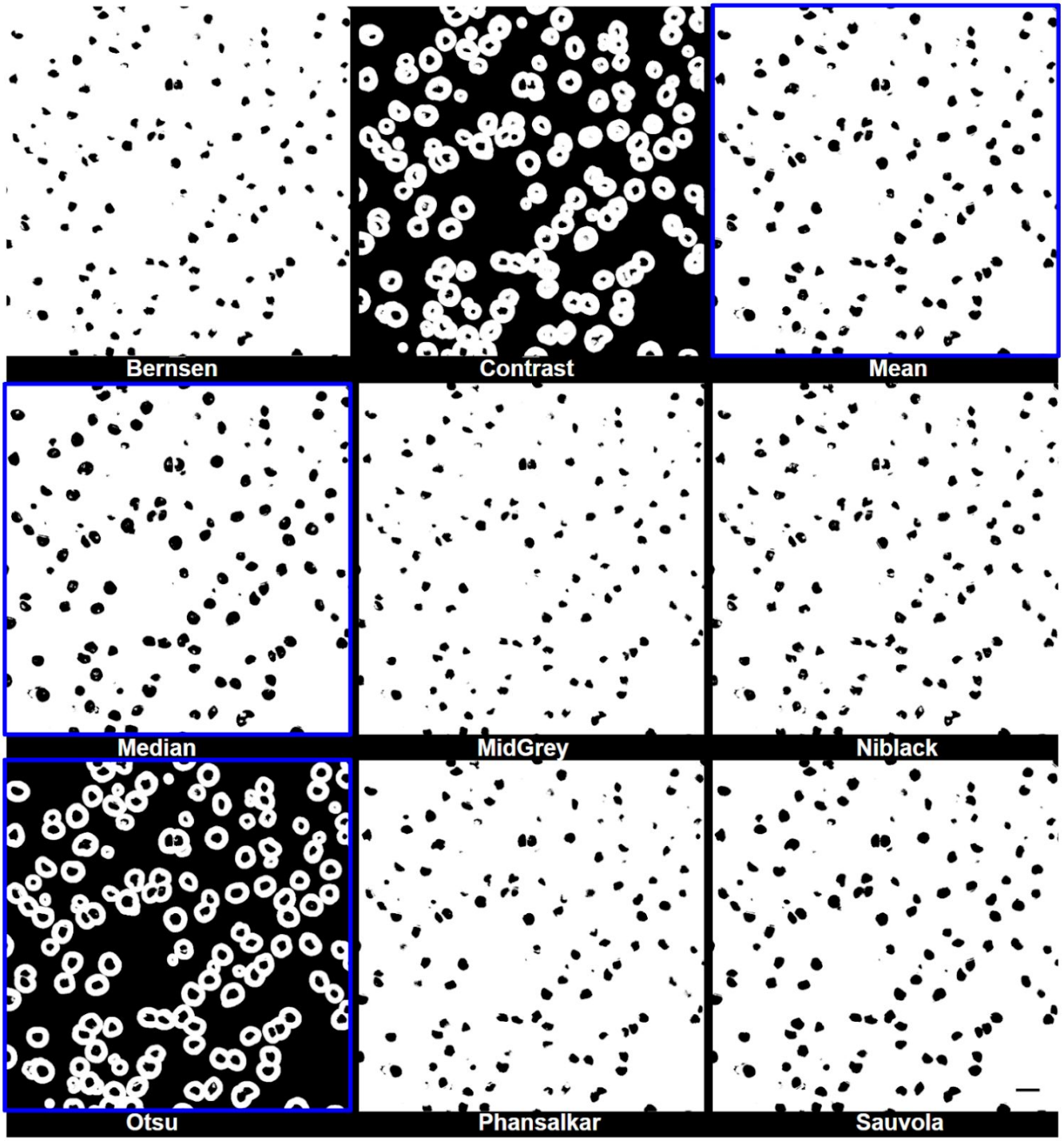


Figure 11: Thresholded images produced by various thresholding algorithms. Algorithm names correspond to the image above. Blue outlines indicate algorithms selected for further analysis. Scale bar corresponds to 20 μm and applies to all images.

Mean, Median, and Otsu algorithms were selected as the most promising local thresholding algorithms, and were further compared by generating ROIs from binary images and overlaying the original image to assess segmentation accuracy. After thresholding and binarisation, lookup table was inverted, and particles were measured with

Analyse Particles to segment ROIs. Figure 12 shows the ROIs generated by using each of these three local thresholding methods overlaid against the raw image (step 5 in Figure 9).

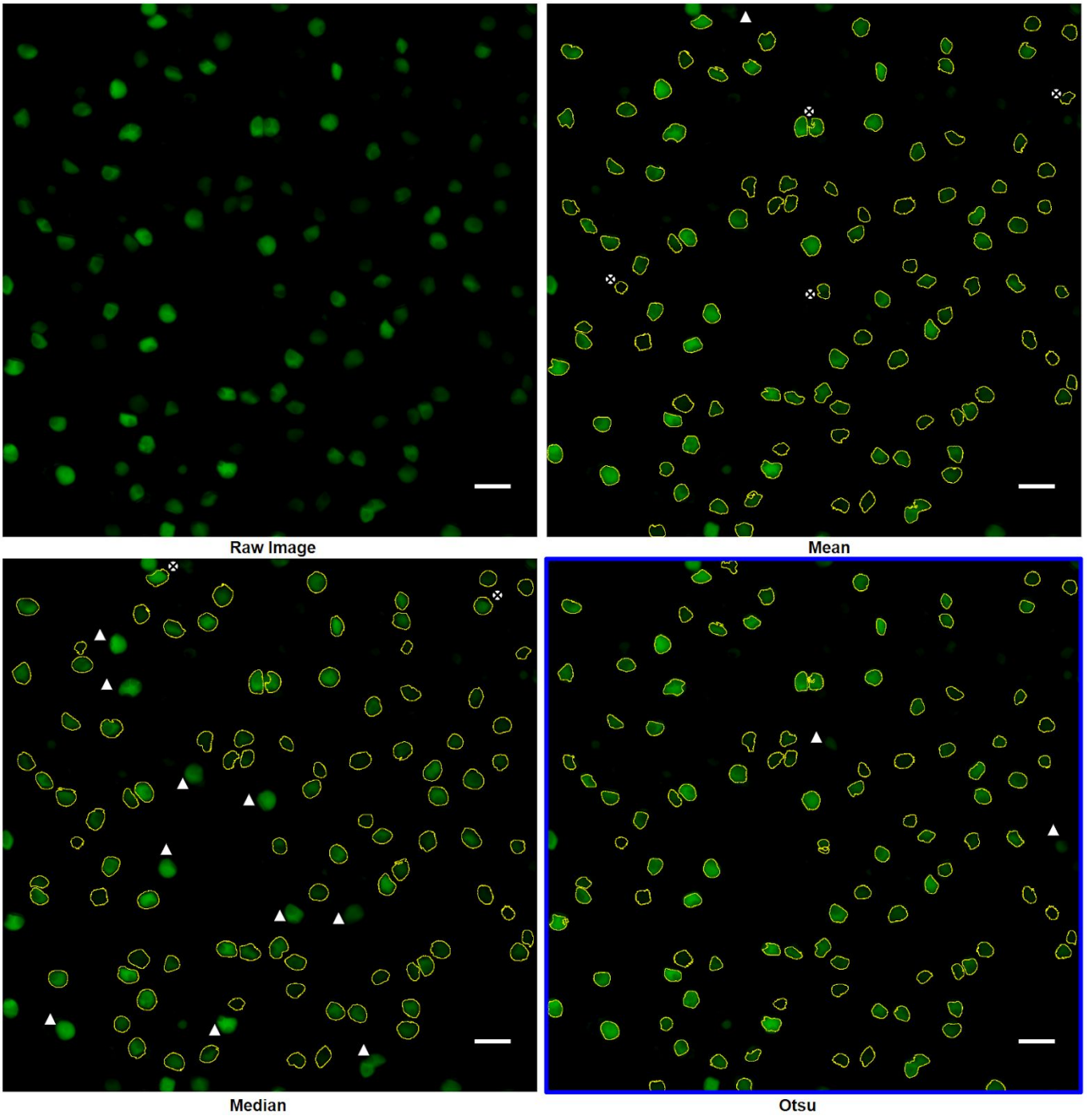


Figure 12: Segmented ROIs overlaid against original image. White triangles highlight missing objects which were identified with other algorithms, circle crosses show problematic segmentation, either by mismatching the underlying shape or segmenting an area not visible in the raw image. Labels correspond to the image above them. Blue outline indicates algorithm selected. Scale bar corresponds to 20 μm.

Of these three algorithms, each had strengths and weaknesses. The Mean algorithm highlighted multiple ROIs not visible in the original image, Median failed to segment ~12% nuclei clearly visible in the original image, and Otsu missed ~2% of nuclei in the original. Mean and Median also often segmented boundaries for nuclei extending beyond the boundary visible in the original image, which could distort mean fluorescence intensity values for segmented ROIs. The Otsu algorithm was selected on the basis of having the fewest mis-segmented ROIs. Otsu thresholding uses the characteristic bimodal distribution of pixel intensity to classify each pixel as a foreground or background pixel, then calculates an ideal threshold that minimises variance within the two classes [182]. Furthermore, it was reasoned that failing to capture a small percentage of the nuclei would bias the results of analysis less than improperly segmenting nuclei. Otsu fails to segment the nuclei with the lowest fluorescence intensity reaching the detector, which may cause this pipeline to overestimate average nuclear fluorescence intensity. However, this effect should apply equally to all transformed plants, and plants should still be comparable relative to each other.

4.4.3 Fluorescence intensity comparison

Once the segmentation pipeline was established, it was performed on images of thallus tissue from stably transformed plants from five independent transformation events each for *Agrobacterium*-mediated and biolistic transformation. ROIs generated by the pipeline described in the previous section were mapped back to the raw images and the mean intensity of each was measured. Images were acquired in one imaging session under identical conditions. 476 nuclei were measured from biolistic transformation, 568 from *Agrobacterium*-mediated transformation.

Data pooled by transformation type

Results are initially compared between two groups: plants transformed with biolistics and plants transformed with *Agrobacterium*. The mean of mean intensities was 43.8 (of 256, arbitrary units) for *Agrobacterium*-mediated transformants and 102.1 for biolistic transformants. The 25th and 75th percentiles for *Agrobacterium*-mediated transformants were 24.4 and 58.5 respectively, compared to 32.3 and 175.5 for biolistic transformants. Violin plots were used to visualise the distribution of mean intensity for the two methods of transformants, and the results are shown in Figure 13.

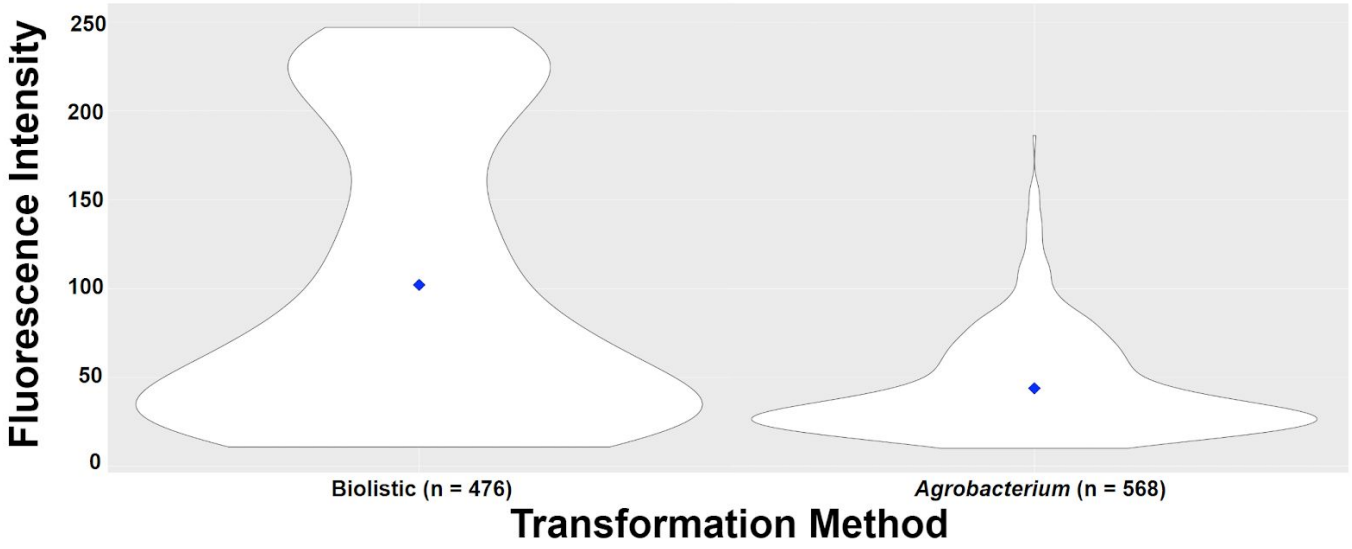


Figure 13: 8-bit mean intensity for nuclei transformed with two different methods. Blue diamond shows the mean of all points, the width of the plot shows a model of the underlying distribution based off a kernel density algorithm. The model distribution is truncated at the most extreme data points actually observed. In the raw data, microscope settings were established such that all values lie above 0 and below 255, preventing signal clipping.

Visual assessment of the violin plots suggests the data are not normally distributed, which was confirmed with the Shapiro-Wilk test of normality ($p < 2.2 \cdot 10^{-22}$ for each) [183]. The data are not normally distributed, hence , the Wilcoxon signed-rank test was used, a nonparametric method of comparing means of distributions [184]. The means differed significantly ($p < 2.2 \cdot 10^{-22}$).

Data for individual transformants

Results are here analysed by each individual independently transformed line, designated Bio1-5 and Agro1-5 after rank-ordering by median fluorescence intensity. Box plots for these are shown in Figure 14. The bimodal density distribution of Biolistic nuclei in Figure 13 is shown to be a result of differing mean fluorescence intensity in independently transformed plants.

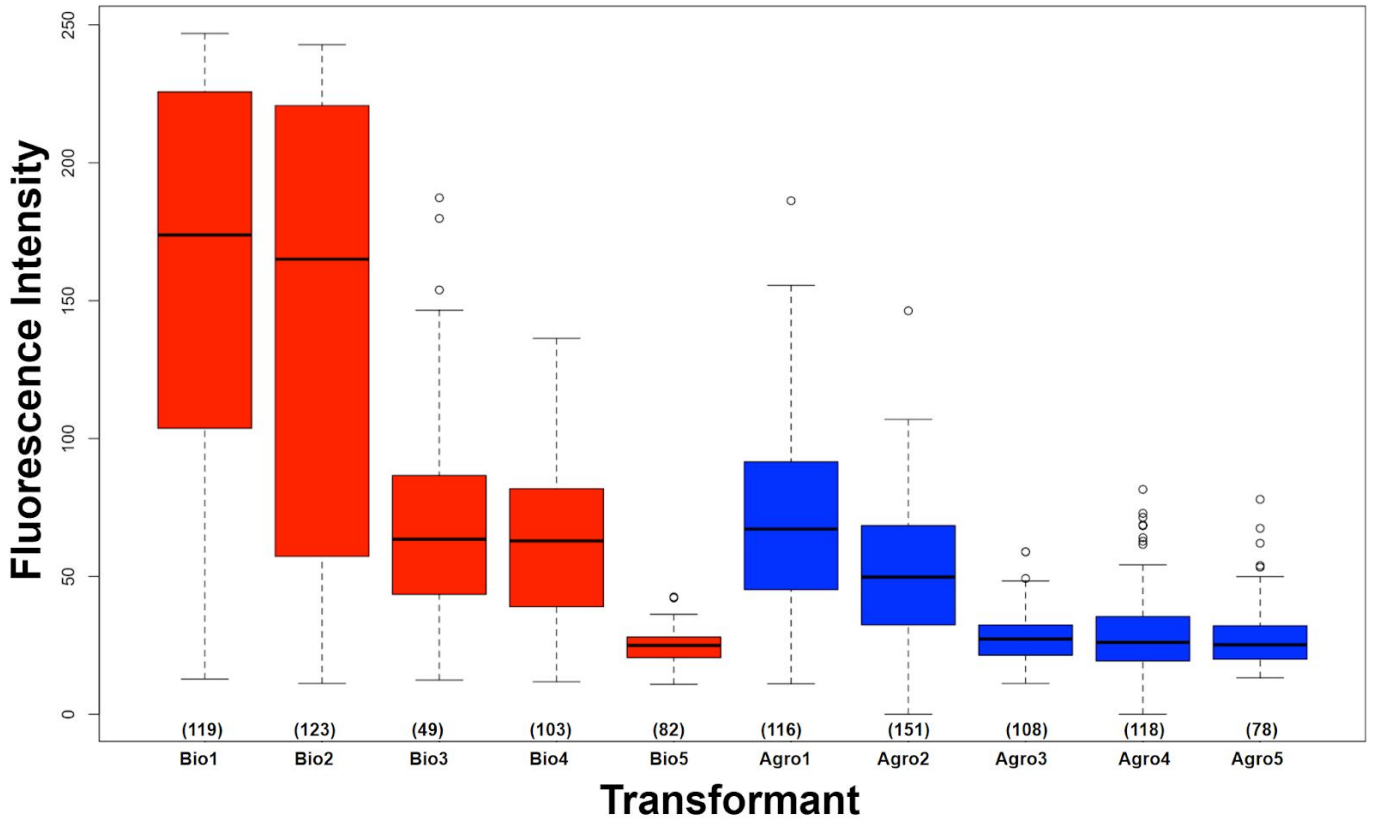


Figure 14: Distribution of mean nuclear intensities for 5 independently transformed plants of *Marchantia* transformed by biolistics and *Agrobacterium* with the plasmid L2_103, shown in red and blue, respectively. Line name designations were chosen in rank order by median mean nuclear intensity. Number in parentheses for each line shows the number of segmented nuclei. Fluorescence intensity is measured at 8-bit depth for a total of 256 relative values.

As can be seen from Figure 14, the higher average fluorescence intensity seen in biolistic transformation is mostly due to two transformant plants which had exceptionally high fluorescence intensity rather than a general trend towards higher expression. Data from 7 of the 10 transformants imaged failed Shapiro-Wilks normality test, so nonparametric methods were used to compare groups. Because a large number of groups are compared, Kruskal-Wallis test [185] was implemented with InStat, which showed significant differences between transformants ($p < 0.0001$). Dunn's post test [186] was used to calculate which transformants differed significantly from each other, results are shown in Appendix C, Table 11.

4.4.4 Discussion

These data show that average fluorescence intensity and therefore average fluorescent protein expression is higher with biolistic transformation than *Agrobacterium*-mediated transformation, at least for plasmid L2_103 (see Table 1) in thallus tissue of *Marchantia*

polymorpha. However, there is also higher variability between different transformation events. Biolistic nuclear transformation is recommended for applications in which a large space of transgene expression levels is useful, and *Agrobacterium*-mediated transformation is recommended for reliable generation of a particular expression level.

4.5 Post-bombardment selection

To the author's knowledge, all experiments with transplastomic *Marchantia* have used Spectinomycin at a concentration of 500 μg mL^{-1}, but there are no published reports comparing tolerance of wild type and transformed spectinomycin resistant plants, raising the possibility that a different concentration might be more useful for selection. In particular, a higher concentration was desired to reduce the time required to differentiate transformants and to achieve homoplasmy. A wide range of concentrations was tested for several *Marchantia* tissue types for wild type and homoplasmic transplastomic plants obtained from Christian Boehm. Because the normal timeline for initial selection is 4-5 weeks, plants were maintained on selection under constant light for 34 days then imaged.

4.5.1 Selection results

Wild type 5-10 day sporelings are the normal input tissue for plastid transformation, and growth after 34 days on selection is shown in Figure 15.

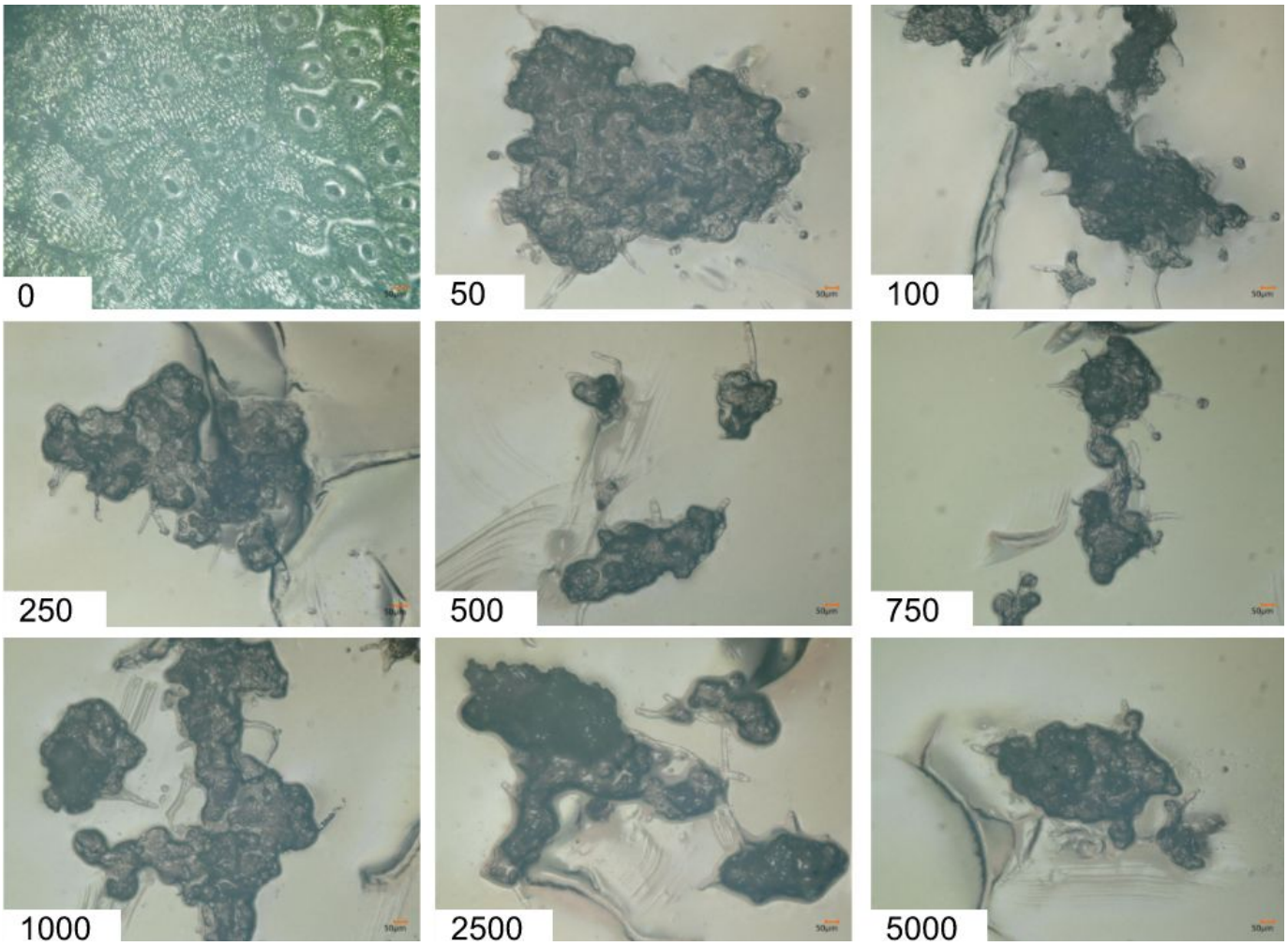


Figure 15: WT sporelings spectinomycin resistance. Growth of Wild type Cam1 sporelings after 34 days on selection. Numbers in the bottom left indicate the concentration of Spectinomycin in μg mL^{-1}. The default selection concentration is 500 μg mL^{-1}. However, sporelings were efficiently bleached by concentrations as low as 50 μg mL^{-1}. Scale bar corresponds to 50 μm.

Gemmae (see section 1.2.1) are an appealing target for transformation due to their isogenic nature and relatively consistent shape, and *Agrobacterium*-mediated nuclear transformation of gemma is now possible [187]. Gemmae were transiently nuclear-genome transformed via biolistics in this project, suggesting that plastid transformation of gemmae may be possible in the future. These data may also be useful for the development of Spectinomycin as a nuclear selectable marker. Wild type gemmae were completely bleached by 50 μg mL^{-1} spectinomycin (see Appendix B, Figure 34) whereas pCS Clo*B gemmae grew at concentrations up 250 μg mL^{-1}, and one of three gemmae survived and grew slightly at 500 μg mL^{-1} (see Appendix B, Figure 35).

Agrobacterium-mediated transformation of *Marchantia* thallus has also been demonstrated, so thallus Spectinomycin tolerance was assessed for wild type and pCS Clo*B plants. In all cases, a thallus fragment approximately 3 mm across was cut from a healthy plant and grown on selection for 34 days before imaging. The results for wild type plants are shown in Figure 16. PCS Clo*B plants are homoplasmic, and no longer require selection to be maintained. Hence, some were available for transfer from Petri dishes containing Spectinomycin 500 µg mL^{-1} while others were taken from Petri dishes without antibiotic. This allows for a comparison of plants with the same genetic levels of resistance but different degrees of antibiotic acclimation. pCS Clo*B thallus maintained on spectinomycin is shown in Figure 17, pCS Clo*B thallus maintained without antibiotics is shown in Figure 18.

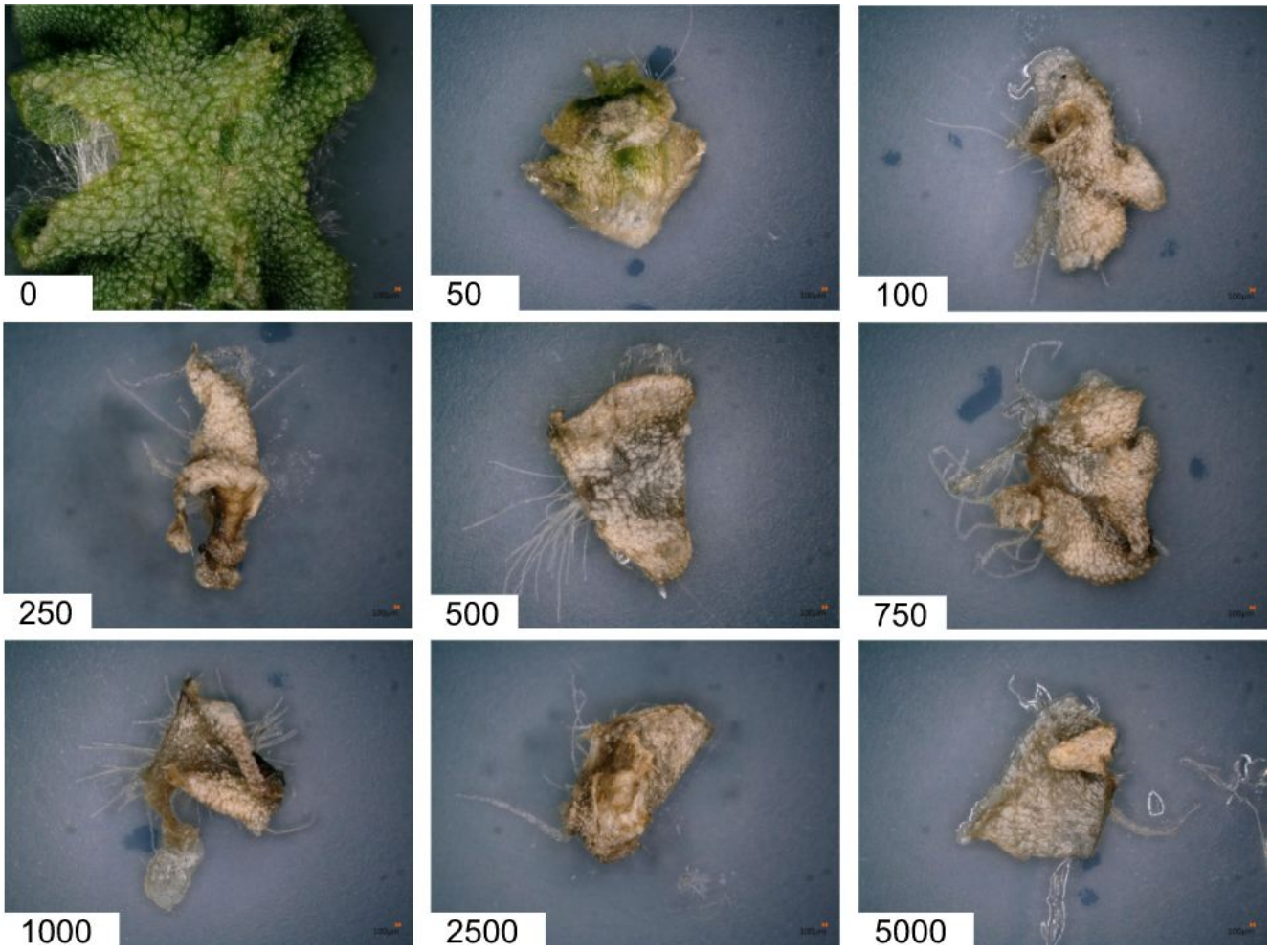


Figure 16: WT thallus spectinomycin resistance. Some green tissue is present at 50 µg mL^{-1} and 100 µg mL^{-1}. Numbers in lower-left of each image represent concentration of spectinomycin in µg mL^{-1}. Images taken after 34 days of selection. Scale bar corresponds to 100 µm.

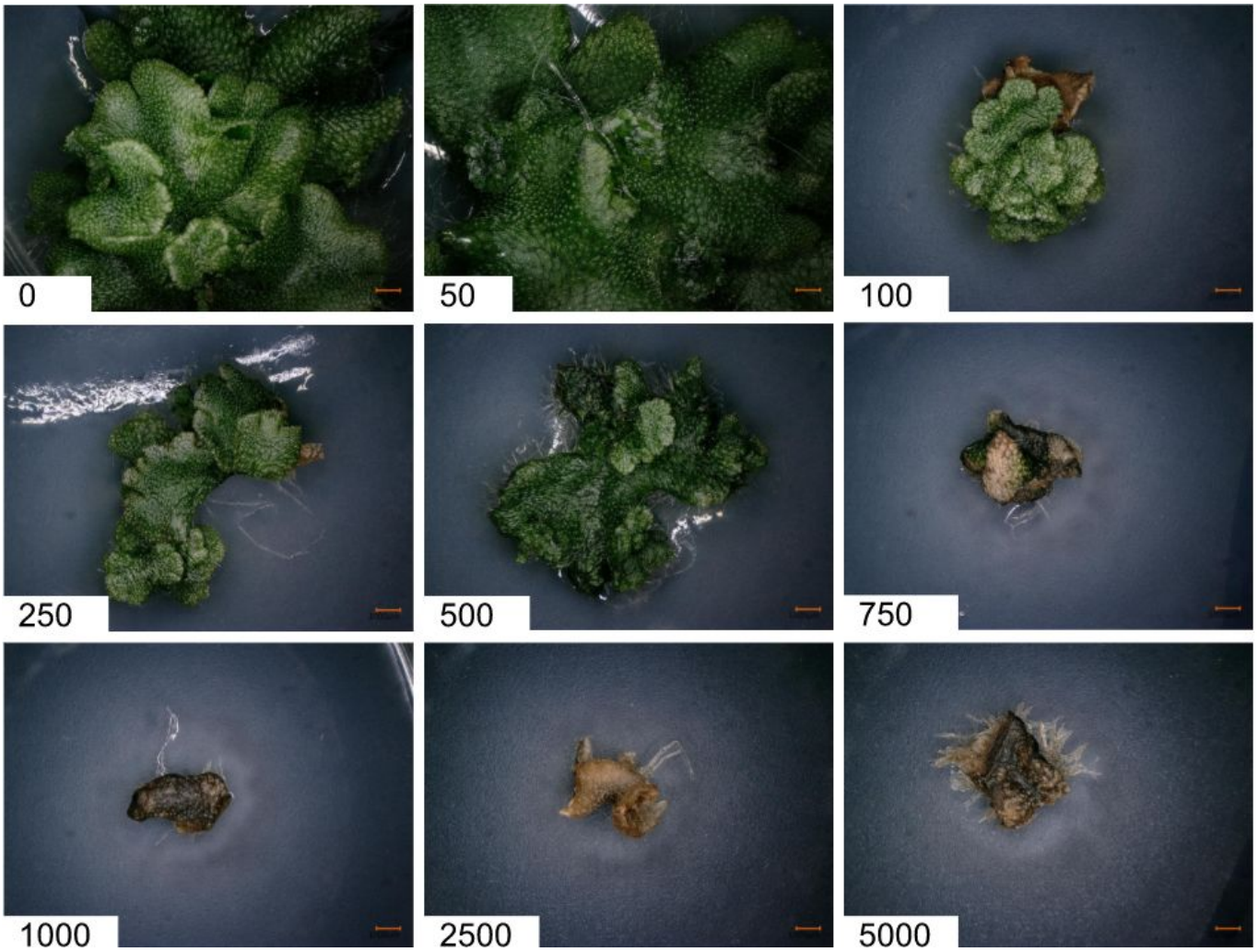


Figure 17: pCS Clo*B thallus maintained without selection.Numbers in lower-left of each image represent concentration of spectinomycin in μg mL^{-1}. Plants transformed by Christian Boehm. Images taken after 34 days of selection. Scale bar corresponds to 1000 μm.

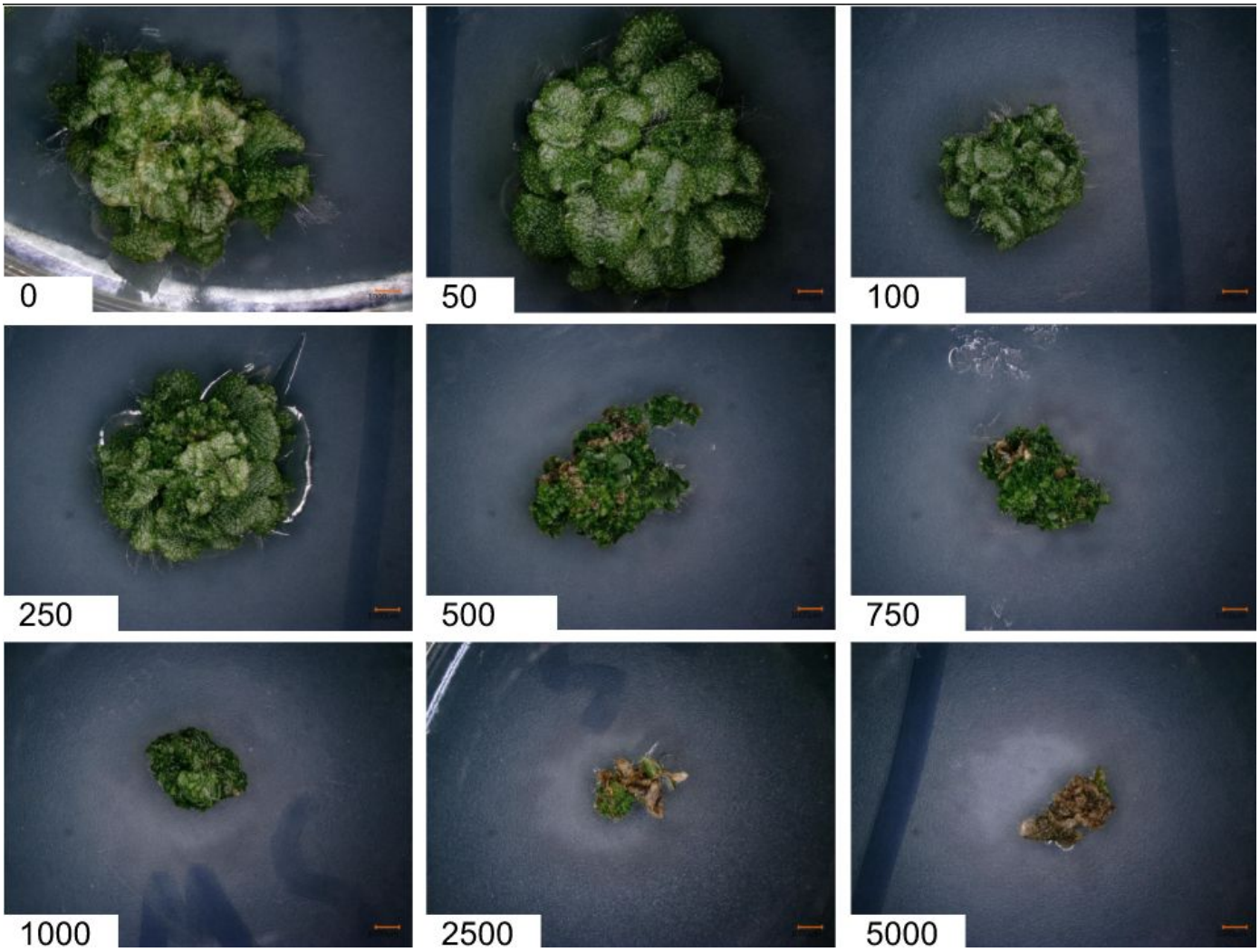


Figure 18: pCS Clo*B thallus from plants maintained on spectinomycin 500 µg mL^{-1}. There is growth up to 1000 µg mL^{-1}, and even at 5000 µg mL^{-1} small sections of living tissue remain. Numbers in lower-left of each image represent concentration of spectinomycin in µg mL^{-1}. Transformed by Christian Boehm. Images taken after 34 days of selection. Scale bar corresponds to 1000 µm.

4.5.2 Discussion

Selection was very efficient on gemmae and sporelings even at 50 µg mL^{-1} spectinomycin, the lowest concentration tested. Resistant plants from previous experiments demonstrate good resistance at 500 µg mL^{-1}, and those which were maintained on selection show resistance at much higher concentrations. Among the thallus experiments, the difference observed between pCS Clo*B thallus fragments maintained on Spectinomycin and those maintained without antibiotics is interesting. Plants which had been maintained on Spectinomycin demonstrated tolerance at more than three times the highest concentration of plants maintained without antibiotics. This is a surprising find in light of the fact that the plants were

previously confirmed to be homoplasmic. These data indicate some degree of environmentally-mediated acclimation in addition to the genetic component of spectinomycin resistance mediated by *aadA*.

For sporelings, these validate 500 μg mL^{-1} as a reasonable selection concentration - resistant transformants are still recovered at an acceptable frequency, but plants are being exposed to at least ten times the lethal concentration, minimising the odds of false positives. However, the complete bleaching of wild type sporelings at much lower concentrations suggests less stringent selection could be used. A tiered approach could also be adopted where a low concentration is used initially to eliminate untransformed plants, then a higher concentration is used to encourage homoplasmy during the later stages of the selection process.

The data generated for thallus and gemmae may prove useful for determining selection conditions for transformations of these tissue types, either through *Agrobacterium*-mediated transformation or biolistic transformation of the nuclear or plastid genome. Because nuclear transformation usually results in lower expression of foreign protein (as discussed in section 1.3.2), a lower concentration spectinomycin selection is advised.

Chapter 5: Transplastomic gene expression

At the beginning of this project, the only available construct for fluorescent protein expression from the *Marchantia* plastome was pCS Clo*B, the construct used in Boehm *et al.* 2016 [112]. However, upon the assembly of the Cam-1/2 plastome, it became clear that the homologous regions used in this plasmid lacked complete homology to the wild type *M. polymorpha* plastome, which may have contributed to the irreproducibility of previous transformations protocols in the laboratory and the lower efficiencies. In particular, the homologous regions used included 4 small insertions or deletions. Of these, one would cause a frameshift mutation in *psbC,* a gene critical for photosystem function. Transformed plants were never sequenced, but this possible mutation could explain the lower fluorescence intensity shown by this construct compared to L2_355, a nearly equivalent construct produced in this project with corrected homologous regions.

A modular toolkit of parts with characterised properties is essential for practical applications of plastid transformation technology. This project focused on the development and characterisation of promoter 5' UTR combinations that result in a range of protein accumulation levels. Such genetic parts are a necessary prerequisite to metabolic engineering applications, and the high degree of plastome conservation among land plants suggests the knowledge gained will readily transfer to crop species.

5.1 Loop Assembly

Type IIS assembly is a relatively new method of DNA assembly that offers substantial benefits over traditional restriction enzyme cloning. The two principal benefits are the use of a one-pot digestion-ligation reaction and the ability to generate arbitrary 'sticky overhangs'. The digestion-ligation reaction eliminates many of the fragment purification steps required in traditional cloning, and the arbitrary nature of the overhangs makes the technique 'quasi-scarless' and has facilitated the proliferation of DNA assembly standards that allow for simple sharing and reuse of parts. Loop Assembly is a particular Type IIS assembly system developed by Bernardo Pollack and Fernan Federici which was used in this project.

5.1.1 Type IIS assembly systems

Restriction enzymes are endonucleases that cleave both pentose phosphate backbone strands of dsDNA in a predictable, sequence-specific fashion. They form the basis for most forms of cloning and DNA assembly (reviewed in [188]). Some cleave each backbone at the same location, resulting in ‘blunt ends’, whereas others cleave one backbone several base pairs offset from the other backbone, resulting in ‘sticky overhangs’. In traditional cloning, these sticky overhangs are comprised of a characteristic sequence, such as $TTAA$ for the enzyme EcoRI [189]. Two DNA sequences cut with the same enzyme have complementary sticky overhangs, which allows for hybridisation by hydrogen bonding if they are mixed under the proper conditions. A ligase can then join the pentose phosphate backbones, resulting in a new sequence comprised of the two sequences joined in series, with a small ‘cloning scar’ whose sequence is determined by the restriction enzyme used. Traditional cloning workflow to join two sequences usually consists of using PCR to add the same restriction enzyme recognition site to each sequence, digestion of both sequences in separate reactions, fragment purification in separate reactions, and mixing the appropriate fragments for ligation. Because only a limited number of restriction enzymes are known, design is constrained by cleavage sequence, resulting in cloning scars.

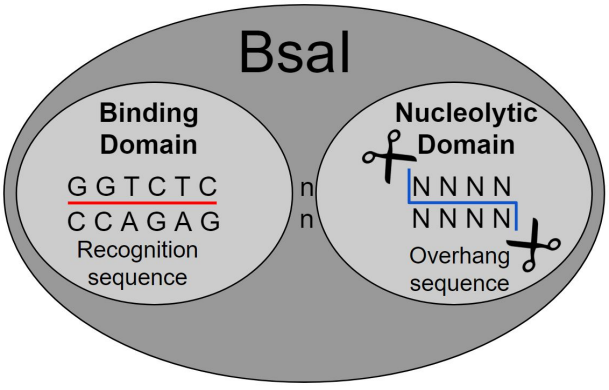


Figure 19: Schematic of a Type IIS restriction enzyme. Here BsaI is used as an example. It binds the sequence underlined in red, and hydrolises both strands of the pentose phosphate backbone along the line shown in blue.

Type IIS restriction enzymes have the special property that their recognition sequence doesn’t overlap with their endonucleolytic function. Instead, they consist of a sequence-specific binding domain and a non sequence-specific nucleolytic domain, which always digests DNA a fixed nucleotide distance from the binding domain recognition sequence, as shown in Figure 19.

In contrast to traditional cloning methods, Type IIS assembly allows for the generation of arbitrary ‘sticky overhangs’, the sequence composition of which is not determined by the restriction enzyme used. As a result, DNA fragments can be joined in a specific order using a simple three step workflow diagrammed in Figure 20: type IIS primer design, traditional PCR and purification, and the one-pot, one-step, digestion-ligation reaction.

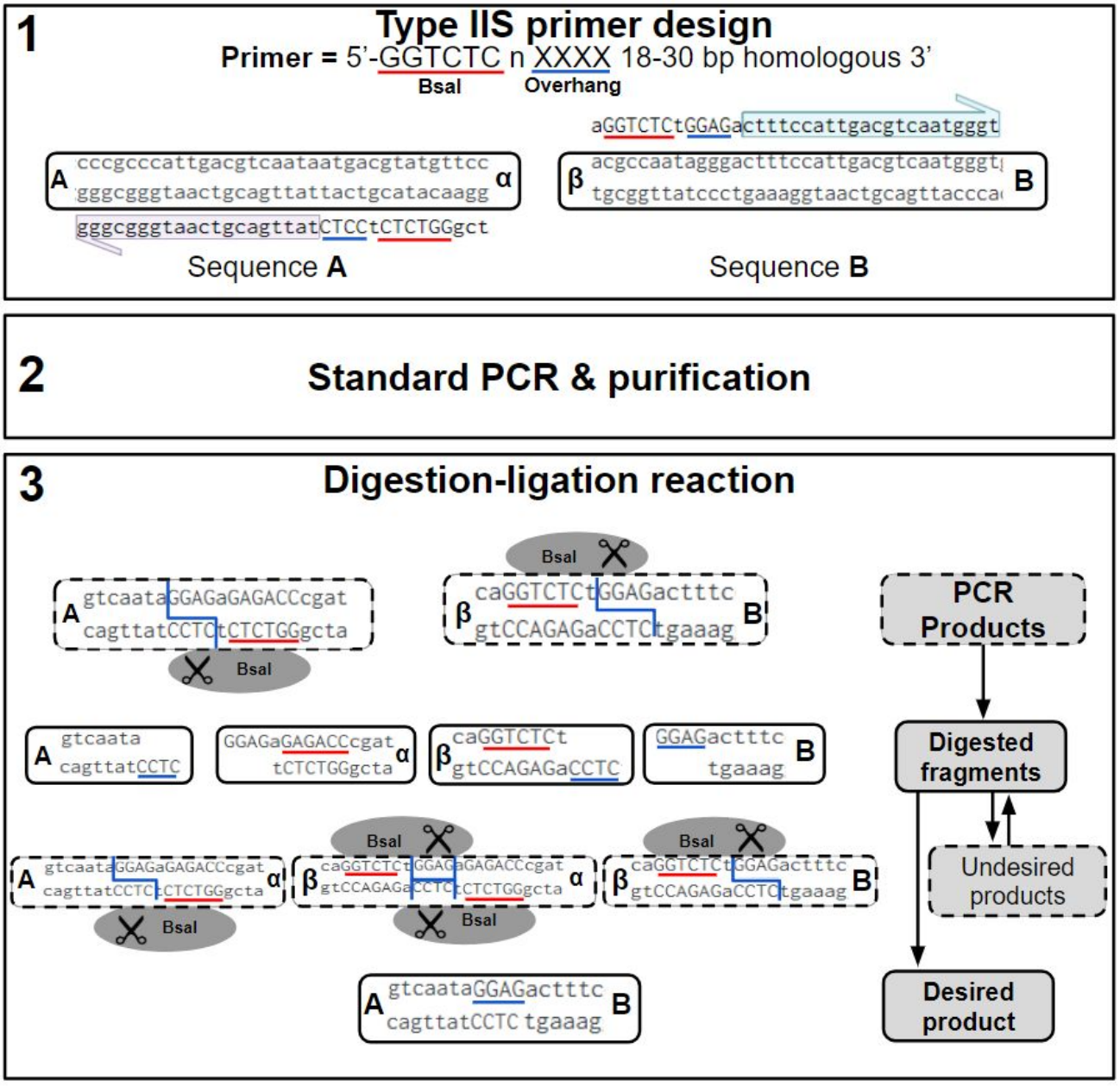


Figure 20: Overview of Type IIS assembly. Here sequence A and sequence B are joined in the order AB.Each sequence is amplified by PCR with primers that contain two extra regions on the 5’ end: a Type IIS enzyme recognition site (here BsaI, recognition sequence shown as red underline) and an arbitrary 4 base ‘sticky overhang’ (here GGAG, blue underline). Standard PCR and PCR product purification is then performed, resulting in dsDNA with these additional regions, designated α and β. Finally, a one-pot, one-step, digestion-ligation reaction is performed. Ligation is thermodynamically favored, and all

undesired ligation products are re-digested, resulting in efficient yield of the desired product.

The use of Type IIS restriction enzymes eliminates the sequence specificity of sticky overhangs, which allows for the design of any cloning scar. Because any scar can be used, Type IIS assembly is called “quasi-scarless.” Many Type IIS enzymes generate 3 or 4 base pair sticky overhangs, for a total of 64 or 256 possible overhangs which can be used. Efficient Type IIS modular cloning systems include Golden Gate [190], MoClo [191], Golden Braid [192], and Loop Assembly [161].

Type IIS assembly systems have facilitated creation of a common standard of sticky overhangs used for particular types of genetic parts, referred to as common syntax [193]. Common syntax means the overhangs used in DNA assembly are determined by the biological function of the DNA. For example, promoters are always cloned with a 5’ overhang of $GGAG$. The common syntax is illustrated in the top three lines of Figure 21. The benefit of such standards is that parts can be easily exchanged and reused. They can also be assembled in a combinatorial fashion - for example, with two promoters, two 5’ UTRs, two coding sequences, and two 3’ UTR/terminators, 16 unique synthetic constructs can be rapidly assembled because parts are natively cross-compatible with each other. This proliferation of standards for exchange of parts is an integral component of the Synthetic Biology movement, which is bringing concepts like standardisation and modularity from engineering disciplines to molecular biology [194].

5.1.2 Loop Assembly

Loop Assembly is a recursive Type IIS assembly system that alternates between two restriction enzymes and two classes of vectors [161]. The restriction enzymes are BsaI which generates four base 5’ overhangs (as shown in Figure 19) and SapI which generates three base 5’ overhangs. Vectors are numbered sequentially starting from 0 by complexity, and are divided into two classes, even and odd. Level 0 parts (line 1 or 2 in Figure 21) are created by digestion with SapI with BsaI recognition sites flanking the part. BsaI can then be used to bring a complete set of level 0 parts (spanning the entirety of line 1 or 2 in Figure 21) into

a Level 1 part, termed a transcriptional unit (line 3 in Figure 21). SapI is used to bring four Level 1 parts into a Level 2 part, the largest used in this work.

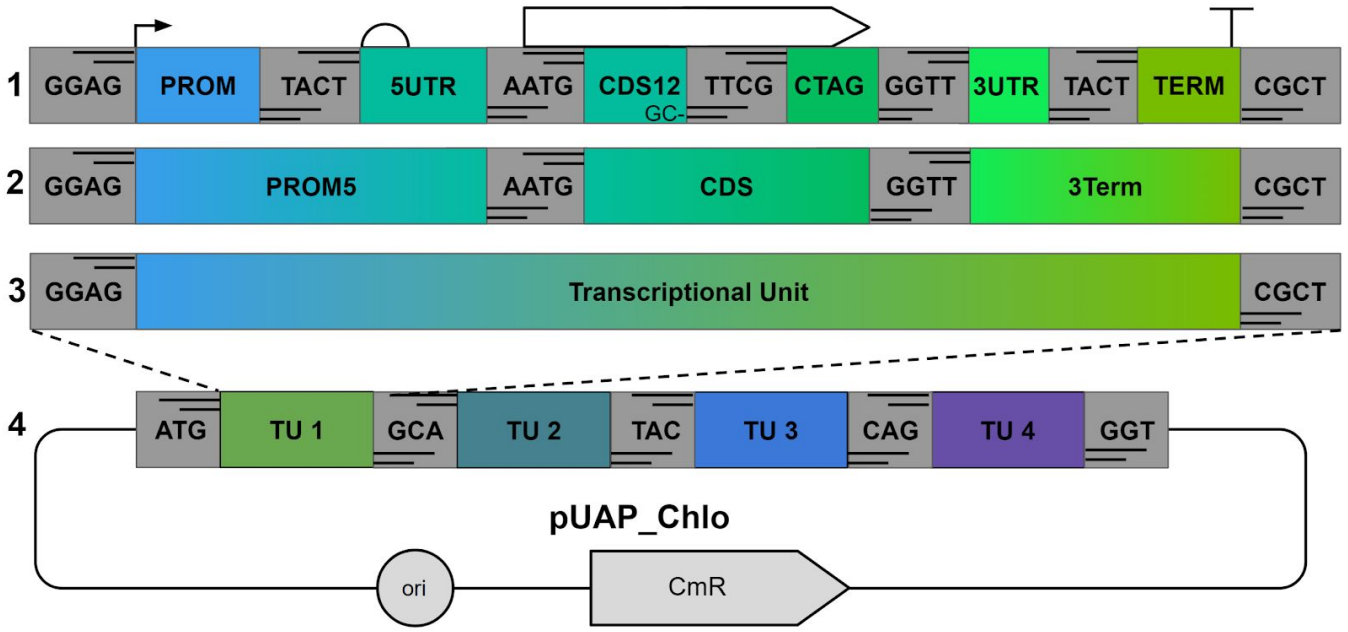


Figure 21: Common syntax and loop overview. Line 1 shows common syntax overhangs as defined in [193], overlaid with Synthetic Biology Open Language Visual symbols for promoter, ribosome binding site, coding sequence, and terminator [195]. Line 2 shows the same with compound parts to reduce complexity. Parts of the size of either of these lines are designated Level 0 parts. Line 3 shows a Level 1 part, or full transcriptional unit. Line 4 shows how four transcriptional units are combined in an ordered fashion to construct a Level 2 part TU 1-4 here correspond to positions 1-4 in Table 1. The backbone pUAP_Chlo was designed and assembled by Eftychios Frangedakis.

The recursive nature of Loop Assembly means that Level 2 parts can be digested with BsaI to combine 16 transcriptional units in a Level 3 part, which could be digested again with SapI to create a Level 4 part with 64 transcriptional units.

5.2 Construct design

A focus of this project was laying a scalable foundation for future engineering efforts, which began with a modular and reconfigurable genetic toolbox. In preparation for possible future experiments with super-transformation, four intergenic regions for insertion were selected based on what had been successfully used in other land plant species (see Figure 4). For each insertion location, two 1.5 kb homologous regions were cloned from the wild type Cam1 plastome, one on each side of the insertion site. These homologous regions were then standardised as transcriptional units 1 and 4 for level 2 assembly based on

the Loop Assembly standards (see Figure 21 for overhang standards and Table 1 for construct details) [161]. Level two plasmids contain four transcriptional units, leaving two open spaces for transgene insertion. Hybrid level vectors would allow for a higher number of transcriptional units to be inserted if desired.

Two *M. polymorpha* plastome codon optimised fluorescent proteins were prepared as Loop-compatible parts, *mturq2cp* and *mVenuscp*, coding for the fluorescent proteins mTurquoise2 [196] and mVenus [197]. These fluorescent markers were selected for their spectral separation, which allows for accurate analysis of each individually in future experiments where both might be integrated. For this project, only constructs expressing *mturq2cp* were used. The final position in the level 2 plasmids is filled with the selectable marker *aadA*, which confers resistance to Spectinomycin. The tRNA-G/tRNA-M intergenic insertion site and the RNA polymerase operon insertion site were prioritised for level 2 construct assembly. For the tRNA-G/tRNA-M insertion site, a variety of promotor/5' UTR constructs were assembled in the *mturq2cp* transcriptional unit. RNA operon insertion site was chosen in an effort to obtain expression of transgenes from a polycistronic mRNA comprised primarily of the native RNA operon, but as proof of principle for that insertion location constructs using the same promoter and 5' UTR elements as were used in the tRNA-G/tRNA-M constructs.

L2_360 and L2_361 (see Table 1) were generated and successfully transformed recently and are not included in the primary analysis. However, preliminary analysis with LSCM indicated that L2_360 produces high levels of cyan fluorescence, as shown in Figure 22.

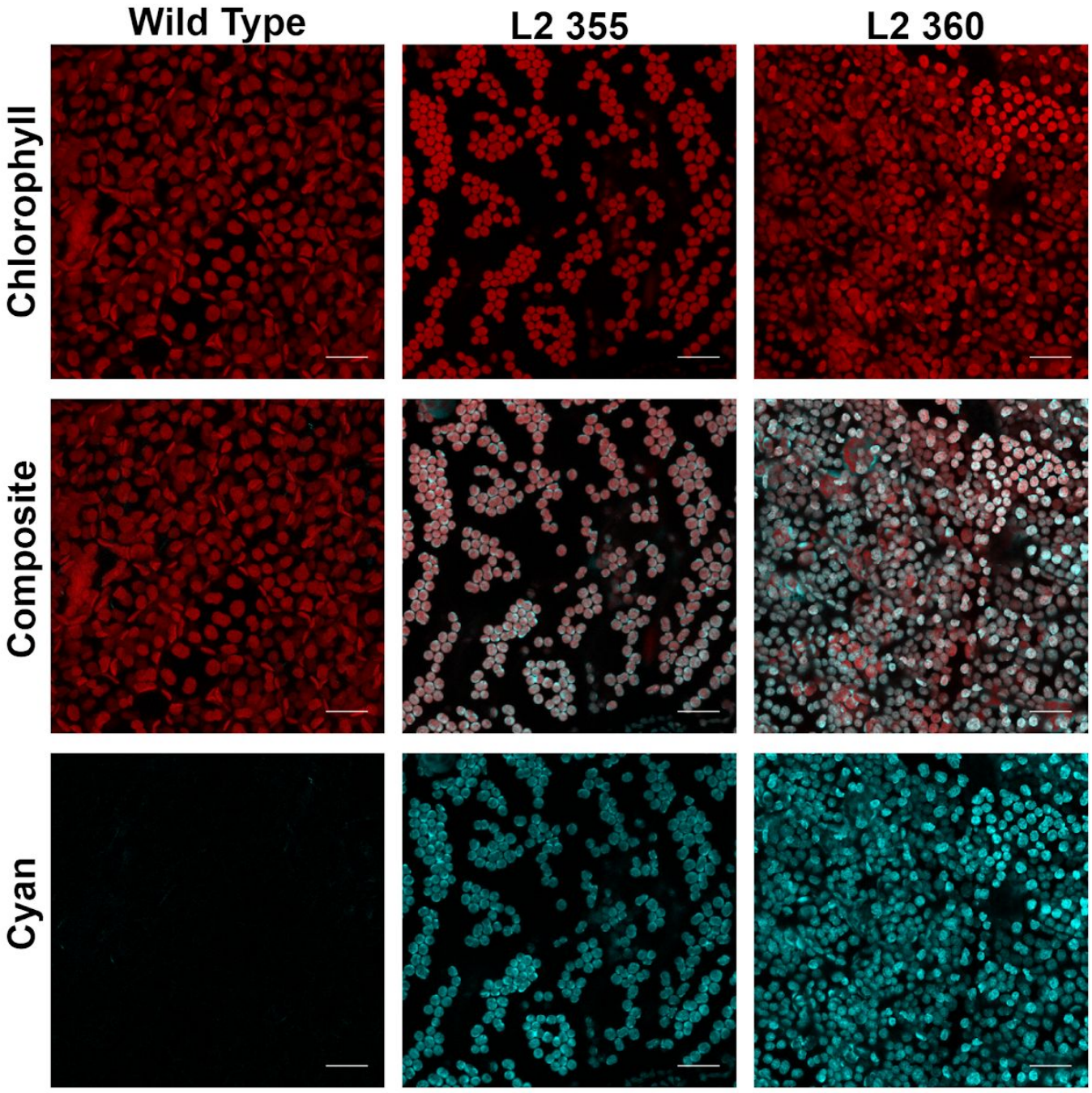


Figure 22**:** Preliminary demonstration of fluorescence in L2_360 as compared to WT plants and L2_355. Instrument settings held constant for all images. Scale bar corresponds to 20 µm.

5.3 Fluorescent protein expression

Fluorescent proteins derived from GFP have been a mainstay of molecular biology research since discovery of the protein in 1962 [198] , and have proven to be useful for a wide variety of experimental techniques [199]. The generation of plastome-expressed fluorescent markers of various strengths was a major focus of

this thesis, both for their instrumental value as markers and for insight into the genetic requirements for high and low protein accumulation in transplastomic *Marchantia*. Here three new promotor-5' UTR combinations are compared against wild type plants and homoplastic PCS Clo*B plants, the fluorescently labeled plants first generated by Boehm *et al.* [112].

Transplastomic plants were grown without selection until the generation of gemma (see section 1.2.1). Three gemmae each per gemma cup were imaged with LSCM for quantification of the fluorescent reporter gene expression. The overall rationale and strategy is the same as outlined in section 4.4.1, though the specific details of the image analysis pipeline are slightly different.

5.3.1 Image analysis optimisation

The pipeline optimised for nuclear fluorescence was initially attempted on these images, but produced unsatisfactory results. In particular, the process failed to segment the majority of the plastids, whether attempted with the chlorophyll autofluorescence channel or the cyan channel, as shown in Figure 23.

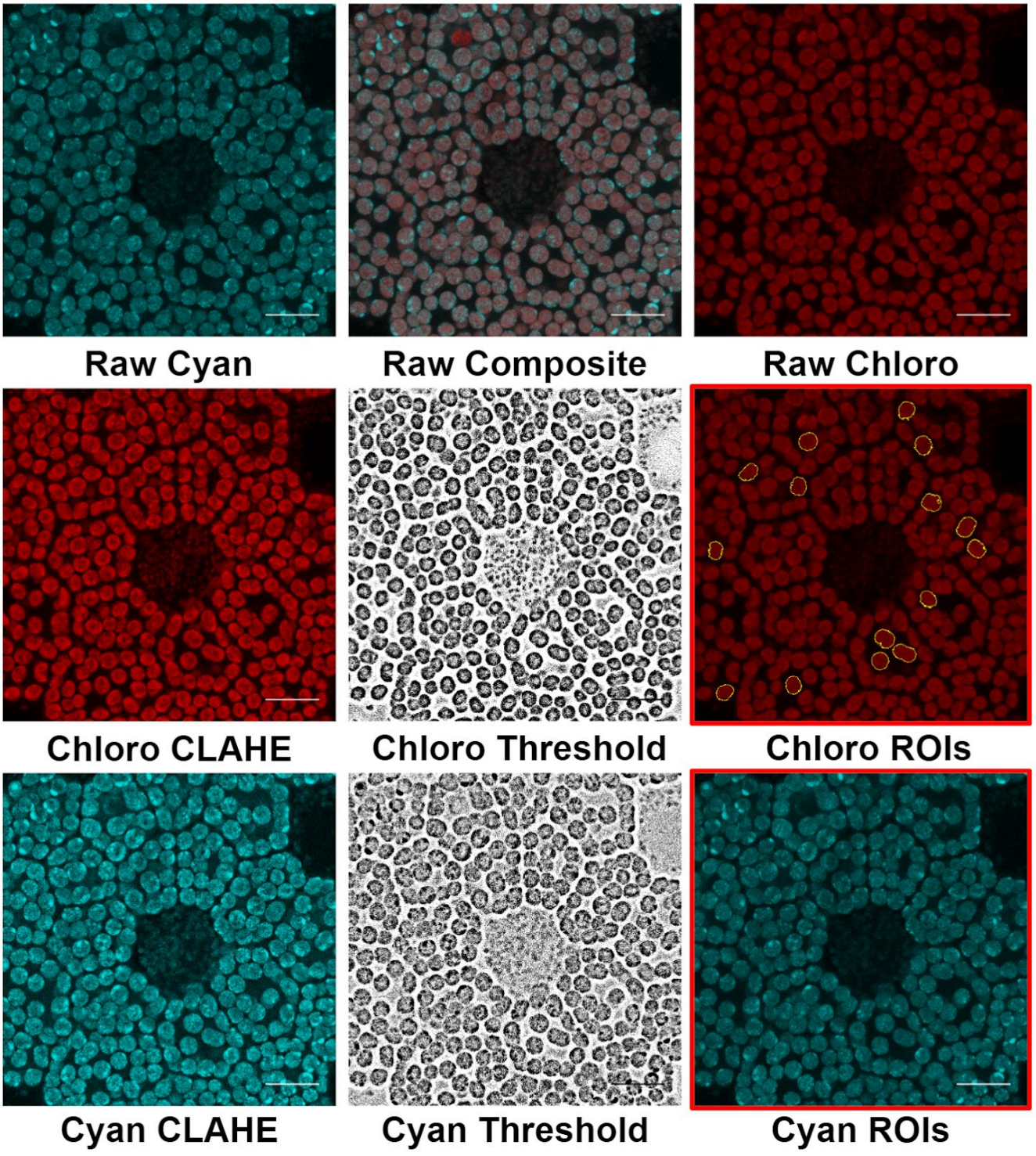


Figure 23: Results of nuclear segmentation pipeline applied to plastid fluorescence images. Top row shows the raw images, middle row shows the chlorophyll autofluorescence channel after local contrast enhancement, local thresholding, and ROIs mapped against the raw image. Bottom row shows cyan fluorescent channel after local contrast enhancement, local thresholding, and ROIs mapped against the original image. Red outline indicates segmented images, neither of which was seen as acceptable. Scale bar corresponds to 20 μm.

This micrograph is of a plant transformed with L2_355, and will serve as the example image for optimising the pipeline. All images in this dataset image are maximum projection LSCM micrographs comprised of two 8-bit channels corresponding to cyan fluorescence and chlorophyll autofluorescence, pseudocoloured blue and red, respectively. They have a resolution of 1024 by 1024 pixels, 15 Z slices, $\Delta Z < 0.8$ µm. The poor performance of the nuclear segmentation pipeline on the plastid fluorescence test image necessitated development of a new pipeline, diagrammed in Figure 24. The optimised pipeline consists of image acquisition, channel separation, image smoothing, local thresholding, Watershed, mapping ROIs back to raw image, and measurement and analysis of ROI characteristics.

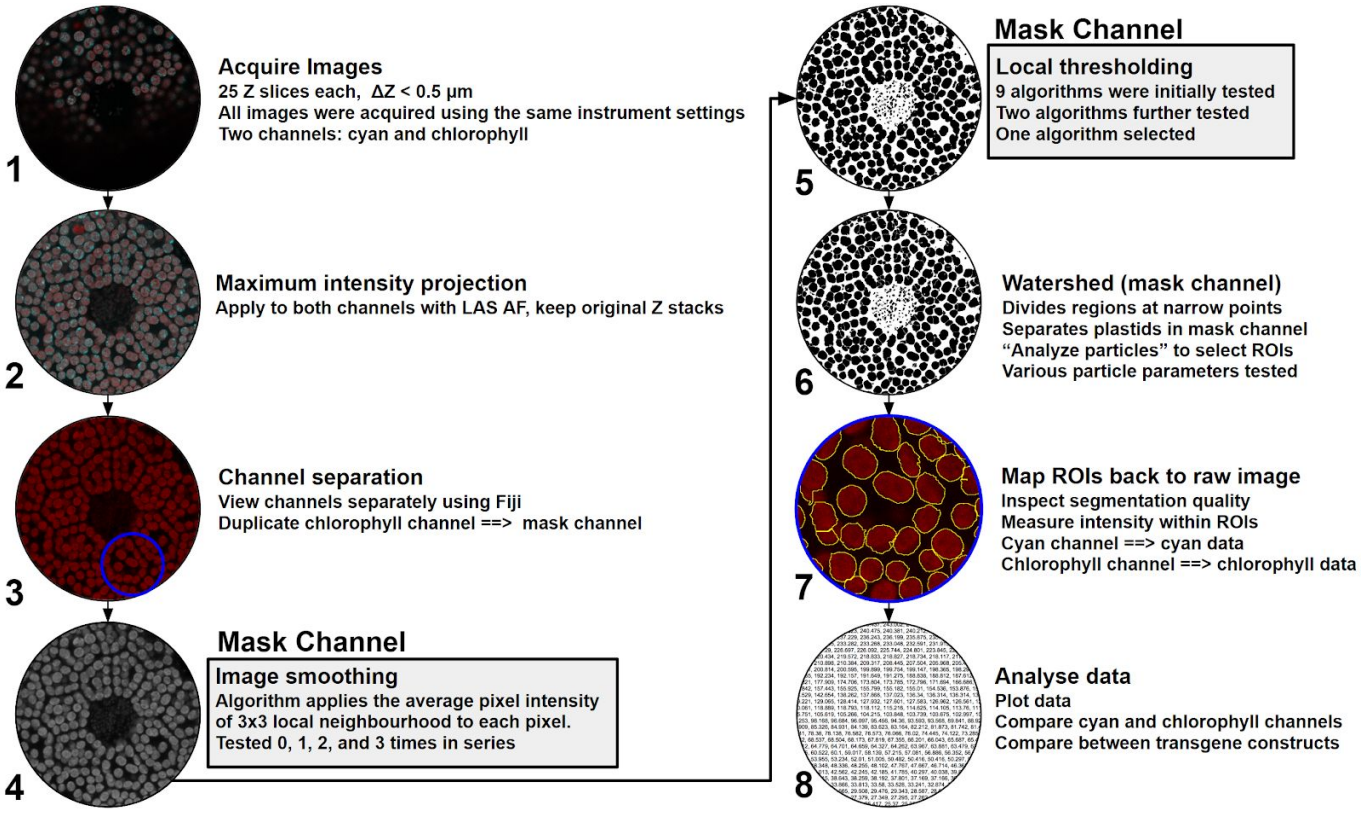


Figure 24: Schematic of plastid segmentation pipeline. Steps 4-6 occur on a duplicate channel which is only used for segmentation and is not directly measured. Step 7 is shown at higher magnification for illustration purposes only, region shown corresponds to the blue circle in step 3.

Experimentation with a variety of segmentation pipelines available through ImageJ functions and plugins revealed the noise in the dataset to be a major limiting factor in analysis, and as such the CLAHE algorithm which increases local noise was eliminated from the pipeline. The signal to noise ratio was higher for the

chlorophyll autofluorescence channel than the cyan channel, so it was used for all further segmentation attempts. The binary outputs of the nine auto local thresholding algorithms in available in Fiji are shown in Figure 25.

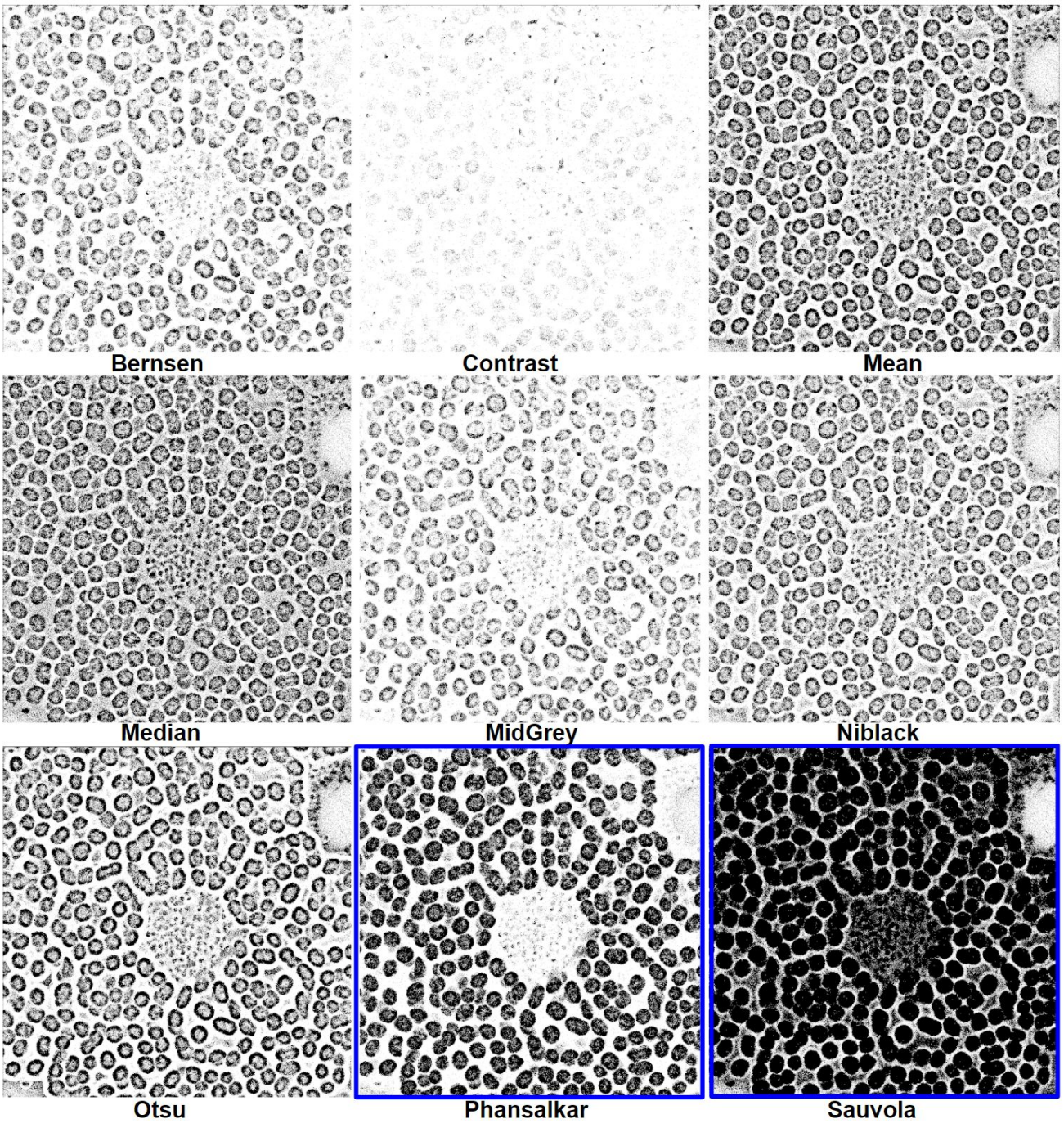


Figure 25: Local thresholding algorithms on the plastid fluorescence micrographs, chlorophyll channel. The 'salt and pepper' thresholding problem is apparent for most algorithms, including Otsu. Blue outline indicates algorithms selected for further analysis.

Thresholded images from every algorithm except Sauvola have a substantial number of blank pixels within plastids, which prevented automatic particle

detection with “Analyze particles”. This prevents successful segmentation because each region of interest is broken into multiple areas. Edge definition and hole-filling functions were explored, but did not resolve the issue. Global thresholding algorithms were also attempted, but none generated improved results.

Sauvola and Phansalkar local thresholding algorithms were selected as the most promising candidate algorithms for plastid segmentation. Sauvola thresholding was originally designed to threshold scanned text documents for archival purposes [200]. It functions by tiling small windows across the image to classify the image into regions of sharply defined edges (text, in the original implementation) and regions of poorly defined edges (non-text, in the original implementation). Areas identified as text have thresholds set in a small local region whereas regions identified as non-text are thresholded based on a global algorithm. Phansalkar thresholding is a modification of Sauvola thresholding that has higher sensitivity for low values of local pixel intensity standard deviation, and was developed specifically for segmentation of fluorescence micrographs [201].

Using these thresholds on raw images for ROI selection resulted in problems due to the small scale spatial noise, so images were pre-filtered with the Smooth function and thresholded with the two candidate algorithms, as shown in Figure 26.

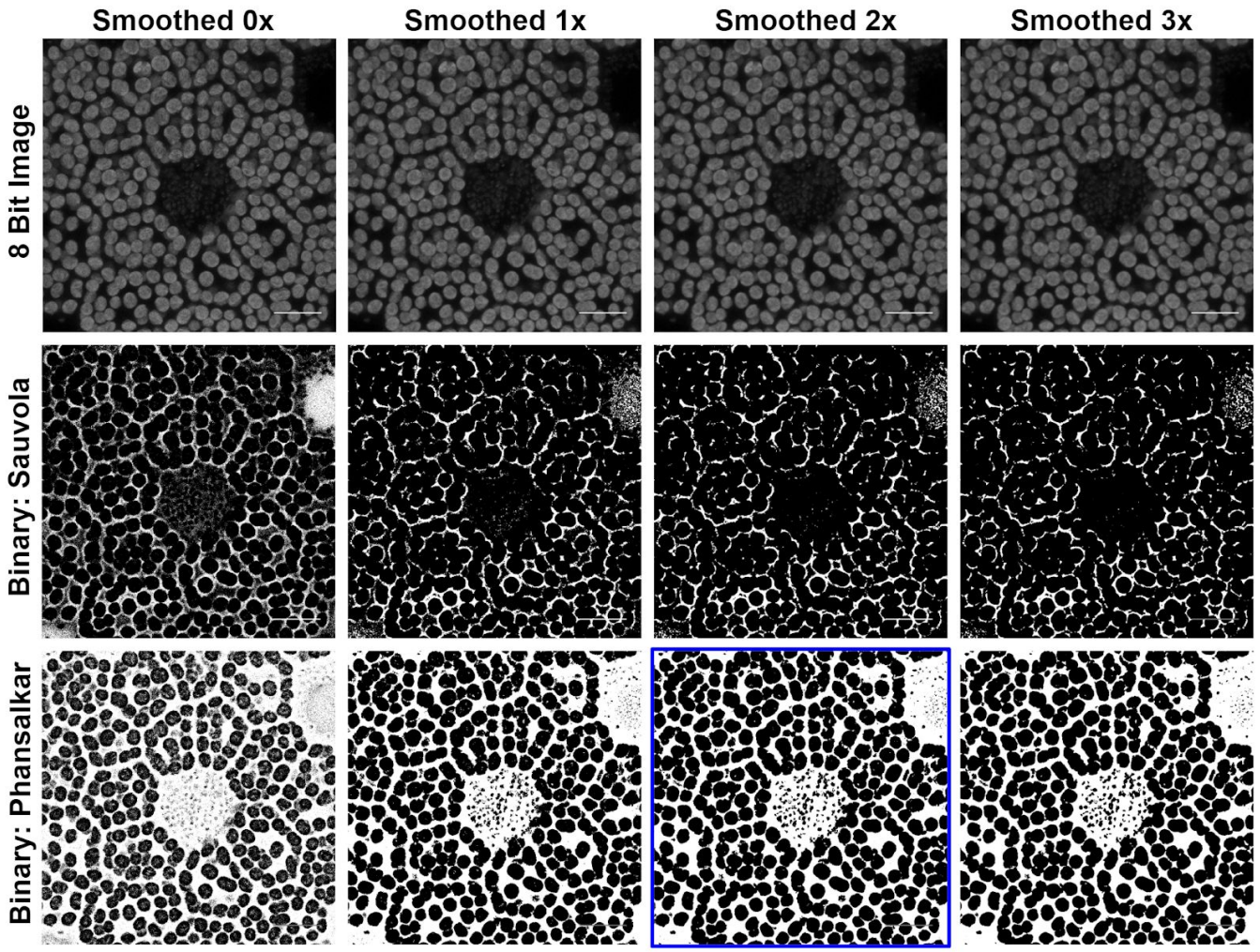


Figure 26: Effects of smoothing on thresholding. Images were smoothed zero, one, two, or three times, then thresholded with either Phansalkar or Sauvola auto local thresholding algorithms. Blue outline indicates parameters selected. Scale bar corresponds to 20 µm.

The Smooth function replaces each pixel with the average value of its $3 \cdot 3$ pixel neighborhood, effectively blurring the image. This processing can be useful for segmentation as long as the $3 \cdot 3$ pixel neighborhood is smaller than the objects of interest, and results in more accurate segmentation. From inspection and comparison to the raw images, two iterations of Smoothing and the Phansalkar algorithm was selected as the most accurate method of segmentation. However, some plastids are merged by this algorithm, necessitating some form of splitting of separate plastids. This was achieved with Watershed segmentation and restriction of possible ROI shapes, as shown in Figure 27. Watershed segmentation functions by seeding local minima of pixel intensity with “water” which is then slowly raised ‘up’ the surrounding values of pixel intensity until all watersheds meet. The lines where different watersheds meet each other are then used for

segmentation [202]. Here it is applied on binary images, which causes the division of regions based on the location of narrow points and inlets, many of which are the locations of two plastids being segmented one hourglass-shaped region.

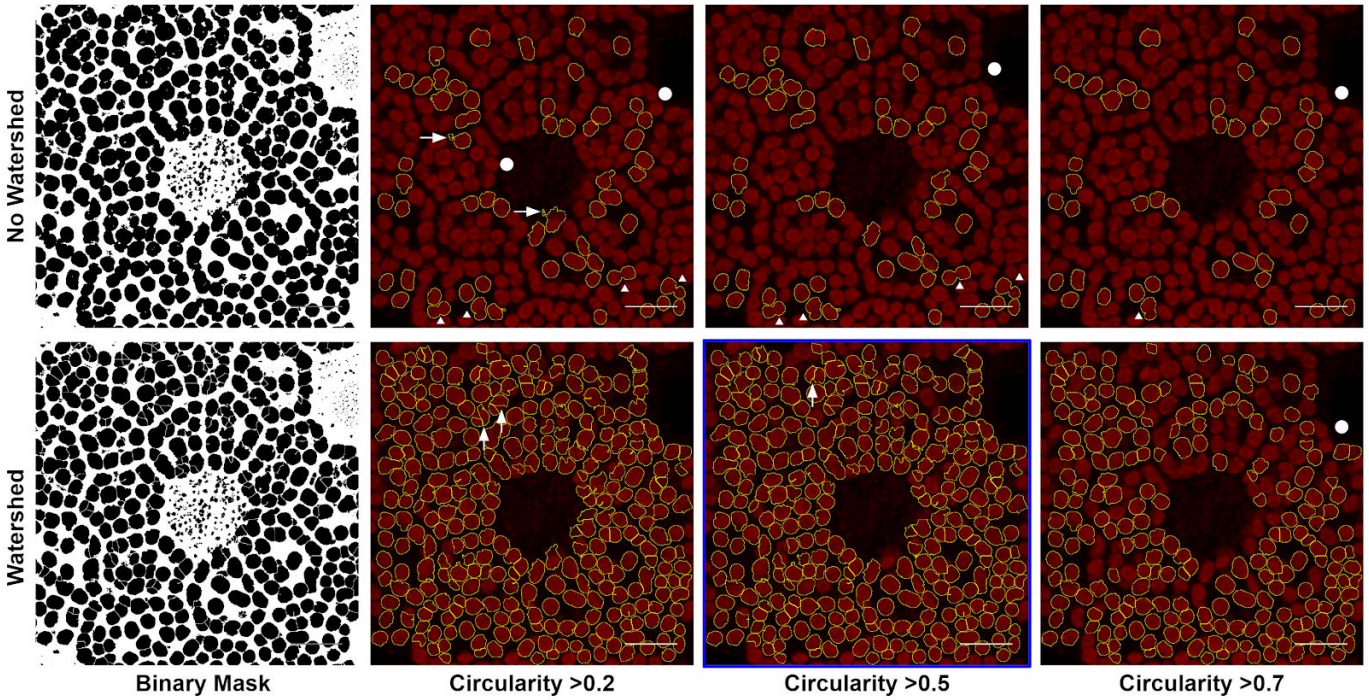


Figure 27: Effect of Watershed algorithm and shape restrictions on ROI generation. Arrows show problematic segmentation, circles prominent plastids not selected as ROIs, triangles show multiple plastids segmented as one ROI. Blue outline indicates parameters selected. Scale bar corresponds to 20 μm.

ROI limits of 10-50 μm^2 and circularity 0.5 - 1.0 were selected for analysis of the whole dataset. Analysis was executed with a custom Fiji macro included in Appendix A. The finalised pipeline is shown in Figure 28, with an example image demonstrating accuracy of segmentation.

1. Acquire two channel LSCM images, cyan fluorescence and chlorophyll autofluorescence
2. Maximum intensity projection with LAS AF, separate channels
3. Duplicate chlorophyll channel for mask preparation
4. Run Fiji image smoothing algorithm twice in series on mask channel
5. Run Phansalkar auto local thresholding algorithm on mask channel
6. Run Watershed algorithm on mask channel
7. Run “Analyze particles” function on mask channel, particle size 10-50 μm^2, circularity 0.5 - 1.0
8. Overlay ROIs identified on both channels of raw image in series, measure mean intensity within ROI

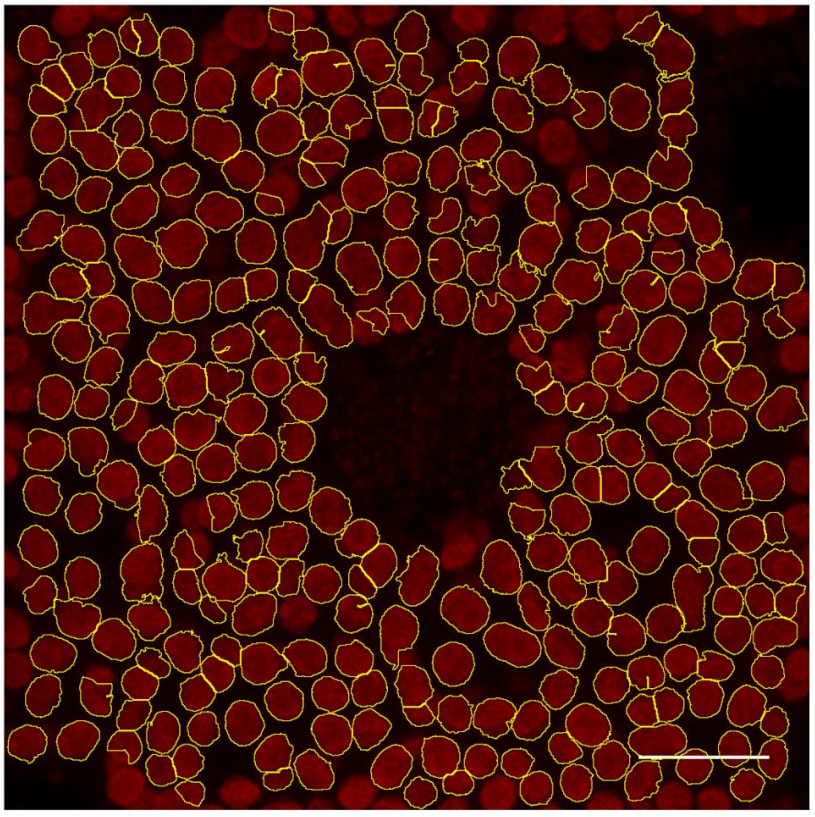


Figure 28: Steps of finalised plastid fluorescence quantification pipeline with example image of step 8. Steps 4-7 are performed on a mask channel generated by duplication of the chlorophyll channel, mask channel is then discarded and raw images used for measurement. Image is a larger version of panel with blue outline in Figure 27. Scale bar corresponds to 20 µm.

5.3.2 Transplastomic construct comparison

Mean cyan fluorescence intensity for all segmented ROIs shown in Figure 29.

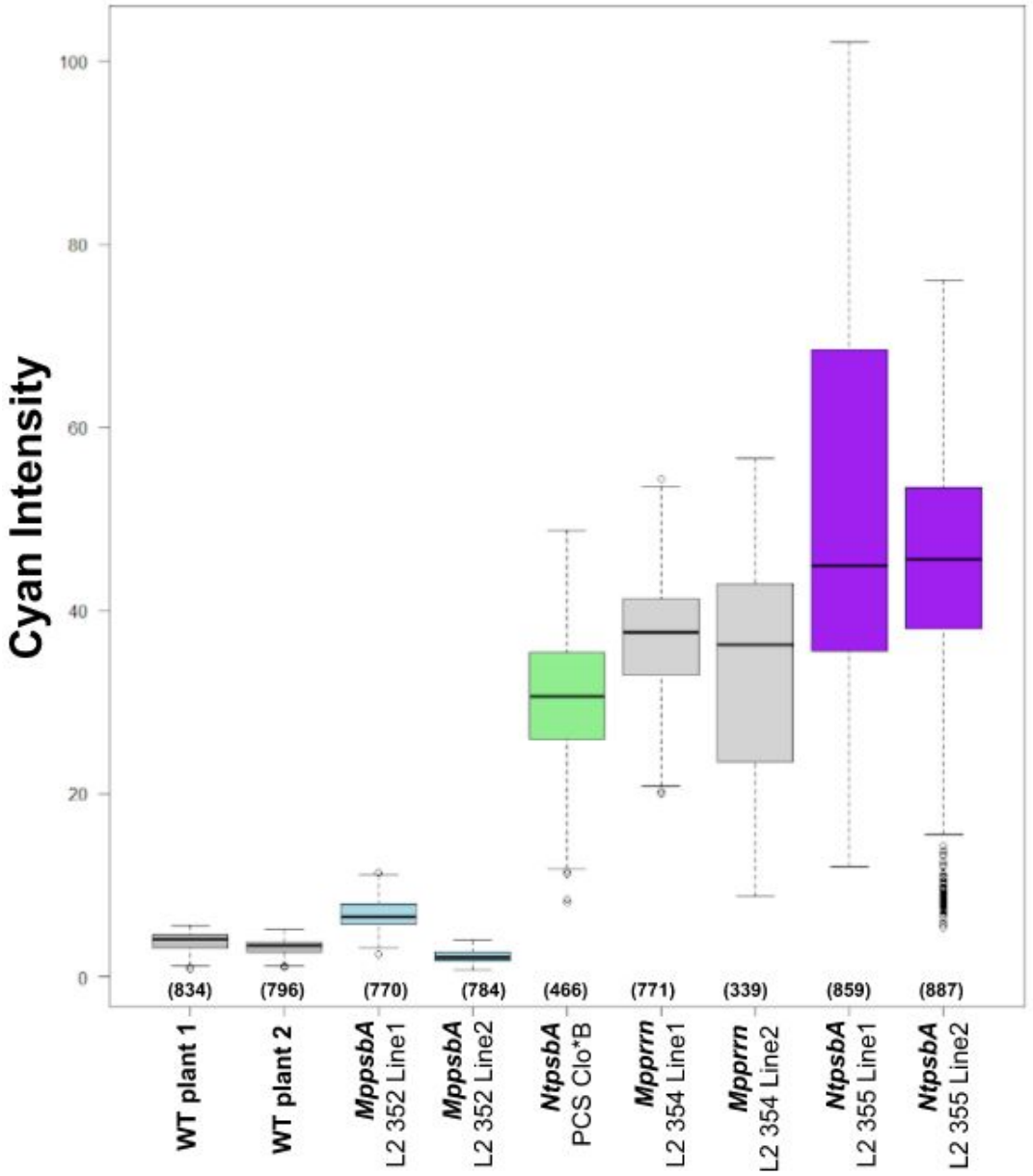


Figure 29: Plastid cyan fluorescence. This is the primary dataset, where for each column three gemmae were imaged from each line, two per construct except in the case of pCS Clo*B, where only one was available. Dark grey, light blue, green, grey, and purple correspond to wild type, L2_352, pCS Clo*B, L2_354, and L2_355 transplastomic plants, respectively. Numbers in parentheses indicate number of ROIs analysed.

The distribution of cyan fluorescence data within each line failed normalcy tests (see appendix C, Table 12 for Shapiro-Wilks normality test statistics [183] and Figure 38 quantile-quantile plots to visually assess normality). Therefore nonparametric statistical tests were used to compare between transformants. Kruskal-Wallis test [185] across all transformants was significant ($p < 0.0001$), so Dunn's post test [186] was implemented using Graphpad InStat. Post test showed statistically significant differences between each construct transformed (all p <

0.001, see Appendix C, Table 13). The analysis was repeated with a new set of micrographs generated from plants still on selection with spectinomycin 100 µg mL^{-1}, with comparable results (for raw fluorescence dataset from plants maintained on spectinomycin 100 µg mL^{-1}, see Figure 40, for box and whisker plot similar to Figure 29, see Figure 39, for table of confidence intervals, see Table 14).

5.3.3 Correlation between chlorophyll and cyan fluorescence

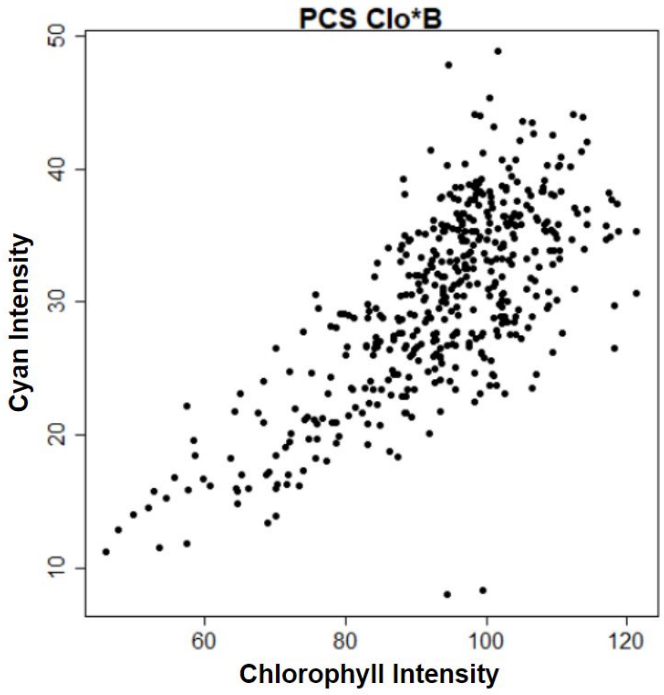


Figure 30: Correlation between chlorophyll intensity and cyan intensity in pCS Clo*B. Each point represents a single segmented plastid. See Appendix B, Figure 37 for similar plots representing other genotypes.

In each individual field of view and each line, there was a positive correlation between chlorophyll autofluorescence and cyan fluorescence, as shown in Figure 30. One explanation is that plastids nearest the objective suffer less attenuation of light through absorption and scattering for both the excitation and emission light, resulting in the plastids in the highest Z planes being imaged with the highest intensity. Alternatively, the correlation could be caused by ROIs that exceed the boundaries of the plastid, artificially decreasing mean intensity for chlorophyll and cyan channels in tandem.

As an exploratory test between these hypotheses, Z stacks of image chlorophyll channels were loaded into Fiji and the lookup table coded to depth using Hyperstack using temporal colour-code. Plastid chlorophyll intensity was then measured and the 10 ROIs with the highest fluorescence intensity and 10 ROIs with the lowest fluorescence intensity were marked. The results are shown in Figure 31.

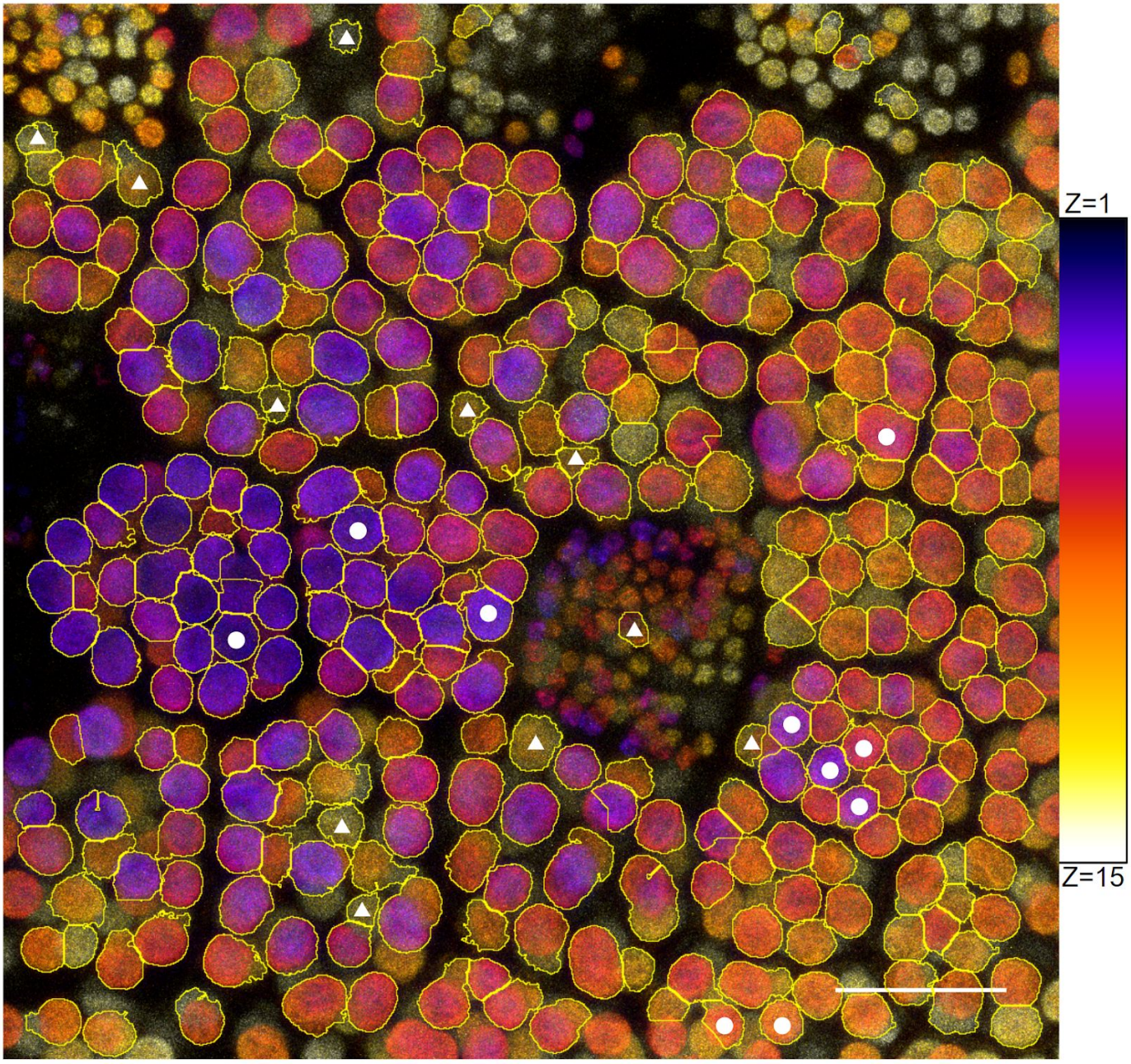


Figure 31: Depth-mapped Z stack with ‘fire’ lookup table, shown on right. Purple indicates highest Z slices closest to the objective, yellow-white indicates the lowest slices furthest from the objective. White circles mark the 10 ROIs with the highest fluorescence intensity as measured by mean pixel intensity, triangles mark the 10 ROIs with the lowest fluorescence intensity. Image is of L2_354, and was selected by eye as having the flattest epithelial surface of the available images. Scale bar corresponds to 20 µm.

The plastids measured as having the highest fluorescence intensity all stand out by having the highest Z position of plastids in their immediate region, indicating these plastids are on the epithelial surface of the sample. The plastids with the lowest measured fluorescence intensity have the opposite tendency, and cluster towards the bottom of the Z stack. This analysis suggests the pipeline could be

improved by either imaging only plastids closest to the surface or by correcting for plastid depth within the sample. This could be accomplished by generating a model for sample depth-associated signal attenuation, which could be derived empirically from studies of average fluorescence by depth, as demonstrated in a proof of principle here.

5.3.4 Discussion

Fluorescence intensity in transplastomic plants with the same fluorescent protein regulatory elements shows some variance between independently transformed plants, and significant variance within a single image. Some of this variance is no doubt biological in nature, and some arises from the process of image acquisition. Comparison of a larger number of independently transformed plants is the best solution to the former, and normalisation for the depth of each plastid within the sample may reduce the latter. Despite the variance, the large sample size enabled by automated image processing allows for strong statistical conclusions to be drawn regarding the relative fluorescence intensities.

L2_355 has the highest fluorescence intensity, markedly higher than pCS Clo*B. This is surprising given that both are driven by the *NtpsbA* promoter and 5' UTR. In light of the discovery of substantial biological variance between transplastomic plants, it is possible that the line of homoplastic pCS Clo*B plants imaged were on the low end of said variance. It is also possible that the mismatches in the homologous regions of pCS Clo*B caused mutations in the plastome which are causing an additional phenotype or compromising plastid transcriptional or translational function. L2_354 consistently shows comparable expression to PCS Clo*B, validating *Mpprrn* promoter 5' UTR combination as the first native regulatory elements reported to be used in transplastomic *Marchantia*. L2_352 results in the lowest expression levels, demonstrating only slightly higher fluorescence than background. *MppsbA* promoter and 5' UTR result in low but detectable protein accumulation, which may be useful for reducing metabolic strain in engineering applications.

Chapter 6: Conclusions

The primary accomplishments of this thesis are the generation of a better *M. polymorpha* plastome assembly, optimisation of plastid transformation protocols, and the *in vivo* analysis of several novel constructs. Each of these will prove useful for future plastome engineering work in the model chassis *M. polymorpha*, and should also have a high degree of transferability to plastid engineering in other plants due to the high levels of plastome conservation.

6.1 Plastome assembly

This project resulted in the first *Marchantia polymorpha* strain Cam plastome assembly, which is annotated, complete, and circular. The assembly has been partially validated by Sanger sequencing, and is closely mirrored by the recently published Kit-2 plastome assembly. This resource is necessary for future efforts to engineer the plastome of the Cam strain of *Marchantia polymorpha*, and has been submitted to GenBank. Due to the gene-dense nature of the plastome, correct genome assemblies are important to prevent the accidental introduction of mutations. Furthermore, the efficiency of homologous recombination is sensitive to even small amounts of non-homology between donor and recipient DNA [159]. Accordingly, homologous regions used in plastid transformation vector design must be prepared from an accurate plastome assembly for the specific strain in question. The plastome assembly validation process also resulted in the cloning of several homologous regions which are now available as Loop-compatible parts, allowing for the generation of plastid transformation vectors which integrate foreign DNA in new loci never reported in *Marchantia*.

6.2 Transformation optimisation

6.2.1 Plastid Transformation

This work presents the most systematic optimisation of *Marchantia* plastid transformation published to date. The result is a simpler and faster protocol, which has already been demonstrated to work both in my hands and in the hands of my colleague Dr. Eftychios Frangedakis, whereas the prior laboratory protocol had not worked for either of us. I developed transient nuclear transformation as a rapid

assay for DNA delivery, and have demonstrated transient transformation of several tissue types which have never been reported to be transformed. The optimised protocol is also substantially faster than the initial protocol taken from Boehm *et al.* 2016 [112], as shown in Figure 32.

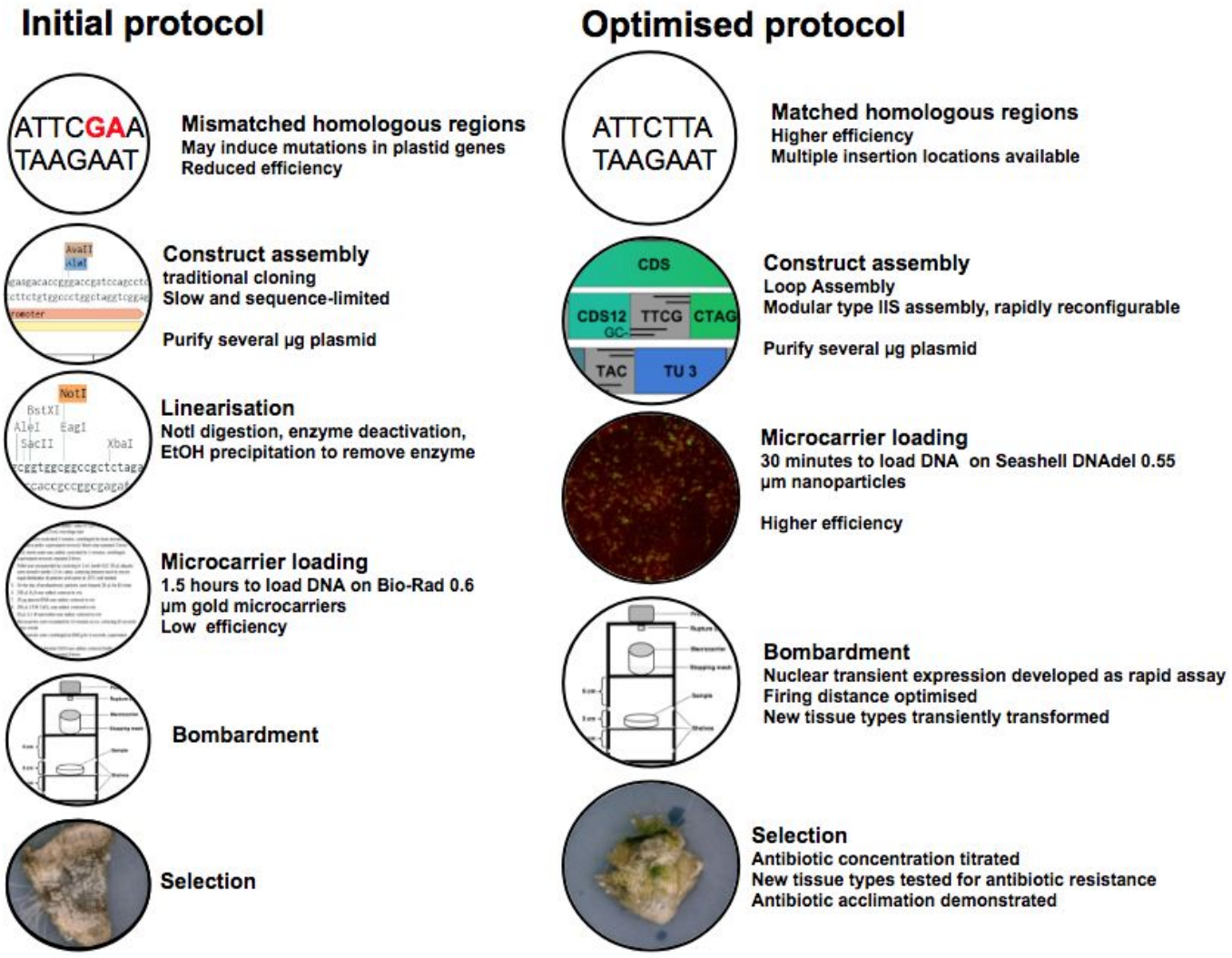


Figure 32: Comparison of initial (left column) and optimised (right column) protocols for *Marchantia* plastid transformation.

6.2.2 Biolistic nuclear transformation

I have compared the phenotypes of *Marchantia* transformed via different techniques in the nuclear genome. The results obtained here suggest that biolistic nuclear transformation results in substantially higher variance in expression levels compared to *Agrobacterium*-mediated transformation. This finding should guide the transformation technique used in future projects, with *Agrobacterium* used

when consistency is desired between independently transformed plants, and biolistic nuclear transformation used to explore a larger phenotypic space.

6.3 Transplastomic gene expression

In collaboration with Eftychios Frangedakis, I generated a toolkit of *Marchantia* plastome-specific genetic Level 0 parts using Type IIS assembly which can be reused and recombined in a combinatorial manner. Included among these parts are two fluorescent markers, one selectable marker, sixteen promoters, three terminators, and homologous regions for four integration loci in the wild type *Marchantia* plastome. From these I assembled twelve Level 2 plastid transformation vectors, of which seven were successfully integrated into *Marchantia* sporelings. Included among these successfully transformed constructs are three which integrate in a new locus in the plastome, which have demonstrated fluorescence but not have been extensively quantified.

I generated a family of constructs that insert at the same locus, and phenotypically characterised the constructs by means of a semi-automated image analysis pipeline. This image analysis pipeline was extensively optimised, and allows for comparison of fluorescent protein expression levels, which can be used in relative terms as a proxy for arbitrary foreign protein accumulation levels for transplastomic genes. Compared to pCS Clo*B, the only previously reported plastome-expressed fluorescent marker in *Marchantia*, new constructs were generated with both higher and lower fluorescence intensity.

6.4 Potential applications

The work described here primarily consists of improving available tools. With the products of this thesis available, *Marchantia* should become a more attractive chassis for chloroplast genetic engineering. *Marchantia* is useful both as a rapid testbed for plastid engineering of higher plants and as a production platform.

Despite the technical feasibility of plastid transformation for nearly three decades, no transplastomic crops have yet become commercially available [203]. The primary reason for this was the recalcitrance to plastid transformation in many crop

species, a limitation that is beginning to be broken down with recent successes in transforming the plastids of maize [204] and rice [205]. These new protocols open the door to commercialisation of transplastomic crops, which may be more acceptable to the public than current transgenic crops due to the transgene containment offered by plastid transformation. However, the transformation protocols for crop species remain technically difficult, low efficiency, and very slow - usually requiring several months of selection. *Marchantia* offers a faster testbed for the analysis of synthetic gene circuit behaviour before investing the time and effort of plastid transformation in crop species. The high degree of plastome sequence conservation suggests insight gained from *Marchantia* plastid engineering will be applicable to higher plant plastid engineering.

Marchantia itself may provide a useful platform for biosynthesis of high value products. The most promising high value compounds to commercially synthesise in *Marchantia* or related plants is terpenoids. Terpenoids comprise a highly diverse group of chemicals primarily synthesised by plants, and have applications as flavours, medicines, fragrances, and renewable polymers (reviewed in [206]). Terpenoids tend to include many chiral carbons, and their stereochemistry is highly specific in nature [207]. The structural variety and stereo-specificity of terpenoids makes them difficult and inefficient to produce with organic chemistry methods [208]. As a result, many commercially available terpenoids such as artemisinin are extracted from wild plants, leading to high and volatile prices [209]. In particular, many valuable terpenoids can be produced through the non-mevalonate pathway, which takes place in plastids [210]. Production of high value products through transplastomic plants is therefore of considerable commercial interest [211].

The methods developed in this thesis simplify the process and increase the efficiency of plastid transformation in *Marchantia,* which will facilitate its use as a rapid-prototyping chassis. Furthermore, this project generated the required genetic components to integrate transgenes in several novel locations, and demonstrated successful transgene integration in one of those. This opens the possibility of stacking multiple transgenic constructs through co-bombardment with a transiently expressed resistance marker [212], co-bombardment with multiple synthetic constructs and selection [213], or selectable marker removal and bombardment of

transplastomic marker-free plants [214]. These techniques are only possible with the availability of homologous regions for multiple insertion locations. To the author's knowledge this is the first report of the generation or use of a genetic toolkit with multiple homologous regions in *Marchantia*.

6.5 Closing remarks

The work described here laid the foundations for plastome engineering in *Marchantia polymorpha* by generating a high-quality plastid genome assembly, optimising a transformation protocol, and developing novel constructs that confer higher expression levels than previously reported. Taken together, these results comprise a significant contribution to the field of plastid engineering in nonvascular plants, a historically neglected taxonomic group with several interesting applications for synthetic biology research.

Appendix A: Selected code

This appendix contains a minimal set of code that should be sufficient to recreate the analyses in this thesis.

R code

```
read.delim(file.choose(), header=T)

#determine quantiles
quantile(data)

# T test to compare samples
t-test(data1, data2)

#generate violin plots
ggplot2.violinplot(data, plot parameters)

# boxplot generation
boxplot(data1,data2, plot parameters)

#variance of data
var(data)
# standard deviation is sqrt of variance
# Shapiro-Wilks test to assess normality
shapiro.test(data)

#Pearson's product-moment correlation
cor.test(data1, data2)

#qqnorm plot matrix
# used to visually assess normality of data
graphics.off()
par(mfrow=c(3,3), mar=c(1,1,1,1), mai = c(0.35, 0.28, 0.2, 0.2))
qqnorm(san[,1], main = 'WT Plant 1', pch= 20, xlab=NA, cex.main=1.5,
cex.lab=1.5)
# repeat above line with other data to fill the remaining 8 plots

#density plot matrix
# kernel density plot combines a histogram with a model of the population
distribution
# it is a model of histogram shape with arbitrarily large sample and arbitrarily small
bins
graphics.off()
par(mfrow=c(3,3), mar=c(1,1,1,1), mai = c(0.35, 0.28, 0.2, 0.2))
plot(density(san$WT_Plant1, na.rm=TRUE), main='WT Plant 1', xlim = c(0,8),
cex.main=2)
#repeat above line with other data to fill the remaining 8 plots

#correlation of cyan and chlorophyll fluorescence intensity paired by plastid
imaged
```

```
# axis limits are not specified, so don't compare between plots too much
graphics.off()
par(mfrow=c(3,3), mar=c(1,1,1,1), mai = c(0.35, 0.28, 0.2, 0.2))
plot(scatters$WT.1.turq, scatters$WT.1.chl, main = 'WT Plant 1', pch=16)
#
repeat above line with other data to fill the remaining 8 plots
```

Fiji Macro for plastid segmentation

```
run("Duplicate...", " ");
run("Smooth");
run("Smooth");
run("Auto Local Threshold", "method=Phansalkar radius=15 parameter_1=0
parameter_2=0 white");
run("Invert LUT");
# LUT = lookup table
run("Watershed");
run("Analyze Particles...", "size=10-50 circularity=0.50-1.00 display exclude clear
add");
run("Clear Results");
close();
roiManager("Measure");
String.copyResults();
run("Clear Results");
```

Appendix B: Supplemental micrographs

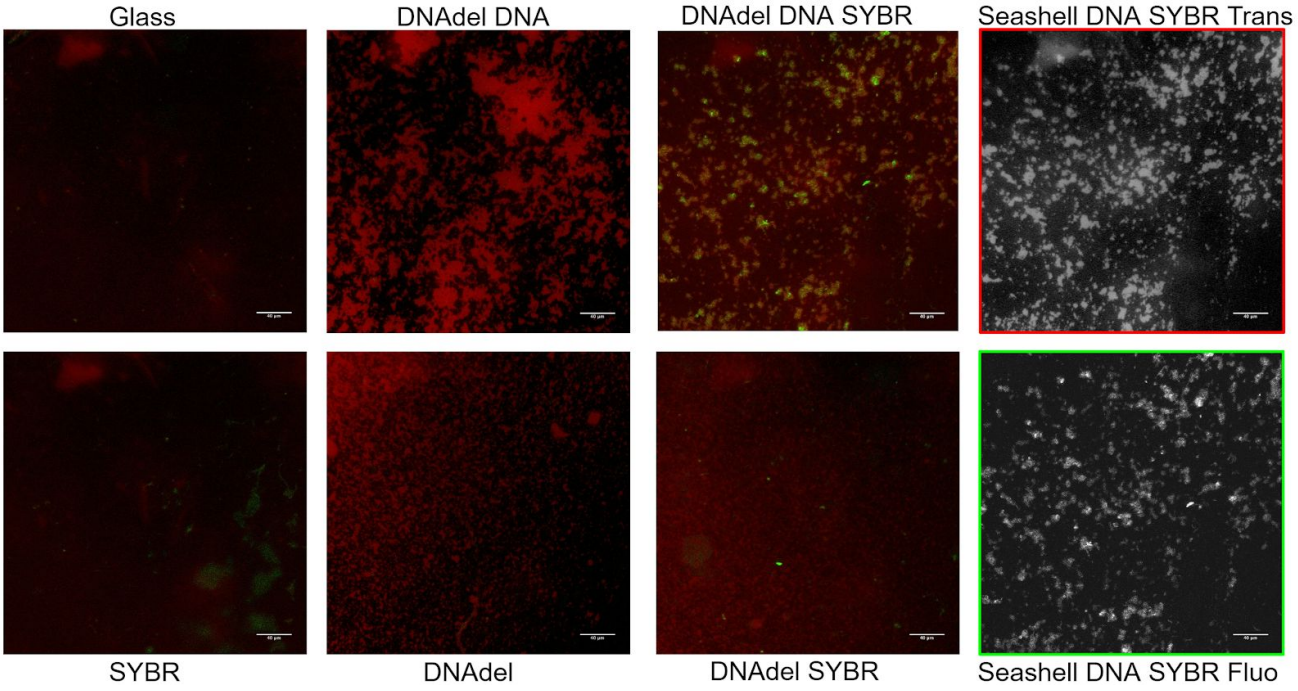


Figure 33: Additional controls for Figure 6, not shown in the main document to allow for better comparison of DNAdel + SYBR and DNAdel + SYBR + DNA. Transmitted light channel is inverted. Where pseudocolouring is used, Glass indicates an empty slide, note the patch of transmitted light in the top left corner which is reproduced in all images and must come from the instrument itself. DNAdel DNA indicates DNAdel particles loaded with DNA. DNAdel DNA SYBR is included in Figure 6, and indicates DNAdel particles loaded with DNA loaded with SYBR safe. The rightmost column shows the two channels separately, with a red outline indicating transmitted light and green outline indicating SYBR safe fluorescence. SYBR indicates SYBR safe stain added to slide alone. DNAdel indicates DNAdel beads added to slide alone. DNAdel SYBR indicates DNADel beads with SYBR safe, its composite channels are shown separately in Figure 6.

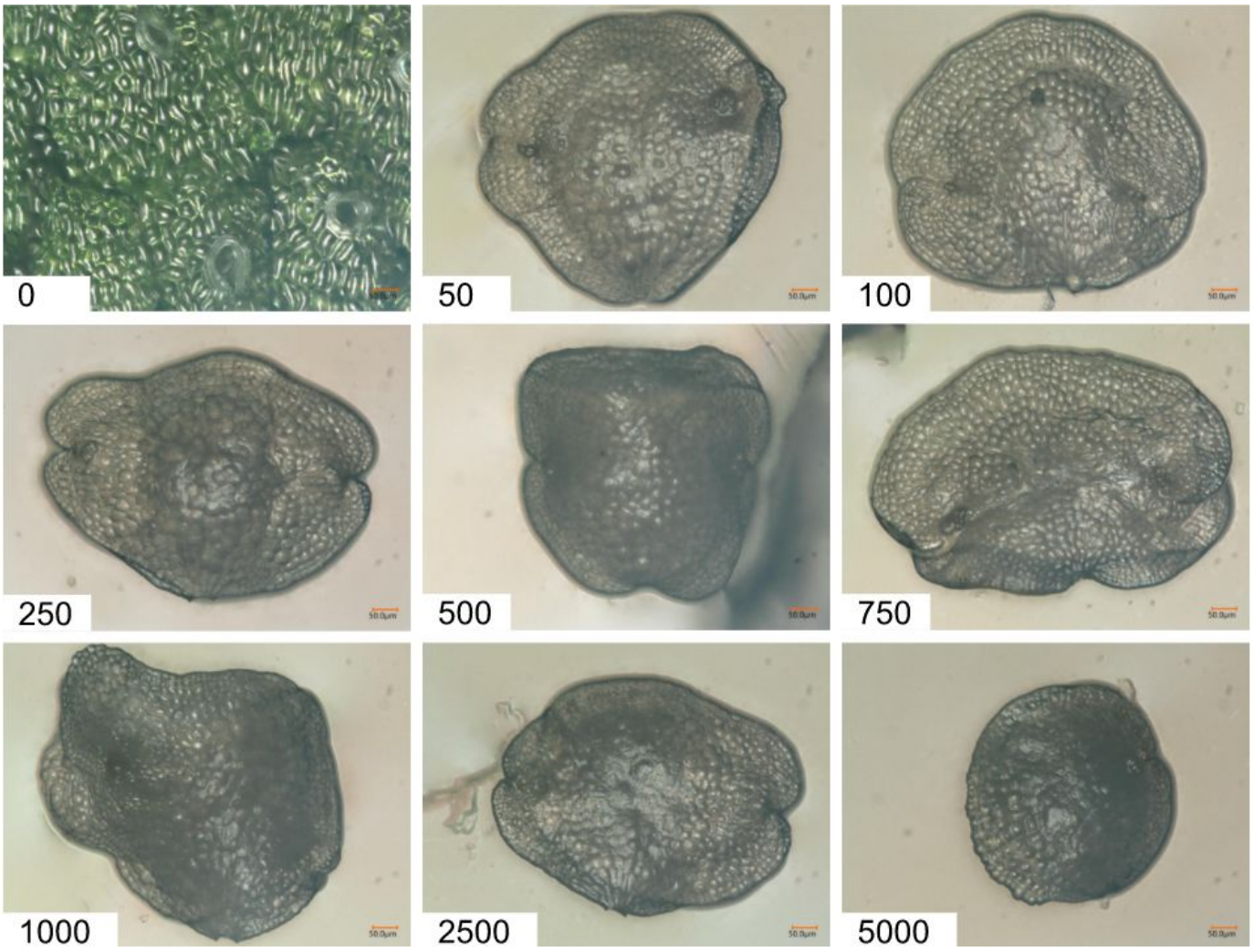


Figure 34: WT gemmae spectinomycin resistance. Numbers in lower-left of each image represent concentration of spectinomycin in μg mL^{-1}. Scale bar corresponds to 50 μm.

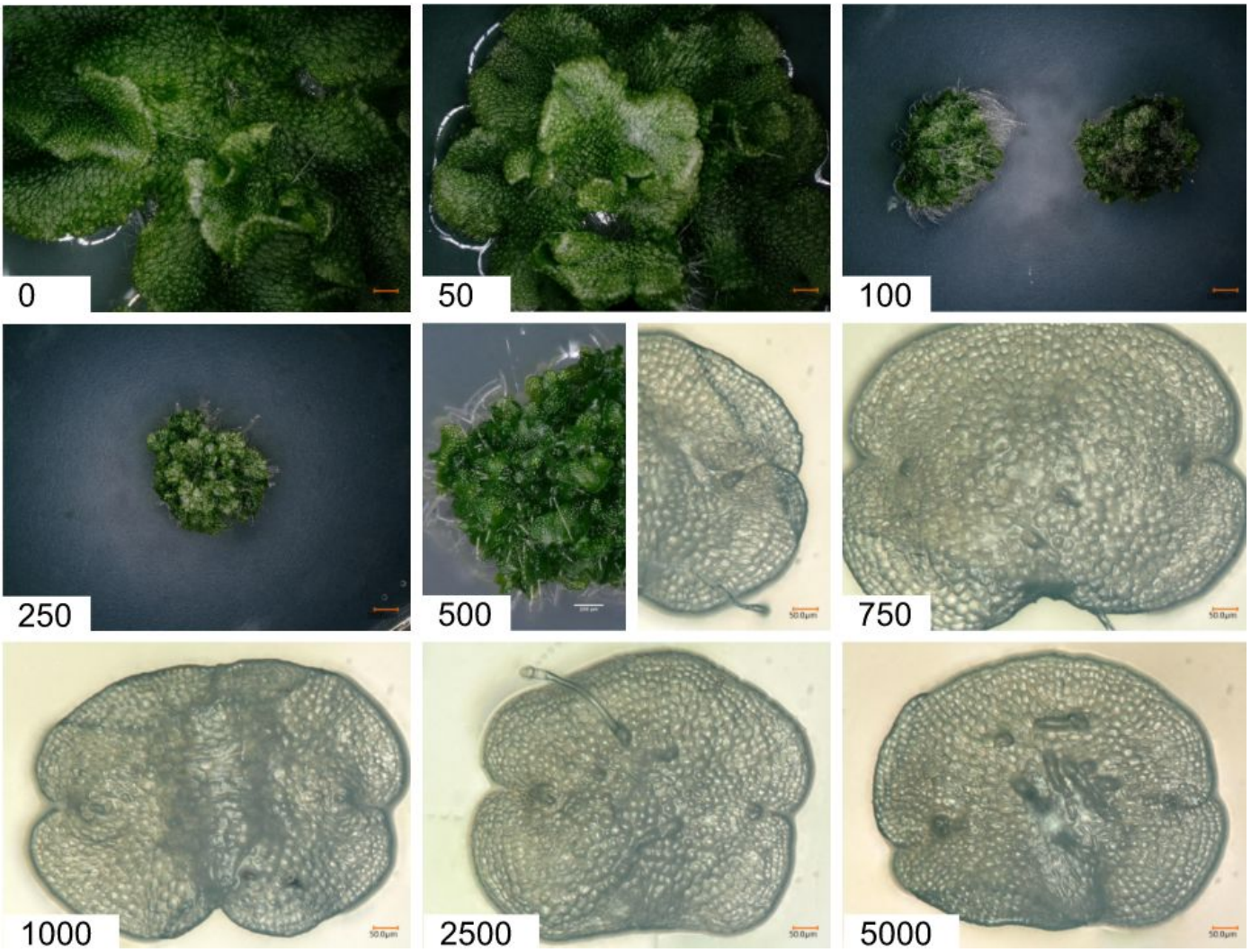


Figure 35: pCS Clo*B gemmae spectinomycin resistance. pCS Clo*B gemmae (generated by Christian R Boehm) after 34 days on selection. Numbers in lower-left of each image represent concentration of spectinomycin in µg mL^{-1}. Images taken after 34 days of selection. Three gemmae were tested at each concentration. At 500 µg mL^{-1}, two died but one grew at a low rate. Plants exhibit a “crunched” phenotype when agitated by Spectinomycin, which is most apparent in the live plant at 500 µg mL^{-1}. Plants vary significantly in size, so different magnifications are present in different panels. For concentrations 0-250 µg mL^{-1}, scale bar corresponds to 1000 µm. For 500 µg mL^{-1}, left scale bar corresponds to 200 µm. All other scale bars correspond to 50 µm.

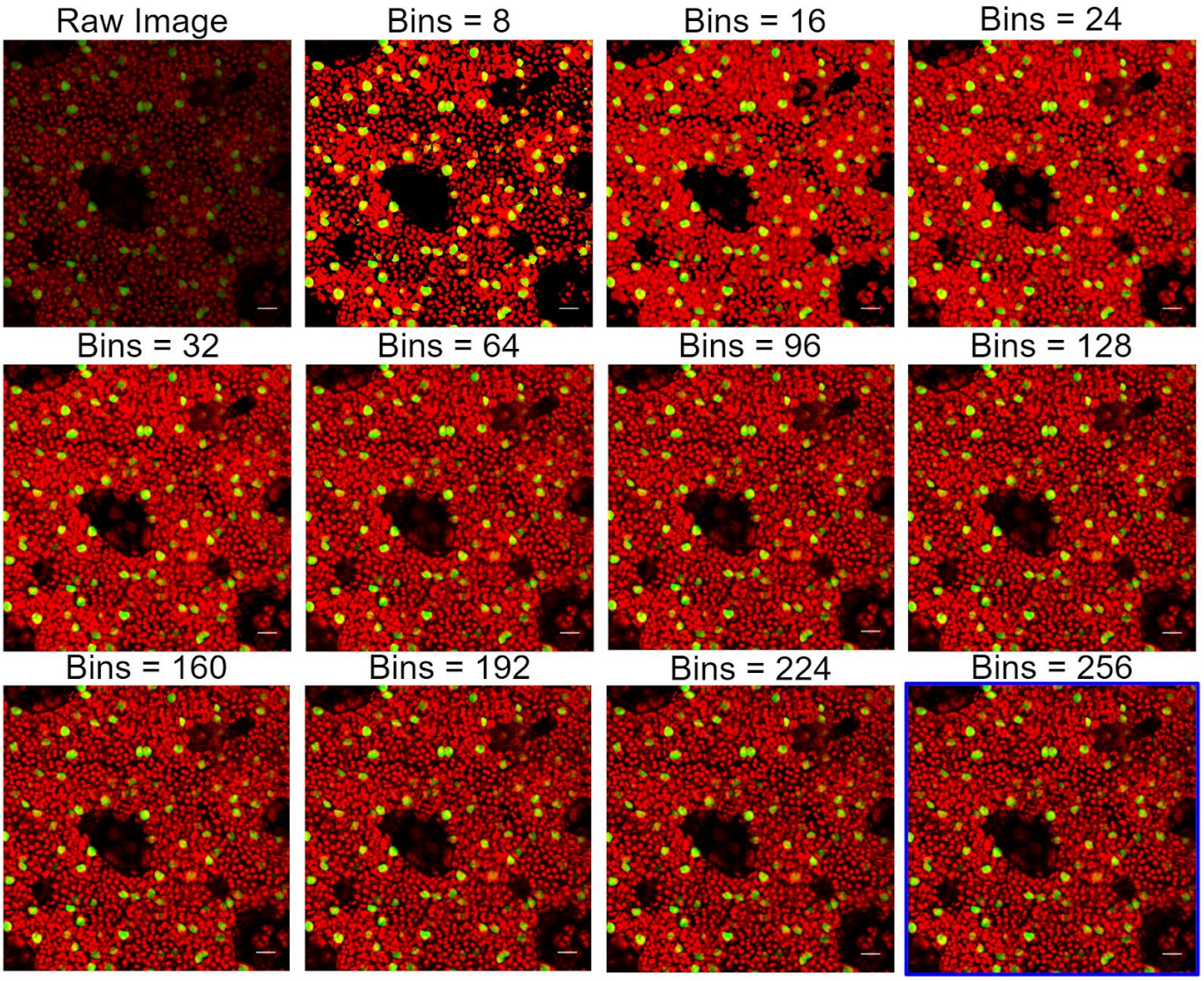


Figure 36: Output images of CLAHE algorithm with a range of histogram bin numbers. Blue outline indicated parameters selected. Scale bar corresponds to 20 μm.

Appendix C: Supplemental data and tables

Table 10: Summary of tissue types transformable with biolistic and *Agrobacterium*-mediated DNA delivery. Single, double, and triple asterisks indicate low, medium, and high efficiency, respectively. Letters indicate first report of transformation type. A = [65], B = [82], C = [84], D = [187], E = [81], F = [215], G = this thesis, transient transformation.

Tissue Type	Nuclear *Agrobacterium*-mediated transformation	Nuclear Biolistic transformation	Plastid transformation
5-8 day sporelings	***A	***B	**C
30 day sporelings	No	**G	No
0-3 day gemmae	**D	*G	No
14 + day thallus	***E	*F	No

Table 11: Statistical differences between fluorescence intensity of individual nuclear transformants. All plants transformed with L2_103, transformants designated Bio1-5 were transformed with biolistics, transformants designated Agro1-5 were transformed with *Agrobacterium*. Significance testing with Dunn's post-test is shown. NS indicates nonsignificant difference, * indicates $p < 0.05$, *** indicates $p < 0.001$.

Versus	Bio1	Bio2	Bio3	Bio4	Bio5	Agro1	Agro2	Agro3	Agro4	Agro5
Bio1	1									
Bio2	NS	1								
Bio3	***	NS	1							
Bio4	***	***	NS	1						
Bio5	***	***	***	***	1					
Agro1	***	NS	NS	NS	***	1				
Agro2	***	***	NS	NS	***	*	1			
Agro3	***	***	***	***	NS	***	***	1		
Agro4	***	***	***	***	NS	***	***	NS	1	
Agro5	***	***	***	***	NS	***	***	NS	NS	1

Table 12: Supplemental statistics for the unselected plastid fluorescence dataset. Coefficient of variance is standard deviation divided by mean, and allows for normalised comparison of the variance between samples. Shapiro-Wilks normality statistic represents the probability of getting a distribution this skewed with random sampling from a normal distribution.

	Sample size	Min	Mean	Median	Max	Standard Deviation	Coefficient of Variance	Shapiro-Wilks normality
WT Plant 1	796	1.779	3.394	3.442	5.106	0.555	0.1635	0.0006075
WT Plant 2	833	0.861	3.86	4.093	5.652	0.958	0.2481	2.20E-16
L2 352 Line 1	785	0.73	2.231	2.171	3.961	0.598	0.2680	2.20E-06
L2 352 Line 2	719	3.349	6.987	6.731	11.297	1.502	0.2149	1.58E-09
PCS Clo*B	466	8.03	30.16	30.68	48.8	7.078	0.2346	7.95E-05
L2 354 Line 1	664	8.729	34.2	35.739	56.6	9.833	0.2875	1.46E-09
L2 354 Line 2	771	19.95	37.35	37.67	54.36	6.109	0.1635	0.06511
L2 355 Line 1	859	12.03	50.4	44.86	102.03	19.403	0.3849	2.20E-16
L2 355 Line 2	887	14.93	45.99	44.99	76.11	11.677	0.2539	2.89E-09

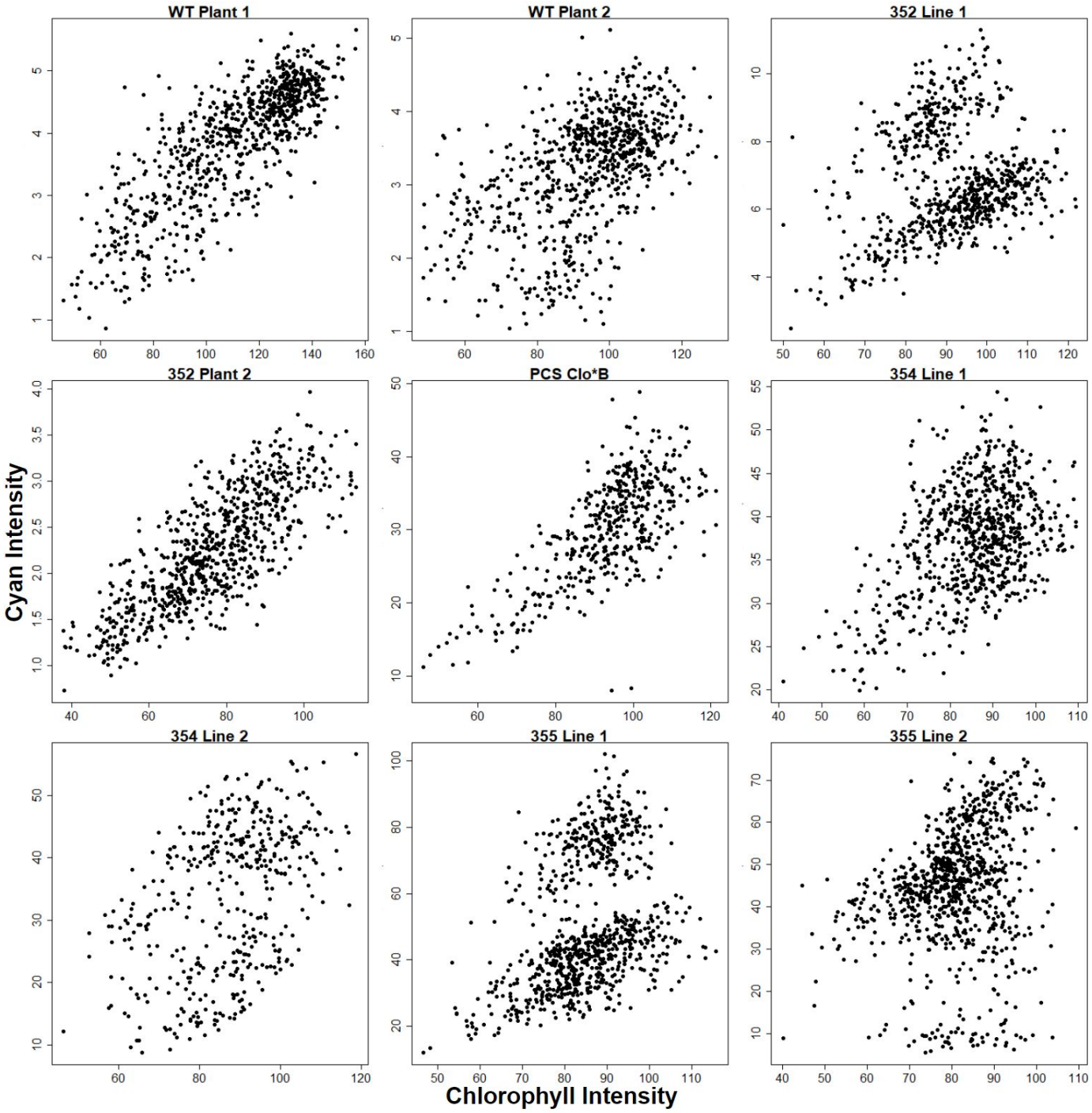


Figure 37: Raw plastid fluorescence data. Each dot on the graphs represents a single plastid, the abscissa shows signal intensity in the chlorophyll autofluorescence channel, the ordinate shows signal intensity in cyan channel. The center panel of this figure is shown on its own as Figure 30.

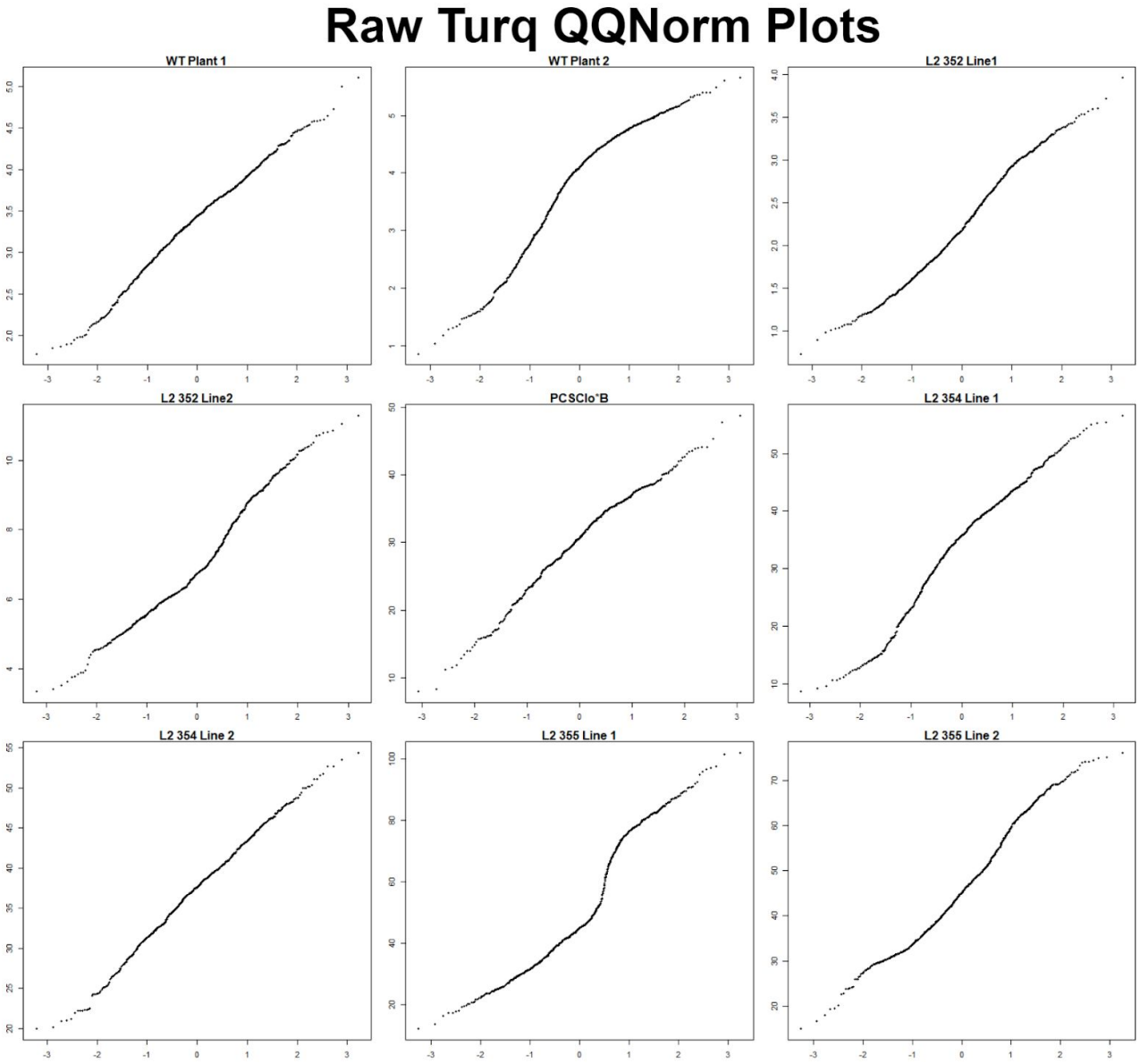


Figure 38: Quantile-quantile plots of cyan fluorescence from the plastid fluorescence dataset. Each line imaged is shown in a different panel. Data from a normal distribution is a straight line. While all distributions except L2 354 line2 fail a standard test of normality (see Table 12), most are not far from a normal distribution.

Table 13: Summary statistics for plastid fluorescence data maintained without selection 95% confidence intervals (CIs) were generated using Graphpad InStat, the results are shown in Table 13.

	Mean	Sample size	Standard deviation	95% CI Lower limit	95% CI Upper limit
WT Plant 1	3.86	833	0.96	3.8	3.93
WT Plant 2	3.2	796	0.78	3.14	3.25
L2 352 Line1	6.84	770	1.56	6.73	6.7
L2 352 Line2	2.23	784	0.60	2.19	2.27
PCS Clo*B	30.16	466	7.08	29.5	30.81
L2 354 Line1	37.35	771	6.11	36.9	37.79
L2 354 Line2	33.59	439	11.65	32.5	34.68
L2 355 Line1	50.4	859	19.40	49.1	51.7
L2 355 Line2	44.65	887	14.52	43.69	45.6

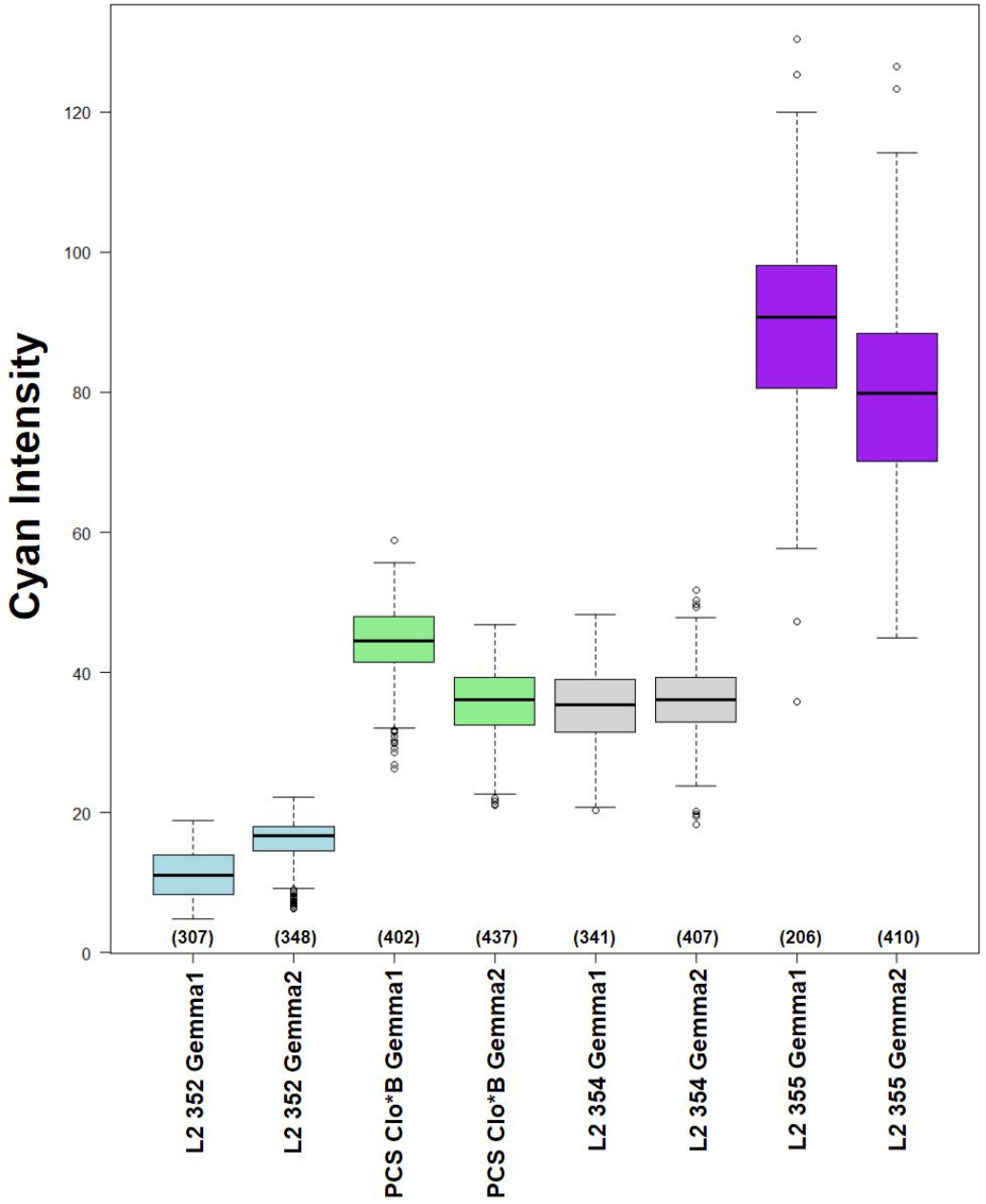


Figure 39: Plastid cyan fluorescence maintained on spectinomycin 100 µg mL^{-1}. Chart on the right is a smaller independent dataset from plants maintained on spectinomycin 100 µg mL^{-1}. Light blue, green, grey, and purple correspond to wild type, L2_352, pCS Clo*B, L2_354, and L2_355 transplastomic plants, respectively. Numbers in parentheses indicate number of ROIs analysed.

Table 14: Summary statistics for the selection plastid fluorescence dataset of plants maintained on spectinomycin 100 µg mL^{-1}

	Mean	Sample size	Standard deviation	95% CI lower limit	95% CI upper limit
L2 352 3	11.14	307	3.32	10.77	11.51
L2 352 4	15.88	348	3.27	15.54	16.23
PCS Clo*B 3	44.28	402	5.36	43.76	44.81

PCS Clo*B 4	35.57	437	5	35.1	36.04
L2 354 3	35.11	341	5.28	34.69	35.81
L2 354 4	36.11	406	5.28	35.6	36.63
L2 355 3	88.96	206	14.36	87	90.92
L2 355 4	79.61	410	12.95	78.36	80.87

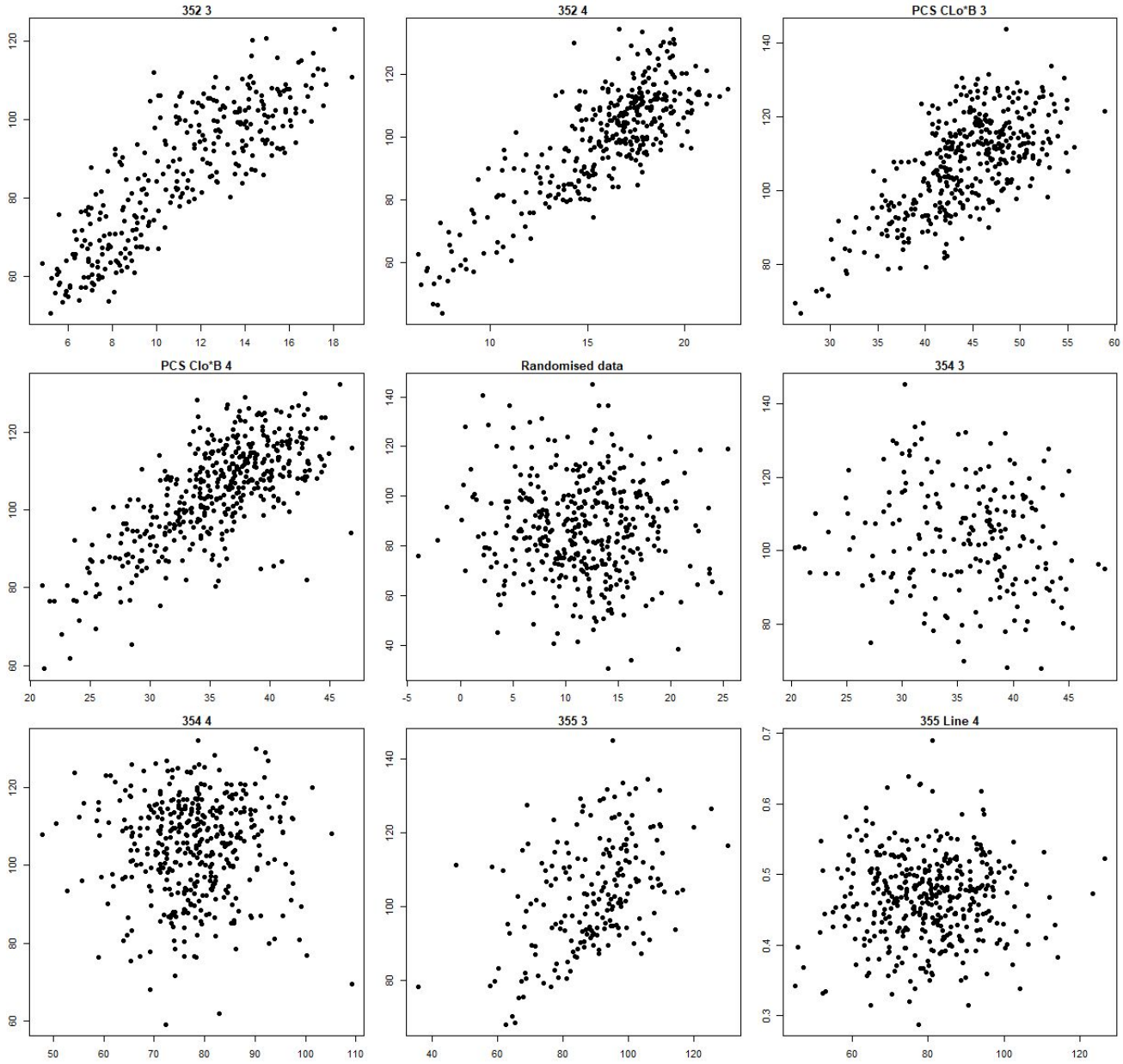


Figure 40: Raw plastid fluorescence data for the plants maintained on spectinomycin 100 µg mL^{-1}. Each dot on the graphs represents a single plastid, the abscissa shows signal intensity in the chlorophyll autofluorescence channel, the ordinate shows signal intensity in cyan channel. Each plot is from a single gemma. The central plot is randomly generated data drawn from two normal distributions for comparison purposes.

References

1. Mackiewicz P, Gagat P. Monophyly of Archaeplastida supergroup and relationships among its lineages in the light of phylogenetic and phylogenomic studies. Are we close to a consensus? Acta Societatis Botanicorum Poloniae; Warsaw. Warsaw, Poland, Warsaw: Polish Botanical Society; 2014;83: 263–280.

2. Turner S, Pryer KM, Miao VP, Palmer JD. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. J Eukaryot Microbiol. 1999;46: 327–338.

3. McFadden GI. PRIMARY AND SECONDARY ENDOSYMBIOSIS AND THE ORIGIN OF PLASTIDS. J Phycol. 2001;37: 951–959.

4. Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D. A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol. 2004;21: 809–818.

5. Palmer JD. A single birth of all plastids? Nature. 2000;405: 32–33.

6. Jensen PE, Leister D. Chloroplast evolution, structure and functions. F1000Prime Rep. 2014;6: 40.

7. Archibald JM. Chapter Three - The Evolution of Algae by Secondary and Tertiary Endosymbiosis. In: Piganeau G, editor. Advances in Botanical Research. Academic Press; 2012. pp. 87–118.

8. Ochoa de Alda JAG, Esteban R, Diago ML, Houmard J. The plastid ancestor originated among one of the major cyanobacterial lineages. Nat Commun. 2014;5: 4937.

9. Stechmann A, Cavalier-Smith T. The root of the eukaryote tree pinpointed. Curr Biol. 2003;13: R665–6.

10. Marin B, Nowack ECM, Melkonian M. A plastid in the making: evidence for a second primary endosymbiosis. Protist. 2005;156: 425–432.

11. Delaye L, Valadez-Cano C, Pérez-Zamorano B. How Really Ancient Is Paulinella Chromatophora? PLoS Curr. 2016;8. doi:10.1371/currents.tol.e68a099364bb1a1e129a17b4e06b0c6b

12. Gibbs SP. The chloroplasts of Euglena may have evolved from symbiotic green algae. Can J Bot. NRC Research Press; 1978;56: 2883–2889.

13. Henze K, Badr A, Wettern M, Cerff R, Martin W. A nuclear gene of eubacterial origin in Euglena gracilis reflects cryptic endosymbioses during protist evolution. Proc Natl Acad Sci U S A. National Acad Sciences; 1995;92: 9122–9126.

14. Harris EH. CHLAMYDOMONAS AS A MODEL ORGANISM. Annu Rev Plant Physiol Plant Mol Biol. annualreviews.org; 2001;52: 363–406.

15. Palenik B, Brahamsha B, Larimer FW, Land M, Hauser L, Chain P, et al. The genome of a motile marine Synechococcus. Nature. nature.com; 2003;424: 1037–1042.

16. Kaneko T, Tabata S. Complete genome structure of the unicellular cyanobacterium Synechocystis sp. PCC6803. Plant Cell Physiol. 1997;38: 1171–1176.

17. Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: organelle

genomes forge eukaryotic chromosomes. Nat Rev Genet. 2004;5: 123–135.

18. Shaw J, Lickey EB, Schilling EE, Small RL. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am J Bot. 2007;94: 275–288.

19. Vasil I. The Molecular Biology of Plastids: Cell Culture and Somatic Cell Genetics of Plants. Academic Press; 2012.

20. Ebert C, Tymms MJ, Schweiger H-G. Homology between 4.3 μm minicircular and plastomic DNA in chloroplasts of Acetabularia cliftonii. Mol Gen Genet. 1985;200: 187–192.

21. Williamson DH, Denny PW, Moore PW, Sato S, McCready S, Wilson RJ. The in vivo conformation of the plastid DNA of Toxoplasma gondii: implications for replication. J Mol Biol. 2001;306: 159–168.

22. Wolfe KH, Morden CW, Palmer JD. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc Natl Acad Sci U S A. 1992;89: 10648–10652.

23. Smith DR, Lee RW. A plastid without a genome: evidence from the nonphotosynthetic green alga Polytomella. Plant Physiol. American Society of Plant Biologists; 2014; 113.233718.

24. Simpson CL, Stern DB. The treasure trove of algal chloroplast genomes. Surprises in architecture and gene content, and their functional implications. Plant Physiol. 2002;129: 957–966.

25. Turmel M, Otis C, Lemieux C. The complete chloroplast DNA sequence of the green alga Nephroselmis olivacea: insights into the architecture of ancestral chloroplast genomes. Proc Natl Acad Sci U S A. 1999;96: 10248–10253.

26. Jansen RK, Wojciechowski MF, Sanniyasi E, Lee S-B, Daniell H. Complete plastid genome sequence of the chickpea (Cicer arietinum) and the phylogenetic distribution of rps12 and clpP intron losses among legumes (Leguminosae). Mol Phylogenet Evol. 2008;48: 1204–1217.

27. Li J, Gao L, Chen S, Tao K, Su Y, Wang T. Evolution of short inverted repeat in cupressophytes, transfer of accD to nucleus in Sciadopitys verticillata and phylogenetic position of Sciadopityaceae. Sci Rep. 2016;6: 20934.

28. Blazier JC, Jansen RK, Mower JP, Govindu M, Zhang J, Weng M-L, et al. Variable presence of the inverted repeat and plastome stability in Erodium. Ann Bot. 2016;117: 1209–1220.

29. Maliga P. Two plastid RNA polymerases of higher plants: an evolving story. Trends Plant Sci. Elsevier; 1998;3: 4–6.

30. Zhelyazkova P, Sharma CM, Förstner KU, Liere K, Vogel J, Börner T. The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. Plant Cell. Am Soc Plant Biol; 2012;24: 123–136.

31. Schweer J, Loschelder H, Link G. A promoter switch that can rescue a plant sigma factor mutant. FEBS Lett. 2006;580: 6617–6622.

32. Hess WR, Börner T. Organellar RNA polymerases of higher plants. Int Rev Cytol. Elsevier; 1999;190: 1–59.

33. Ortelt J, Link G. Plastid gene transcription: promoters and RNA polymerases. Methods

Mol Biol. 2014;1132: 47–72.

34. Germain A, Hotto AM, Barkan A, Stern DB. RNA processing and decay in plastids. Wiley Interdiscip Rev RNA. 2013;4: 295–316.

35. Gilbert W. Genes-in-pieces revisited. Science. go.galegroup.com; 1985;228: 823–824.

36. Erickson JM, Rahire M, Rochaix JD. Chlamydomonas reinhardii gene for the 32 000 mol. wt. protein of photosystem II contains four large introns and is located entirely within the chloroplast inverted repeat. EMBO J. Wiley Online Library; 1984;3: 2753–2762.

37. Saldanha R, Mohr G, Belfort M, Lambowitz AM. Group I and group II introns. FASEB J. FASEB; 1993;7: 15–24.

38. Cech TR. Self-splicing of group I introns. Annu Rev Biochem. annualreviews.org; 1990;59: 543–568.

39. Huang H-R, Rowe CE, Mohr S, Jiang Y, Lambowitz AM, Perlman PS. The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. Proc Natl Acad Sci U S A. 2005;102: 163–168.

40. Lambowitz AM, Zimmerly S. Group II introns: mobile ribozymes that invade DNA. Cold Spring Harb Perspect Biol. 2011;3: a003616.

41. Zoschke R, Nakamura M, Liere K, Sugiura M, Börner T, Schmitz-Linneweber C. An organellar maturase associates with multiple group II introns. Proc Natl Acad Sci U S A. National Acad Sciences; 2010;107: 3245–3250.

42. Soltis P, Doyle JJ. Molecular Systematics of Plants II: DNA Sequencing. Springer Science & Business Media; 2012.

43. Kuhsel MG, Strickland R, Palmer JD. An ancient group I intron shared by eubacteria and chloroplasts. Science. science.sciencemag.org; 1990;250: 1570–1573.

44. Rubin BE, Wetmore KM, Price MN, Diamond S, Shultzaberger RK, Lowe LC, et al. The essential gene set of a photosynthetic organism. Proc Natl Acad Sci U S A. 2015;112: E6634–43.

45. Sagan L. On the origin of mitosing cells. J Theor Biol. 1967;14: 255–274.

46. Zimorski V, Ku C, Martin WF, Gould SB. Endosymbiotic theory for organelle origins. Curr Opin Microbiol. 2014;22: 38–48.

47. Barkan A. Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. EMBO J. 1988;7: 2637–2644.

48. Barkan A, Walker M, Nolasco M, Johnson D. A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J. 1994;13: 3170–3181.

49. Hirose T, Sugiura M. Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast ndhD mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. EMBO J. 1997;16: 6804–6811.

50. Zhou F, Karcher D, Bock R. Identification of a plastid intercistronic expression element (IEE) facilitating the expression of stable translatable monocistronic mRNAs from operons. Plant J. 2007;52: 961–972.

51. Ichinose M, Sugita M. RNA Editing and Its Molecular Mechanism in Plant Organelles. Genes . 2016;8. doi:10.3390/genes8010005

52. Shikanai T. RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci. Springer; 2006;63: 698–708.

53. Bock R. Sense from nonsense: How the genetic information of chloroplastsis altered by RNA editing. Biochimie. Elsevier; 2000;82: 549–557.

54. Shikanai T. RNA editing in plants: Machinery and flexibility of site recognition. Biochim Biophys Acta. 2015;1847: 779–785.

55. Knie N, Grewe F, Fischer S, Knoop V. Reverse U-to-C editing exceeds C-to-U RNA editing in some ferns--a monilophyte-wide comparison of chloroplast and mitochondrial RNA editing suggests independent evolution of the two processes in both organelles. BMC Evol Biol. BioMed Central; 2016;16: 134.

56. Soll J, Schleiff E. Protein import into chloroplasts. Nat Rev Mol Cell Biol. nature.com; 2004;5: 198–208.

57. Woodson JD, Chory J. Coordination of gene expression between organellar and nuclear genomes. Nat Rev Genet. 2008;9: 383–395.

58. Schmitz-Linneweber C, Small I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci. 2008;13: 663–670.

59. Barkan A, Small I. Pentatricopeptide repeat proteins in plants. Annu Rev Plant Biol. annualreviews.org; 2014;65: 415–442.

60. Pfalz J, Bayraktar OA, Prikryl J, Barkan A. Site-specific binding of a PPR protein defines and stabilizes 5’ and 3' mRNA termini in chloroplasts. EMBO J. 2009;28: 2042–2052.

61. Moller SG. Annual Plant Reviews, Plastids. John Wiley & Sons; 2009.

62. Link G. Plastid Differentiation: Organelle Promoters and Transcription Factors. Plant Promoters and Transcription Factors. Springer Berlin Heidelberg; 1994. pp. 65–85.

63. Leech RM, Rumsby MG, Thomson WW. Plastid differentiation, acyl lipid, and Fatty Acid changes in developing green maize leaves. Plant Physiol. 1973;52: 240–245.

64. Harris WM, Spurr AR. CHROMOPLASTS OF TOMATO FRUITS. II. THE RED TOMATO. Am J Bot. 1969;56: 380–389.

65. Ishizaki K, Chiyoda S, Yamato KT, Kohchi T. Agrobacterium-mediated transformation of the haploid liverwort Marchantia polymorpha L., an emerging model for plant biology. Plant Cell Physiol. academic.oup.com; 2008;49: 1084–1091.

66. Era A, Tominaga M, Ebine K, Awai C, Saito C, Ishizaki K, et al. Application of Lifeact reveals F-actin dynamics in Arabidopsis thaliana and the liverwort, Marchantia polymorpha. Plant Cell Physiol. academic.oup.com; 2009;50: 1041–1048.

67. McDaniel SF, Perroud P-F. Invited perspective: bryophytes as models for understanding the evolution of sexual systems. Bryologist. The American Bryological and Lichenological Society, Inc.; 2012;115: 1–11.

68. Shaw AJ, Szövényi P, Shaw B. Bryophyte diversity and evolution: windows into the early evolution of land plants. Am J Bot. 2011;98: 352–369.

69. Puttick MN, Morris JL, Williams TA, Cox CJ, Edwards D, Kenrick P, et al. The Interrelationships of Land Plants and the Nature of the Ancestral Embryophyte. Curr

Biol. Elsevier; 2018;28: 733–745.e2.

70. Cox CJ, Li B, Foster PG, Embley TM, Civán P. Conflicting phylogenies for early land plants are caused by composition biases among synonymous substitutions. Syst Biol. 2014;63: 272–279.

71. Qiu Y-L, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, et al. The deepest divergences in land plants inferred from phylogenomic evidence. Proc Natl Acad Sci U S A. 2006;103: 15511–15516.

72. Qiu YL, Cho Y, Cox JC, Palmer JD. The gain of three mitochondrial introns identifies liverworts as the earliest land plants. Nature. 1998;394: 671–674.

73. Mirbel M. Researches anatomiques et physiologiques sur le Marchantia polymorpha. Mém Acad Roy Sc Inst France. 1835;13: 337–436.

74. Beauverie J-J. Etudes des modifications morphologiques et anatomiques de thalles de Marchantia et de Lunularia obtenues expérimentalement. Publications de la Société Linnéenne de Lyon. Persée - Portail des revues scientifiques en SHS; 1898;44: 57–69.

75. Prior PV, Brown PR. Germination and Early Growth Responses of Marchantia Gemmae to Some Growth Regulators. Bryologist. American Bryological and Lichenological Society; 1970;73: 687–691.

76. Benson-Evans K. Physiology of the Reproduction of Bryophytes. Bryologist. American Bryological and Lichenological Society; 1964;67: 431–445.

77. Daniell H. Universal chloroplast integration and expression vectors, transformed plants and products thereof [Internet]. US Patent. 7129391, 2006. Available: https://patentimages.storage.googleapis.com/0c/d1/0b/86018aa8155bf2/US7129391.pdf

78. Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, et al. Chloroplast gene organization deduced from complete sequence of liverwort Marchantia polymorpha chloroplast DNA. Nature. 1986;322: 572.

79. Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, et al. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. 1986;5: 2043–2049.

80. Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, et al. Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. Cell. Elsevier; 2017;171: 287–304.e15.

81. Kubota A, Ishizaki K, Hosaka M, Kohchi T. Efficient Agrobacterium-mediated transformation of the liverwort Marchantia polymorpha using regenerating thalli. Biosci Biotechnol Biochem. Taylor & Francis; 2013;77: 167–172.

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

83. Chiyoda S, Linley PJ, Yamato KT, Fukuzawa H, Yokota A, Kohchi T. Simple and efficient plastid transformation system for the liverwort Marchantia polymorpha L. suspension-culture cells. Transgenic Res. Springer; 2007;16: 41–49.

84. Chiyoda S, Yamato KT, Kohchi T. Plastid transformation of sporelings and suspension-cultured cells from the liverwort Marchantia polymorpha L. Methods Mol Biol. Springer; 2014;1132: 439–447.

85. Blakeslee AF. Differentiation of Sex in Thallus Gametophyte and Sporophyte. Bot Gaz. University of Chicago Press; 1906;42: 161–178.

86. Voth PD. Effects of Nutrient-Solution Concentration on the Growth of Marchantia polymorpha. Bot Gaz. 1943;104: 591–601.

87. Katoh K, Ishikawa M, Miyake K, Ohta Y, Hirose Y, Iwamura T. Nutrient utilization and requirement under photoheterotrophic growth of Marchantia polymorpha: Improvement of the culture medium. Physiol Plant. 1980;49: 241–247.

88. TAKEUCHI, M. Long-term freeze-preservation of protoplasts of carrot and Marchantia. CryoLetters. ci.nii.ac.jp; 1980;1: 519–524.

89. Tanaka D, Ishizaki K, Kohchi T, Yamato KT. Cryopreservation of Gemmae from the Liverwort Marchantia polymorpha L. Plant Cell Physiol. academic.oup.com; 2016;57: 300–306.

90. Barampuram S, Zhang ZJ. Recent Advances in Plant Transformation. In: Birchler JA, editor. Plant Chromosome Engineering: Methods and Protocols. Totowa, NJ: Humana Press; 2011. pp. 1–35.

91. Gelvin SB. Agrobacterium-Mediated Plant Transformation: the Biology behind the “Gene-Jockeying” Tool. Microbiol Mol Biol Rev. 2003;67: 16–37.

92. Ziemienowicz A. Agrobacterium-mediated plant transformation: Factors, applications and recent advances. Biocatal Agric Biotechnol. 2014;3: 95–102.

93. Rech EL, Vianna GR, Aragão FJL. High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. Nat Protoc. 2008;3: 410–418.

94. Kikkert JR, Vidal JR, Reisch BI. Stable Transformation of Plant Cells by Particle Bombardment/Biolistics. In: Peña L, editor. Transgenic Plants: Methods and Protocols. Totowa, NJ: Humana Press; 2004. pp. 61–78.

95. Rivera AL, Gómez-Lim M, Fernández F, Loske AM. Physical methods for genetic plant transformation. Phys Life Rev. 2012;9: 308–345.

96. Verma D, Daniell H. Chloroplast vector systems for biotechnology applications. Plant Physiol. 2007;145: 1129–1143.

97. Bendich AJ. Why do chloroplasts and mitochondria contain so many copies of their genome? Bioessays. 1987;6: 279–282.

98. Ashraf S, Singh PK, Yadav DK, Shahnawaz M, Mishra S, Sawant SV, et al. High level expression of surface glycoprotein of rabies virus in tobacco leaves and its immunoprotective activity in mice. J Biotechnol. 2005;119: 1–14.

99. Hood EE, Bailey MR, Beifuss K, Magallanes-Lundback M, Horn ME, Callaway E, et al. Criteria for high-level expression of a fungal laccase gene in transgenic maize. Plant Biotechnol J. Wiley Online Library; 2003;1: 129–140.

100. Staub JM, Garcia B, Graves J, Hajdukiewicz PT, Hunter P, Nehra N, et al. High-yield production of a human therapeutic protein in tobacco chloroplasts. Nat Biotechnol. 2000;18: 333–338.

101. De Cosa B, Moar W, Lee SB, Miller M, Daniell H. Overexpression of the Bt cry2Aa2 operon in chloroplasts leads to formation of insecticidal crystals. Nat Biotechnol. 2001;19: 71–74.

102. Oey M, Lohse M, Kreikemeyer B, Bock R. Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. Plant J.

2009;57: 436–445.

103. Wurbs D, Ruf S, Bock R. Contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. Plant J. Wiley Online Library; 2007;49: 276–288.

104. Kumar S, Dhingra A, Daniell H. Stable transformation of the cotton plastid genome and maternal inheritance of transgenes. Plant Mol Biol. Springer; 2004;56: 203–216.

105. Daniell H. Molecular strategies for gene containment in transgenic crops. Nat Biotechnol. 2002;20: 581–586.

106. Birky CW Jr. Uniparental inheritance of organelle genes. Curr Biol. 2008;18: R692–5.

107. Smith SE. Biparental inheritance of organelles and its implications in crop improvement. Plant Breeding Reviews, Volume 6. Wiley Online Library; 1989; 361–393.

108. Svab Z, Maliga P. Exceptional transmission of plastids and mitochondria from the transplastomic pollen parent and its impact on transgene containment. Proc Natl Acad Sci U S A. 2007;104: 7003–7008.

109. Ruf S, Karcher D, Bock R. Determining the transgene containment level provided by chloroplast transformation. Proc Natl Acad Sci U S A. National Acad Sciences; 2007;104: 6998–7002.

110. Daniell H. Transgene containment by maternal inheritance: effective or elusive? Proc Natl Acad Sci U S A. 2007;104: 6879–6880.

111. Daniell H, Khan MS, Allison L. Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. Trends Plant Sci. Elsevier; 2002;7: 84–91.

112. Boehm CR, Ueda M, Nishimura Y, Shikanai T, Haseloff J. A Cyan Fluorescent Reporter Expressed from the Chloroplast Genome of Marchantia polymorpha. Plant Cell Physiol. academic.oup.com; 2016;57: 291–299.

113. Iida A, Seki M, Kamada M, Yamada Y, Morikawa H. Gene delivery into cultured plant cells by DNA-coated gold particles accelerated by a pneumatic particle gun. Theor Appl Genet. 1990;80: 813–816.

114. Carrer H, Hockenberry TN, Svab Z, Maliga P. Kanamycin resistance as a selectable marker for plastid transformation in tobacco. Mol Gen Genet. 1993;241: 49–56.

115. Barone P, Zhang X-H, Widholm JM. Tobacco plastid transformation using the feedback-insensitive anthranilate synthase [α]-subunit of tobacco (ASA2) as a new selectable marker. J Exp Bot. Oxford University Press; 2009;60: 3195–3202.

116. Maliga P. Engineering the plastid genome of higher plants. Curr Opin Plant Biol. 2002;5: 164–172.

117. Prikryl J, Rojas M, Schuster G, Barkan A. Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. Proc Natl Acad Sci U S A. 2011;108: 415–420.

118. Eibl C, Zou Z, Beck A, Kim M, Mullet J, Koop HU. In vivo analysis of plastid psbA, rbcL and rpl32 UTR elements by chloroplast transformation: tobacco plastid gene expression is controlled by modulation of transcript levels and translation efficiency.

Plant J. 1999;19: 333–345.

119. Boynton JE, Gillham NW, Harris EH, Hosler JP, Johnson AM, Jones AR, et al. Chloroplast transformation in Chlamydomonas with high velocity microprojectiles. Science. science.sciencemag.org; 1988;240: 1534–1538.

120. Svab Z, Hajdukiewicz P, Maliga P. Stable transformation of plastids in higher plants. Proc Natl Acad Sci U S A. National Acad Sciences; 1990;87: 8526–8530.

121. Scharff LB, Bock R. Synthetic biology in plastids. Plant J. 2014;78: 783–798.

122. Heifetz PB. Genetic engineering of the chloroplast. Biochimie. 2000;82: 655–666.

123. Maliga P, Svab Z. Engineering the Plastid Genome of Nicotiana sylvestris, a Diploid Model Species for Plastid Genetics. In: Birchler JA, editor. Plant Chromosome Engineering: Methods and Protocols. Totowa, NJ: Humana Press; 2011. pp. 37–50.

124. Lee SM, Kang K, Chung H, Yoo SH, Xu XM, Lee S-B, et al. Plastid transformation in the monocotyledonous cereal crop, rice (Oryza sativa) and transmission of transgenes to their progeny. Mol Cells. 2006;21: 401–410.

125. Ruf S, Hermann M, Berger IJ, Carrer H, Bock R. Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. Nat Biotechnol. nature.com; 2001;19: 870–875.

126. Sidorov VA, Kasten D, Pang S-Z, Hajdukiewicz PTJ, Staub JM, Nehra NS. Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. Plant J. Wiley Online Library; 1999;19: 209–216.

127. Singh AK, Verma SS, Bansal KC. Plastid transformation in eggplant (Solanum melongena L.). Transgenic Res. Springer; 2010;19: 113–119.

128. Dufourmantel N, Pelissier B, Garçon F, Peltier G, Ferullo J-M, Tissot G. Generation of fertile transplastomic soybean. Plant Mol Biol. 2004;55: 479–489.

129. Cui C, Song F, Tan Y, Zhou X, Zhao W, Ma F, et al. Stable chloroplast transformation of immature scutella and inflorescences in wheat (Triticum aestivum L.). Acta Biochim Biophys Sin . academic.oup.com; 2011;43: 284–291.

130. Yu Q, Lutz KA, Maliga P. Efficient Plastid Transformation in Arabidopsis. Plant Physiol. 2017;175: 186–193.

131. Eeckhaut T, Lakshmanan PS, Deryckere D, Van Bockstaele E, Van Huylenbroeck J. Progress in plant protoplast research. Planta. Springer; 2013;238: 991–1003.

132. Golds T, Maliga P, Koop H-U. Stable Plastid Transformation in PEG-treated Protoplasts of Nicotiana tabacum. Biotechnology. Nature Publishing Company; 1993;11: 95.

133. O'Neill C, Horváth GV, Horváth E, Dix PJ, Medgyesy P. Chloroplast transformation in plants: polyethylene glycol (PEG) treatment of protoplasts is an alternative to biolistic delivery systems. Plant J. 1993;3: 729–738.

134. Huang F-C, Klaus SMJ, Herz S, Zou Z, Koop H-U, Golds TJ. Efficient plastid transformation in tobacco using the aphA-6 gene and kanamycin selection. Mol Genet Genomics. 2002;268: 19–27.

135. Lelivelt CLC, McCabe MS, Newell CA, deSnoo CB, van Dun KMP, Birch-Machin I, et al. Stable plastid transformation in lettuce (Lactuca sativa L.). Plant Mol Biol. Springer; 2005;58: 763–774.

136. Sugiura C, Sugita M. Plastid transformation reveals that moss tRNAArg-CCG is not essential for plastid function. Plant J. Wiley Online Library; 2004;40: 314–321.

137. Nugent GD, Coyne S, Nguyen TT, Kavanagh TA, Dix PJ. Nuclear and plastid transformation of Brassica oleracea var. botrytis (cauliflower) using PEG-mediated uptake of DNA into protoplasts. Plant Sci. Elsevier; 2006;170: 135–142.

138. Knoblauch M, Hibberd JM, Gray JC, van Bel AJ. A galinstan expansion femtosyringe for microinjection of eukaryotic organelles and prokaryotes. Nat Biotechnol. nature.com; 1999;17: 906–909.

139. Kindle KL, Richards KL, Stern DB. Engineering the chloroplast genome: techniques and capabilities for chloroplast transformation in Chlamydomonas reinhardtii. Proc Natl Acad Sci U S A. 1991;88: 1721–1725.

140. James C. Global status of commercialized biotech/GM crops: 2014. ISAAA brief. isaaa.org; 2015; Available: http://www.isaaa.org/resources/publications/briefs/43/download/isaaa-brief-43-2011.pdf

141. Daniell H, Datta R, Varma S, Gray S, Lee SB. Containment of herbicide resistance through genetic engineering of the chloroplast genome. Nat Biotechnol. 1998;16: 345–348.

142. Kota M, Daniell H, Varma S, Garczynski SF, Gould F, Moar WJ. Overexpression of the Bacillus thuringiensis (Bt) Cry2Aa2 protein in chloroplasts confers resistance to plants against susceptible and Bt-resistant insects. Proc Natl Acad Sci U S A. 1999;96: 1840–1845.

143. Tabashnik BE, Brévault T, Carrière Y. Insect resistance to Bt crops: lessons from the first billion acres. Nat Biotechnol. 2013;31: 510–521.

144. Streatfield SJ. Approaches to achieve high-level heterologous protein production in plants. Plant Biotechnol J. 2007;5: 2–15.

145. Walmsley AM, Arntzen CJ. Plants for delivery of edible vaccines. Curr Opin Biotechnol. 2000;11: 126–129.

146. Daniell H, Lee SB, Panchal T, Wiebe PO. Expression of the native cholera toxin B subunit gene and assembly as functional oligomers in transgenic tobacco chloroplasts. J Mol Biol. 2001;311: 1001–1009.

147. Koya V, Moayeri M, Leppla SH, Daniell H. Plant-based vaccine: mice immunized with chloroplast-derived anthrax protective antigen survive anthrax lethal toxin challenge. Infect Immun. Am Soc Microbiol; 2005;73: 8266–8274.

148. Stephanopoulos G. Metabolic engineering. Biotechnol Bioeng. Wiley Online Library; 1998;58: 119–120.

149. Galanie S, Thodey K, Trenchard IJ, Filsinger Interrante M, Smolke CD. Complete biosynthesis of opioids in yeast. Science. science.sciencemag.org; 2015;349: 1095–1100.

150. Nakagawa A, Matsumura E, Koyanagi T, Katayama T, Kawano N, Yoshimatsu K, et al. Total biosynthesis of opiates by stepwise fermentation using engineered Escherichia coli. Nat Commun. 2016;7: 10390.

151. Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, et al. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. Nat Biotechnol. nature.com; 2005;23: 482–487.

152. Hasunuma T, Miyazawa S-I, Yoshimura S, Shinzaki Y, Tomizawa K-I, Shindo K, et al. Biosynthesis of astaxanthin in tobacco leaves by transplastomic engineering. Plant J. 2008;55: 857–868.

153. Jayaraj J, Devlin R, Punja Z. Metabolic engineering of novel ketocarotenoid production in carrot plants. Transgenic Res. 2008;17: 489–501.

154. Munekage Y, Hashimoto M, Miyake C, Tomizawa K-I, Endo T, Tasaka M, et al. Cyclic electron flow around photosystem I is essential for photosynthesis. Nature. 2004;429: 579–582.

155. Kuroda H, Maliga P. The plastid clpP1 protease gene is essential for plant development. Nature. nature.com; 2003;425: 86–89.

156. Flores-Pérez Ú, Sauret-Güeto S, Gas E, Jarvis P, Rodríguez-Concepción M. A Mutant Impaired in the Production of Plastome-Encoded Proteins Uncovers a Mechanism for the Homeostasis of Isoprenoid Biosynthetic Enzymes in Arabidopsis Plastids. Plant Cell. American Society of Plant Biologists; 2008;20: 1303–1315.

157. Kode V, Mudd EA, Iamtham S, Day A. The tobacco plastid accD gene is essential and is required for leaf development. Plant J. Wiley Online Library; 2005;44: 237–244.

158. Fleischmann TT, Scharff LB, Alkatib S, Hasdorf S, Schöttler MA, Bock R. Nonessential plastid-encoded ribosomal proteins in tobacco: a developmental role for plastid translation and implications for reductive genome evolution. Plant Cell. 2011;23: 3137–3155.

159. Watt VM, Ingles CJ, Urdea MS, Rutter WJ. Homology requirements for recombination in Escherichia coli. Proc Natl Acad Sci U S A. 1985;82: 4768–4772.

160. Hanahan D, Jessee J, Bloom FR. [4] Plasmid transformation of Escherichia coli and other bacteria. Methods in Enzymology. Academic Press; 1991. pp. 63–113.

161. Pollak B, Cerda A, Delmans M, Álamos S, Moyano T, West A, et al. Loop Assembly: a simple and open system for recursive fabrication of DNA circuits [Internet]. bioRxiv. 2018. p. 247593. doi:10.1101/247593

162. Kahl L. Draft OpenMTA 11 June 2017, for comment. dspace.mit.edu; 2017; Available: https://dspace.mit.edu/handle/1721.1/111012?show=full

163. Boisnard-Lorig C, Colon-Carmona A, Bauch M, Hodge S, Doerner P, Bancharel E, et al. Dynamic analyses of the expression of the HISTONE::YFP fusion protein in arabidopsis show that syncytial endosperm is divided in mitotic domains. Plant Cell. 2001;13: 495–509.

164. Benchling (Biology Software) [Internet]. [cited 2018]. Available: www.benchling.com

165. Multiple Primer Analyzer [Internet]. [cited 2018]. Available: https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html

166. Breslauer KJ, Frank R, Blöcker H, Marky LA. Predicting DNA duplex stability from the base sequence. Proc Natl Acad Sci U S A. 1986;83: 3746–3750.

167. Gamborg OL, Miller RA, Ojima K. Nutrient requirements of suspension cultures of soybean root cells. Exp Cell Res. Elsevier; 1968;50: 151–158.

168. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nature Methods. 2012;9:

676–682.

169. Kijak H, Rurek M, Nowak W, Dabert M, Odrzykoski IJ. Resequencing the “classic” Marchantia polymorpha chloroplast genome [Internet]. researchgate.net; 2016 [cited 10 Jul 2018]. Available: https://www.researchgate.net/profile/Hania_Kijak/publication/308699734_Resequencing_the_'classic'_Marchantia_polymorpha_chloroplast_genome/links/57ebf43d08ae93b7fa9578b5.pdf

170. Svab Z, Maliga P. High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. PNAS (USA). 1993;90: 913–917.

171. Klaus SMJ, Huang F-C, Eibl C, Koop H-U, Golds TJ. Rapid and proven production of transplastomic tobacco plants by restoration of pigmentation and photosynthesis. Plant J. 2003;35: 811–821.

172. Lohse M, Drechsel O, Kahlau S, Bock R. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res. Oxford University Press; 2013;41: W575–W581.

173. Lu X-M, Yin W-B, Hu Z-M. Chloroplast Transformation. In: Loyola-Vargas VM, Vázquez-Flota F, editors. Plant Cell Culture Protocols. Totowa, NJ: Humana Press; 2006. pp. 285–303.

174. Shao M, Kumar S, Thomson JG. Precise excision of plastid DNA by the large serine recombinase Bxb1. Plant Biotechnol J. 2014;12: 322–329.

175. Barrera D, Gimpel J, Mayfield S. Rapid Screening for the Robust Expression of Recombinant Proteins in Algal Plastids. In: Maliga P, editor. Chloroplast Biotechnology: Methods and Protocols. Totowa, NJ: Humana Press; 2014. pp. 391–399.

176. Toyoda H, Yamaga T, Matsuda Y, Ouchi S. Transient expression of the β-glucuronidase gene introduced into barley coleoptile cells by microinjection. Plant Cell Rep. Springer; 1990;9: 299–302.

177. Tzfira T, Citovsky V. Agrobacterium-mediated genetic transformation of plants: biology and biotechnology. Curr Opin Biotechnol. 2006;17: 147–154.

178. Irshad H, Veillard A, Roux L, Racoceanu D. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. IEEE Rev Biomed Eng. IEEE; 2014;7: 97–114.

179. Gonzales-Barron U, Butler F. A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumb features by digital image analysis. J Food Eng. 2006;74: 268–278.

180. Krämer P, Boto F, Wald D, Bessy F, Paloc C, Callol C, et al. Comparison of Segmentation Algorithms for the Zebrafish Heart in Fluorescent Microscopy Images. Advances in Visual Computing. Springer Berlin Heidelberg; 2009. pp. 1041–1050.

181. ZUIDERVELD, K. Contrast Limited Adaptive Histogram Equalization. Graphics Gems. Academic Press; 1994; 474–485.

182. Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern. IEEE; 1979;9: 62–66.

183. Shapiro SS, Wilk MB. An Analysis of Variance Test for Normality (Complete Samples). Biometrika. [Oxford University Press, Biometrika Trust]; 1965;52: 591–611.

184. Wilcoxon F. Individual Comparisons by Ranking Methods. Biometrics Bulletin. [International Biometric Society, Wiley]; 1945;1: 80–83.

185. Kruskal WH, Wallis WA. Use of Ranks in One-Criterion Variance Analysis. J Am Stat Assoc. Taylor & Francis; 1952;47: 583–621.

186. Dunn OJ. Multiple Comparisons Using Rank Sums. Technometrics. 1964;6: 241.

187. Tsuboyama-Tanaka S, Kodama Y. AgarTrap-mediated genetic transformation using intact gemmae/gemmalings of the liverwort Marchantia polymorpha L. J Plant Res. 2015;128: 337–344.

188. Nathans D, Smith HO. Restriction endonucleases in the analysis and restructuring of dna molecules. Annu Rev Biochem. annualreviews.org; 1975;44: 273–293.

189. Chang S, Cohen SN. In vivo site-specific genetic recombination promoted by the EcoRI restriction endonuclease. Proc Natl Acad Sci U S A. National Acad Sciences; 1977;74: 4811–4815.

190. Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS One. journals.plos.org; 2011;6: e16765.

191. Werner S, Engler C, Weber E, Gruetzner R, Marillonnet S. Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. Bioeng Bugs. Taylor & Francis; 2012;3: 38–43.

192. Sarrion-Perdigones A, Falconi EE, Zandalinas SI, Juárez P, Fernández-del-Carmen A, Granell A, et al. GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. PLoS One. journals.plos.org; 2011;6: e21622.

193. Patron NJ, Orzaez D, Marillonnet S, Warzecha H, Matthewman C, Youles M, et al. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. New Phytol. Wiley Online Library; 2015;208: 13–19.

194. Andrianantoandro E, Basu S, Karig DK, Weiss R. Synthetic biology: new engineering rules for an emerging discipline. Mol Syst Biol. msb.embopress.org; 2006;2: 2006.0028.

195. Quinn J, Beal J, Bhatia S, Cai P, Chen J, Clancy K, et al. Synthetic biology open language visual (SBOL visual), version 1.0. 0 [Internet]. dspace.mit.edu; 2013. Available: https://dspace.mit.edu/bitstream/handle/1721.1/78249/BBFRFC93.pdf?sequence=1

196. Goedhart J, von Stetten D, Noirclerc-Savoye M, Lelimousin M, Joosen L, Hink MA, et al. Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. Nat Commun. 2012;3: 751.

197. Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat Biotechnol. 2002;20: 87–90.

198. Shimomura O, Johnson FH, Saiga Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. J Cell Comp Physiol. Wiley Online Library; 1962;59: 223–239.

199. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. Science. science.sciencemag.org; 1994;263: 802–805.

200. Sauvola J, Pietikäinen M. Adaptive document image binarization. Pattern Recognit. 2000;33: 225–236.

201. Phansalkar N, More S, Sabale A, Joshi M. Adaptive Local Thresholding for Detection of Nuclei in Diversely Stained Cytology Images. Available: https://it.mathworks.com/matlabcentral/answers/uploaded_files/23812/ICCSP_thresholding_2.pdf

202. Chen Q, Yang X, Petriu EM. Watershed segmentation for binary images with different distance transforms. Proceedings of the 3rd IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications. andrei.clubcisco.ro; 2004. pp. 111–116.

203. Wani SH, Sah SK, Sági L, Solymosi K. Transplastomic plants for innovations in agriculture. A review. Agron Sustain Dev. 2015;35: 1391–1430.

204. Sidorov V, Staub JM, Wan Y, Ye G. Plastid Transformation of Maize [Internet]. US Patent. 20170314032:A1, 2017. Available: https://patentimages.storage.googleapis.com/32/33/81/d1445d31b72f8f/US20170314032A1.pdf

205. Wang Y, Wei Z, Xing S. Stable plastid transformation of rice, a monocot cereal crop. Biochem Biophys Res Commun. 2018; doi:10.1016/j.bbrc.2018.06.164

206. Singh B, Sharma RA. Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. 3 Biotech. 2015;5: 129–151.

207. Dickschat JS. Isoprenoids in three-dimensional space: the stereochemistry of terpene biosynthesis. Nat Prod Rep. 2011;28: 1917–1936.

208. Maimone TJ, Baran PS. Modern synthetic efforts toward biologically active terpenes. Nat Chem Biol. 2007;3: 396–407.

209. Kindermans J-M, Pilloy J, Olliaro P, Gomes M. Ensuring sustained ACT production and reliable artemisinin supply. Malar J. 2007;6: 125.

210. Lichtenthaler HK. THE 1-DEOXY-D-XYLULOSE-5-PHOSPHATE PATHWAY OF ISOPRENOID BIOSYNTHESIS IN PLANTS. Annu Rev Plant Physiol Plant Mol Biol. 1999;50: 47–65.

211. Barretto S, Michoux F, Nixon PJ. Temporary Immersion Bioreactors for the Contained Production of Recombinant Proteins in Transplastomic Plants. Methods Mol Biol. Springer; 2016;1385: 149–160.

212. Klaus SMJ, Huang F-C, Golds TJ, Koop H-U. Generation of marker-free plastid transformants using a transiently cointegrated selection gene. Nat Biotechnol. nature.com; 2004;22: 225–229.

213. Krech K, Fu H-Y, Thiele W, Ruf S, Schöttler MA, Bock R. Reverse genetics in complex multigene operons by co-transformation of the plastid genome and its application to the open reading frame previously designated psbN. Plant J. 2013;75: 1062–1074.

214. Day A, Goldschmidt-Clermont M. The chloroplast transformation toolbox: selectable markers and marker removal. Plant Biotechnol J. Wiley Online Library; 2011;9: 540–553.

215. Takenaka M, Yamaoka S, Hanajiri T, Shimizu-Ueda Y, Yamato KT, Fukuzawa H, et al. Direct transformation and plant regeneration of the haploid liverwort Marchantia polymorpha L. Transgenic Res. Springer; 2000;9: 179–185.