Alteration of potato plastids as platforms for biotechnology and synthetic biology

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Alexander Carl Pfotenhauer
August 2021

Acknowledgements

I would like to say thank you to my mentor Scott Lenaghan for encouraging me to pursue graduate school in a field I knew nothing about. I will always appreciate him taking a chance on someone who was completely unfamiliar with the material, and his countless hours spent discussing science and other topics with me. I also want to thoroughly acknowledge Alessandro Occhialini, whose experience and training helped to answer many of my questions. Thank you to Neal Stewart for his comprehension of plant biotechnology and support throughout graduate school. Thank you to my other committee members, David White and Curtis Luckett for your insight and questions throughout graduate school. Thank you to all of the other lab members who helped me over the years. I would especially like to thank my girlfriend Alyssa Farr, who encouraged me to continue my education, as well as my parents and sister. Final thanks goes to my two dogs Rowdy and Ramsey, and cat Callie, for always being a source of happiness at home.

Abstract

Plastids represent a unique opportunity for plant biotechnology and synthetic biology. Their fundamental uses for photosynthesis, starch storage, and other processes can be exploited to benefit food production and other industries. In this work, we attempted to modulate these plastids in order to produce potato plants with beneficial characteristics for further use in synthetic biology. The first chapter of this dissertation involves genome-editing of the *FtsZ1* gene, which is involved in plastid division. Plants were generated with large plastids and starch granules, and similar methods could be used to produce these plants without foreign DNA integration. The second chapter involves generating similarly large plastid plants that were used for chloroplast transformation. These plants could be potentially useful for the installation of large constructs via microinjection. The final chapter involves the generation of an alternative chloroplast transformation method, where plasmids were maintained within the organelle without integration to the plastome. Next generations of this method could be used to integrate specified genes, while allowing others to be only transiently expressed.

Table of Contents

List of Tables

List of Figures

Introduction

Current status of genome-editing

Biotechnology has allowed for direct manipulation of plant genomes, and a popular tool since 2013 is the clustered regularly interspaced short palindromic repeats (CRISPR) system [1]. CRISPR was originally discovered in bacteria as an adaptive immune system, in which short fragments of invading viral DNA are integrated into the bacterial genome. These fragments are used as a sort of memory by the bacterium [2]. Once transcribed, the RNA fragment forms a ribonucleoprotein with an endonuclease, Cas9, to cleave foreign DNA. The fragments are used as sequence-targeting machinery for the Cas9 protein. Researchers have since exploited this machinery by designing guide RNAs (gRNAs), and expressing them with Cas9 protein to cleave complementary DNA of interest. The double strand breaks resulting from Cas9 can be useful for gene knockout due to the error prone non-homologous end joining DNA repair pathway executed by plants [3]. CRISPR/Cas9 has been widely implemented in plants because of its ease of experimental implementation [4]. Its relative simplicity allows for target genes of any interest to not only be targeted for editing, but also for modulating expression patterns [5]. Fusing a transcriptional activator or repressor to an edited “dead”Cas9 allows for the gRNA targeting tool to either up or down regulate genes of interest. There have been many revisions to the CRISPR systems over the years, and likely improvements will continue to occur [6].

Traditional plant biotechnology typically uses *Agrobacterium tumefaciens* to deliver and integrate plasmid DNA into the plant genome. The transfer “t” DNA is integrated in a random fashion, and multiple insertions of intact or fragmented portions is typical [7-8]. Transgene expression levels vary widely depending on the location and number of insertions. Additionally, this integration commonly results in plants being considered genetically engineered by regulatory agencies [9]. Subsequent backcrossing or other breeding techniques are required to remove the integrated DNA. In March 2018, the USDA announced that plants that have been genetically engineered without DNA incorporation into the plant genome need not enter their regulatory framework; they would be treated by the agency as conventionally bred plants [10]. In order to circumvent these potential obstacles, CRISPR/Cas9 reagents can be expressed transiently in protoplasts in the form of mRNA or pre-assembled ribonucleoprotein [11]. The CRISPR/Cas9 reagents perform their function in the plant cells, and are subsequently degraded without any foreign DNA integration. Plants can then be regenerated from these edited protoplasts, ultimately producing plants with edited phenotypes without any transgene insertion. A variety of plants have been recovered in this manor, including potato [12-14].

The role of plastids for plant biotechnology

Plants contain organelles known as plastids, of which there are a variety of subcategories [15]. All plastids arise from undifferentiated proplastids, and develop into specialized forms such as chloroplasts during cell development. Each specialized plastid performs a specific function, such as photosynthesis in chloroplasts, and starch synthesis and storage in amyloplasts. Plastids are thought to have originated from enveloped bacteria, and they share many similar characteristics to their progenitor microorganisms [16]. They possess their own genome, termed plastome, which interacts with the nucleus through retrograde signaling [17]. Also similar to bacteria, plastids divide through binary fission, which is mediated through the FtsZ protein family [18]. The expression level and stoichiometry of FtsZ proteins has shown to be important for proper plastid division in several plants [19-22]. Specifically for potato, both an increase and decrease of FtsZ1 levels disrupted plastid division, resulting in plastids that continued to grow without dividing [23]. These potato plants contained plastids that were much larger than wild-type, termed “macro-plastids”. Because starch granules are formed and stored in amyloplasts, and each amyloplast typically contains one starch granule, these plants contained starch granules that were also larger than wild-type. The large starch granule tubers contained more phosphorus, and displayed higher viscosity parameters, which could be industrially useful for both food and non-food purposes [24].

As previously stated, plastids contain their own non-nuclear genomes. Though most plant biotechnology focuses on transient expression or transformation of nuclear genomes, there are significant advantages to transforming the plastome [25]. The lack of gene silencing, the precise method of homologous-recombination mediated transformation, and the ability to accumulate large amounts of foreign protein are only some of these favorable distinctions. Plastids, typically chloroplasts, are currently only transformed through physical methods [26]. Though PEG mediated and microinjection have been used successfully, the most commonly used method is through a gene gun. The gene gun is a relatively simple tool, which builds and releases helium pressure in a directed manner. Researchers bind construct DNA onto inert metal microbeads, and then use the gene gun to accelerate the microbeads into plant leaf tissue. The microbeads must permeate the chloroplasts in order to release DNA directly into the organelle. If successful, transgenes are integrated into the plastome in a site-specific manner due to homologous regions present on the transformation vectors. This relatively crude and entirely physical method of transferring DNA into the plastome results in very low levels of transformation.

Additional problems arise from using the most common antibiotic for chloroplast transformation, spectinomycin. This antibiotic works by inhibiting protein synthesis from ribosomes found in the chloroplast [27]. Since 1990, chloroplast transformation vectors have usually contained a selection cassette that expresses the aminoglycoside adenyltransferase (*aadA*) resistance gene [28-29]

[30]. *aadA* had previously been shown functional in several bacterial species, and was quickly adopted for chloroplasts [31]. Because tissue culture media is supplemented with large amounts of sucrose, plants can survive with this alternative source of carbon. The result is a sublethal selection, where leaf tissue in plants that are not expressing *aadA* bleach, but are nonetheless able to survive. Furthermore, nearly all monocot species have natural resistance to spectinomycin [27, 32]. Though a few papers have indicated that rice chloroplasts can be physically transformed, monocots in general are considered recalcitrant to transformation [33-34]. The purely physical transformation method, long tissue culture process following DNA bombardment, and poor selective markers result in chloroplast transformation being an inefficient and lengthy task.

Plastome architecture and method of replication

Plastomes' structures are debated, though the method of DNA integration is relatively straightforward [35]. Homologous "arms" flank transgene cassettes, and once introduced into the plastome these arms are used for precise insertion into predetermined sites via homologous recombination. Since the first tobacco plastome was sequenced in 1986, over 800 complete plastomes have been sequenced and made publically available [36]. Most chloroplast genomes are comprised of four parts, the small single copy (SSC), the large single copy (LSC), and two inverted repeats (IR). Because of the relatively small size, (potato plastome is ~150 kb), it is straightforward and possible to choose a variety of transgene insertion sites [37]. There are several well characterized integration sites, and vectors available for immediate use [38-39]. Researchers can choose to integrate in the IR regions and theoretically achieve twice as much expression as one would achieve in a single copy region. Alternatively, one can choose to insert in the SSC or LSC regions in a variety of loci for predetermined reasons. Nonetheless, the most common method for transgene expression in the chloroplast is through DNA insertion mediated by homologous recombination.

As with the plastome structure, the method of replication is also somewhat undetermined. Nonetheless, there are several factors that are mostly agreed upon. Plastome replication is not likely linked to cell cycle, and therefore many copies of the plastome are created and maintained within the organelles [40]. Many of the genes necessary for replication are expressed by the nucleus, though certain regions in the plastome must also be present. Some of these regions are origins of replications, "oris", that are responsible for the initiation of replication [41-42]. Oris are small regions of DNA that are recognized and bound by other proteins, and are essential for replication in many different microorganisms [43]. Though the exact manner of DNA replication is unclear, many agree that it relies at least partially on these oris.

Scientific gaps

Gap 1. Plant biotechnology research rarely reaches the market

Genetic transformation of plants allows for direct manipulation for predetermined traits, and has rapidly accelerated our understanding of fundamental and applied plant sciences. Regardless, many governing bodies require extensive consideration before transgenics reach the market, and very few crops are ever approved for release [44]. Fortunately in the United States, the USDA announced in 2018 that they would not consider plants that have been gene-edited as GMO, and these plants would circumvent the regulatory bottlenecks [45]. While this announcement is especially important for the agricultural industry, research institutions still primarily conduct experiments through the generation of transgenic plants. If useful phenotypes can ultimately be achieved through genome-editing, this would allow for the creation of many more plants to feed the world.

Gap 2. Chloroplast transformation is currently limited to inefficient physical methods

Regardless of the method used (PEG, microinjection, or gene gun mediated), transforming chloroplasts is extremely inefficient. The unpredictable efficiency of a gene gun leads to many hours of tissue culture in order to select and screen for transformed plants. Monocot plants are naturally resistant to the most commonly used antibiotic for selection, spectinomycin. These and other factors result in chloroplast transformation being largely time-consuming and inefficient. If any of the bottlenecks associated with chloroplast transformation can be eliminated, researchers can potentially use these very valuable organelles to help progress plant biotechnology.

Gap 3. Chloroplast biotechnology currently relies on one method of transgene expression

Complex gene circuits require coordinated and precise methods of control. Currently, transgene expression in chloroplasts is limited to two levels of control. The first is through modulating promoters, UTRs, intercistronic elements, and other regulatory sequences within their transgene cassettes. One problem with this level of control is a lack of research dedicated to quantifying and ranking various combinations of these regulatory sequences. Without personal experience of how to build transgene cassettes, one is left to choose from only a few well-characterized sequences. This lack of publicly available databases severely limits the ability to rationally design transformation vectors. The second level of control involves transformation insertion sites. Theoretically, one can choose to integrate in any site of a sequenced and publicly available plastome. However, there are few commonly used sites, and vectors are immediately available for these. Because these few sites have been validated for use, researchers rarely venture out to new loci on the plastome. The same problem that plagues regulatory element control affects the gene insertion site. Despite

the large potential for inserting anywhere in the plastome, researchers typically limit themselves to several more characterized locations [38-39]. Multiple layers of control will allow for coordinated gene expression through rationally designed transformation vectors.

Research objectives

Objective 1. Knockout *FtsZ1* to produce large starch granule potatoes
The purpose of this study was to generate a large starch granule potato plant through CRISPR/Cas9 mediated genome editing. In this study, we used single gRNAs and a relatively simple construct design that could be later applied in a DNA-free genome-editing manner. Further experiments could be conducted with our molecular design to produce non-GMO potato plants with large starch granules.

Objective 2. Generate macro-chloroplast potato for chloroplast transformation
The purpose of this study was to generate macro-chloroplast potato plants for chloroplast transformation. We predicted that using macro-chloroplast potato plants as tissue for transformation would increase the transformation efficiency. Further experiments could use these plants for other methods of transformation such as microinjection that would likely be necessary for the installation of large constructs.

Objective 3. Develop an alternative approach to transform chloroplasts
The purpose of this study was to generate an alternative approach for transforming chloroplasts. As opposed to the traditional approach of DNA integration for chloroplast transformation, we predicted that transformation could occur through the replication and persistence of replicating plasmids. By including documented chloroplast oris on our transformation vectors, we envisioned that these plasmids could be maintained as a separate DNA entity within the chloroplast. Further generations of our constructs could expand the level of gene expression and allow for another degree of metabolic control.

References

[1] Yin, K., Gao, C., Qiu, J. L., Progress and prospects in plant genome editing. *Nat Plants* 2017, *3*, 17107.
[2] Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R.*, et al.*, The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Res* 2011, *39*, 9275-9282.
[3] Gorbunova, V., Levy, A. A., Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Res* 1997, *25*, 4650-4657.
[4] Afzal, S., Sirohi, P., Singh, N. K., A review of CRISPR associated genome engineering: application, advances and future prospects of genome targeting tool for crop improvement. *Biotechnol Lett* 2020, *42*, 1611-1632.
[5] Moradpour, M., Abdulah, S. N. A., CRISPR/dCas9 platforms in plants: strategies and applications beyond genome editing. *Plant Biotechnol J* 2020, *18*, 32-44.
[6] Huang, T. K., Puchta, H., Novel CRISPR/Cas applications in plants: from prime editing to chromosome engineering. *Transgenic Res* 2021.
[7] Chilton, M. D., Que, Q., Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration. *Plant Physiol* 2003, *133*, 956-965.
[8] Ohba, T., Yoshioka, Y., Machida, C., Machida, Y., DNA rearrangement associated with the integration of T-DNA in tobacco: an example for multiple duplications of DNA around the integration target. *Plant J* 1995, *7*, 157-164.
[9] Kanchiswamy, C. N., Malnoy, M., Velasco, R., Kim, J. S., Viola, R., Non-GMO genetically edited crop plants. *Trends Biotechnol* 2015, *33*, 489-491.
[10] Press, U., Secretary Perdue issues USDA statement on plant breeding innovation. 2018, *0070.18*.
[11] Kanchiswamy, C. N., DNA-free genome editing methods for targeted crop improvement. *Plant Cell Rep* 2016, *35*, 1469-1474.
[12] Andersson, M., Turesson, H., Olsson, N., Fält, A. S.*, et al.*, Genome editing in potato via CRISPR-Cas9 ribonucleoprotein delivery. *Physiol Plant* 2018, *164*, 378-384.
[13] Makhotenko, A. V., Khromov, A. V., Snigir, E. A., Makarova, S. S.*, et al.*, Functional Analysis of Coilin in Virus Resistance and Stress Tolerance of Potato Solanum tuberosum using CRISPR-Cas9 Editing. *Dokl Biochem Biophys* 2019, *484*, 88-91.
[14] Zhao, X., Jayarathna, S., Turesson, H., Fält, A. S.*, et al.*, Amylose starch with no detectable branching developed through DNA-free CRISPR-Cas9 mediated mutagenesis of two starch branching enzymes in potato. *Sci Rep* 2021, *11*, 4311.
[15] Lopez-Juez, E., Pyke, K. A., Plastids unleashed: their development and their integration in plant development. *Int J Dev Biol* 2005, *49*, 557-577.
[16] Zimorski, V., Ku, C., Martin, W. F., Gould, S. B., Endosymbiotic theory for organelle origins. *Curr Opin Microbiol* 2014, *22*, 38-48.

[17] Nott, A., Jung, H. S., Koussevitzky, S., Chory, J., Plastid-to-nucleus retrograde signaling. *Annu Rev Plant Biol* 2006, *57*, 739-759.
[18] Osteryoung, K. W., Stokes, K. D., Rutherford, S. M., Percival, A. L., Lee, W. Y., Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial ftsZ. *Plant Cell* 1998, *10*, 1991-2004.
[19] Chikkala, V. R., Nugent, G. D., Stalker, D. M., Mouradov, A., Stevenson, T. W., Expression of Brassica oleracea FtsZ1-1 and MinD alters chloroplast division in Nicotiana tabacum generating macro- and mini-chloroplasts. *Plant Cell Rep* 2012, *31*, 917-928.
[20] Schmitz, A. J., Glynn, J. M., Olson, B. J., Stokes, K. D., Osteryoung, K. W., Arabidopsis FtsZ2-1 and FtsZ2-2 are functionally redundant, but FtsZ-based plastid division is not essential for chloroplast partitioning or plant growth and development. *Mol Plant* 2009, *2*, 1211-1222.
[21] Yoder, D. W., Kadirjan-Kalbach, D., Olson, B. J., Miyagishima, S. Y*., et al.*, Effects of mutations in Arabidopsis FtsZ1 on plastid division, FtsZ ring formation and positioning, and FtsZ filament morphology in vivo. *Plant Cell Physiol* 2007, *48*, 775-791.
[22] Yun, M. S., Kawagoe, Y., Septum formation in amyloplasts produces compound granules in the rice endosperm and is regulated by plastid division proteins. *Plant Cell Physiol* 2010, *51*, 1469-1479.
[23] de Pater, S., Caspers, M., Kottenhagen, M., Meima, H*., et al.*, Manipulation of starch granule size distribution in potato tubers by modulation of plastid division. *Plant Biotechnol J* 2006, *4*, 123-134.
[24] Liu, Q. W., E. Currie, V. Yada, R. , Physicochemical properties of starches during potato growth. *Carbohydrate Polymers* 2003, *51*.
[25] Bogorad, L., Engineering chloroplasts: an alternative site for foreign genes, proteins, reactions and products. *Trends Biotechnol* 2000, *18*, 257-263.
[26] Verma, D., Samson, N. P., Koya, V., Daniell, H., A protocol for expression of foreign genes in chloroplasts. *Nat Protoc* 2008, *3*, 739-758.
[27] Fromm, H., Edelman, M., Aviv, D., Galun, E., The molecular basis for rRNA-dependent spectinomycin resistance in Nicotiana chloroplasts. *Embo j* 1987, *6*, 3233-3237.
[28] Svab, Z., Harper, E. C., Jones, J. D., Maliga, P., Aminoglycoside-3''-adenyltransferase confers resistance to spectinomycin and streptomycin in Nicotiana tabacum. *Plant Mol Biol* 1990, *14*, 197-205.
[29] Goldschmidt-Clermont, M., Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker of site-directed transformation of chlamydomonas. *Nucleic Acids Res* 1991, *19*, 4083-4089.
[30] Svab, Z., Maliga, P., High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. *Proc Natl Acad Sci U S A* 1993, *90*, 913-917.
[31] Chinault, A. C., Blakesley, V. A., Roessler, E., Willis, D. G*., et al.*, Characterization of transferable plasmids from Shigella flexneri 2a that confer resistance to trimethoprim, streptomycin, and sulfonamides. *Plasmid* 1986, *15*, 119-131.

[32] Staub, J. M., Transformation of the Plastid Genome in Tobacco Suspension Cell Cultures. *Methods Mol Biol* 2021, *2317*, 167-175.
[33] Khan, M. S., Maliga, P., Fluorescent antibiotic resistance marker for tracking plastid transformation in higher plants. *Nat Biotechnol* 1999, *17*, 910-915.
[34] Lee, S. M., Kang, K., Chung, H., Yoo, S. H.*, 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.
[35] Bendich, A. J., Circular chloroplast chromosomes: the grand illusion. *Plant Cell* 2004, *16*, 1661-1666.
[36] Daniell, H., Lin, C. S., Yu, M., Chang, W. J., Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol* 2016, *17*, 134.
[37] Heinhorst, S. G., G.C. Galun, E. Kenschaft, L. Weissbach, A., Clone bank and physical and genetic map of potato chloroplast DNA. *Theoretical and applied genetics* 2987, *75*.
[38] Vafaee, Y., Staniek, A., Mancheno-Solano, M., Warzecha, H., A modular cloning toolbox for the generation of chloroplast transformation vectors. *PLoS One* 2014, *9*, e110222.
[39] Occhialini, A., Piatek, A. A., Pfotenhauer, A. C., Frazier, T. P.*, et al.*, MoChlo: A Versatile, Modular Cloning Toolbox for Chloroplast Biotechnology. *Plant Physiol* 2019, *179*, 943-957.
[40] Nielsen, B. L., Cupp, J. D., Brammer, J., Mechanisms for maintenance, replication, and repair of the chloroplast genome in plants. *J Exp Bot* 2010, *61*, 2535-2537.
[41] Palmer, J. D., Comparative organization of chloroplast genomes. *Annu Rev Genet* 1985, *19*, 325-354.
[42] Meeker, R., Nielsen, B., Tewari, K. K., Localization of replication origins in pea chloroplast DNA. *Mol Cell Biol* 1988, *8*, 1216-1223.
[43] Chiu, W. L., Sears, B. B., Electron microscopic localization of replication origins in Oenothera chloroplast DNA. *Mol Gen Genet* 1992, *232*, 33-39.
[44] Lucht, J. M., Public Acceptance of Plant Biotechnology and GM Crops. *Viruses* 2015, *7*, 4254-4281.
[45] Waltz, E., Gene-edited CRISPR mushroom escapes US regulation. *Nature* 2016, *532*, 293.

Chapter I
Generation of enlarged starch granules with increased viscosity in potato through genome-editing of *FtsZ1*

A version of this chapter will be published by Alexander C. Pfotenhauer, Alessandro Occhialini, Stacee A. Harbison, Li Li, Agnieszka A. Piatek, Curtis Luckett, C. Neal Stewart Jr., and Scott C. Lenaghan. A.C.P., A.A.P., and S.C.L. designed the strategy; A.C.P., A.O., S.A.H., L.L., and C.L. collected the data; A.C.P. and C.L. analyzed data; A.C.P. wrote the article.

Abstract

Genome-editing has enabled rapid crop improvement with staple food crops, such as potato, a key beneficiary of the technology. In potato, starch contained within tubers represents the primary product for use in food and non-food industries. Starch granules are produced in the plastids of tubers with plastid size correlated with the size of starch grana. Division of plastids is controlled and maintained by proteins including the tubulin-like GTPase FtsZ1. Altered expression of *FtsZ1* has been shown to alter or inhibit plastid division, leading to the production of “macro-plastid” containing plants. These macro-chloroplast plants are characterized by cells containing fewer and enlarged plastids. In this work, we utilize CRISPR/Cas9 to generate *FtsZ1* edited potato lines to demonstrate that genome-editing can be used to increase the size of starch granules in tubers. Altered plastid morphology was comparable to overexpression of *FtsZ1* in previous work in potato and other crops. As a result, several lines were generated with up to a 1.97-fold increase in starch granule size that was otherwise phenotypically indistinguishable from wild-type plants. Starch paste from one line showed a 2.07-fold increase in final viscosity. The advantages of enlarged starch granules and potential of CRISPR/Cas9 based technologies for food crop improvement are further discussed.

Introduction

Potato (*Solanum tuberosum*) is the most consumed non-grain food and has an important position in maintaining global food security for a rapidly expanding population [1-3]. Starch is the predominant storage carbohydrate in potato tubers and is utilized in both food and non-food applications [4-7]. Potato is one of the top sources of industrial starch, thus modification of the starch content and composition in potato would significantly impact a variety of industries [8-10]. In 2011, the first sequenced potato genome was publicly released, providing key information to enable CRISPR/Cas9 mediated genome-editing for crop improvement [11-12]. CRISPR/Cas9 has demonstrated numerous successes in potato, including increasing amylopectin content, eliminating steroidal glycoalkaloids, and overcoming self-incompatibility [13-17]. Despite the potential of gene editing for agricultural advancements, there is still some public concern regarding the safety of genetically modified crops. To address this concern, non-transgenic gene knockout potato lines have been engineered through transient expression of CRISPR/Cas9 [18-20]. Expression of CRISPR/Cas9 by plasmid, mRNA, or ribonucleoprotein in protoplasts can allow for gene editing without

foreign DNA integration. Potato lines have been successfully established from these edited protoplasts.

The starch produced in the amyloplasts is the most prominent constituent in potatoes [21-22]. Individual amyloplasts in potato are typically composed of one starch granule [23]. Starch is comprised of composed of amylose and amylopectin and stored in semi-crystalline granules of various sizes. Potato granules tend to be larger than those of cereal starches and the granule size is a key parameter for industrial processing. Because of the wide variety of starch applications, there is a desire for granules of differing properties, including increased size [24-28]. Some applications, such as noodle production favor small granules, though larger granules may increase yield during processing [29-30]. Additionally, it has been shown in sweet potato that smaller starch granules degrade faster than large granules, and large granule tubers may be beneficial for storage [31]. Amyloplasts, like all plastids, are organelles that arise from undifferentiated proplastids, and through complex signaling differentiate to perform a specific function [32-34]. In the case of chloroplasts the function is photosynthesis, while in amyloplasts it is starch storage. Previous studies have demonstrated that plastid division occurs through binary fission with the formation of an electron dense ring responsible for constriction [35-37]. This ring is composed of the Fts (filamentous temperature sensitive) proteins, which are homologous to the same proteins responsible for cell division in *Escherichia coli* [38-41]. In plants there are two distinct *FtsZ* gene families; *FtsZ1* and *FtsZ2*, both of which are required for effective plastid division [42]. The expression ratio of the genes is speculated to have an important effect on chloroplast division mechanisms [43]. Both increasing and decreasing expression levels of *FtsZ1* and *FtsZ2* result in altered plastid division, potentially from only one and not both proteins accumulating [44]. Knockdown of either of these in *Arabidopsis* resulted in one or few enlarged chloroplasts, termed “macro-chloroplasts”, as opposed to wild-type cells reported to contain ~100 chloroplasts [36, 44]. Regardless of the dramatic change in organelle size and abundance, the phenotypes of resulting plants were similar to wild-type. Specifically for potato, it has been shown that decreased expression of *FtsZ1* through RNAi reduces plastid number, while overexpression of *FtsZ1* results in few and large plastids [45]. The macro-plastid lines resulted in larger starch granules, and increased tuber phosphate. Additionally, the viscosity of the starch paste after cooling was higher in macro-plastid lines. In rice endosperms, knockdown of *FtsZ1* produced larger pleomorphic amyloplasts that began division and expanded, but did not complete plastid division [46]. Recently, we produced potato lines overexpressing *Arabidopsis FtsZ1*, and proved their capability for chloroplast transformation [47]. These macro-chloroplast lines performed similarly to wild-type as material for chloroplast transformation, but were delayed in growth and yielded less tuber biomass.

Here, CRISPR/Cas9 mediated genome-editing was used for crop improvement of potato, to specifically augment starch granule size by the formation of macro-plastid lines. Plants were generated that contained larger starch granules, but otherwise were similar to wild-type in regards to growth and nutritional profile. One line contained starch with increased viscosity characteristics. The use of CRISPR/Cas9 allows for the potential to generate non-transgenic gene edited potato plants with large starch granule tubers and increased viscosity attributes. Transient expression of this system in protoplasts along with plant regeneration could lead to similar lines to the ones described.

Materials and Methods

FtsZ CDS cloning/sequencing

Solanum tuberosum var. 'Desirée' were grown in Magenta GA7 boxes with MS Reg media [47]. RNA was extracted from ~1 month-old tissue using Molecular Research Center TRI Reagent (Cincinnati, Ohio, USA). RNA was cleaned with a Zymo Research RNA Clean and Concentrator kit (Irvine, California, USA). cDNA was synthesized according to protocol using a Thermo Fisher Scientific SuperScript III First-Strand Synthesis System (Waltham, Massachusetts, USA). The *FtsZ1* coding sequence was amplified using primers 1 and 2 (all primers used in this study can be found in Supplementary Table 1.1) (all tables and figures can be found in the appendix). The *FtsZ1* amplicon and pUC19 were digested with New England Biolabs XbaI and HindIII (Ipswich, Massachusetts, USA). One microliter of New England Biolabs calf intestinal alkaline phosphatase was added to the pUC19 digestion for dephosphorylation, and the amplicon was ligated into the vector. The vector was Sanger sequenced with M13 forward and reverse primers.

Vector construction

pKSE401 was obtained through Addgene (plasmid number 62202) (Watertown, Massachusetts, USA) [48]. The vector was modified to have hygromycin plant selection, as opposed to kanamycin. Briefly, a hygromycin fragment from plasmid pMDC32 was amplified using primers 3 and 4 [49]. The hygromycin amplicon and pKSE401 were digested with New England Biolabs NcoI and SacII. One microliter of New England Biolabs calf intestinal phosphatase was added to the pKSE401 digestion, and the hygromycin fragment was ligated into the vector, now named pKSE401-Hyg. Five guide RNAs (gRNAs) were designed on two different *FtsZ1* exons by considering GC content, targeting unique restriction enzyme recognition sites, and considering off-target effect by using the tool CRISPOR (Supplementary Table 1.1) [50]. gRNAs were designed to target Cas9 to a restriction enzyme recognition site. Upon Cas9 cleavage and subsequent DNA repair, indels will destroy these recognition sites. Undigested PCR fragments are indicative of mutagenesis. gRNAs were designed with BsaI overhangs and ordered as complementary single stranded oligonucleotides from Integrated DNA Technologies (Coralville, Iowa, USA). gRNAs were annealed at

95°C for two minutes in oligo annealing buffer (10 mM Tris HCl, pH 8.0; 50 mM NaCl; 1 mM EDTA) and cloned into pKSE401-Hyg vector using BsaI sites.

Plant transformation

pKSE401-Hyg constructs were transformed into *Agrobacterium tumefaciens* strain LBA4404 by the freeze-thaw method [51]. One-month old *in vitro* potato cultures were used for plant transformation as previously described [52]. Selective media was supplemented with 20 mg/l hygromycin. Fifty-six putative transgenic plantlets were chosen for genomic DNA extraction based on their ability to form roots in media containing hygromycin.

Genomic DNA extraction and molecular analysis

DNA was extracted from the 56 putative transgenic plantlets with the CTAB method [53]. PCR was performed with primers 5 and 6 on 20 ng genomic DNA to check for the presence of Cas9 DNA. All 56 lines were then subjected to the restriction enzyme digest assay. Twenty nanogram of genomic DNA from plantlets generated with gRNAs 1 and 2 were amplified with primers 7 and 8; while plantlets generated with gRNAs 3, 4, and 5 were amplified with primers 9 and 10. Following amplification, the PCR products were cleaned with a Zymo Research DNA Clean & Concentrator kit. PCR products were diluted to 900 ng and digested with appropriate restriction enzymes overnight (Supplementary Table 1.1). The digestions were ran on an agarose gel and checked for undigested bands. All undigested bands were cut out and cleaned with a Qiagen QIAquick Gel Extraction Kit (Hilden, Germany).

Cloning and sequencing of mutated amplicons

pUC19 was digested with 1 µl New England Biolabs HincII and 1 µl calf intestinal phosphatase. The undigested cleaned DNA bands from 16 lines' restriction digest assay were re-amplified using the same phosphorylated primers and ligated into the vector. Following colony PCR, 5 putative positive colonies (A-E) per transgenic line were used for plasmid extraction and Sanger sequenced using M13 forward and reverse primers.

RT PCR and q-RT PCR

cDNA was synthesized as before for all 16 downselected lines. RT-PCR was conducted using primers 11 and 12 to amplify the entire 1260 bp *FtsZ1* coding sequence, and 15 and 16 to amplify a portion of *EF1*α as a control. Primers 13 and 14, and 15 and 16 were used to amplify a portion of *FtsZ1* or *EF1*α, respectively, for qRT-PCR. Primers 13 and 14 were designed on a portion of *FtsZ1* removed following CRISPR/Cas9 mutagenesis in all lines. Therefore, amplification of cDNA will not occur on removed sequence in mutated lines. All qRT-PCR reactions were conducted with Thermo Fisher Scientific PowerUp SYBR Green Master Mix on a Thermo Fisher Scientific QuantStudio 6 Flex. The data is expressed as $2^{-\Delta CT}$ of *FtsZ1* vs. *EF1*α. Results are shown as mean ± standard deviation of 3 technical replicates.

Starch granule size determination
Microtubers were grown on plates containing microtuber induction media (2.89 g/l Murashige-Skoog Salts (without vitamins); 1.04 g/l B5 salts (Gamborg B5 basal); 5 ml “Complete” vitamin stock solution (400 g/l glycine; 100 mg/l nicotinic acid; 100 mg/l pyridoxine HCl; 100 mg/l thiamine HCl); 330 µg/l folic acid; 500 µg/l d-biotin; 10 mg/l kinetin; 100 µg/l indole acetic acid; 100 mg/l inositol; 80 g/l sucrose; 6 g/l agar; pH 5.7)). ~One centimeter internodes with leaves removed were placed on media that was kept in the dark at 17°C. After one month of growth, microtubers were scraped onto microscope slides, and stained with a 1:1 mixture of Lugol’s solution (6.7 g/l potassium iodide; 3.3 g/l iodine) to glycerol. Stained starch granules were characterized using light microscopy. Granule area was measured using ImageJ 1.41o from the National Institute of Health (n in Supplementary Table 1.3) (Bethesda, Maryland, USA).

Confocal microscopy
One-month-old fully expanded leaves were analyzed using an Olympus Fv1000 confocal microscope (Tokyo, Japan). Chlorophyll was excited using a 543 nm helium-neon laser and was detected with an emission wavelength of 667 nm. Images were taken with the same image setting parameters using Olympus FV10-ASW Viewer software Ver.4.2a. The images were processed with software ImageJ 1.41o from the National Institute of Health.

Growth studies
~One month-old *in vitro* potato tissue was propagated from the 16 gene-edited lines and Desirée wild-type into MS Reg media [47]. After roots emerged ~2 weeks later, 5 apical meristems from each line and wild-type were transferred to Griffin Greenhouse Supplies Pro-Mix BK25 potting mix (Knoxville, TN, USA) and grown under fluorescent lights at ambient temperature in 6 X 6 X 9 cm pots. Six weeks later the plants were transferred 11.4 l volume pots and finished growing until senescence in a greenhouse. Tubers were collected and washed with deionized water.

Triplicate replicates of plants MacroGranule1, MacroGranule2, and wild-type were propagated and transferred to potting mix as before. They were grown in growth chambers with 16/8 hours of light/dark cycle and a constant temperature of 22 °C. For all lines, pictures and all analysis other than tuber weight and above ground dry weight were taken simultaneously as flowers began to emerge. Plant height was measured. CO_2 assimilation values per unit of leaf area (µmol m^{-2} s^{-2}) were measured with a LI-COR Biosciences LI-6800 Portable Photosynthesis System (Lincoln, Nebraska, USA) with atmospheric CO_2 conditions (400 µmol mol air^{-1}), constant irradiance (1000 µmol photons m^{-2} s^{-2}), 23 °C, and a vapor pressure deficit of 0.8-1.2 kPa, and a flow rate of 200 µmol s^{-1}. Two readings were taken from each plant replicate. An Opti-Sciences CCM-200plus chlorophyll meter (Hudson, New Hampshire, USA) was used to measure chlorophyll content

index (CCI). Three readings from three different leaves (9 total) were taken from each plant replicate. Tubers were collected, cleaned with deionized water, dried at ambient temperature, and weighed. All tissue above the potting mix line was collected, dried at 50 °C, and weighed. Results are expressed as mean ± standard deviation of measurements taken from the three biological replicates.

Tuber analysis

Tubers from the 16 edited lines and wild-type were sent to Eurofins Food Testing Services (USA) for determination of total fat, phosphorus, phosphate, nitrogen, and protein content. Association of Official Agricultural Chemists methods 990.03, 992.15, 954.02, modified 984.27, modified 927.02, modified 985.01, and modified 965.17 were conducted. Single values are given for individual lines other than for compiled gRNA groups, which are shown as mean ± standard deviation. Tubers were also lyophilized and milled in order to determine total starch and amylose content with Megazyme kits K-TSTA-100A and K-AMYL (Wicklow, Ireland). Results are shown as mean ± standard deviation of 3 technical replicates.

Starch pasting properties

Starch was extracted from MacroGranule1, MacroGranule2, and wild-type tubers as previously described [54]. Pasting properties were analyzed with previously described methods on a Discovery HR-2 Rheometer using a 40mm Peltier steel parallel plate (New Castle, Delaware, USA) [55]. Results are shown as mean ± standard deviation of 3 technical replicates.

Statistical analysis

Means were compared with ANOVA and when significant analyzed by Tukey-Kramer ($p < 0.01$) using CoStat software (United Kingdom).

Results

Generation of potato lines with reduced FtsZ1 expression levels

Fifty-six transgenic lines from five different gRNA constructs were produced from *Agrobacterium tumefaciens* transformation. Six transgenic lines were produced with gRNA1, 12 with gRNA2, 18 with gRNA3, 8 with gRNA4, and 12 with gRNA5. A restriction digest assay was performed to identify mutated lines (Supplementary Figure 1.1). Six plants were mutated from gRNA1 (100%), 11 from gRNA2 (92%), 13 from gRNA3 (72%), 8 from gRNA4 (100%), and 0 from gRNA5 (0%) (38 total). No lines were generated with only undigested bands. Three or more lines from each gRNA were chosen for sequencing to characterize the level of mutagenesis caused by Cas9 cleavage, and all 16 lines' sequencing show mutations (Figure 1.1A, B). Lines produced with gRNA1 showed deletions ranging from -11 to -167 base pairs. Every line produced with gRNA2 showed a -65 base pair deletion. Lines produced with gRNA3 showed deletions and insertions ranging from -/+1 to -394 or +777 base pairs. Interestingly, the +412

and +777 fragments inserted in line 3.16 align with 100% homology to an unannotated mitochondrial DNA sequence (Supplementary Table 1.2). This insert appears to be non-coding sequence. Lines produced with gRNA4 showed deletions ranging from -3 to -266 base pairs, but only a maximum of +9 base pairs insertion. No lines that were carried through had synonymous mutations such as indels of 3 or 6 base pairs, lending confidence that the genome-editing should have an effect on *FtsZ1* function. Coding sequence deletions or insertions will result in impaired transcripts and subsequent peptides. Deletions or insertions in intron spanning sequence will similarly disrupt transcription, and affect FtsZ1 synthesis. PCR amplification of genomic DNA *FtsZ1* fragments for the 16 downselected lines match with sequencing results, and show a detectable level of mutation (Supplementary Figure 1.2).

cDNA was produced from the 16 lines and analyzed by RT-PCR and qRT-PCR. Full-length 1260 bp *FtsZ1* amplification showed many lines with multiple band amplifications or very low amplification (Supplementary Figure 1.3). Lower molecular weight amplicons corroborate sequencing data that indicate large portions of *FtsZ1* were deleted following DNA repair). Primers for qRT-PCR were designed on conserved deleted regions between all 16 lines (Figure 1.1B). qRT-PCR resulted in all lines having a decrease in *FtsZ1* expression compared to wild-type, ranging from a 1.3- to 115-fold reduction (Figure 1.1C). Line 4.1 showed an undetectable transcript using the primers for qRT-PCR. These primers anneal on sequence that was completely deleted after DNA repair mutagenesis. Decreased ability to detect transcript through qRT-PCR in generated lines is due to the sequence complementary to either primer 13 or 14 being completely destroyed.

The reduction of FtsZ1 expression produced potato tubers with increased starch granule size without comprising nutritional quality

Microtubers were generated from the 16 lines. Starch granule size from these microtubers was determined by light microscopy (Figure 1.2A). Because tuber and starch granule size is correlated, only microtubers ranging from 10 to 30 mg were used (Supplementary Table 1.3) [56]. Three wild-type controls between 13.6 and 29.7 g were included, and none were significantly different from one another. Lines 3.3 and 3.9 contain starch granules significantly larger than wild-type or any other transgenic lines (Figure 1.2B). Because of this result, lines 3.3 and 3.9 were renamed MacroGranule1 and MacroGranule2, respectively. Leaf tissue from MacroGranule1 and MacroGranule2 was analyzed by confocal microscopy (Supplementary Figure 1.4). Both lines appear to contain fewer and larger chloroplasts as compared to wild-type. The morphology of the chloroplasts, especially for MacroGranule2, also appears to be less consistent than the typical oval shape shown in wild-type.

All 16 lines were grown in a greenhouse for the production of tubers (Supplementary Figure 1.5). Tubers were analyzed by a variety of methods to

determine total phosphorus, phosphate, fat, nitrogen, protein, starch, and amylose content (Table 1.1). All lines showed the same nutritional content to wild-type and all other lines with the exception of Lines 4.1 and 4.3, which had a significant decrease in starch content. Lines were also compiled by gRNA to compare to wild-type, and no compiled gRNA groups were significantly different from each other. These results suggest that there is not a nutritional penalty for the investigated macro and micronutrients following gene editing of *FtsZ1*.

The phenotypes and growth characteristics of MacroGranule1 and MacroGranule2 were comparable to wild-type

MacroGranule1, MacroGranule2, and wild-type plants were chosen for a second growth study to analyze additional phenotypic parameters (Figure 1.3). Both transgenic lines reached the same height, contained the same level of chlorophyll, and conducted gas exchange the same as wild-type. No differences in tuber yield or above ground dry weight were found either (Supplementary Figure 1.6). In contrast to our previous overexpressing *FtsZ1* potato lines, MacroGranule1 and MacroGranule2 did not display any delay in growth [47]. These results lend confidence that *FtsZ1* mutants can be successfully grown for the production of macro-plastid potatoes.

MacroGranule2 starch displays a higher viscosity level than MacroGranule1 and wild-type

MacroGranule1, MacroGranule2, and wild-type tubers' starch was used to determine pasting properties through the use of a rheometer. All three lines' viscosity began increasing at around 65°C (Figure 1.4). MacroGranule1 and wild-type reached a maximum viscosity during the 90°C hold phase of 0.322 (± 0.04) and 0.356 (± 0.05) Pa.s, respectively. Interestingly, MacroGranule2 reached a maximum viscosity during the hold phase of 0.629 (± 0.05) Pa.s, a 1.77-fold increase over wild-type. MacroGranule1 and wild-type reached a final viscosity of 0.544 (± 0.05) and 0.593 (± 0.06) Pa.s, respectively. MacroGranule2 reached a final viscosity of 1.225 (± 0.07) Pa.s, a 2.07-fold increase over wild-type.

Discussion

We show here that a reduction in *FtsZ1* expression mediated by CRISPR/Cas9 can produce macroplastid lines that contain larger starch granules, particularly lines MacroGranule1 and MacroGranule2 (Figure 1.2 and Supplementary Figure 1.3). Regardless of constitutive Cas9 expression, no lines indicated homozygous mutations as determined by genomic DNA PCR (Supplementary Figure 1.2). This may be due to allelic differences in *FtsZ1*. gRNAs may not target all four potato alleles, allowing for one or several wild-type copies to persist. Though not homozygous, several of the lines potentially arose from the same piece of transformed callus. Lines 3.12, 3.13, and 3.16 all show the same distinct amplification pattern, and are likely genetically indistinguishable (Supplementary Figure 1.2). Several large amplicon insertions detected through sequencing of

line 3.16 could probably be likewise detected in lines 3.12 and 3.13 with additional technical replicates. The large 412 and 777 base pair insertion detected in line 3.16 is particularly interesting. These sequences align with 100% homology to unannotated potato mitochondrial DNA. Mitochondrial DNA has been previously reported to insert into nuclear genomes in eukaryotes including plants, with up to a 620 kilobase insertion in *Arabidopsis* [57-59]. Mitochondrial DNA insertion is thought to occur following double-strand breaks, and to the best of our knowledge this is the first instance mediated through cleavage via CRISPR/Cas9 or any other gene-editor [60].

Line 4.1 produced an undetectable level of transcript for qRT-PCR. This line may contain infrequent mutations not detected by our sequencing that prevent the qRT-PCR primers from binding. Alternatively, the mutations may result in the production of rapidly degraded unstable transcripts. Full-length coding sequence primers 9 and 10 still retained the ability to produce an amplicon, though at a reduced molecular weight (Supplementary Figure 1.3). It is interesting that large deletions occurred regardless that only one gRNA was used for each transgenic line. Perhaps the persistence of Cas9 activity due to its constitutive expression led to eventual large deletions.

Previous work using RNAi demonstrated that disruption of *FtsZ1* can affect plastid morphology and division patterns, however this is the first study to prove that CRISPR/Cas9 mediated genome-editing can result in larger starch granule potato tubers [45-46]. Our preceding attempt to generate macro-plastid lines through overexpression of *Arabidopsis FtsZ1* produced lines that were delayed in growth and yielded less tuber biomass [47]. MacroGranule1 and MacroGranule2 produced the same tuber biomass as wild-type without a delay in growth (Figure 1.3). All other measured growth patterns were comparable to wild-type. Additionally, nutritional profile for macro and micronutrients of tubers from these two lines were also without difference. Interestingly, there was no increase in phosphate content as was found previously [45]. The lack of delayed growth patterns observed in these lines may be due to differences in expression levels of *FtsZ1* as compared to previous overexpressing lines [61]. Overexpression of *FtsZ1* and its affect on chloroplast division has shown to be dose dependent, and perhaps a moderate decrease in expression can result in plants with intermediate sized plastids [44]. MacroGranule1 and MacroGranule2 may fall into this category of macro-plastid plants, retaining adequate growth patterns while still producing larger starch granules. Potato amyloplasts typically contain one starch granule, and the enlarged plastids in lines MacroGranule1 and MacroGranule2 contained larger granules (Figure 1.2). Previous research has shown that suspensions of larger starch granules exhibit increased viscosity characteristics [62-64]. MacroGranule2 contained starch granules significantly larger than wild-type, and this likely led to the increased final viscosity (Figure 1.4). As MacroGranule2's starch paste reached a maximum viscosity twice as high as wild-type, it could potentially be used in smaller amounts. Additionally, blends of

different starch sources are commonly used to fit a specific purpose, and tubers from MacroGranule2 may be useful to fit a distinct need [65-67]. Potato starch paste has a very high level of clarity, and this benefit may increase the usefulness of starch from MacroGranule2 [68].

CRISPR/Cas9 has revolutionized plant biotechnology, and will continue to be applied and amended for new and potentially undiscovered potentials [69]. However, there is still public skepticism regarding genetically engineered crops, and a significant hurdle to bringing transgenic crops to market [70-72]. A potential mitigation to address both public concern and also avoid bureaucratic bottlenecks is to produce genome-edited crops that were created without any foreign DNA integration. CRISPR/Cas9 is most frequently expressed through plasmid DNA. However, plant transformation with plasmids results in random integration of DNA into the genome, which results in significant line-to-line variation [73]. Additionally, these plants are likely now subject to regulatory scrutiny that can prevent them from ever reaching the market. An attractive alternative is to produce plants through transient genome-editing, through the use of preassembled ribonucleoproteins or mRNA [74-76]. Delivery of CRISPR/Cas9 cargo through either of these methods using protoplasts or biolistics ensures that DNA editing can occur without the possibility of foreign DNA integration. These methods could potentially be applied with our gRNA to generate lines similar to MacroGranule1 and MacroGranule2 for public use of a large starch granule potato line.

Here we have described the generation of larger starch granule potato tubers through the use of CRISPR/Cas9 mediated genome-editing. Two lines produced, MacroGranule1 and MacroGranule2, grew without a fitness penalty, and produced tubers with larger starch granules but otherwise similar phenotype and nutritional profile. MacroGranule2 produced starch with increased viscosity characteristics. Our design could be taken forward in to DNA-free genome-editing platform in order to produce larger starch granule potatoes suitable for the modern market.

References

[1] Barrell, P. J., Meiyalaghan, S., Jacobs, J. M., Conner, A. J., Applications of biotechnology and genomics in potato improvement. *Plant Biotechnol J* 2013, *11*, 907-920.
[2] Lutaladio, N. C., L., Potato: The hidden treasure. *Journal of Food Composition and Analysis* 2009, *22*, 491-493.
[3] Devaux, A. K., P. Ortiz, O., Potatoes for Sustainable Global Food Security. *Potato Research* 2014, *57*, 185-199.
[4] Chakraborty, R., Kalita, P., Sen, S., Natural Starch in Biomedical and Food Industry: Perception and Overview. *Curr Drug Discov Technol* 2019, *16*, 355-367.
[5] Ellis, R. P. C., M.P. Dale, M.F.B. Duffus, C.M. Lynn, A. Morrison, I.M. Prentice, R.D.M. Swanston, J.S. Tiller, S.A. , Starch production and industrial use. *Journal of the Science of Food and Agriculture* 1999, *77*, 289-311.
[6] Kraak, A., Industrial applications of potato starch products. *Industrial Crops and Products* 1992, *1*, 107-112.
[7] Burrell, M. M., Starch: the need for improved quality or quantity--an overview. *J Exp Bot* 2003, *54*, 451-456.
[8] Jobling, S., Improving starch for food and industrial applications. *Current Opinion in Plant Biology* 2004, *7*, 210-218.
[9] Tharanathan, R. N., Starch--value addition by modification. *Crit Rev Food Sci Nutr* 2005, *45*, 371-384.
[10] Davis, J. P. S., N. Khandelwal, R.L. Chibbar, R.N., Synthesis of novel starches *in planta*: opportunities and challenges. *Starch* 2003, *55*, 107-120.
[11] Xu, X., Pan, S., Cheng, S., Zhang, B.*, et al.*, Genome sequence and analysis of the tuber crop potato. *Nature* 2011, *475*, 189-195.
[12] Cong, L., Ran, F. A., Cox, D., Lin, S.*, et al.*, Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013, *339*, 819-823.
[13] Butler, N. M., Atkins, P. A., Voytas, D. F., Douches, D. S., Generation and Inheritance of Targeted Mutations in Potato (Solanum tuberosum L.) Using the CRISPR/Cas System. *PLoS One* 2015, *10*, e0144591.
[14] Wang, S., Zhang, S., Wang, W., Xiong, X.*, et al.*, Efficient targeted mutagenesis in potato by the CRISPR/Cas9 system. *Plant Cell Rep* 2015, *34*, 1473-1476.
[15] Andersson, M., Turesson, H., Nicolia, A., Fält, A. S.*, et al.*, Efficient targeted multiallelic mutagenesis in tetraploid potato (Solanum tuberosum) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Rep* 2017, *36*, 117-128.
[16] Nakayasu, M., Akiyama, R., Lee, H. J., Osakabe, K.*, et al.*, Generation of α-solanine-free hairy roots of potato by CRISPR/Cas9 mediated genome editing of the St16DOX gene. *Plant Physiol Biochem* 2018, *131*, 70-77.
[17] Enciso-Rodriguez, F., Manrique-Carpintero, N. C., Nadakuduti, S. S., Buell, C. R.*, et al.*, Overcoming Self-Incompatibility in Diploid Potato Using CRISPR-Cas9. *Front Plant Sci* 2019, *10*, 376.

[18] Veillet, F., Perrot, L., Chauvin, L., Kermarrec, M. P.*, et al.*, Transgene-Free Genome Editing in Tomato and Potato Plants Using Agrobacterium-Mediated Delivery of a CRISPR/Cas9 Cytidine Base Editor. *Int J Mol Sci* 2019, *20*.
[19] Johansen, I. E., Liu, Y., Jørgensen, B., Bennett, E. P.*, et al.*, High efficacy full allelic CRISPR/Cas9 gene editing in tetraploid potato. *Sci Rep* 2019, *9*, 17715.
[20] Tuncel, A., Corbin, K. R., Ahn-Jarvis, J., Harris, S.*, et al.*, Cas9-mediated mutagenesis of potato starch-branching enzymes generates a range of tuber starch phenotypes. *Plant Biotechnol J* 2019, *17*, 2259-2271.
[21] Naeem, M. T., I.J. Emes, M.J., Starch synthesis in amyloplasts purified from developing potato tubers. *The Plant Journal* 2002, *11*, 1095-1103.
[22] Smith, A. M., Denyer, K., Martin, C., THE SYNTHESIS OF THE STARCH GRANULE. *Annu Rev Plant Physiol Plant Mol Biol* 1997, *48*, 67-87.
[23] Kram, A. M., Oostergetel, G. T., Van Bruggen, E., Localization of Branching Enzyme in Potato Tuber Cells with the Use of Immunoelectron Microscopy. *Plant Physiol* 1993, *101*, 237-243.
[24] Ji, Q., Oomen, R. J., Vincken, J. P., Bolam, D. N.*, et al.*, Reduction of starch granule size by expression of an engineered tandem starch-binding domain in potato plants. *Plant Biotechnol J* 2004, *2*, 251-260.
[25] Lindeboom, N. C., P.R. Tyler, R.T., Analytical, biochemical and physicochemical aspects of starch granule size, with emphasis on small granule starches: a review. *Starch* 2004, *56*, 89-99.
[26] Park, S.-H. W., J.D. Seabourn, B.W., Starch granule size distribution of hard red winter and hard red spring wheat: its effects on mixing and breadmaking quality. *Journal of Cereal Science* 2009, *49*, 98-105.
[27] Smith, A. M. M., C. , Starch biosynthesis and the potential for its manipulation, in: Grierson, D. (Ed.), *Biosynthesis and Manipulation of Plant Products*, Springer, Dordrecht 1993, pp. 1-54.
[28] Stoddard, F. L., Genetics of starch granule size distribution in tetraploid and hexaploid wheats. *Australian Journal of Agricultural Research* 2003, *54*, 637-648.
[29] Chen, Z. S., H.A. Voragen, A.G.J., Starch granule size strongly determines starch noodle processing and noodle quality. *Journal of Food Science* 2006, *68*, 1584-1589.
[30] Gutierrez, O. A. C., M.R. Glover, D.V., Starch particle volume in single- and double-mutant maize endosperm genotypes involving the soft starch (h) gene. *Crop Science* 2002, *42*, 355-359.
[31] Niu, S. L., X-Q. Tang, Ruimin. Zhang, G. Li, X. Cui, B. Mikitzel, L. Haroon, M., Starch granule sizes and degradation in sweet potatoes during storage. *Postharvest Biology and Technology* 2019, *150*, 137-147.
[32] Liebers, M., Grübler, B., Chevalier, F., Lerbs-Mache, S.*, et al.*, Regulatory Shifts in Plastid Transcription Play a Key Role in Morphological Conversions of Plastids during Plant Development. *Front Plant Sci* 2017, *8*, 23.
[33] Pyke, K. A., Plastid division and development. *Plant Cell* 1999, *11*, 549-556.
[34] Osteryoung, K. W., Pyke, K. A., Division and dynamic morphology of plastids. *Annu Rev Plant Biol* 2014, *65*, 443-472.

[35] Vitha, S., McAndrew, R. S., Osteryoung, K. W., FtsZ ring formation at the chloroplast division site in plants. *J Cell Biol* 2001, *153*, 111-120.
[36] Osteryoung, K. W., Stokes, K. D., Rutherford, S. M., Percival, A. L., Lee, W. Y., Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial ftsZ. *Plant Cell* 1998, *10*, 1991-2004.
[37] Strepp, R., Scholz, S., Kruse, S., Speth, V., Reski, R., Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proc Natl Acad Sci U S A* 1998, *95*, 4368-4373.
[38] Donachie, W. D., The cell cycle of Escherichia coli. *Annu Rev Microbiol* 1993, *47*, 199-230.
[39] Margolin, W., FtsZ and the division of prokaryotic cells and organelles. *Nat Rev Mol Cell Biol* 2005, *6*, 862-871.
[40] TerBush, A. D., Yoshida, Y., Osteryoung, K. W., FtsZ in chloroplast division: structure, function and evolution. *Curr Opin Cell Biol* 2013, *25*, 461-470.
[41] Yoshida, Y., Mogi, Y., TerBush, A. D., Osteryoung, K. W., Chloroplast FtsZ assembles into a contractible ring via tubulin-like heteropolymerization. *Nat Plants* 2016, *2*, 16095.
[42] Schmitz, A. J., Glynn, J. M., Olson, B. J., Stokes, K. D., Osteryoung, K. W., Arabidopsis FtsZ2-1 and FtsZ2-2 are functionally redundant, but FtsZ-based plastid division is not essential for chloroplast partitioning or plant growth and development. *Mol Plant* 2009, *2*, 1211-1222.
[43] McAndrew, R. S., Froehlich, J. E., Vitha, S., Stokes, K. D., Osteryoung, K. W., Colocalization of plastid division proteins in the chloroplast stromal compartment establishes a new functional relationship between FtsZ1 and FtsZ2 in higher plants. *Plant Physiol* 2001, *127*, 1656-1666.
[44] Stokes, K. D., McAndrew, R. S., Figueroa, R., Vitha, S., Osteryoung, K. W., Chloroplast division and morphology are differentially affected by overexpression of FtsZ1 and FtsZ2 genes in Arabidopsis. *Plant Physiol* 2000, *124*, 1668-1677.
[45] de Pater, S., Caspers, M., Kottenhagen, M., Meima, H.*, et al.*, Manipulation of starch granule size distribution in potato tubers by modulation of plastid division. *Plant Biotechnol J* 2006, *4*, 123-134.
[46] Yun, M. S., Kawagoe, Y., Septum formation in amyloplasts produces compound granules in the rice endosperm and is regulated by plastid division proteins. *Plant Cell Physiol* 2010, *51*, 1469-1479.
[47] Occhialini, A., Pfotenhauer, A. C., Frazier, T. P., Li, L.*, et al.*, Generation, analysis, and transformation of macro-chloroplast Potato (Solanum tuberosum) lines for chloroplast biotechnology. *Sci Rep* 2020, *10*, 21144.
[48] Xing, H. L., Dong, L., Wang, Z. P., Zhang, H. Y.*, et al.*, A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol* 2014, *14*, 327.
[49] Curtis, M. D., Grossniklaus, U., A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* 2003, *133*, 462-469.

[50] Haeussler, M., Schonig, K., Eckert, H., Eschstruth, A.*, et al.*, Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* 2016, *17*, 148.
[51] Weigel, D., Glazebrook, J., Transformation of agrobacterium using the freeze-thaw method. *CSH Protoc* 2006, *2006*.
[52] Chronis, D., Chen, S., Lu, S., Hewezi, T.*, et al.*, A ubiquitin carboxyl extension protein secreted from a plant-parasitic nematode Globodera rostochiensis is cleaved in planta to promote plant parasitism. *Plant J* 2013, *74*, 185-196.
[53] Occhialini, A., Piatek, A. A., Pfotenhauer, A. C., Frazier, T. P.*, et al.*, MoChlo: A Versatile, Modular Cloning Toolbox for Chloroplast Biotechnology. *Plant Physiol* 2019, *179*, 943-957.
[54] Liu, Q. W., E. Currie, V. Yada, R., Physicochemical properties of starches during potato growth. *Carbohydrate Polymers* 2003, *51*, 213-221.
[55] Mendez-Montealvo, G. W., Y-J. Campbell, M., Thermal and rheological properties of granular waxy maize mutant starches after β-amylase modification. *Carbohydrate Polymers* 2011, *83*, 1106-1111.
[56] Christensen, D. H. M., M.H., Changes in potato starch quality during growth. *Potato Research* 1996, *39*, 43-50.
[57] Stupar, R. M., Lilly, J. W., Town, C. D., Cheng, Z.*, et al.*, Complex mtDNA constitutes an approximate 620-kb insertion on Arabidopsis thaliana chromosome 2: implication of potential sequencing errors caused by large-unit repeats. *Proc Natl Acad Sci U S A* 2001, *98*, 5099-5103.
[58] Lough, A. N., Roark, L. M., Kato, A., Ream, T. S.*, et al.*, Mitochondrial DNA transfer to the nucleus generates extensive insertion site variation in maize. *Genetics* 2008, *178*, 47-55.
[59] Michalovova, M., Vyskot, B., Kejnovsky, E., Analysis of plastid and mitochondrial DNA insertions in the nucleus (NUPTs and NUMTs) of six plant species: size, relative age and chromosomal localization. *Heredity (Edinb)* 2013, *111*, 314-320.
[60] Puertas, M. J., González-Sánchez, M., Insertions of mitochondrial DNA into the nucleus-effects and role in cell evolution. *Genome* 2020, *63*, 365-374.
[61] McAndrew, R. S., Olson, B. J., Kadirjan-Kalbach, D. K., Chi-Ham, C. L.*, et al.*, In vivo quantitative relationship between plastid division proteins FtsZ1 and FtsZ2 and identification of ARC6 and ARC3 in a native FtsZ complex. *Biochem J* 2008, *412*, 367-378.
[62] Okechukwu, P. E. R., M.A., Influence of granule size on viscosity of cornstarch suspension. *Journal of Texture Studies* 1995, *26*, 501-516.
[63] Rao, M. A. T., J., Granule size and rheological behavior of heated tapioca starch dispersions. *Carbohydrate Polymers* 1999, *38*, 123-132.
[64] Kumar, R. K., B.S., Thermal, pasting and morphological properties of starch granules of wheat (Triticum aestivum L.) varieties. *Journal of Food Science and Technology* 2017, *54*, 2403-2410.
[65] Waterschoot, J. G., S.V. Fierens, E. Delcour, J.A., Starch blends and their physicochemical properties. *Starch* 2014, *67*, 1-13.

[66] Zhu, F. H., Y. Li, G., Physicochemical properties of potato, sweet potato and quinoa starch blends. *Food Hydrocolloids* 2020, *100*.
[67] Waterschoot, J. G., S.V. Willebrords, J.K. Fierens, E. Delcour, J.A., Pasting properties of blends of potato, rice and maize starches. *Food Hydrocolloids* 2014, *41*, 298-308.
[68] Chung, K. M. M., T.W. Kim, H. Chun, J.K., Physicochemical properties of sonicated mung bean, potato, and rice starches. *Cereal Chemistry* 2002, *79*, 631-633.
[69] Jaganathan, D., Ramasamy, K., Sellamuthu, G., Jayabalan, S., Venkataraman, G., CRISPR for Crop Improvement: An Update Review. *Front Plant Sci* 2018, *9*, 985.
[70] Scott, S. E., Inbar, Y., Wirz, C. D., Brossard, D., Rozin, P., An Overview of Attitudes Toward Genetically Engineered Food. *Annu Rev Nutr* 2018, *38*, 459-479.
[71] Lusk, J. L. M., B.R. Wilson, N., Do consumers care how a genetically engineered food was created or who created it? *Food Policy* 2018, *78*, 81-90.
[72] Menz, J., Modrzejewski, D., Hartung, F., Wilhelm, R., Sprink, T., Genome Edited Crops Touch the Market: A View on the Global Development and Regulatory Environment. *Front Plant Sci* 2020, *11*, 586027.
[73] Gase, K., Weinhold, A., Bozorov, T., Schuck, S., Baldwin, I. T., Efficient screening of transgenic plant lines for ecological research. *Mol Ecol Resour* 2011, *11*, 890-902.
[74] Kim, S., Kim, D., Cho, S. W., Kim, J., Kim, J. S., Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* 2014, *24*, 1012-1019.
[75] Zhang, Y., Liang, Z., Zong, Y., Wang, Y*., et al.*, Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat Commun* 2016, *7*, 12617.
[76] Si, X., Zhang, H., Wang, Y., Chen, K., Gao, C., Manipulating gene translation in plants by CRISPR-Cas9-mediated genome editing of upstream open reading frames. *Nat Protoc* 2020, *15*, 338-363.

Appendix

Table 1.1. Nutritional analysis of *FtsZ1* edited tubers

Line	Phosphorus %	Phosphate %	Fat %	Nitrogen %	Protein %	Starch %	Amylose %
1.2	0.11	0.34	0.36	0.61	3.81	59.05 (± 0.89)	29.86 (± 0.50)
1.3	0.11	0.34	0.15	0.61	3.81	55.44 (± 1.22)	30.28 (± 0.32)
1.5	0.10	0.31	0.32	0.50	3.13	59.82 (± 1.74)	29.99 (± 1.18)
2.6	0.11	0.33	0.10	0.50	3.13	57.79 (± 2.38)	29.10 (± 0.47)
2.7	0.12	0.35	0.16	0.58	3.63	57.63 (± 1.96)	25.58 (± 0.265)
2.9	0.12	0.36	0.18	0.54	3.38	58.95 (± 1.12)	27.64 (± 0.58)
MG1	0.12	0.36	0.27	0.55	3.44	55.80 (± 1.07)	27.10 (± 0.50)
3.8	0.12	0.38	0.15	0.64	4.00	58.75 (± 0.41)	28.89 (± 0.53)
MG2	0.13	0.41	0.10	0.61	3.81	57.83 (± 0.63)	26.16 (± 0.41)
3.12	0.11	0.34	0.15	0.55	3.44	56.89 (± 0.83)	27.04 (± 0.42)
3.13	0.11	0.35	0.10	0.56	3.50	56.51 (± 0.37)	30.04 (± 0.56)
3.16	0.11	0.33	0.26	0.54	3.38	58.76 (± 0.45)	29.12 (± 0.45)
3.18	0.11	0.33	0.30	0.53	3.31	57.69 (± 0.73)	29.09 (± 0.40)
4.1	0.11	0.33	0.26	0.58	3.63	** 50.39 (± 0.84)	27.63 (± 0.38)
4.2	0.12	0.38	0.27	0.71	4.44	-	-
4.3	0.13	0.39	0.29	0.73	4.56	** 48.74 (± 0.37)	28.33 (± 0.58)
WT	0.12	0.37	0.19	0.67	4.19	54.14 (± 0.54)	28.70 (± 0.56)
Compiled gRNA1	0.11 (± 0.01)	0.33 (± 0.02)	0.28 (± 0.11)	0.57 (± 0.06)	3.58 (± 0.39)	58.10 (± 2.34)	30.04 (± 0.21)
Compiled gRNA2	0.11 (± 0.01)	0.35 (± 0.01)	0.15 (± 0.04)	0.54 (± 0.04)	3.38 (± 0.25)	58.12 (± 0.72)	27.44 (± 1.77)
Compiled gRNA3	0.12 (± 0.01)	0.36 (± 0.03)	0.19 (± 0.08)	0.57 (± 0.04)	3.55 (± 0.25)	57.46 (± 1.12)	28.21 (± 1.43)
Compiled gRNA4	0.12 (± 0.01)	0.37 (± 0.03)	0.27 (± 0.02)	0.67 (± 0.08)	4.21 (± 0.51)	49.56 (± 1.17)	27.98 (± 0.49)

Tubers from greenhouse-grown plants were analyzed for determination of total phosphorus, phosphate, fat, nitrogen, protein, starch, and amylose. Single values are shown for individual lines for phosphorus, phosphate, fat, nitrogen, and protein. Results for starch and amylose for individual lines, and all parameters for compiled gRNAs are expressed as mean ± standard deviation. Means were compared by ANOVA and when significant analyzed by Tukey-Kramer (* = $p < 0.05$, ** = $p < 0.01$ different from all other lines).MG1 = MacroGranule1. MG2 = MacroGranule2. WT = wild-type. gRNA = guide RNA.

Supplementary Table 1.1. Primers and gRNAs

Primers used 5' >> 3'	Sequence	Purpose	gRNAs used 5' >> 3'	Sequence	GC content	Off-target effect	Corresponding restriction enzyme
1	AATATTAAGCTTATGGCGATTTTAGGGCTCTCA	Clone FtsZ cDNA	gRNA1-Fw	attgAAACACGGATGCTCAAGCAC	50	0-0-0-1-12	BtsIMutI
2	GTATGCTCTAGACTAAAAGAACAGCCTTCGAGT	Clone FtsZ cDNA	gRNA1-Rv	aaacGTGCTTGAGCATCCGTGTTT		0-0-0-0-0	
3	CAGTGGCCATGGATGAAAAAGCCTGAACTCACC	Clone hyg cassette	gRNA2-Fw	attgATTGGAGAACTTCTGACTCG	45	0-0-1-0-17	BssSaI
4	CGATTTTGAAACCGCGGTGA	Clone hyg cassette	gRNA2-Rv	aaacCGAGTCAGAAGTTCTCCAAT		0-0-0-0-0	
5	CCCTGAAGTCAAAGCTCGTC	Check plants for Cas9	gRNA3-Fw	attgGCTGCCTGTTCCCCTAAAAG	55	0-0-0-5-10	EcoNI
6	GGCGATCATCTTCCTCACAT	Check plants for Cas9	gRNA3-Rv	aaacCTTTTAGGGGAACAGGCAGC		0-0-0-0-0	
7	GGATTTTCTATCAGTTGGCT	gRNA 2, 3 RED	gRNA4-Fw	attgTTTAGGGGAACAGGCAGCGG	60	0-0-0-0-7	MspA1I
8	AAGAGGATTGCCACCAGTAC	gRNA 2, 3 RED	gRNA4-Rv	aaacCCGCTGCCTGTTCCCCTAAA		0-0-0-0-0	
9	CTTCTGACTCGTGGGCTTGG	gRNA 4, 5, 6 RED	gRNA5-Fw	attgGCAGGAATGGGTGGAGGTAC	60	0-0-2-5-36	RsaI
10	CCTGCACAGATCTTTTACGTCC	gRNA 4, 5, 6 RED	gRNA5-Rv	aaacGTACCTCCACCCATTCCTGC		0-0-0-0-1	
11	ATGGCGATTTTAGGGCTCTC	FtsZ CDS RT-PCR					
12	CTAAAAGAACAGCCTTCGAGTAGG	FtsZ CDS RT-PCR					
13	ACAGTCTGCTGCCGAGA	FtsZ qRT-PCR					
14	CTCCGCTGCCTGTTCCC	FtsZ qRT-PCR					
15	GATGGTCAGACACGTGAACA	EF1α qRT-PCR					
16	CCTTGGAGTACTTGGGGGTG	EF1α qRT-PCR					

Primers and gRNAs used in this study are shown. gRNAs off-target effect top numbers refer to the number of off-targets per number of mismatches. For examples 0-12-24-36-48 means 0 off-targets with 0 mismatches, 12 off-targets with 1 mismatch, 24 off-targets with 2 mismatches, etc. Bottom numbers are the off-targets with mismatches in the 12 bp flanking the PAM sequence, which are the most likely off-targets. Restriction enzymes targeted by the gRNAs for use in the restriction enzyme digest assay are shown. gRNA = guide RNA. PAM = protospacer adjacent motif.

Supplementary Table 1.2. Mitochondrial DNA insertion in line 3.16

412 basepair insertion	777 basepair insertion
CTCTACCGACAGATGCGCGAATTGCACTG	AGAAGAAAGTATTCTGAAGCCAAATGCTCTTGGATGGGTAGTTGAATAGGGGGATAAAT
GTCATTGTGCGGAATGCCCTCTAAGCCAGT	CCTTTTTTCTCCTTTATATTGAATGTAGTCAATCATTTTGAGGTCATGAGAACAAAGAGTC
CAGTCAATGGAGGGACCGGGGCGCTAGGT	ATAGAGATGTGAGAAGTTTTGCTTTCGCTATCGACTGACCGTCCTGTTAGCAGTGCTATA
ATTTTCTTTCTTCTGGCCGTAGGTGTGATC	CTAGATTGGCACCCTAACTGGTACCTATACAGCTACTCCTTGTCAGTTGCTATATTTTTTC
AATGCCATTGAGCTTCGGTACTCTAAAATA	ACCTTGAGGCTGAGAGGTTAACTCCTTCCACTCTCAACCGCCTTTGCTTATCTTACTAG
GTAGTCGGAATTCGCTTCTGAGTGAGCTTC	TAGCTCTATCTATTGGTACCAGCCAGGGAACTCCCTCTTTGCGAGCTACTTACTTTATCT
CTTCCTTATGCTCTGGCCTGACCTTCACGA	GAACCCTCTACCGACAGATGCGCGAATTGCACTGGTCATTGTGCGGAATGCCCTCTAA
ACAGCTTTCTTTCAGCTACCAACTTCAAAG	GCCAGTCAGTCAATGGAGGGACCGGGGCGCTAGGTATTTTCTTTCTTCTGGCCGTAGG
CTTGATCTCATCTCTCTTCTTTTGAAAGTTT	TGTGATCAATGCCATTGAGCTTCGGTACTCTAAAATAGTAGTCGGAATTCGCTTCTGAGT
AGTACTATTATTGTTGGCATGGAAGGAGTT	GAGCTTCCTTCCTTATGCTCTGGCCTGACCTTCACGAACAGCTTTCTTTCAGCTACCAA
CATCTTGAATTGCCAGTACAGTGATTGGAG	CTTCAAAGCTTGATCTCATCTCTCTTCTTTTGAAAGTTTAGTACTATTATTGTTGGCATGG
GCCTTCCTCACCTGTTCTAGTGACATAGGC	AAGGAGTTCATCTTGAATTGCCAGTACAGTGATTGGAGGCCTTCCTCACCTGTTCTAGT
CCATCCGGTCGAGAGAAGAGACTTGAGAT	GACATAGGCCCATCCGGTCGAGAGAAGAGACTTGAGATTAGCCACTAGGTGAAGTACC
TAGCCACTAGGTGAAGTACCAAGG	AAGG

The 412 and 777 base pair insertion in line 3.16 aligns with 100% homology to *S. tuberosum* var. ‘Desirée’. BLAST sequence ID: MN104801.1.

Supplementary Table 1.3. Microtuber weight and mean area

Line	Tuber weight (mg)	Granule mean area (μm^2)	Standard deviation	n
1.2	10.3	157.38	135.53	1740
1.3	19.7	162.68	150.10	1428
1.5	10.6	151.70	134.24	1943
2.6	10.4	123.83	91.77	4480
2.7	17.5	116.37	76.24	3147
2.9	15.3	107.47	71.55	8150
3.3	23.6	197.56	167.06	1069
3.8	10	175.44	179.68	461
3.9	24.6	284.90	171.97	765
3.12	12.1	157.56	128.40	2879
3.13	21	159.43	123.31	2172
3.16	28.4	148.80	131.19	733
3.18	11.5	159.51	141.12	707
4.1	25.5	165.38	136.89	1175
4.2	10.5	110.27	97.58	1529
4.3	11.2	166.59	127.23	827
WT-A	13.6	134.33	100.28	4062
WT-B	18.2	137.92	108.53	1805
WT-C	29.7	143.95	120.28	861

Lines of similar weight were chosen in order to avoid the possibility that larger tubers would contain larger starch granules. Weight, mean area, standard deviation, and number of granules measured by ImageJ 1.41o from the National Institute of Health are shown. WT = wild-type.

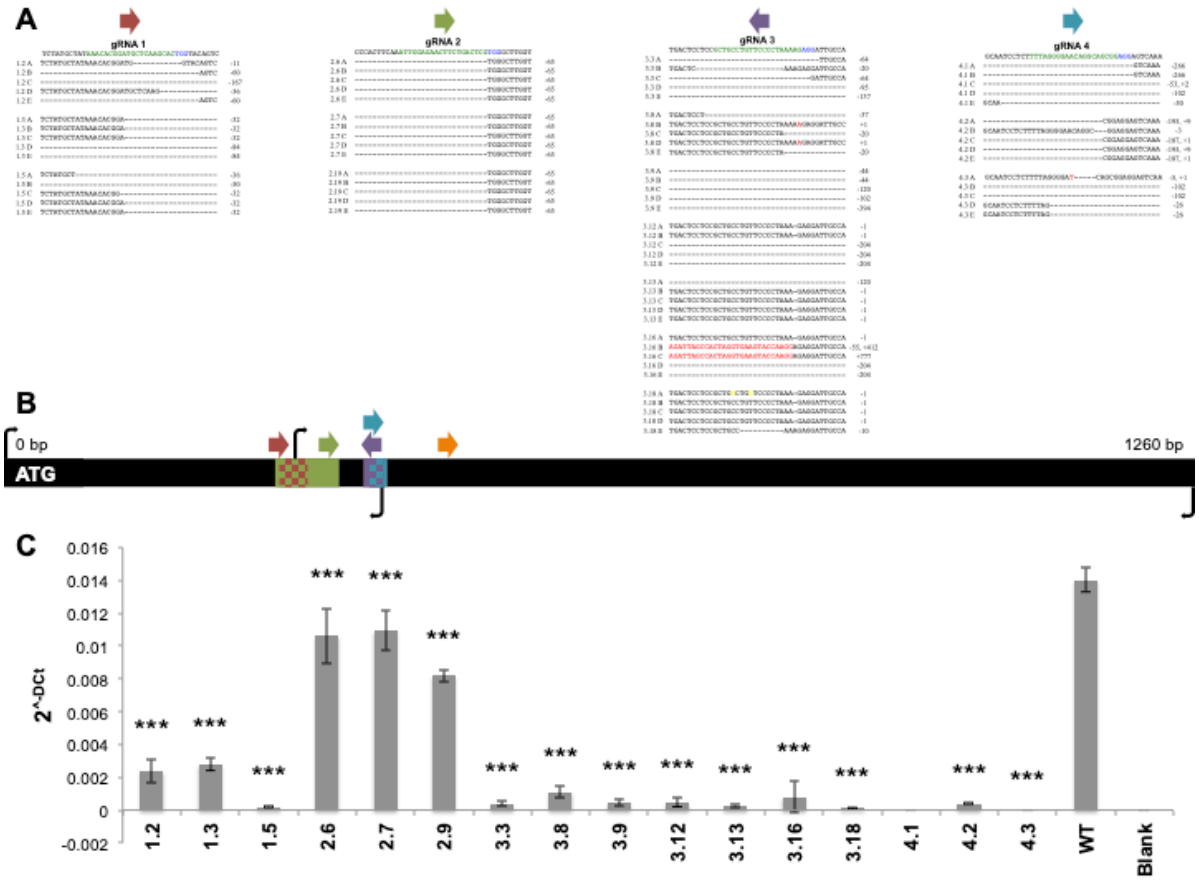


Figure 1.1. CRISPR/Cas9 mediated genome-editing leads to reduced expression of potato *FtsZ1*. Several lines from each gRNA were selected for sequencing (A). The sequence under each gRNA corresponds to WT sequence. The green and blue highlighted sequence corresponds to the gRNA and PAM, respectively. Characterization of each mutation is annotated on the right of each sequence. Schematic of gRNAs, deleted DNA regions, and primers on *FtsZ1* coding sequence (B). The entire black bar is representative of the 1260 bp coding sequence of *FtsZ1*. Colored arrows represent gRNAs. Colored rectangles on the *FtsZ1* black bar represent deleted regions of DNA, checkerboard rectangles represent areas deleted in multiple lines. Black arrows represent primers used for RT and qRT-PCR. qRT-PCR was conducted on *FtsZ1* vs. *EF1* (C). Results are expressed as $2^{\wedge\text{-DCt}}$ mean ± standard deviation of the three experiments. Means were compared with ANOVA and when significant analyzed by Tukey-Kramer (*** = $p < 0.001$ different from WT). WT = wild-type. gRNA = gRNA. Bp = base pair. PAM = protospacer adjacent motif.

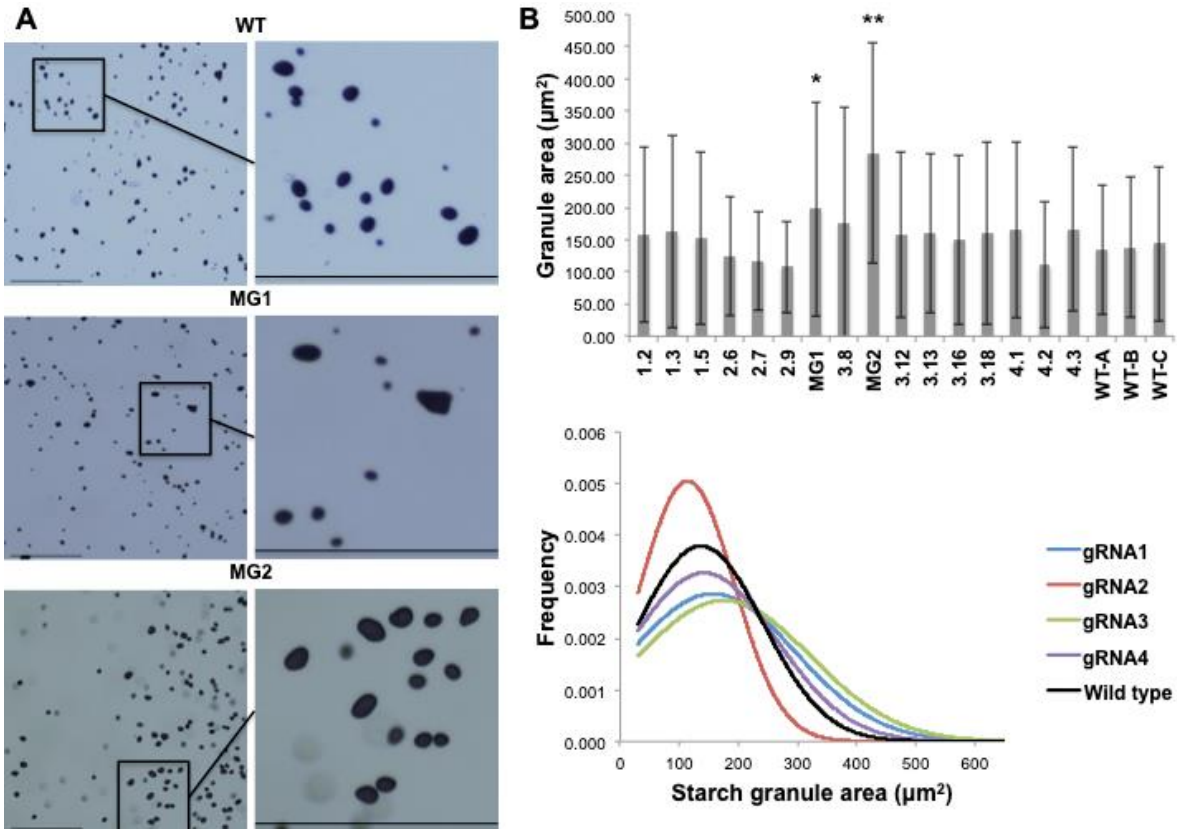


Figure 1.2. Reduced *FtsZ1* expression leads to the production of larger starch granules. Microtuber starch granules were analyzed by light microscopy (A). Tubers were scraped onto slides and stained. Black scale bar = 275 µM. Granule area (μm^2) mean ± standard deviation of the granule area (B). Means were compared with ANOVA and when significant analyzed by Tukey-Kramer (* = $p < 0.05$, ** = $p < 0.01$ different from all other lines). Distribution of granule area for compiled lines from each gRNA (C). WT = wild-type. MG1 = MacroGranule1. MG2 = MacroGranule2.

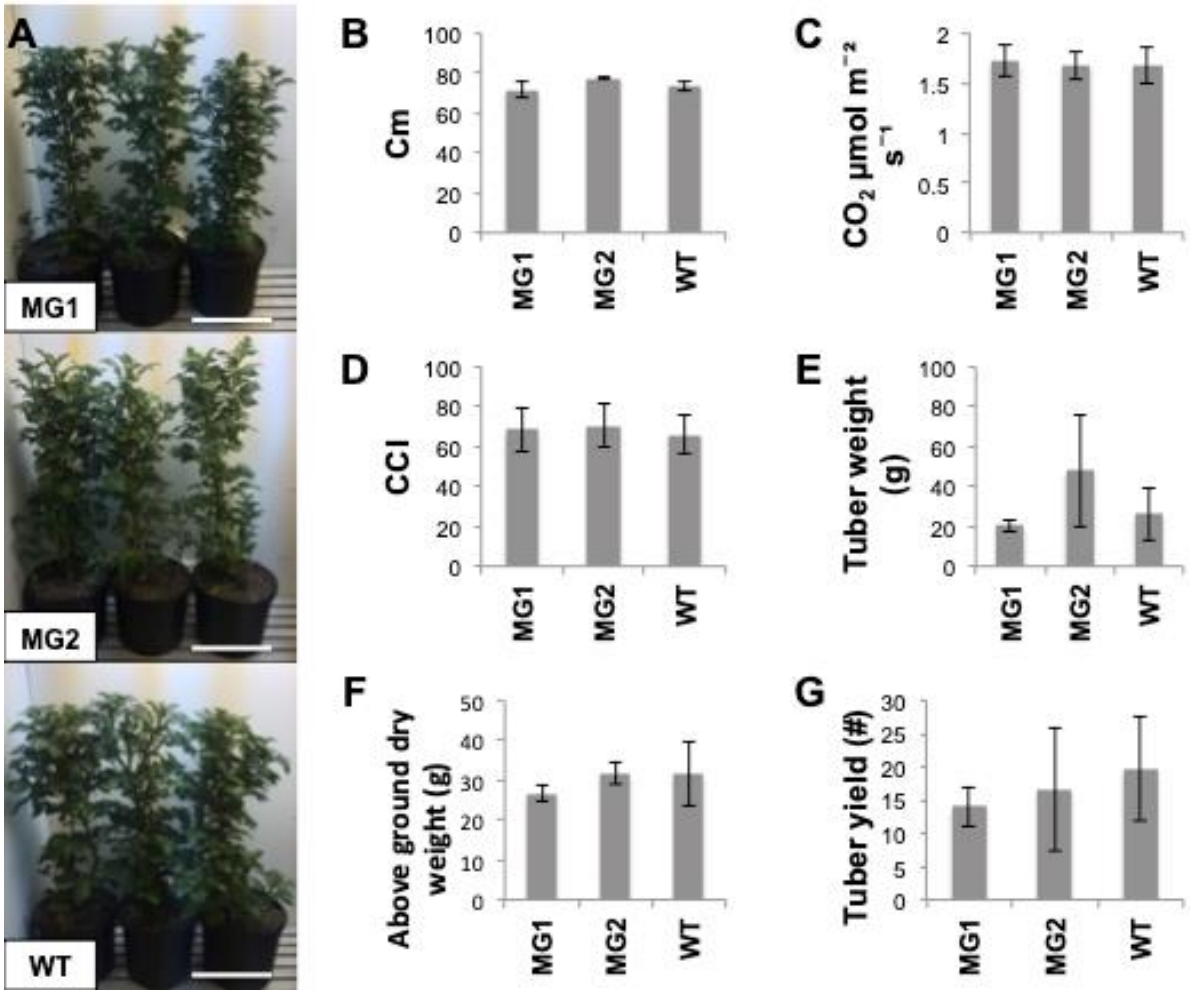


Figure 1.3. Larger starch granule potato lines can grow similarly to wild-type. Pictures of each line growing in an 11.4 l volume bucket shortly before bolting (A). White scale bar = 30 cm. Ability to accumulate mass (B) and (C). Plant height (D). Leaf CO_2 assimilation (E). Chlorophyll content (F). Total tuber weight (G). Above ground tissue dry weight (H). Results are expressed as mean ± standard deviation. Means were compared with ANOVA and when significant analyzed by Tukey-Kramer (* = $p < 0.05$, ** = $p < 0.01$ different from WT). WT = wild-type. MG1 = MacroGranule1. MG2 = MacroGranule2.

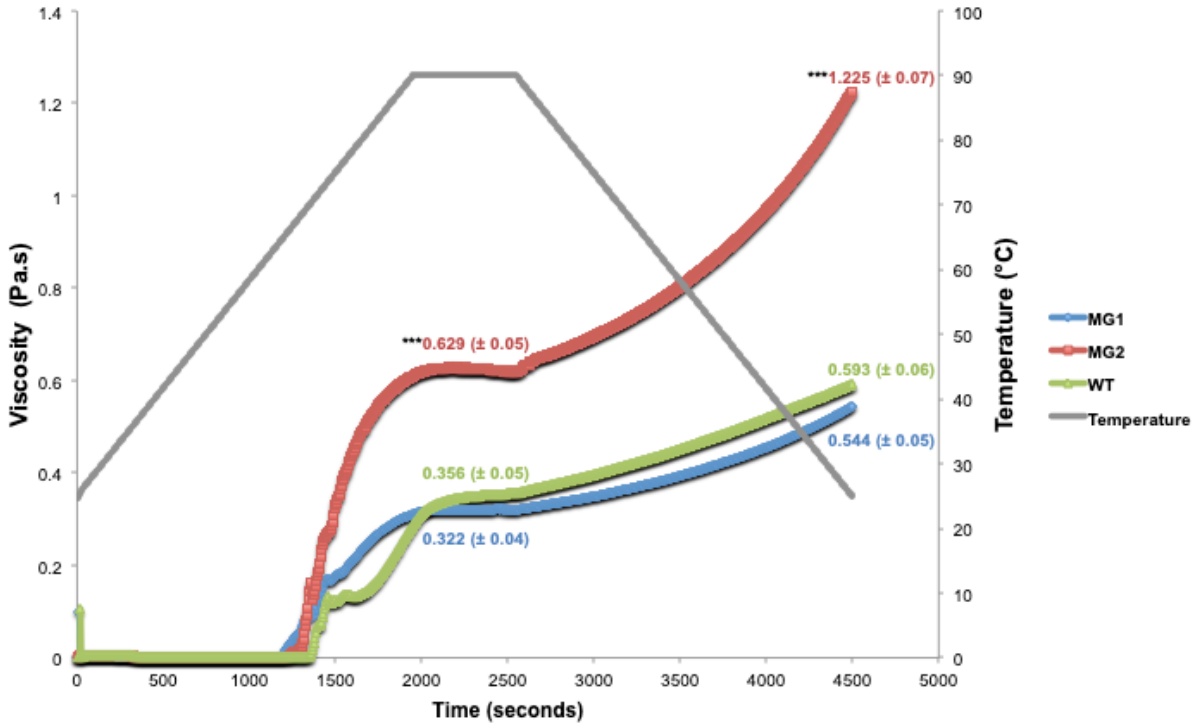
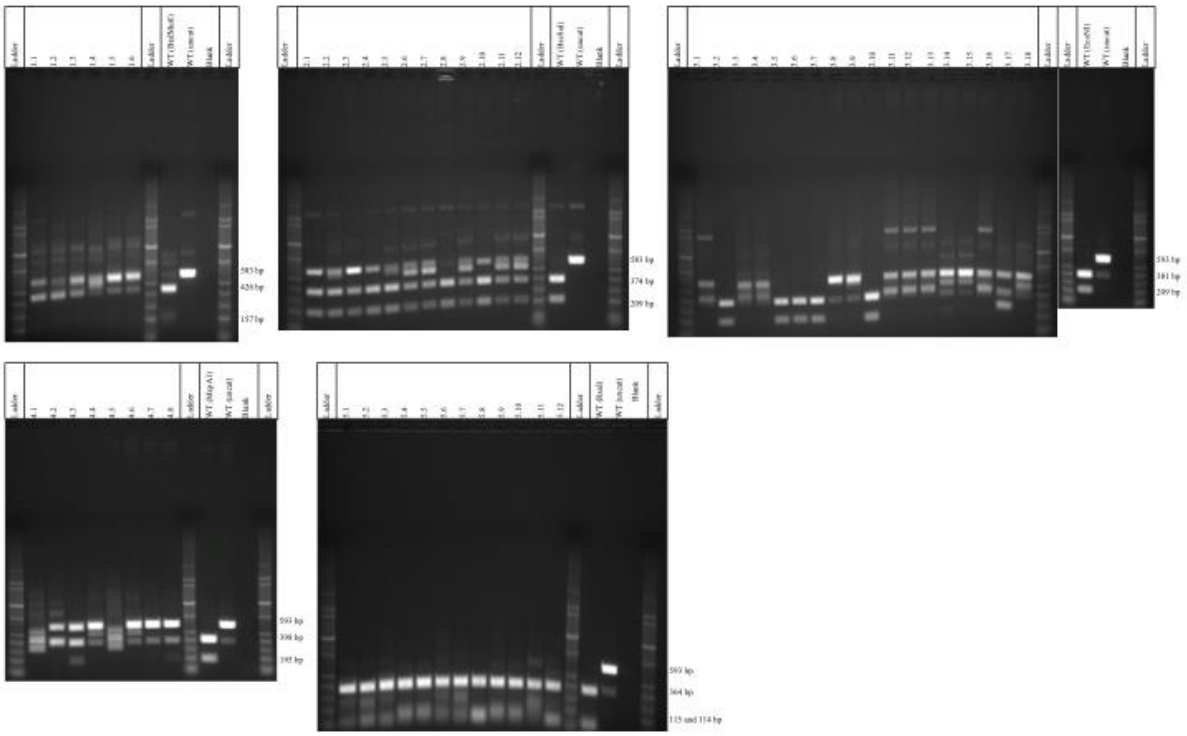
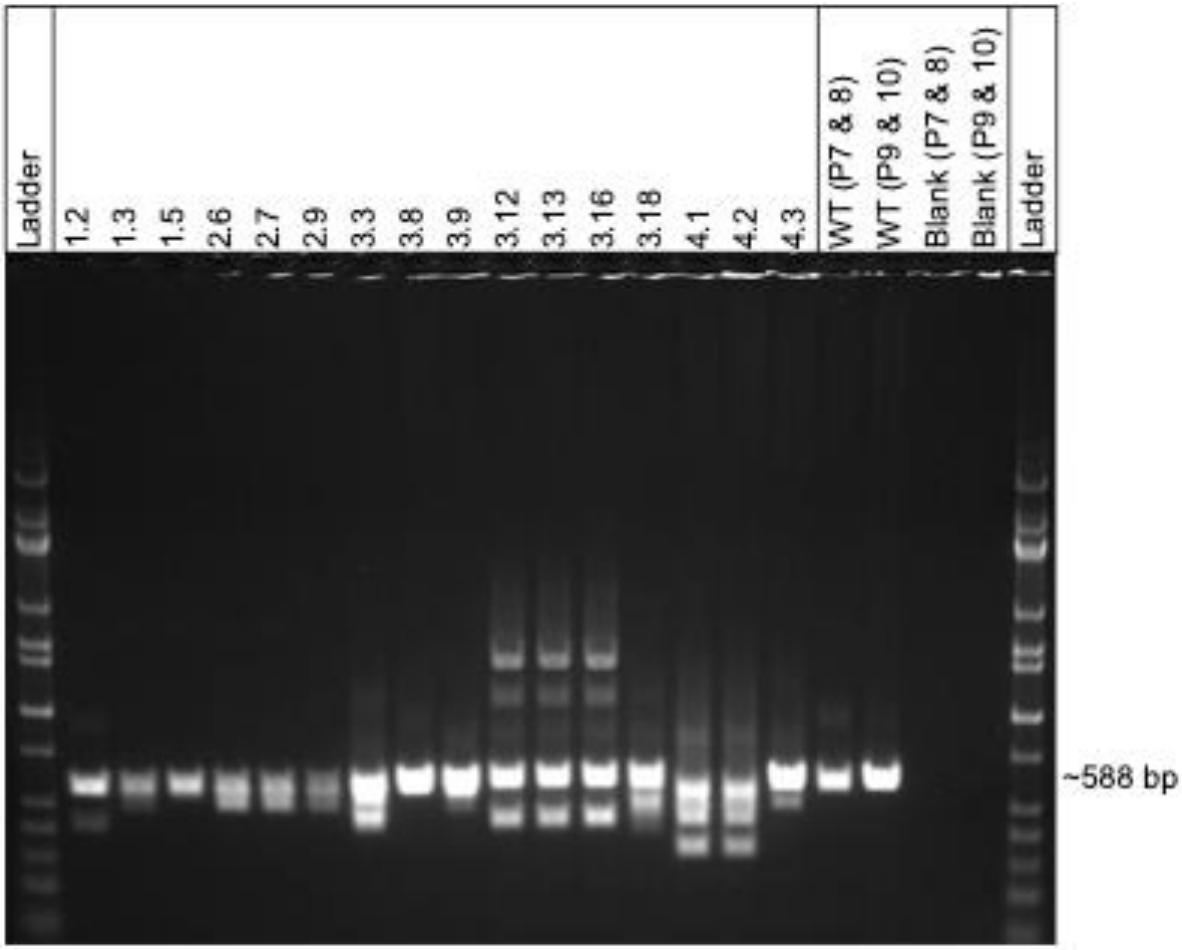


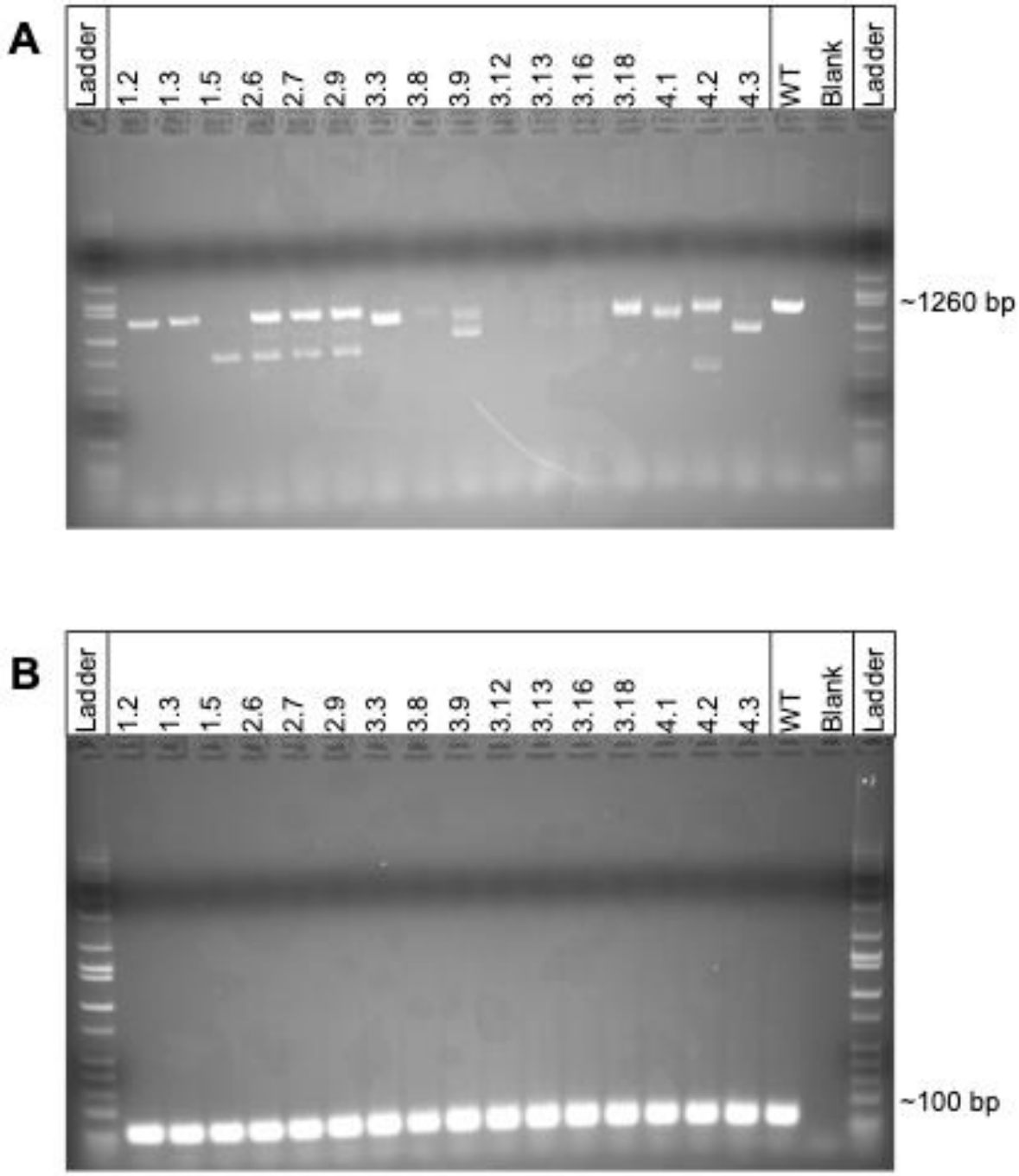
Figure 1.4. MacroGranule2 starch granules show increased viscosity compared to MacroGranule1 and wild-type. The left y-axis indicates viscosity measured in pascal-seconds, the right y-axis indicates temperature in °C, and the x-axis indicates time in seconds. Values indicate maximum viscosity during the 90°C holding phase or final viscosity, and colors are maintained for the three lines. Results are expressed as mean ± standard deviation. Means were compared with ANOVA and when significant analyzed by Tukey-Kramer (*** = $p < 0.001$ different from WT). WT = wild-type. MG1 = MacroGranule1. MG2 = MacroGranule2.



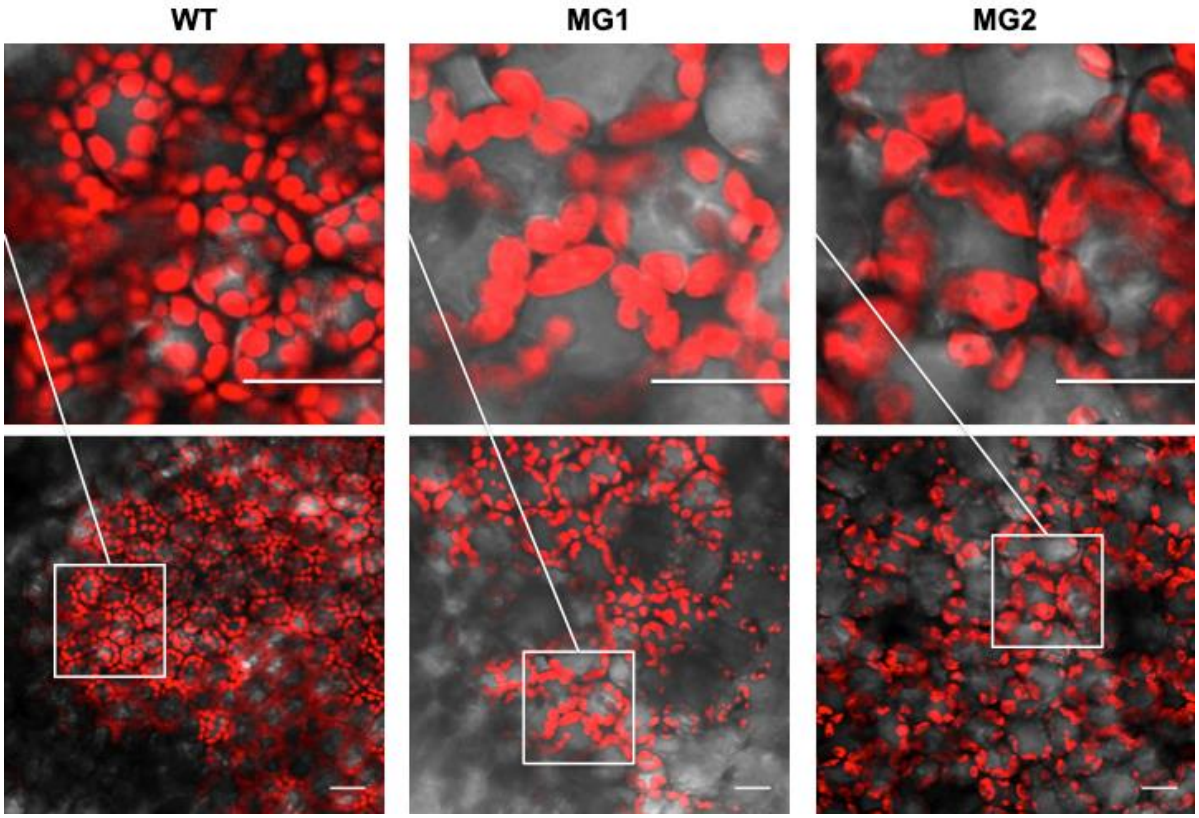
Supplementary Figure 1.1. Restriction digest assay. Genomic DNA from 56 putative transgenic plants was extracted and PCR amplified. Fragments were then digested with appropriate restriction enzymes overnight and ran on an agarose gel. A higher molecular weight band is indicative of Cas9 activity, subsequent DNA repair, and loss of a restriction enzyme recognition site. WT = wild-type.



Supplementary Figure 1.2. PCR amplification of *FtsZ1* fragments. PCR was conducted on the 16 downselected lines with primers 7 & 8 or 9 & 10. WT = wild-type. P = primer.

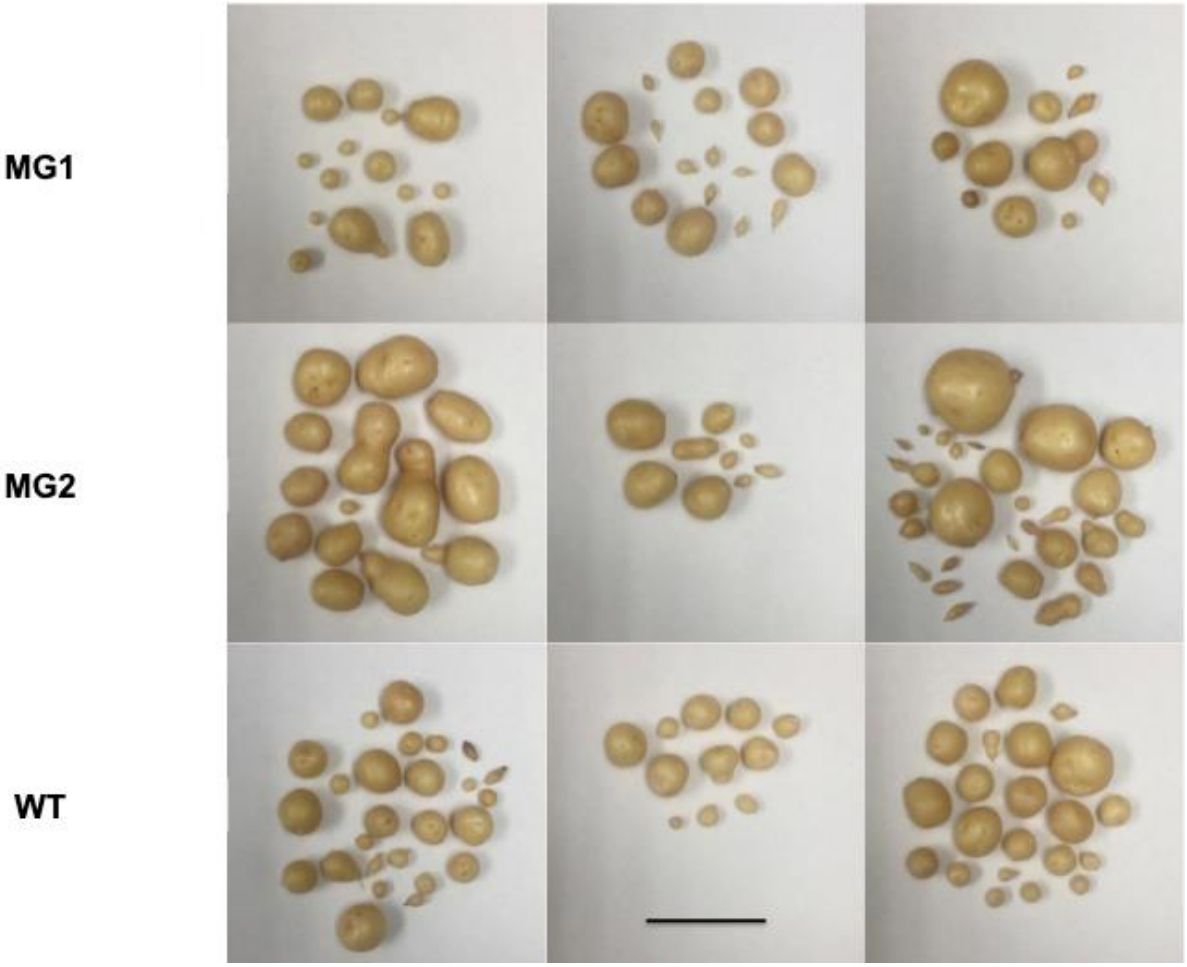


Supplementary Figure 1.3. *FtsZ1* expression patterns. RT-PCR was conducted on full length 1260 bp coding sequence of *FtsZ1* (A). RT-PCR was conducted on 100 bp fragment of *EF1α* as a control (B). WT = wild-type.



Supplementary Figure 1.4. Chloroplast morphology by confocal microscopy of large starch granule lines. Leaf tissue from lines MacroGranule1, MacroGranule2, and wild-type were screened for chloroplast abnormalities. Scale bar = 30 μm. WT = wild-type. MG1 = MacroGranule1. MG2 = MacroGranule2.

Supplementary Figure 1.5. Greenhouse production of *FtsZ1* mutant potato tubers. Plants were grown in tubers until senescence for the production of tubers.



Supplementary Figure 1.6. Tubers produced from MacroGranule1, MacroGranule2, and WT. Plants were grown in a greenhouse until senescence to produce tubers. Tubers from three plant replicates were produced and are shown in the different panels. Scale bar = 5 cm. WT = wild-type. MG1 = MacroGranule1. MG2 = MacroGranule2.

Chapter II
Generation, analysis, and transformation of macro-chloroplast potato (*Solanum tuberosum*) lines for chloroplast biotechnology

A version of this chapter was published in 2020 by Alessandro Occhialini, Alexander C. Pfotenhauer, Taylor P. Frazier, Stacee A. Harbison, Andrew J. Lail, Zachary Mebane, Agnieszka A. Piatek, Stephen B. Rigoulot, Henry Daniell, C. Neal Stewart, and Scott C. Lenaghan. (2020) *Scientific Reports*. 10, 10.1038/s41598-020-78237-x. S.C.L. and A.O. designed the strategy; A.O., A.C.P., T.P.F., L.L., S.A.H., A.J.L., Z.M., A.A.P., and S.B.R. collected the data; A.O. analyzed data; H.D. designed the selection cassette and advised on chloroplast transformation; A.O., C.N.S., and S.C.L. wrote the article.

Abstract

Chloroplast biotechnology is a route for novel crop metabolic engineering. The potential bio-confinement of transgenes, the high protein expression and the possibility to organize multi-gene pathways into operons represent considerable advantages that make chloroplasts valuable targets in agricultural biotechnology and synthetic biology. In the last 3 decades, chloroplast genomes (plastomes) from a few economically important crops have been successfully transformed. The main bottlenecks that prevent efficient transformation in a greater number of crops include the dearth of proven selectable marker gene-selection combinations and tissue culture methods for efficient regeneration of transplastomic plants. The prospects of increasing organelle size are attractive from several perspectives, including an increase in size of potential targets per cell and the decreased of number of targets. As a proof-of-concept, we generated *Solanum tuberosum* (potato) macro-chloroplast lines overexpressing the tubulin-like GTPase protein gene *FtsZ1* from *Arabidopsis thaliana*. Despite their slower growth relative to wild-type, these lines accumulated higher biomass per leaf area at maturity. Macro-chloroplasts were successfully transformed by biolistic DNA-delivery and efficiently regenerated into homoplasmic transplastomic lines. We also demonstrated that macro-chloroplasts accumulate the same amount of heterologous protein than wild-type organelles, confirming efficient usage in plastid engineering. Advantages and limitations of using enlarge compartments in chloroplast biotechnology are discussed.

Introduction

Chloroplast biotechnology has long been recognized as a relatively untapped resource for enabling metabolic engineering and crop improvement (Bock 2013; Bock 2015; Jin and Daniell 2015; Maliga and Bock 2011). While metabolic engineering in plastids has been used for bioproduction of carotenoids (Wurbs, et al. 2007), terpenoids (Pasoreck, et al. 2016), vaccines (Daniell, et al. 2019a; Daniell, et al. 2019b), pharmaceutical drugs (Boyhan and Daniell 2011; Daniell, et al. 2020; Oey, et al. 2009; Park, et al. 2020) and valuable enzymes of industrial interest (Daniell, et al. 2019c; Kumari, et al. 2019), no commercial transplastomic crops have been introduced in the market. Even fewer studies have focused on plastid engineering for crop improvement, such as increasing

photosynthetic efficiency (Lin, et al. 2014b; Occhialini, et al. 2016) or mitigating biotic and abiotic stresses (Chen, et al. 2014; Jin, et al. 2011; Rey, et al. 2013). Given the inherent bioconfinement of transgenes in plastids (Arlen, et al. 2007; Daniell 2007), the lack of transgene position effects (Bock 2015; Schindel, et al. 2018), absence of gene silencing mechanisms, extremely high heterologous protein production (Valkov, et al. 2009; Verma and Daniell 2007; Verma, et al. 2008), and ability to coordinate gene expression in operons (Bock 2013; Lu, et al. 2013), plastid engineering appears to be an attractive strategy. Nonetheless, there are some technological barriers that are impediments in need of addressing in order to widely implement chloroplast transformation for crop agriculture.

Together with the model species tobacco and more recently, Arabidopsis (Ruf et al. 2019; Yu, et al. 2019), chloroplasts of few agronomically important dicotyledon crops including tomato, potato and lettuce (Fuentes, et al. 2016; Lu, et al. 2013; Su, et al. 2015; Yu, et al. 2017) have been successfully transformed and regenerated to homoplasmic transplastomic lines. All monocotyledon cereals have proven recalcitrant to current methods to generate stable transplastomic plants. Only few examples of transient chloroplast transformation in rice (*Oryza sativa*) have been published (Lee, et al. 2006; Wang, et al. 2018). Successful chloroplast transformation in key crops, including cereals, will depend on the ability to find more efficient selection (along with selectable marker genes) and tissue culture methods to efficiently regenerate transplastomic plants.

The ability to significantly increase plastid size provides an intriguing avenue with regards to plastid biotechnology. While the small size of plastids has been shown to not be a limiting factor for biolistic transformation, assuming a strong selection system (Langbecker, et al. 2004), for alternative methods of plastid transformation (e.g. microinjection), an increase in plastid size would be beneficial if not necessary. For example, the installation of large DNA molecules (synthetic plastid genomes) will likely require alternative transformation strategies. The most common strategy to increase chloroplast size is through modification of FtsZ1, a tubulin-like GTPase that regulates chloroplast division (McAndrew, et al. 2001; Mori, et al. 2001; Osteryoung, et al. 1998; Vitha, et al. 2001). Overexpression of *FtsZ1* in Arabidopsis was shown to alter chloroplast division resulting in a reduced number of very large chloroplasts, termed macro-chloroplasts (Stokes, et al. 2000). Surprisingly, despite this extreme chloroplast phenotype, macro-chloroplast lines were indistinguishable from wild-type lines with regards to plant growth and morphology (Osteryoung, et al. 1998). However, a follow-up study in *Nicotiana tabacum* (tobacco) a significant growth reduction was observed under high- and low-light irradiation in lines with a severe macro-chloroplast phenotype (defined as 1-3 large macro-chloroplasts) per cell (Jeong, et al. 2002). In medium-light irradiation, the growth of all macro-chloroplast lines were indistinguishable from controls (Jeong, et al. 2002). In potato, plastid size engineering experiments produced effects in both chloroplasts and amyloplasts (de Pater, et al. 2006). In that research, macro-chloroplast plants had

morphological change in starch granule size as well as increased phosphorous content in tubers compared to wild-type lines, indicating biochemical change (de Pater, et al. 2006). The ability to alter the size and content of non-green plastids is particularly relevant to plastid biotechnology for food crop improvement, where non-green plastids represent a primary source of nutrition (e.g. chromoplasts of tomato).

While methods to generate macro-chloroplasts have been established for decades, there are few examples of chloroplast engineering using these altered plastids. The first successful transformation of macro-chloroplast tobacco demonstrated a 40% increase in chloroplast transformation efficiency relative to wild-type controls (Jeong et al. 2006). In addition, researchers demonstrated that homoplasmic macro-chloroplast plants could be regenerated in tissue culture where they accumulated a similar amount of foreign protein compared to traditional chloroplast engineering (Chikkala, et al. 2012). Based on this foundational work, we sought to develop macro-chloroplast potato lines specifically for the purposes of evaluating their potential for chloroplast biotechnology. We hypothesized that the use of macro-chloroplast containing leaves would increase the transformation efficiency of potato, providing a superior method for potato chloroplast engineering. Secondarily, we hypothesized that the use of macro-chloroplasts may increase the capacity for protein overproduction without negatively affecting chloroplast function. This would be particularly useful for production of plant bioreactors for high production of foreign proteins.

Specifically, in this work we engineered *AtFtsZ1* overexpressing potato lines and selected for macro-chloroplast phenotypes for use in chloroplast biotechnology. The phenotypes of macro-chloroplast lines were characterized, as well as, the efficiency of chloroplast transformation, selection, and regeneration in select lines.

Materials and Methods

Plant growth conditions and growth experiments

Solanum tuberosum (potato) var. ‘Desirée’ wild-type and *AtFtsZ1* overexpressing lines were grown in Magenta boxes containing MS Reg media (4.33 g/l MS basal salt mixture; 25 g/l sucrose; 100 mg/l myo-inositol; 170 mg/l sodium phosphate monobasic monohydrate; 440 mg/l calcium chloride dihydrate; 0.9 mg/l thiamine-HCl; 2 mg/l glycine; 0.5 mg/l nicotinic acid; 0.5 mg/l pyridoxine-HCl; 1x MS vitamins; 3 g/l phytagel; pH 5.8). Transplastomic lines of both genotypes (wild-type and macro) were grown in selective MS rooting media (Valkov, et al. 2014).

For determination of growth characteristics, apical shoots (same length and age) from both transgenic and wild-type plants were *in vitro* propagated. After 2 weeks, small plantlets with roots were transferred on soil and kept growing for 9

weeks before harvesting. Plant height (cm), fresh and dry weight (g) of leaves and entire plants were measured. For determination of dry weight, plant tissue was dried for 1 week at 50°C. The results were expressed as mean ± standard deviation (sd) of four biological replicates for each genotype (*AtFtsZ1* independent lines (macro 1 and 2) along with wild-type control). ANOVAs with post-hoc Tukey statistical analysis (IBM SPSS software) ($p<0.05$) were performed to determine if differences were statistically significant different among wild-type and *AtFtsZ1* lines. Plants grown *in vitro* and in pots were subjected to congruent environmental conditions, 16/8 hours of light and dark respectively, and the temperature was kept constant at 24°C.

Construction of transformation vectors

The vector pAP202 containing the full-length cDNA of *AtFtsZ1* was used to generated *AtFtsZ1* overexpressing lines as published previously (Stokes, et al. 2000). For the construction of pIR and pSSC, the tobacco (*Nicotiana tabacum*) plastome IR (*trnI/trnA*; 102623-105457 bp and 105458-110067 bp; GenBank: KU199713.1) and SSC (*ndhG/ndhI*; 119184-120988 bp and 120989-126029 bp; GenBank: KU199713.1) regions, respectively, were synthesized by GeneArt (Thermo Fisher Scientific, Waltham, MA, USA). The two homologous regions were modified including several SNPs to facilitate cloning. Before introducing the selection cassette between homologous arms, the *trnI/trnA* and *ndhG/ndhI* regions were cloned into pMK vector (Thermo Fisher Scientific, Waltham, MA, USA). A chloroplast selection cassette (*Prrn-SD::aadA::5'UTR::GFP::psbA3'UTR*) was PCR amplified from the pLD-PTD-GFP plasmid using 1Fw/1Rv primers and cloned into the *Pme*I site of *trnI/trnA* and *ndhG/ndhI* sites generating pIR and pSSC plasmids, respectively. A list of primers used in this work is shown in Supplementary Table 2.1.

Generation of transgenic lines

Potato *AtFtsZ1* overexpressing lines (macro-chloroplast lines) were generated by *Agrobacterium*-mediated transformation. The vector pAP202 (Stokes, et al. 2000) was introduced into *Agrobacterium tumefaciens* LBA4404 by the freeze thaw method (Weigel and Glazebrook 2006). For each transformation, 20 internodes of ~1 cm-length from 1-month-old *in vitro* potato plants were used. Internodes were then transformed via *Agrobacterium* and transgenic plants were regenerated in selective media, as previously described (Chronis, et al. 2013). Putative transgenic lines were screened by fluorescent microscopy for clearly-defined enlarged chloroplasts. Thereafter, selected lines were genetically characterized for construct integration and transgene expression.

The PDS-1000/He biolistics device (Bio-Rad, Hercules, CA, USA) was subsequently used to transform chloroplasts of both macro-chloroplast and wild-type potato plants (Occhialini, et al. 2019). Macro-chloroplast line 1 was used for chloroplast transformation. Per each transformation ~6 cm^2 of leaf tissue from 1-month-old *in vitro* potato plants were used. Per each shoot 0.3 mg of gold-

particles (0.6 µm in diameter) binding 1 µg of plasmid were used following the manufacturer protocol (Seashell Technology, La Jolla, CA, USA). Transplastomic plants of both genotypes (wild-type and macro-chloroplast) were then regenerated from transformed leaf material incubated in selective media as described previously (Valkov, et al. 2014). The second round of transplastomic plants were obtained by applying the same protocol of tissue culture/selection/regeneration (Valkov, et al. 2014). All lines were genetically characterized for the presence of the selection cassette integrated and GFP expression.

Total DNA extraction and PCR analysis

Total genomic-DNA preparations from leaves were obtained using ~50 mg of tissue and the CTAB-based procedure of extraction (Occhialini, et al. 2019). PCR (25 cycles) was performed in 25 µl reaction-volume by using 5 ng of total genomic DNA and the DreamTaq Green PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA). For molecular characterization of *AtFtsZ1* overexpressing lines, the pairs of primers 2Fw/2Rv and 3Fw/3Rv were used to check for *nptII/AtFtsZ1* cassette integration and the internal control *actin* (GeneID-102593904), respectively. For genotyping of transplastomic lines, the two pairs of primers 4Fw/4Rv and 5Fw/5Rv were used to verify cassette integration in the *trnI/trnA* and *ndhG/ndhI* sites of plastome, respectively. The two pair of primers 6Fw/6Rv and 7Fw/7Rv were used to amplify full-length *GFP* (NCBI ID: AEX93343.1) and *aadA* (NCBI ID: AAR14532.1) genes, respectively. The primers 8Fw/8Rv were used to amplify an internal portion of *rbcL* (NCBI ID: 4099985).

Total RNA extraction and RT-PCR

Total RNA extraction was performed using Tri-Reagent (Molecular Research Center, Inc, Cincinnati, OH, USA) according to manufacturer's protocol. Per each extraction ~50 mg healthy leaf tissue was used. Total RNA preparations were cleaned and subjected to DNase treatment using the RNA Clean & Concentrator Kit (Zymogen, Irvine, CA, USA) according to manufacturer's instruction. The cDNA synthesis was performed using the Super Script III Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's instruction. cDNA was quantified using a NanoDrop (Thermo Fisher Scientific, Waltham, MA, USA), and several serial cDNA dilutions were tested by PCR (as described before) for the presence of the transgene *AtFtsZ*, whereas the two internal controls *rbcL* (plastome) and *ef1α* (nuclear) were used for normalization. The pair of primers 9Fw/9Rv, 10Fw/10Rv and 11Fw/11Rv were used to amplify and internal fragment (~100 bp) of *AtFtsZ* (Tair ID: AT5G55280), *rbcL* (NCBI ID: 4099985) and *ef1α* (NCBI ID: NM_001288491.1), respectively.

Southern blot analysis

DNA probes for detection of pIR and pSSC constructs integrated into potato plastome (GenBank: NC_008096.2) were designed on IR (104457-104978 bp)

and SSC (120269-120790 bp) regions, respectively. The PCR DIG Probe Synthesis Kit (Roche, Indianapolis, IN, USA) was used to synthesize digoxigenin(DIG)-sUTP-labelled IR and SSC DNA-probes using the pair of primers 12Fw/12Rv and 13Fw/13Rv, respectively. Total genomic DNA from leaf tissue was extracted using CTAB, as described above. After quantification, 1 µg of DNA for each pIR and pSSC sample was digested using *Kas*I/*Hind*III and *Fsp*I/*Sca*I restriction enzymes, respectively. The DNA fragments were separated on 0.9% agarose gel, and after that, the gel was depurinated, denatured and transferred on a nylon membrane as described previously (Occhialini, et al. 2016). The membrane was then incubated with the DIG-labelled probe and detected using the anti- digoxigenin-AP Fab fragments detection kit (Roche Indianapolis, IN, USA) accordingly to the manufacturer's protocol.

GFP quantification

The amount of GFP accumulated in leaf tissue of transplastomic lines was quantified by using the Fluorometric GFP Quantification Kit (Cell Biolabs, Inc., San Diego, CA, USA) according to the manufacturer's protocol (Wamboldt, et al. 2009). The results were expressed as mean ± standard deviation (ng GFP/mg fresh weight) of two indipendent experiments, resulting in a total of 2 biological replicates and 2 technical replicates per biological replicate for each independent transplastomic line. We analyzed 15 pIR and 8 pSSC lines for each genotype, normal and macro-chloroplast.

Confocal microscopy and chloroplast size determination

Leaf tissue from 3-week-old *in vitro* plants was imaged using an Olympus Fv1200 confocal microscope (Olympus, Center Valley, PA, USA). GFP was excited at 488 nm and detected at an emission wavelength of 509 nm. Chlorophyll autofluorescence was visualized using an excitation wavelength of 543 nm and an emission wavelength of 667 nm. Confocal images were acquired using the manufacturer's Olympus FV10-ASW Viewer software Ver.4.2a (Olympus, Center Valley, PA). Confocal images were processed using the online software ImageJ 1.41o (National Institute of Health, Bethesda, MD, USA). The same ImageJ software was also used to estimate chloroplast size in both *AtFtsZ1* overexpressing lines and wild-type controls. The results were expressed as mean ± standard deviation and the statistical analysis was performed using SPSS statistics 25 software (IBM).

Transmission electron microscope (TEM)

Leaf tissue from 3-week-old *in vitro* plants were chemically fixed in glutaraldehyde/ paraformaldehyde along with an osmium tetroxide solution and embedded in EMBed-821 resin (Electron Microscopy Sciences, Hatfield, PA, USA) as described previously (Lin, et al. 2014a). Ultrathin sections (60-100 nm) of embedded tissue were cut using a diamond knife and a Leica EM UC7 ultramicrotome (Leica, Buffalo Grove, IL, USA). Ultrathin sections were post strained using aqueous solutions of uranyl acetate and lead citrate as described

previously (Lin, et al. 2014a). Images were obtained using a JEOL 1400 transmission electron microscope operating at 80kV (JEOL, Peabody, MA, USA) equipped with a Gatan OneView camera (Gatan, Pleasanton, CA, USA). TEM micrographs were processed using the software ImageJ 1.41o (National Institute of Health, Bethesda, MD, USA).

Fluorescence imaging by the Fluorescence-Inducing Laser Projector (FILP)
Plant fluorescence patterns were characterized using the Fluorescence-Inducing Laser Projector (FILP), a custom-built instrument for imaging fluorescence in plants. FILP imaging was performed on 3-week-old transplastomic lines in pots (Rigoulot, et al. 2019). GFP fluorescence in plants was acquired at 150 ms exposure time using 465 nm excitation and 525 nm emission notch (50 nm) filter. Standoff detection was performed at 3 m from the laser source. FILP images were processed using ImageJ 1.41o software (National Institute of Health, Bethesda, MD, USA).

Results

Generation of *AtFtsZ1* transgenic potato lines with enlarged chloroplast morphology

Nuclear genome-transformed potato lines overexpressing *AtFtsZ1* under the control of the CaMV 35S promoter were generated by *Agrobacterium*-mediated transformation. In a first screen, six independent lines were chosen for in-depth analysis from > 200 putative transgenic lines using microscopic analysis of chloroplast morphology in leaf mesophyll-derived protoplasts (Supplementary Figure 2.1). Thereafter, we focused on two independent lines that had large chloroplasts (*AtFtsZ1* lines 1 and 2, Supplementary Figure 2.1), which appeared to be vigorous in tissue. These two lines were PCR-positive for the transgene (Fig 1). The presence of the expression cassette integrated into total genomic DNA preparations was verified by amplifying the upstream *nptII* selection gene along with the downstream *AtFtsZ1* transgene (~2 kb DNA bands; Fig 1a). No DNA bands were detected in wild-type samples. In both wild-type and macro-chloroplast lines, DNA bands at very similar intensity were obtained for the endogenous loading control *actin* indicating unaltered amount of genomic DNA in all samples (~0.25 kb DNA bands; Fig 1a). RT-PCRs using primers specific for *AtFtsZ1* confirmed transgene overexpression in the two macro-chloroplast lines (Fig 1b). The DNA gel was normalized using RT-PCR reactions for plastome and nuclear internal controls, *rbcL* and *ef1α*, respectively (Fig 1b).

Top leaves of the canopy from 3-week-old *in vitro* grown macro-chloroplast lines 1, 2 and wild-type controls were observed using confocal microscopy. In leaf mesophyll cells of transgenic lines 1 and 2, significantly larger and elongated chloroplasts were observed relative to wild-type lines (Fig 1c-e, g-i and k-m, respectively; Supplementary Fig S2). The same lines were used for ultrastructural characterization of chloroplast morphology by transmission

electron microscopy (TEM). Micrographs of ultrathin sections from mesophyll leaf cells confirmed a larger and elongated chloroplast morphology in *AtFtsZ1* transgenic potato lines (Fig 1f, j) compared to wild-type control plants (Fig 1n). Internal chloroplast structures, including thylakoid membranes, appeared to be morphologically normal without any apparent structural alterations (Supplementary Fig S3).

The enlarged chloroplast phenotypes in *AtFtsZ1* transgenic lines was confirmed by measurement of organelle size from single plane confocal images from isolated protoplasts (Fig 2). Compared to chloroplast size in wild-type potato leaves (6.7 ± 2.1 µm), the organelle size in macro-chloroplast lines 1 and 2 were substantially enlarged (17.3 ± 11.9 µm and 16.8 ± 9.6 µm, respectively) ($p < 0.05$). The chloroplast size in *AtFtsZ1* transgenic lines 1 and 2 were not statistically different ($p > 0.05$).

Phenotype and growth characteristics of macro-chloroplast lines

The phenotype of macro-chloroplast lines along with wild-type controls was investigated by growing plants in pots in a controlled environment. Two-week-old *in vitro* plants were grown in potting mix for 9 weeks prior to data collection (Fig 3). In comparison with wild-type plants, the two transgenic lines were ~40% shorter (Fig 3a, b) and had lower total fresh weights per plant (~20 and 48% reduction in line 1 and 2, respectively; Fig 3c). Macro line 2 also had ~43% less total dry mass per plant whereas the values of line 1 were not statistically different compared to wild-type (Fig 3d). Despite the similarity in height for the macro-chloroplast lines, line 1 had ~35 and 32% more total fresh and dry weight per plant than line 2, respectively (Fig 3c, d). Furthermore, transgenic line 1 had higher total fresh and dry weights per unit of plant height compared to wild-type plants (~26 and 29% increase, respectively), whereas the values of line 2 were not statistically different to the control plants (Fig 3e and f, respectively). According to these data, macro-chloroplast line 1 had higher values of fresh and dry leaf weight per unit of leaf area (~17 and 29% increase, respectively), while line 2 had only a higher amount of dry weight per leaf area (~15% increase; Fig 3g and h). These growth parameters suggest that macro-chloroplast lines grow slower than wild-type controls. However, the two macro-chloroplast lines accumulated more biomass per unit of leaf area compared to wild-type. Despite the delay in growth, the shoot morphology in macro-chloroplast lines (stem and leaves) were similar to wild-type plants (Fig 3a).

Biolistic transformation of chloroplast genomes in macro-chloroplast lines

The possibility to transform macro-chloroplasts was investigated by using chloroplast transformation vectors delivered by biolistics. Two vectors, pIR and pSSC specific for recombination into the IR (*trnI*/*trnA*) or SSC (*ndhG*/*ndhI*) regions of plastome were used (Fig 4a and b, respectively). Considering the chloroplast morphology and growth parameters of the two *AtFtsZ1* transgenic lines were very similar (Fig 1-3), only the macro line 1 was used for biolistic

transformation. Wild-type potato leaves were used as positive controls for plastome transformation. A total of 25 wild-type and 35 macro pIR transplastomic lines along with 16 wild-type and 18 macro pSSC lines were regenerated, respectively (Supplementary Fig S4 and S5). In wild-type potato, the transformation efficiency for both constructs was higher than macro-chloroplast lines (Table 2.1).

Primers specific for the *trnI*/*trnA* or *ndhG*/*ndhI* integration site and Southern blots were used to check operon integration in the second cycle of selection/regeneration of pIR and pSSC transplastomic lines, respectively (Fig 4c-f and Supplementary Fig S4 and S5). PCRs specific for the two transgenes (*aadA* and *GFP*) along with the internal plastome control *rbcL* were conducted (Fig 4c, d). The presence of 2.6 kb DNA-bands at the predicted molecular weight indicated correct operon integration in both plastome regions (*trnI*/*trnA* or *ndhG*/*ndhI*) in all wild-type and macro-chloroplast transplastomic lines (Fig 4c, d). The presence of bands at the predicted molecular weight in Southern blots for the *trnI*/*trnA* and *ndhG*/*ndhI* integration sites with the cassette integrated (3.8 kb and 4.5 kb, respectively Fig 4e, f) confirmed correct integration. In Southern blots for both pIR and pSSC lines, the absence of wild-type bands (1.6 kb and 2.3 kb, respectively, Fig 4e, f) indicates 100% homoplasmy. Compared to normal and macro-chloroplast transplastomic lines, there was no difference in the level of homoplasmy for either integration site (*trnI*/*trnA* or *ndhG*/*ndhI*, Fig 4e, f).

GFP accumulation in macro-chloroplast lines

GFP fluorescence was detected using the confocal microscope in chloroplasts in wild-type and macro-chloroplast lines transformed with pIR or pSSC constructs. These fluorescent profiles confirmed GFP production and correct subcellular localization in all transplastomic lines (Fig 5a-d). The GFP amount produced in leaf tissue was also investigated by fluorometry (Fig 5e). Quantification of GFP indicated that both normal and macro-chloroplast transplastomic lines accumulated the same amount of protein (ng of GFP) per mg of leaf fresh weight (Fig 5e). No statistically significant difference was observed in GFP accumulation in both genotypes when compared to transplastomic plants transformed with the same construct, pIR or pSSC. Considering the IR region is present in two copies in the chloroplast genome, both wild-type and macro-chloroplast transplastomic lines transformed with pIR accumulated more protein then pSSC plants.

In order to investigate if these levels of GFP expression were enough for efficient standoff detection, transplastomic pIR and pSSC lines of both genotypes were imaged using the recently developed FILP (Fluorescence-Inducing Laser Projector) system (Rigoulot, et al. 2019) (Fig 5f-g). Fluorescent images of normal and macro-chloroplast lines transformed with pIR or pSSC indicated bright green fluorescence detected in engineered plant canopies at a 3-m standoff (Fig 5f and g, respectively).

Discussion

The ultrastructural characterization of leaf tissue from potato *AtFtsZ1* lines using TEM showed the presence of large and elongated chloroplasts throughout all analyzed cells (Fig 1, 2 and Supplementary Fig S2, 3). This is in accordance with previous results demonstrating that the overexpression of *FtsZ1* affects the normal formation of the z-ring, resulting in inhibition of division and excessive chloroplast growth (Chikkala, et al. 2012; de Pater, et al. 2006; Stokes, et al. 2000). Analyzing TEM micrographs, we did not observe any evidence for ultrastructural differences of thylakoid membranes in macro-chloroplasts compared to wild-type organelles (Fig 1, and Supplementary Fig S3).

The results from this work suggest that macro-chloroplast lines have a delay in growth compared to wild-type controls, although they accumulate more biomass per unit of leaf area over the same time period (Fig 3). It is possible that macro-chloroplast lines reach full maturity (flowering) slower than wild-type plants, which was observed in a separate growth study (data not shown). Despite altered growth, the macro-chloroplast plants had no apparent changes in leaf and stem morphology compared to wild-type controls (Fig 3a). Altered growth characteristics may be the result of altered leaf optical properties in the wild-type and macro-chloroplast phenotypes. It has been demonstrated in tobacco *AtFtsZ* overexpressing lines, the particular chloroplast phenotype and cell rearrangement can lead to either decreased light absorbance (at low-light irradiation) or photodamage of the photosynthetic machinery (at high-light irradiation), resulting in reduced plant growth in both cases (Jeong, et al. 2002).

The prospect of using agronomically-relevant plants with large organelles is an attractive addition to the plastid biotechnology toolkit. Macro-chloroplasts from *AtFtsZ1* transgenic lines can be transformed using biolistics and transplastomic lines can be regenerated using standard tissue culture (Valkov, et al. 2014). We anticipated that increasing the size of the target organelle may provide an advantage in transformation efficiency in potato as it has been previously demonstrated in tobacco (Jeong, et al. 2006). Considering the transformation efficiencies of potato chloroplasts reported in the literature (Nguyen, et al. 2005; Sidorov, et al. 1999; Valkov, et al. 2011), our macro-chloroplast lines can be efficiently transformed, yielding approximately 1 transplastomic plant per each plate bombarded (0.9 and 1.1 for pIR and pSSC construct, respectively, Table 2.1). However, the transformation efficiency of transgenic lines was not greater than the efficiency obtained in our laboratory for wild-type potato leaves. For both constructs tested (pIR and pSSC) about 4 positive plants per each plate bombarded (3.5 and 3.9, respectively, Table 2.1) were obtained. This observation suggests that there is not a clear advantage of having larger organelles using biolistics. This low transformation efficiency could be due to the particular spatial distribution of macro-chloroplasts in leaf mesophyll cells. Compared to wild-type, macro-chloroplasts seem not to be regularly distributed

throughout the cytoplasm and are instead pushed to the plasma membrane. This particular distribution may reduce the probability of gold particles hitting the target organelle.

Nevertheless, this result does not exclude the potential for macro-chloroplasts to be desirable targets for other gene-delivery methods in which the transformable surface is an important limiting factor. For almost 3 decades, the gene gun has been the most efficient and reproducible method used for chloroplast transformation (Svab, et al. 1990; Svab and Maliga 1993). This is an invasive physical method that could induce the shearing of large vectors and, thus, is not particularly suited for transformation of large, chromosome-sized constructs. The effective transformation of large DNA molecules, such as an entire synthetic plastome, may necessitate alternative methods of transformation. Many methods including micro-injection (Knoblauch, et al. 1999; Verhoeven and Blaas 1992), polyethylene glycol (PEG) transformation (Golds, et al. 1993; Koop and Kofer 1995) and electroporation (To, et al. 1996; Zhang, et al. 2014) have been used to transform chloroplasts. Different kinds of commercial carbon-nanotubes have also been used to deliver DNA into plants cells (Burlaka, et al. 2015; Kwak, et al. 2019; Yuan, et al. 2011), and these delivery techniques could also represent an alternative to deliver vectors more efficiently into macro-chloroplasts. Thus, macro-chloroplasts could provide considerable advantages in increasing transformation efficiency using these methods.

Compared to conventional transplastomic lines, macro-chloroplast lines do not differ in the level of homoplasmy for both *trnI*/*trnA* and *ndhG*/*ndhI* integration sites (Fig 4). Two rounds of selection in tissue culture were required to get pIR and pSSC homoplasmic lines (*trnI*/*trnA* and *ndhG*/*ndhI*) in both macro- and wild-type chloroplast-engineered plants. This result suggests that, if there is a difference in plastome copy number in the two genotypes (Osteryoung and Pyke 2014), such differences do not influence the level of homoplasmy after two generations of transplastomic lines. This is also supported by protein quantification of GFP in wild-type and macro-chloroplast transplastomic lines (Fig 5e). Moreover, as observed for wild-type chloroplasts, the level of GFP accumulation obtained in macro-chloroplast lines was enough for efficient standoff imaging of whole plant canopy using the FILP system (Fig 5f, g). This is of particular interest in a context of modern agriculture, where phenomics has the prospect to bring together high-dimensional phenotypic and molecular data together with their interaction with environmental stimuli (Houle, et al. 2010; Rigoulot, et al. 2019).

References

Arlen PA, Falconer R, Cherukumilli S, Cole A, Cole AM, Oishi KK, Daniell H (2007) Field production and functional evaluation of chloroplast-derived interferon-alpha2b. Plant Biotechnol J 5:511-525

Bock R (2013) Strategies for metabolic pathway engineering with multiple transgenes. Plant Mol Biol 83:21-31

Bock R (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. Ann Rev Plant Biol 66:211-241

Boyhan D, Daniell H (2011) Low-cost production of proinsulin in tobacco and lettuce chloroplasts for injectable or oral delivery of functional insulin and C-peptide. Plant Biotechnol J 9:585-598

Burlaka OM, Pirko YV, Yemets AI, Blume YB (2015) Plant genetic transformation using carbon nanotubes for DNA delivery. Tsitol Genet 49:349-357

Chen PJ, Senthilkumar R, Jane WN, He Y, Tian Z, Yeh KW (2014) Transplastomic *Nicotiana benthamiana* plants expressing multiple defence genes encoding protease inhibitors and chitinase display broad-spectrum resistance against insects, pathogens and abiotic stresses. Plant Biotechnol J 12:503-515

Chikkala VRN, Nugent GD, Stalker DM, Mouradov A, Stevenson TW (2012) Expression of *Brassica oleracea FtsZ1-1* and *MinD* alters chloroplast division in *Nicotiana tabacum* generating macro- and mini-chloroplasts. Plant Cell Rep 31:917-928

Chronis D, Chen S, Lu S, Hewezi T, Carpenter SC, Loria R, Baum TJ, Wang X (2013) A ubiquitin carboxyl extension protein secreted from a plant-parasitic nematode *Globodera rostochiensis* is cleaved in planta to promote plant parasitism. Plant J 74:185-196

Daniell H (2007) Transgene containment by maternal inheritance: effective or elusive? Proc Natl Acad Sci USA 104:6879-6880

Daniell H, Kulis M, Herzog RW (2019a) Plant cell-made protein antigens for induction of oral tolerance. Biotechnol Adv 37:107413

Daniell H, Mangu V, Yakubov B, Park J, Habibi P, Shi Y, Gonnella PA, Fisher A, Cook T, Zeng L, Kawut SM, Lahm T (2020) Investigational new drug enabling angiotensin oral-delivery studies to attenuate pulmonary hypertension. Biomaterials 233:119750

Daniell H, Rai V, Xiao Y (2019b) Cold chain and virus-free oral polio booster vaccine made in lettuce chloroplasts confers protection against all three poliovirus serotypes. Plant Biotechnol J 17:1357-1368

Daniell H, Ribeiro T, Lin S, Saha P, McMichael C, Chowdhary R, Agarwal A (2019c) Validation of leaf and microbial pectinases: commercial launching of a new platform technology. Plant Biotechnol J 17:1154-1166

de Pater S, Caspers M, Kottenhagen M, Meima H, ter Stege R, de Vetten N (2006) Manipulation of starch granule size distribution in potato tubers by modulation of plastid division. Plant Biotechnol J 4:123-134

Fuentes P, Zhou F, Erban A, Karcher D, Kopka J, Bock R (2016) A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. eLife 5: e13664
Golds T, Maliga P, Koop H-U (1993) Stable plastid transformation in PEG-treated protoplasts of *Nicotiana tabacum*. Bio/Technology 11:95-97
Houle D, Govindaraju DR, Omholt S (2010) Phenomics: the next challenge. Nat Rev Genet 11:855-866
Jeong WJ, Min SR, Liu JR (2006) Enhancement of chloroplast transformation frequency by using mesophyll cells containing a few enlarged chloroplasts from nuclear transformed plants in tobacco. J Plant Biotechnol 34:271-275
Jeong WJ, Park YI, Suh K, Raven JA, Yoo OJ, Liu JR (2002) A large population of small chloroplasts in tobacco leaf cells allows more effective chloroplast movement than a few enlarged chloroplasts. Plant Physiol 129:112-121
Jin S, Daniell H (2015) The engineered chloroplast genome just got smarter. Trends Plant Sci 20:622-640
Jin S, Kanagaraj A, Verma D, Lange T, Daniell H (2011) Release of hormones from conjugates: chloroplast expression of β-glucosidase results in elevated phytohormone levels associated with significant increase in biomass and protection from aphids or whiteflies conferred by sucrose esters. Plant Physiol 155:222-235
Knoblauch M, Hibberd JM, Gray JC, van Bel AJ (1999) A galinstan expansion femtosyringe for microinjection of eukaryotic organelles and prokaryotes. Nat Biotechnol 17:906-909
Koop H-U, Kofer W (1995) Plastid transformation by polyethylene glycol treatment of protoplasts and regeneration of transplastomic tobacco plants. In: Potrykus I, Spangenberg G (eds) Gene Transfer to Plants. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 75-82
Kumari U, Singh R, Ray T, Rana S, Saha P, Malhotra K, Daniell H (2019) Validation of leaf enzymes in the detergent and textile industries: launching of a new platform technology. Plant Biotechnol J 17:1167-1182
Kwak SY, Lew TTS, Sweeney CJ, Koman VB, Wong MH, Bohmert-Tatarev K, Snell KD, Seo JS, Chua NH, Strano MS (2019) Chloroplast-selective gene delivery and expression in planta using chitosan-complexed single-walled carbon nanotube carriers. Nat Nanotechnol 14:447-455
Langbecker CL, Ye GN, Broyles DL, Duggan LL, Xu CW, Hajdukiewicz PT, Armstrong CL, Staub JM (2004) High-frequency transformation of undeveloped plastids in tobacco suspension cells. Plant Physiol 135:39-46
Lee SM, Kang K, Chung H, Yoo SH, Xu XM, Lee SB, Cheong JJ, Daniell H, Kim M (2006) Plastid transformation in the monocotyledonous cereal crop, rice (*Oryza sativa*) and transmission of transgenes to their progeny. Mol Cells 21:401-410
Lin MT, Occhialini A, Andralojc PJ, Devonshire J, Hines KM, Parry MA, Hanson MR (2014a) β-Carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts. Plant J 79:1-12

Lin MT, Occhialini A, Andralojc PJ, Parry MA, Hanson MR (2014b) A faster Rubisco with potential to increase photosynthesis in crops. Nature 513:547-550
Lu Y, Rijzaani H, Karcher D, Ruf S, Bock R (2013) Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. Proc Natl Acad Sci USA 110:E623-632
Maliga P, Bock R (2011) Plastid biotechnology: food, fuel, and medicine for the 21st century. Plant Physiol 155:1501-1510
McAndrew RS, Froehlich JE, Vitha S, Stokes KD, Osteryoung KW (2001) Colocalization of plastid division proteins in the chloroplast stromal compartment establishes a new functional relationship between *FtsZ1* and *FtsZ2* in higher plants. Plant Physiol 127:1656-1666
Mori T, Kuroiwa H, Takahara M, Miyagishima SY, Kuroiwa T (2001) Visualization of an FtsZ ring in chloroplasts of *Lilium longiflorum* leaves. Plant Cell Physiol 42:555-559
Nguyen TT, Nugent G, Cardi T, Dix PJ (2005) Generation of homoplasmic plastid transformants of a commercial cultivar of potato (*Solanum tuberosum* L.). Plant Sci 168:1495-1500
Occhialini A, Lin MT, Andralojc PJ, Hanson MR, Parry MA (2016) Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO_2. Plant J 85:148-160
Occhialini A, Piatek AA, Pfotenhauer AC, Frazier TP, Stewart CN, Jr., Lenaghan SC (2019) MoChlo: a versatile, modular cloning toolbox for chloroplast biotechnology. Plant Physiol 179:943-957
Oey M, Lohse M, Scharff LB, Kreikemeyer B, Bock R (2009) Plastid production of protein antibiotics against pneumonia via a new strategy for high-level expression of antimicrobial proteins. Proc Natl Acad Sci USA 106:6579-6584
Osteryoung KW, Pyke KA (2014) Division and dynamic morphology of plastids. Ann Rev Plant Biol 65:443-472
Osteryoung KW, Stokes KD, Rutherford SM, Percival AL, Lee WY (1998) Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial *FtsZ*. Plant Cell 10:1991-2004
Park J, Yan G, Kwon KC, Liu M, Gonnella PA, Yang S, Daniell H (2020) Oral delivery of novel human IGF-1 bioencapsulated in lettuce cells promotes musculoskeletal cell proliferation, differentiation and diabetic fracture healing. Biomaterials 233:119591
Pasoreck EK, Su J, Silverman IM, Gosai SJ, Gregory BD, Yuan JS, Daniell H (2016) Terpene metabolic engineering via nuclear or chloroplast genomes profoundly and globally impacts off-target pathways through metabolite signalling. Plant Biotechnol J 14:1862-1875
Rey P, Sanz-Barrio R, Innocenti G, Ksas B, Courteille A, Rumeau D, Issakidis-Bourguet E, Farran I (2013) Overexpression of plastidial thioredoxins f and m differentially alters photosynthetic activity and response to oxidative stress in tobacco plants. Front Plant Sci 4:390

Rigoulot SB, Schimel TM, Lee JH, Brabazon H, Meier KA, Schmid MJ, Seaberry EM, Poindexter MR, Layton JS, Brabazon JW, Madajian JA, Finander MJ, DiBenedetto J, Occhialini A, Lenaghan SC, Stewart CN (2019) Fluorescence-based whole plant imaging and phenomics. bioRxiv (preprint) 865428, 10.1101/865428
Ruf, S., Forner, J., Hasse, C., Kroop, X., Seeger, S., Schollbach, L., Schadach, A. and Bock, R. (2019) High-efficiency generation of fertile transplastomic Arabidopsis plants. Nat Plants 5(3):282-289.
Schindel HS, Piatek AA, Stewart CN, Jr., Lenaghan SC (2018) The plastid genome as a chassis for synthetic biology-enabled metabolic engineering: players in gene expression. Plant Cell Rep 37:1419-1429
Sidorov VA, Kasten D, Pang SZ, Hajdukiewicz PT, Staub JM, Nehra NS (1999) Technical Advance: stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker. Plant J 19:209-216
Stokes KD, McAndrew RS, Figueroa R, Vitha S, Osteryoung KW (2000) Chloroplast division and morphology are differentially affected by overexpression of *FtsZ1* and *FtsZ2* genes in *Arabidopsis*. Plant Physiol 124:1668-1677
Su J, Zhu L, Sherman A, Wang X, Lin S, Kamesh A, Norikane JH, Streatfield SJ, Herzog RW, Daniell H (2015) Low cost industrial production of coagulation factor IX bioencapsulated in lettuce cells for oral tolerance induction in hemophilia B. Biomaterials 70:84-93
Svab Z, Hajdukiewicz P, Maliga P (1990) Stable transformation of plastids in higher plants. Proc Natl Acad Sci USA 87:8526-8530
Svab Z, Maliga P (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. Proc Natl Acad Sci USA 90:913-917
To KY, Cheng MC, Chen LF, Chen SC (1996) Introduction and expression of foreign DNA in isolated spinach chloroplasts by electroporation. Plant J 10:737-743
Valkov VT, Gargano D, Manna C, Formisano G, Dix PJ, Gray JC, Scotti N, Cardi T (2011) High efficiency plastid transformation in potato and regulation of transgene expression in leaves and tubers by alternative 5' and 3' regulatory sequences. Transgenic Res 20:137-151
Valkov VT, Gargano D, Scotti N, Cardi T (2014) Plastid transformation in potato: *Solanum tuberosum*. Methods Mol Biol (Clifton, NJ) 1132:295-303
Valkov VT, Scotti N, Kahlau S, Maclean D, Grillo S, Gray JC, Bock R, Cardi T (2009) Genome-wide analysis of plastid gene expression in potato leaf chloroplasts and tuber amyloplasts: transcriptional and posttranscriptional control. Plant Physiol 150:2030-2044
Verhoeven HA, Blaas J (1992) Direct cell to cell transfer of organelles by microinjection. Plant Cell Rep 10:613-616
Verma D, Daniell H (2007) Chloroplast vector systems for biotechnology applications. Plant Physiol 145:1129-1143
Verma D, Samson NP, Koya V, Daniell H (2008) A protocol for expression of foreign genes in chloroplasts. Nat Protoc 3:739-758

Vitha S, McAndrew RS, Osteryoung KW (2001) FtsZ ring formation at the chloroplast division site in plants. J Cell Biol 153:111-120
Wamboldt Y, Mohammed S, Elowsky C, Wittgren C, de Paula WB, Mackenzie SA (2009) Participation of leaky ribosome scanning in protein dual targeting by alternative translation initiation in higher plants. Plant Cell 21:157-167
Wang Y, Wei Z, Xing S (2018) Stable plastid transformation of rice, a monocot cereal crop. Biochem Biophys Res Commun 503:2376-2379
Weigel D, Glazebrook J (2006) Transformation of *agrobacterium* using the freeze-thaw method. CSH Protoc 2006: 10.1101/pdb.prot4666
Wurbs D, Ruf S, Bock R (2007) Contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. Plant J 49:276-288
Yu Q, Barkan A, Maliga P (2019) Engineered RNA-binding protein for transgene activation in non-green plastids. Nat Plants 5:486-490
Yu Q, Lutz KA, Maliga P (2017) Efficient Plastid Transformation in *Arabidopsis*. Plant Physiol 175:186-193
Yuan H, Hu S, Huang P, Song H, Wang K, Ruan J, He R, Cui D (2011) Single walled carbon nanotubes exhibit dual-phase regulation to exposed *Arabidopsis* mesophyll cells. Nanoscale Res Lett 6:44
Zhang R, Patena W, Armbruster U, Gang SS, Blum SR, Jonikas MC (2014) High-throughput genotyping of green algal mutants reveals random distribution of mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. Plant Cell 26:1398-1409

Appendix

Table 2.1. Transformation efficiency of chloroplasts from wild-type potato and macro-chloroplast lines transformed with either pIR or pSSC vector

	Wild-type chloroplasts			Macro-chloroplasts		
Construct	**Plates bombarded**	**Positive plants**	**Plants/plates bombarded**	**Plates bombarded**	**Positive plants**	**Plants/plates bombarded**
pIR	**18**	**64**	**3.5**	**26**	**23**	**0.9**
pSSC	**20**	**78**	**3.9**	**26**	**28**	**1.1**

The number of plates bombarded, the total number of positive plants and the number of plants per each plate bombarded are indicated.

Supplementary Table 2.1. Primers used in this study

id	name	sequence (5'-3')
1 Fw	Selectio-Cassette-1-Fw	CAATGTGAGTTTTTGTAGTTGGATTTGCTCC
2 Fw	nptII/AtFtsZ1-Cass-Fw	TCGGGGTAGGTGTTTCTT
3 Fw	Actin-P-Fw	CGGAGCGTGGTTACTCATTT
4 Fw	trnA-Fw	CAGTAGAGTCTTTCAGTGGCACGTT
5 Fw	SSC2-Fw	CCCCCTAATATAAGACCCGACCC
6 Fw	mGFP-full-Fw	ATGAGTAAAGGAGAAGAACTTT
7 Fw	SmR-full-Fw	ATGGCAGAAGCGGTGATC
8 Fw	rbcL-P-Fw	GCTGCCGAATCTTCTACTGG
9 Fw	AtFtsZ-qPCR Fw	AGCGGTTTACAGAGTGTTGA
10 Fw	rbcL-q-Fw	AGATCTGCGAATCCCTGTTG
11 Fw	ef1-alpha-q-Fw	ATTGGAAACGGATATGCTCCA
12 Fw	IR-probe-Fw	GATATAGCTCAGTTGGTAGAGCTCCGCTCT
13 Fw	SSC-probe-Fw	CAACCACTAGTTTGAATTGCCCAAGCAAAA
1 Rv	Selectio-Cassette-1-Rv	CTGCAGCCCAAACAAATACAAAATCAAAATAGA
2 Rv	nptII/AtFtsZ1-Cass-Rv	TTCAGTGACAACGTCGAGCA
3 Rv	Actin-P-Rv	GCAGCTTCCATTCCAATCAT
4 Rv	trnI-Rv	GCCAGGGTAAGGAAGAAGGGG
5 Rv	SSC2-Rv	CCGAATTACGAAGGCTTAGTTCGG
6 Rv	mGFP-full-Rv	TTATTTGTATAGTTCATCCATGCC
7 Rv	SmR-full-Rv	TTATTTGCCGACTACCTTGGT
8 Rv	rbcL-P-Rv	CAGGGCTTTGAACCCAAATA
9 Rv	AtFtsZ-qPCR Rv	AGCCCACGAGTTAAAAGTTC
10 Rv	rbcL-q-Rv	CAGGGGACGACCATACTTGT
11 Rv	ef1-alpha-q-Rv	TCCTTACCTGAACGCCTGTCA
12 Rv	IR-probe-Rv	GCGGACAGCTAATGCGTTCCACTTATTGAA
13 Rv	SSC-probe-Rv	ATGATTACCCTGTCCCACGCAAATCGTTTA

The primer id, the full name and the nucleotide sequence (from 5’ to 3’) are indicated in the table. The primers are subdivided in forward (1 Fw-13 Fw) and reverse (1 Rv-13 Rv) primers.

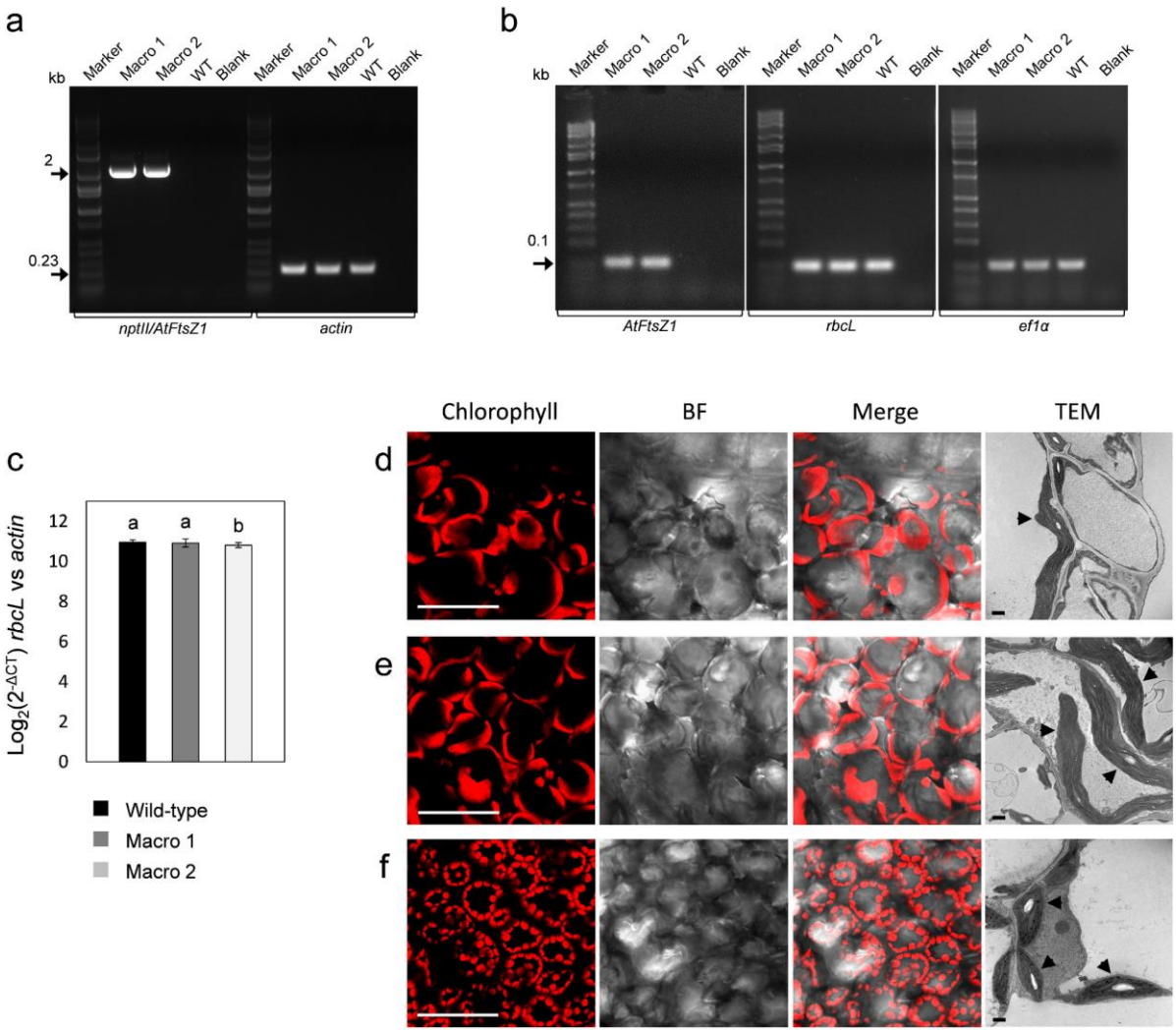


Figure 2.1. Genetic and microscopy characterization of *AtFtsZ1* over-expressing potato lines. (a) PCRs using a pair of primers specific for the cassette *nptII/AtFtsZ1* confirmed the presence of transgenes integrated (2 kb DNA bands) in the genome of *AtFtsZ1*-line 1 (Macro 1) and 2 (Macro 2). Primers specific for the *actin* gene were used as loading controls (0.23 kb DNA bands). (b) PCR reactions using cDNA preparations of macro lines (Macro 1 and 2) and primers specific for *AtFtsZ1* confirmed transgenes expression in both transgenic lines (0.1 kb DNA bands). Primers for *rbcL* and *ef1α* were used as loading controls (0.1 kb DNA bands) for plastome and nuclear genome, respectively. Wild-type samples (WT), blanks and molecular-weight markers are also shown in the gels. (c-n) Chloroplast morphology in *AtFtsZ1* lines and wild-type plants (WT). Confocal stack images of leaf mesophyll cells showing large chloroplasts in *AtFtsZ1* over-expressing lines (c-e: Macro line 1; g-i: Macro line 2) and normal chloroplast morphology in the WT control (k-m) are shown. Electron micrographs of ultrathin sections showing mesophyll cells from leaf tissue prepared by chemical fixation are also shown (f, j, n). Ultra-structures of macro-chloroplast morphology in *AtFtsZ1*-line 1 and 2 (f, j) along with organelles with normal

morphology in wild type plants (n) are indicated with black arrows. Chlorophyll (red: c, g and k), bright-field (gray: d, h and l) and merge images (e, i and m) are shown. Scale bars: 50 µm (c-e, g-i and k-m); 1 µm (f, j and n).

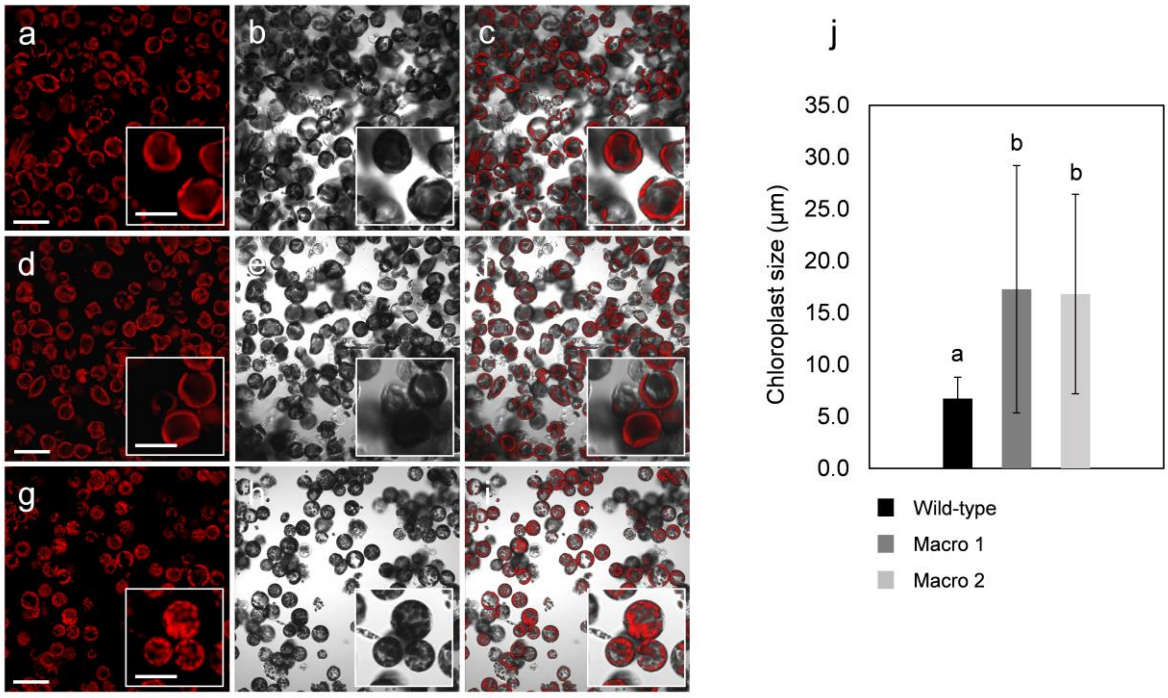


Figure 2.2. Chloroplasts size in *AtFtsZ1* over-expressing lines and wild-type controls. Confocal images of protoplasts from leaf mesophyll cells showing enlarge chloroplast phenotype in *AtFtsZ1* over-expressing lines (a-c: Macro line 1; d-f: Macro line 2) and normal chloroplast morphology in the wild-type control (g-i) are shown. Chlorophyll (red: a, d and g), bright-field (gray: b, e and h) and merge images (c, f and i) are shown. Scale bars: 50 µm (a-c, d-f and g-i); 25 µm (insert a-c, d-f and g-i). Graph representing the chloroplasts size (µm) in macro-chloroplast lines (Macro 1 and 2) and wild-type potato (j). Confocal images of isolated leaf protoplasts from transgenic lines (Macro 1 and 2) and wild-type plants have been evaluated. The results are expressed as mean ± standard deviation of 1382, 720 and 786 number (n) of chloroplasts analyzed in wild-type, Macro 1 and 2, respectively. Means were compared using ANOVA and when significant, mean separations were analyzed using Tukey HSD ($p<0.05$). Statistical significance is indicated by letters (a or b).

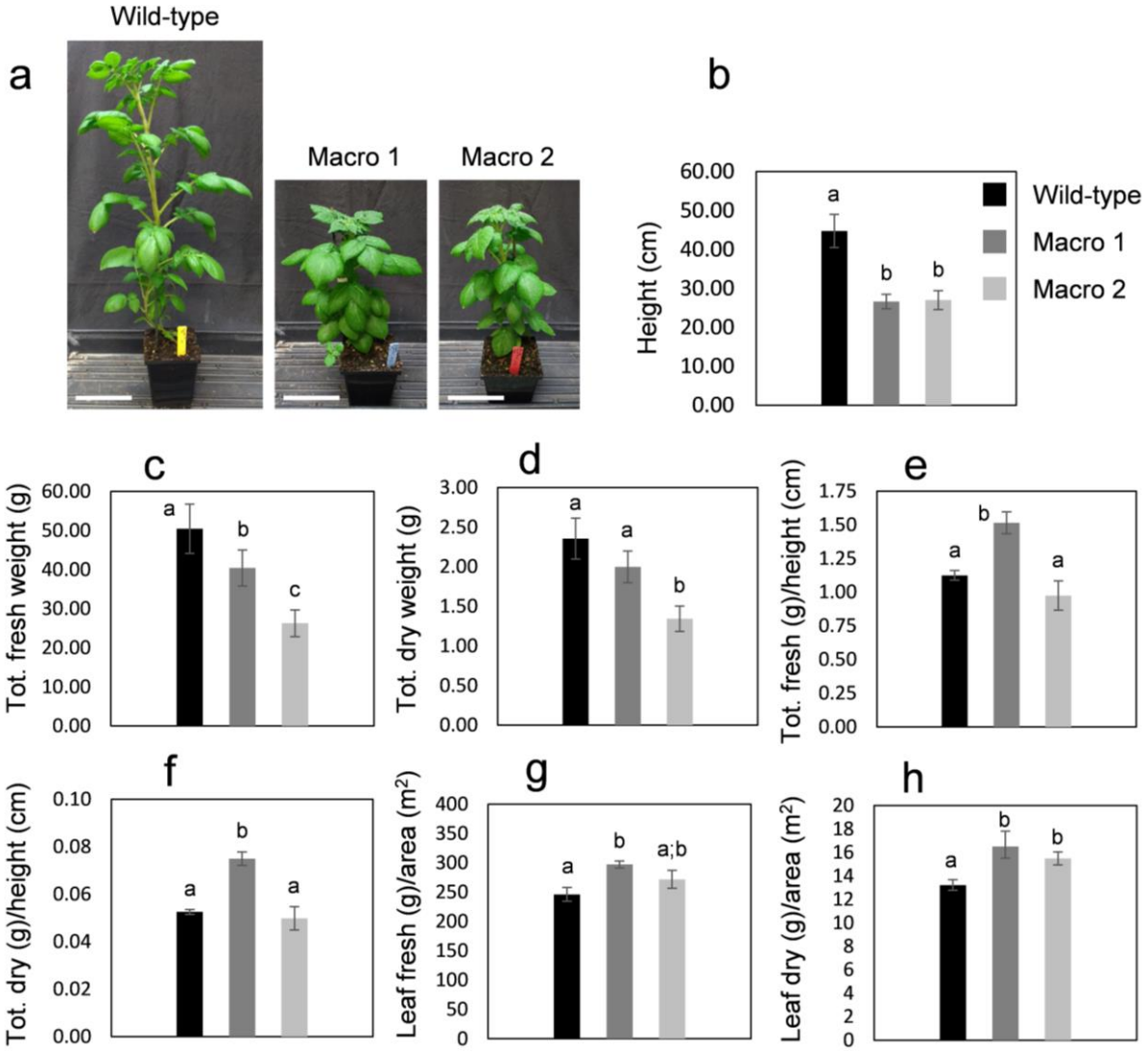


Figure 2.3. Growth characteristics of *AtFtsZ1* over-expressing potato lines. (a) Images showing wild-type potato plants and two independent *AtFtsZ1* lines (Macro lines 1 and 2) grown in pots for 9 weeks. Histograms represent various plant characteristics: (b) height; (c) total fresh weight; (d) total dry weight; (e) ratio of total fresh weight to height; (f) ratio of total dry weight to height; (g) ratio of fresh weight of leaves to their foliar area; (h) ratio of leaf dry weight to foliar area. The results as expressed as mean ± standard deviation of four plants per each genotype. Means were compared using ANOVA and when significant, mean separations were analyzed using Tukey HSD ($p<0.05$). Statistical significance is indicated by letters (a, b or c). Scale bars: 10 cm (a).

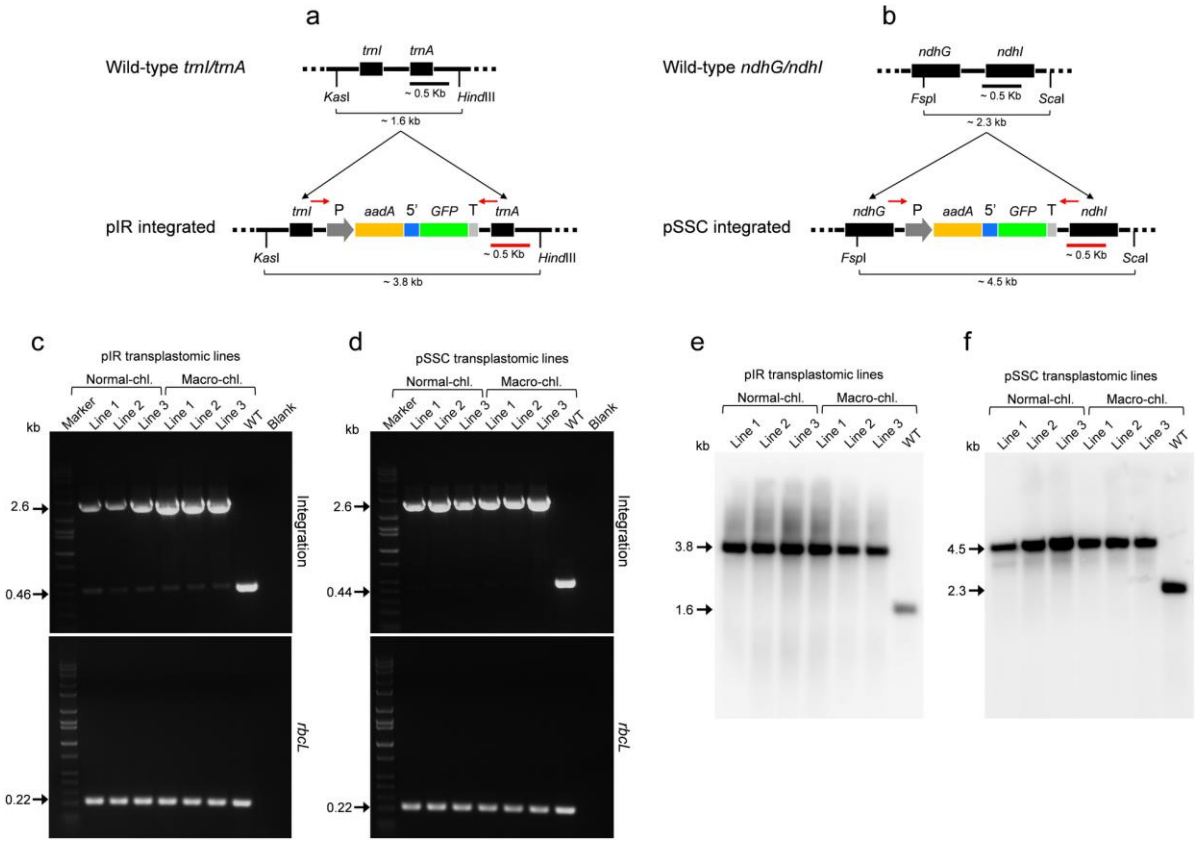


Figure 2.4. Genotyping of second round of transplastomic lines. Schematic representation of pIR and pSSC vectors (a and b) integrated into *trnI/trnA* site of the inverted repeat region and *ndhG/ndhI* site of the short single-copy regions of the plastome, respectively. The dual marker operon construct is identical between vectors and includes: *rrn* promoter along with a synthetic RBS (P); *aadA* gene for spectinomycin selection; 5' untranslated region (5'); a GFP reporter gene; and a terminator/ 3' untranslated region (T). Restriction enzymes used for Southern blots of pIR and pSSC lines (*Kas*I/*Hind*III and *Fsp*I/*Sca*I, respectively) along with the sizes of expected DNA fragments detected by a ~0.5 kb probe (red bar) are indicated. Oligonucleotide used to check integration into either *trnI/trnA* or *ndhG/ndhI* site are indicated (red arrows). Pairs of primers specific for the *trnI/trnA* or *ndhG/ndhI* integration sites were used to check integration in pIR and pSSC lines, respectively (c and d, respectively). Three lines (1-3) for each construct (pIR or pSSC) and genotype (normal or macro) were tested (more lines are shown in Supplementary Fig S4 and S5). DNA bands of 2.6 kb indicate correct integration of both constructs, whereas 0.46 and 0.44 kb-bands indicate the presence of wild-type *trnI/trnA* or *ndhG/ndhI* integration sites, respectively. PCR bands of 0.72 and 0.79 kb confirm the presence of the two transgenes, *GFP* and *aadA*, respectively. PCRs specific for the *rbcL* gene (0.22 kb) were used as loading controls. Wild-type samples, blanks and molecular-weight markers are also shown in the gels. DNA blot analysis of genomic DNA from pIR and pSSC lines (normal and macro) along with wild-type controls (WT) digested

with the indicated restriction enzymes (a and b) are shown in e and f, respectively. DNA fragments at 3.8 kb and 4.5 kb indicate the presence of the transplastomic plastome in pIR and pSSC lines, respectively, were fragments at 1.6 kb and 2.3 kb indicate the presence of wild-type DNA only in wild-type controls.

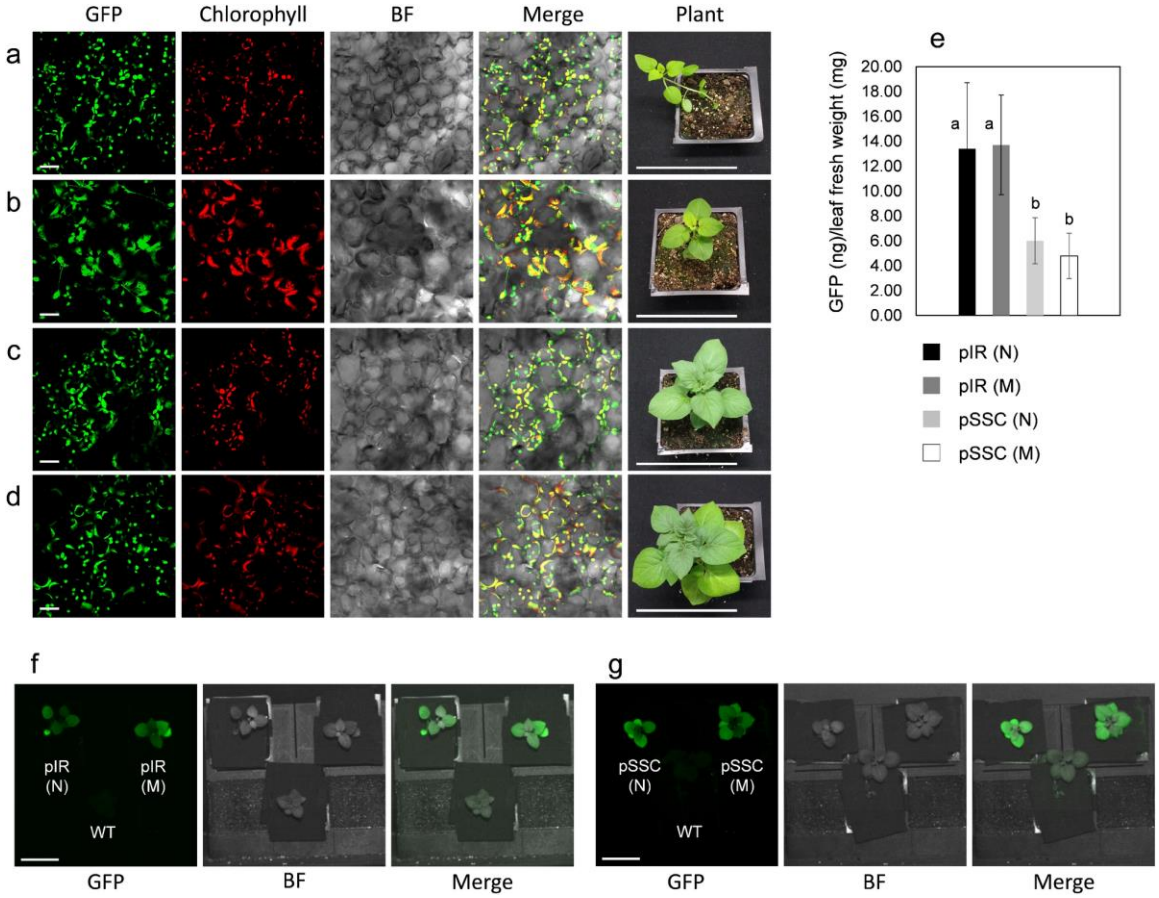
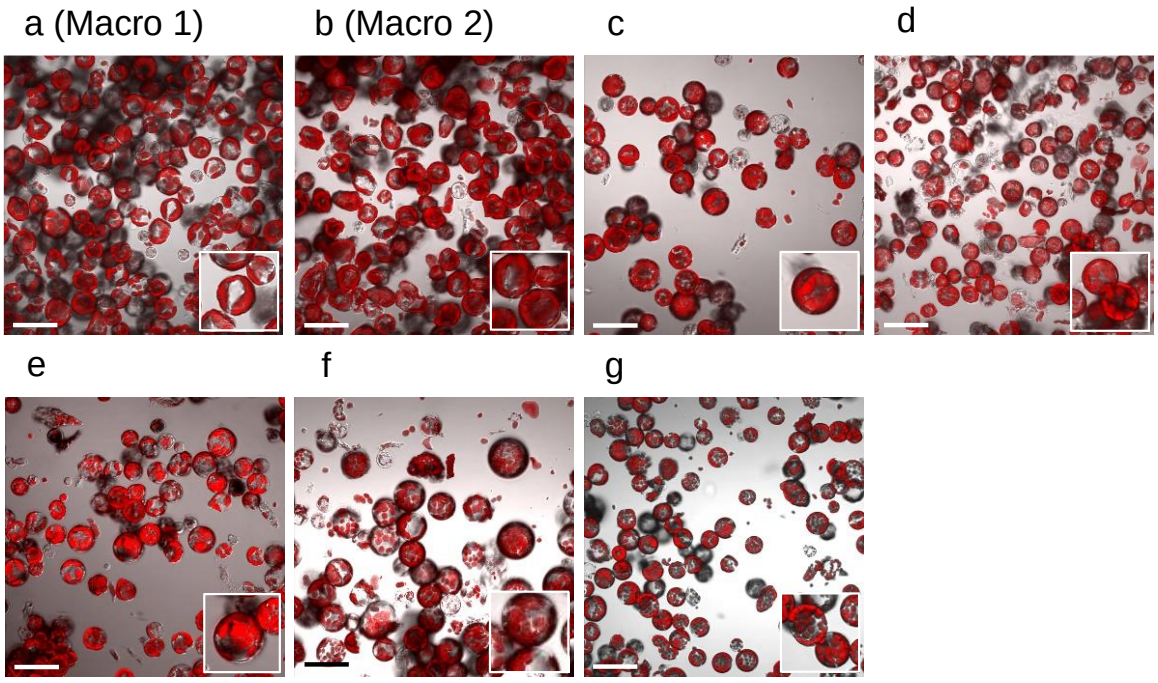


Figure 2.5. GFP accumulation into chloroplasts of normal transplastomic and macro-chloroplast lines. Confocal images showing GFP localization into the chloroplast stroma of leaf mesophyll cells from normal and macro transplastomic lines, along with pictures showing 3-weeks-old *in vitro* plants are indicated: Normal pIR (a); macro-chloroplasts pIR (b); normal pSSC (c); macro-chloroplasts pSSC (d). GFP (green), chlorophyll (red), bright-field (BF, gray) and merge images (GFP/chlorophyll/bright-field) are indicated. Per each genotype (a-d), images of indicated transplastomic plants are shown. Scale bars: 20 µm (a-d, in GFP, chlorophyll, BF and merge); 10 mm (in plants a-d). (u) Graph indicating the GFP amount (ng) per unit of leaf fresh weight (mg). Normal (N) and macro-chloroplast (M) lines transformed with either pIR or pSSC constructs are indicated. The results are expressed as mean ± standard deviation of 2 biological replicate and 2 technical replicate per biological replicate per each independent line. A number of 15 pIR and 8 pSSC lines were used per each genotype, respectively (N: normal; M: macro). Means were compared using ANOVA and when significant, mean separations were analyzed using Tukey HSD ($p<0.05$). Statistical significance are indicated by numerical letter (a or b). Images acquired using the FILP (Fluorescence-Inducing Laser Projector) stand-off detection system of transplastomic lines along with wild-type controls (WT) are shown: 3-week-old pIR and pSSC lines of both genotypes (N and M) on potting are shown in f and g, respectively. GFP signal, bright-field (BF) and merge images are

indicated in f and g. The images were acquired at 150 ms exposure time. Scale bar: 50 mm (f and g).

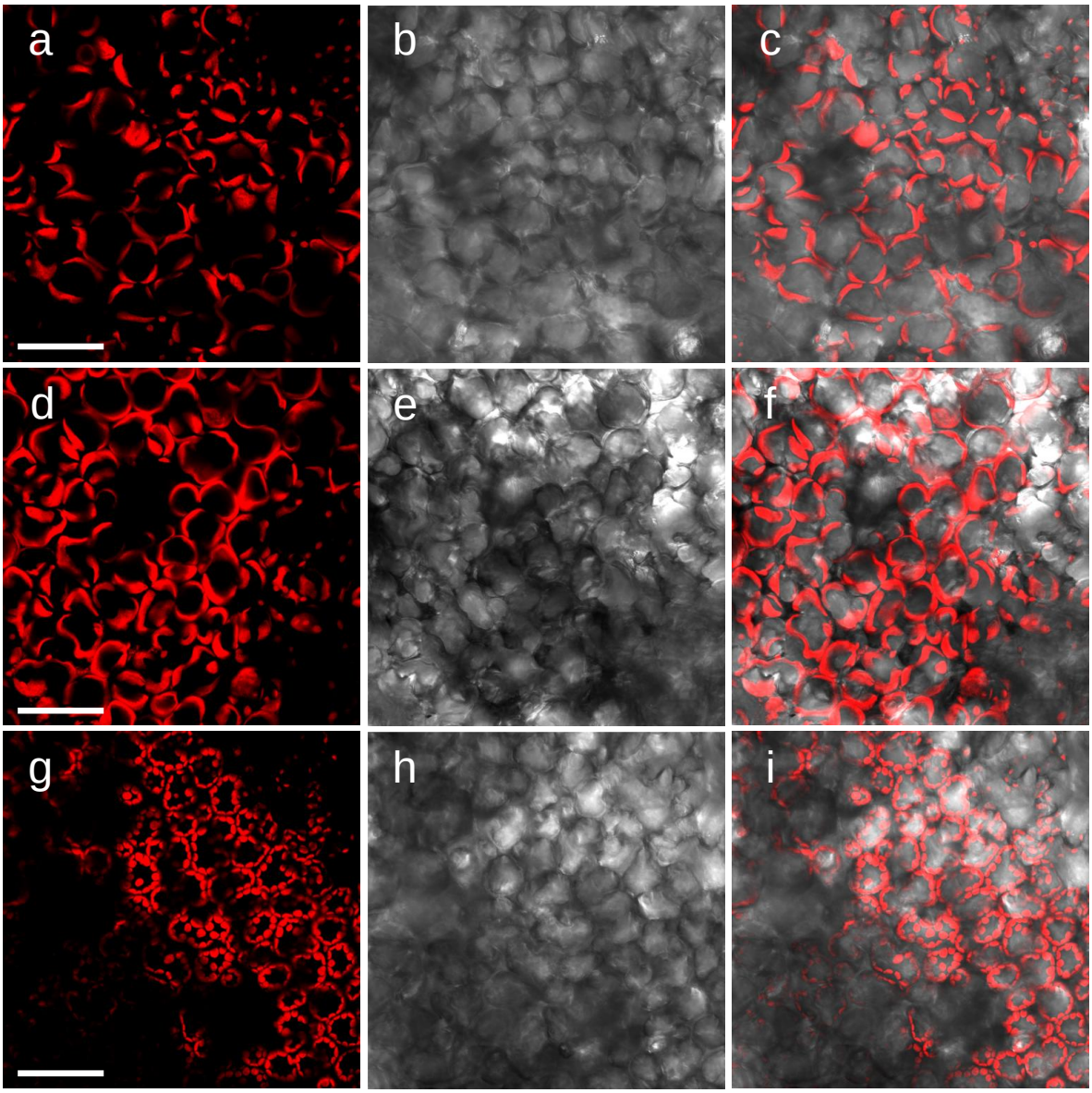


h

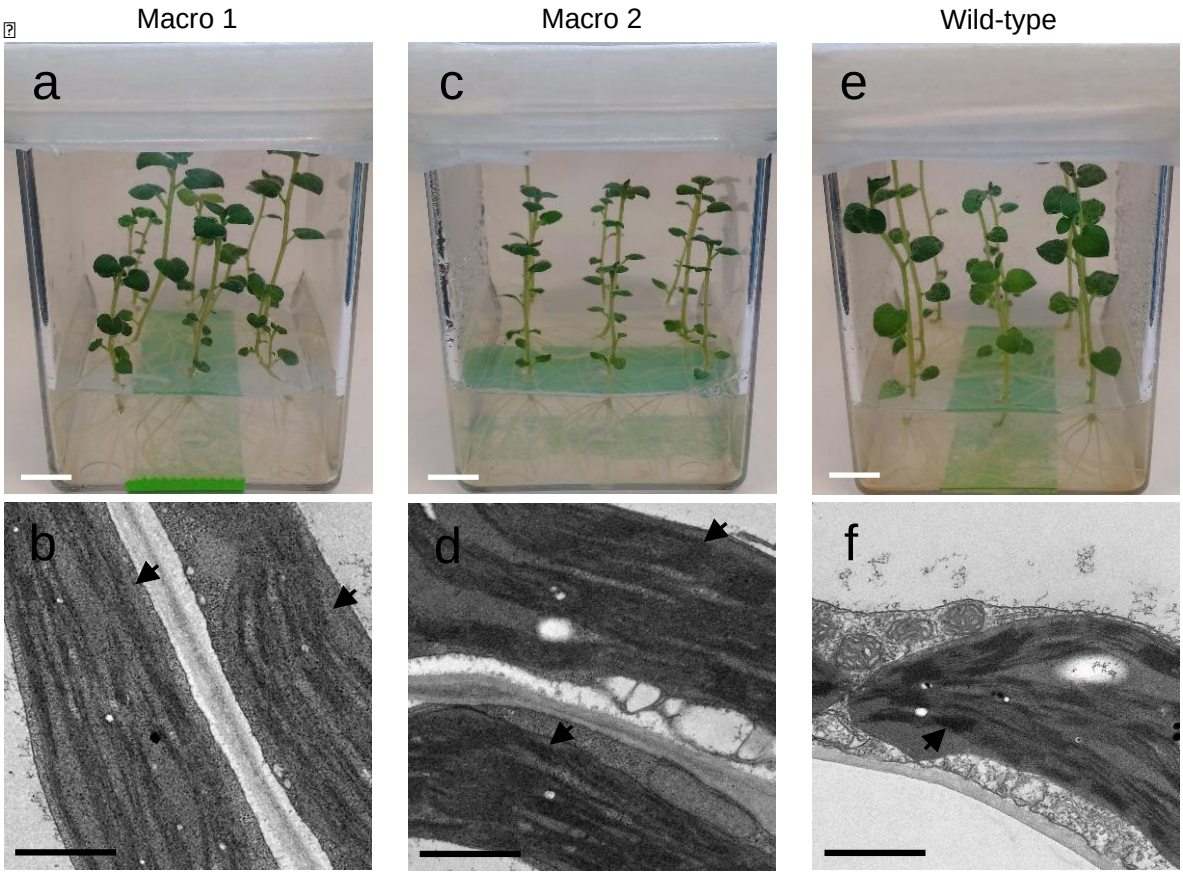
Line	Mean (µm)	SD	N. measurements
a (Macro 1)	17.27 (a)	11.9	720
b (Macro 2)	16.81 (a)	9.61	786
c	11.17 (b)	5.43	1123
d	12.83 (c)	7.96	1649
e	14.82 (d)	9.06	843
f	14.38 (d)	8.66	938
g (Wild-type)	6.73 (e)	2.06	1382

Supplementary Figure 2.1. Confocal images showing chloroplast morphology in protoplasts from *AtFtsZ1* potato lines. Confocal images (including insert images) showing chloroplasts into leaf protoplasts of six independent *AtFtsZ1* lines (a-f: Macro line 1-6, respectively), along with protoplasts from wild-type potato leaves (g). The six *AtFtsZ1* lines show enlarged chloroplast size (µm) comparing to wild-type potato plants (h). Results are expressed as mean ± SD (standard deviation) of the indicated number of measurements (N). Means were compared using ANOVA and when significant, mean separations were analyzed using Tukey HSD ($p<0.05$). Statistical significance is indicated by letters (a, b, c, d or e). Macro 1 and 2 (a and b, respectively) have been chosen for the following experiments. Confocal images

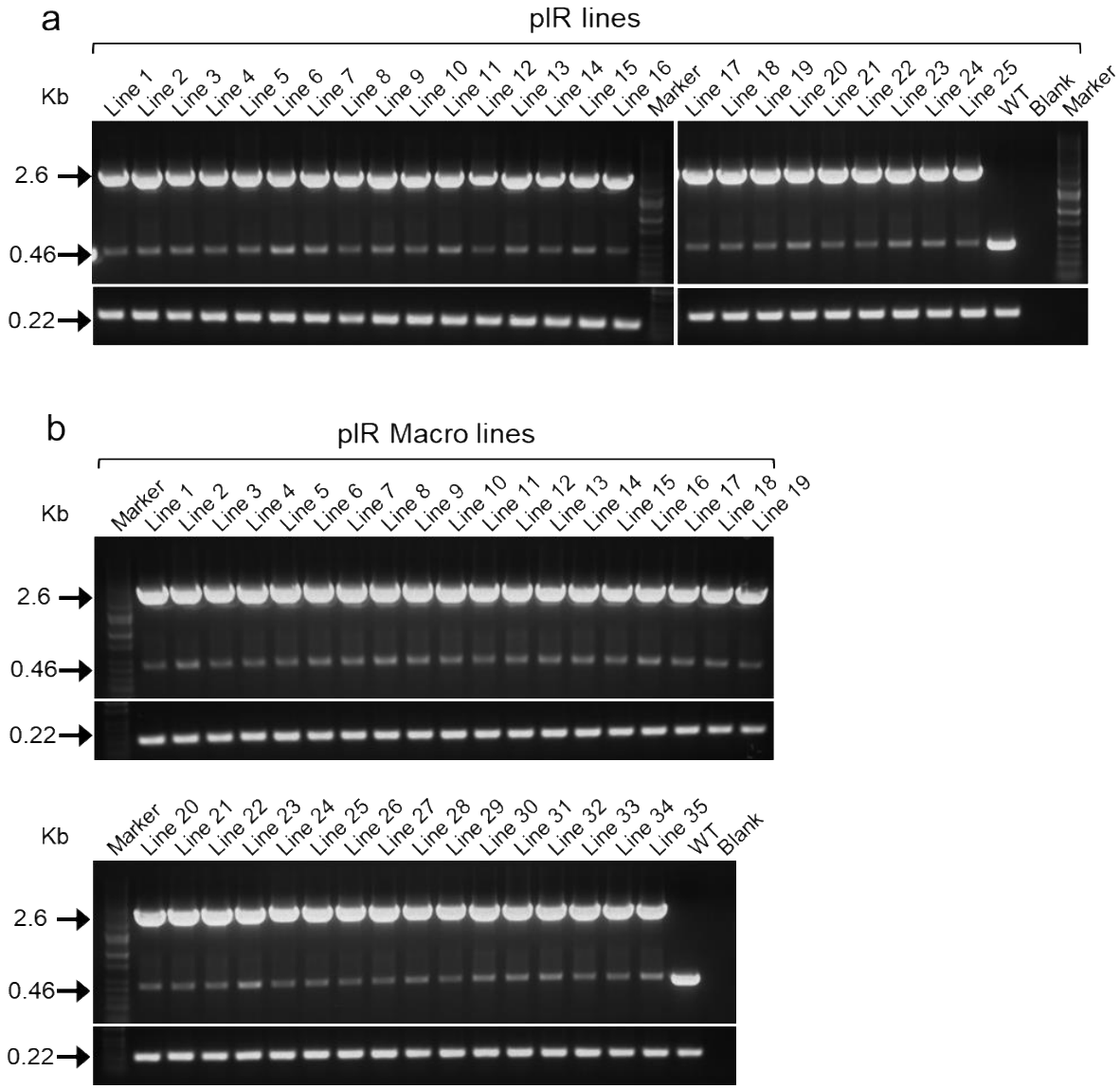
are merge images of chlorophyll signal (red) and bright-field (gray). Scale bars: 50 µm; insert size: 50 µm.



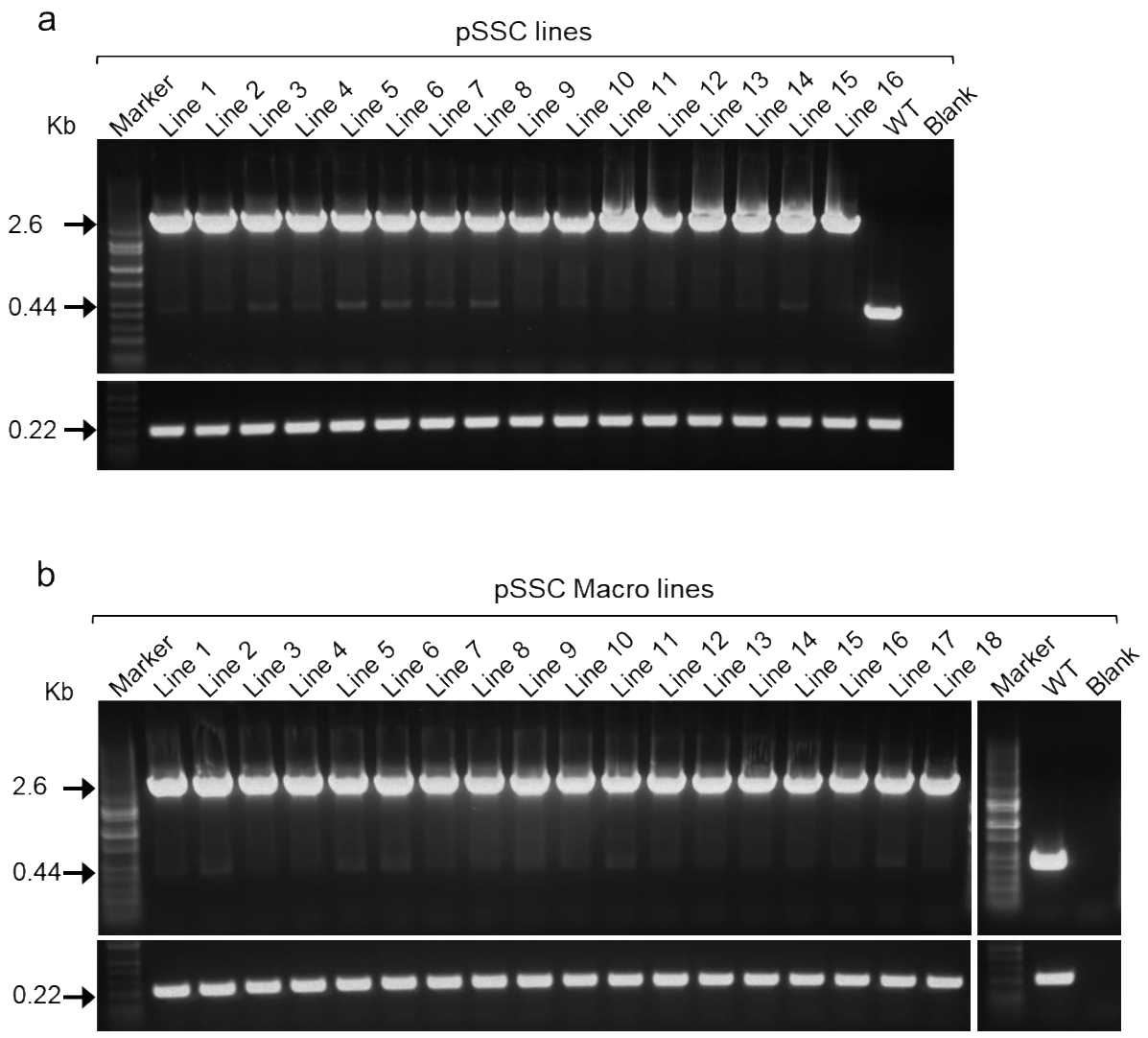
Supplementary Figure 2.2. Confocal images showing chloroplast morphology in leaf from *AtFtsZ1* potato lines. High magnification stack images taken using the confocal microscope of leaf mesophyll cells showing chloroplast morphology in *AtFtsZ1* over-expressing lines (a-c: Macro line 1; d-f: Macro line 2) and wild-type controls (g-i). Chlorophyll (red: a, d and g), bright-field (gray: b, e and h) and merge images (c, f and i) are shown. Scale bars: 50 µm (a-c; d-f; g-i).



Supplementary Figure 2.3. Chloroplast morphology in *AtFtsZ1* lines and wild-type potato. Three-week-old *AtFtsZ1* lines (Macro line 1 and 2) and wild-type control plants along with the ultra-structures of chloroplasts are shown (a-b, c-d and e-f, respectively). Thylakoids from macro-chloroplasts (b, d) and from wild-type controls (f) are indicated with black arrows. Electron micrographs of ultrathin sections showing mesophyll cells from leaf tissue prepared by chemical fixation. Scale bars: 10 mm (a, c, e); Scale bars = 1 µm (b, d, f).



Supplementary Figure 2.4. PCR screening of the second round of pIR transplastomic lines. Integration of the pIR vector in the plastome of normal and macro-chloroplasts lines, respectively (a and b). Pairs of primers specific for the *trnI/trnA* site were used to check integration in 25 normal pIR and 35 macro-chloroplasts pIR lines, respectively (a and b). DNA bands of 2.6 kb indicate correct integration, whereas 0.46 kb-bands indicate the presence of wild-type *trnI/trnA* site. PCR reactions specific for the *rbcL* gene (0.22 kb) were used as loading controls. Wild-type samples, blanks and molecular-weight markers are also shown in the gels.



Supplementary Figure 2.4. PCR screening of the second round of pIR transplastomic lines. Integration of the pIR vector in the plastome of normal and macro-chloroplasts lines, respectively (a and b). Pairs of primers specific for the *trnI/trnA* site were used to check integration in 25 normal pIR and 35 macro-chloroplasts pIR lines, respectively (a and b). DNA bands of 2.6 kb indicate correct integration, whereas 0.46 kb-bands indicate the presence of wild-type *trnI/trnA* site. PCR reactions specific for the *rbcL* gene (0.22 kb) were used as loading controls. Wild-type samples, blanks and molecular-weight markers are also shown in the gels.

Chapter III
Mini-synplastomes for plastid genetic engineering

A version of this chapter will be published by Alessandro Occhialini, Alexander C. Pfotenhauer, Li Li, Stacee A. Harbison, Andrew J Lail, Jason N. Burris, Cristiano Piasecki, Agnieszka A. Piatek, Henry Daniell, C. Neal Stewart, and Scott C. Lenaghan. At the time of writing this dissertation the article was under review at *Plant Biotechnology Journal*. A.O., A.A.P., C.N.S., and S.C.L. designed the strategy; A.O., A.C.P., L.L., S.A.H., A.J.L., J.N.B., and C.P. collected data; A.O. and A.C.P. analyzed data; H.D. designed the selection cassette and advised on chloroplast transformation; A.O., A.C.P., A.A.P., H.D., C.N.S., and S.C.L. wrote the article.

Abstract

In the age of synthetic biology, plastid engineering requires a nimble platform to introduce novel synthetic circuits in plants. While effective for integrating relatively small constructs into the plastome, plastid engineering via homologous recombination of transgenes is over thirty-years-old. Here we show the design-build-test of a novel synthetic genome structure that does not disturb the native plastome: the “mini-synplastome.” The mini-synplastome was inspired by dinoflagellate plastome organization, which is comprised of numerous minicircles residing in the plastid instead of a single organellar genome molecule. The first mini-synplastome in plants was developed *in vitro* to meet the following criteria: 1) episomal replication in plastids; 2) facile cloning; 3) predictable transgene expression in plastids; 4) non-integration of vector sequences into the endogenous plastome and; 5) autonomous persistence in the plant over generations in the absence of exogenous selection pressure. Mini-synplastomes are anticipated to revolutionize chloroplast biotechnology, enable facile marker-free plastid engineering, and provide an unparalleled platform for one-step metabolic engineering in plants.

Introduction

With the emergence of plant synthetic biology, the need to secure the global food supply for a rapidly growing population, and the increased interest in biopharming, chloroplast engineering has once again emerged as a potential solution to meet these demands. For example, in the context of the current SARS-CoV-2 pandemic, angiotensin converting enzyme 2 expressed in lettuce chloroplasts has advanced to the clinic to treat COVID-19 patients, after completion of toxicology and pharmacokinetic studies required for regulatory approval (Daniell *et al.,* 2020). Therapeutic proteins expressed in edible plant cells have already been approved by FDA for both injectable and oral delivery (Fox, 2012; Vickery *et al.,* 2018), paving the way for advancing protein drugs made in chloroplasts. In addition, chloroplast-engineered molecules have been validated for developing oral booster vaccines to prolong immunity against infectious diseases, eliminating the need for cold storage and transportation (Chan *et al.,* 2016; Daniell *et al.,* 2019a; Xiao and Daniell, 2017). Further,

chloroplast biotechnology offers very high production ceilings for heterologous proteins (up to 70% by total leaf protein) (Ruhlman *et al.,* 2010), far surpassing the potential of traditional genetic engineering of the plant nuclear genome. Recent advances in chloroplast biotechnology has enhanced the ability of plastids to express large, complex eukaryotic and viral proteins through the development of new software and algorithms (Kwon *et al.,* 2016; Kwon *et al.,* 2018). For more complex metabolic engineering, the prokaryotic transcription/translation machinery housed in plastids enables coordinated expression of multiple genes by a single promoter, simplifying the control of complex transgene cassettes (De Cosa *et al.,* 2001; Kumar *et al.,* 2012; Malhotra *et al.,* 2016).

A hallmark of chloroplast biotechnology is the use of homologous recombination (HR) to insert transgenes, with a selection cassette, into the native plastome enabling the generation of stable transplastomic lines (Daniell *et al.,* 1998; Kota *et al.,* 1999; Lu *et al.,* 2013; Svab *et al.,* 1990; Svab and Maliga, 1993; Zhang *et al.,* 2015). Marker free transplastomic lines have also been created in crops such as soybean (Dufourmantel *et al.,* 2007) and lettuce (Daniell *et al.,* 2020; Daniell *et al.,* 2019b; Kumari *et al.,* 2019; Park *et al.,* 2020). Currently there is no alternative to HR for chloroplast genome engineering. Episomally-replicating plasmids may provide an alternative engineering platform that leaves the native plastome untouched, with no foreign DNA integrated into native molecules. Indeed, the very first report of expressing foreign genes in chloroplasts demonstrated the value of autonomously-replicating chloroplast vectors containing a fully characterized chloroplast origin of replication D-loops (*ori*) (Daniell *et al.,* 1990; Meeker *et al.,* 1988; Nielsen and Tewari, 1988). Chloroplast vectors containing the *ori* showed prolonged expression in bombarded cells when compared to vectors without the chloroplast *ori* (Daniell *et al.,* 1990). Although prolonged expression of foreign genes was observed in chloroplast vectors containing a chloroplast *ori*, this was a short-term study. This study was followed by subsequent reports of the presence of an 868 bp *Nic*otiana plastid extrachromosomal element, NICE1, as an unexpected product of HR (Staub and Maliga, 1994; Staub and Maliga, 1995). While the mechanism of replication of NICE1-containing plasmids still remains unclear (Mühlbauer *et al.,* 2002; Staub and Maliga, 1994), the NICE1 sequence was devoid of a chloroplast *ori* sequence. Shuttle vectors utilizing NICE1 were developed and proven capable of replication in *Escherichia coli*, and able to generate transplastomic tobacco through expression of *aadA* on the episomal construct (Staub and Maliga, 1994; Staub and Maliga, 1995). Other chloroplast transformation constructs, with promoter and terminator sequences homologous to host plastid DNA, have also been shown to generate episomal plasmids that persist at least to the T3 generation (Min *et al.,* 2015b). As an alternative approach, a previous attempt was made to develop an episomal plastid transformation system in tobacco that was bioinspired by the dinoflagellate plastome organization (Min *et al.,* 2015a). While the angiosperm plastome is composed of a single circular molecule

ranging in size from 107 to 218 kb (Daniell *et al.,* 2016; Shinozaki *et al.,* 1986), in some marine dinoflagellates the plastome is significantly reduced and composed of multiple small minicircles of 2-3 kb (Howe *et al.,* 2003; Howe *et al.,* 2008; Koumandou and Howe, 2007). The core regions of each minicircle contains an *ori* and, in most cases, a coding region comprised of a single gene (Barbrook *et al.,* 2012; Barbrook and Howe, 2000; Koumandou and Howe, 2007). Using a dinoflagellate *ori*, it was possible to engineer transplastomic tobacco with extrachromosomal circular DNA, but after relaxing the selective pressure, the plants rapidly lost the episome. Additionally, the recovery of GFP-positive plants without the extrachromosomal DNA indicated that integration into the native plastome was occurring, preventing achievement of the ultimate goal (Min *et al.,* 2015a).

The objective of this work, was to develop a minicircle architecture, herein termed the mini-synplastome, with the ability to express transgenes for selection and visualization without the need for integration of heterologous sequence into the native plastome. Further, the mini-synplastome must persist for multiple generations without selection. In this way mini-synplastomes differ from spontaneous minicircles in that the architecture is stable and meets the goals for chloroplast biotechnology and synthetic biology. To validate the system, potato was chosen, as this is a relevant food crop that would significantly benefit from plastid engineering with regards to biofortification.

Materials and Methods

Construction of transformation vectors

For construction of Gen1 and pSSC plasmids, synthetic sequences homologous to the IR (*trnI*/*trnA*; 102623-105457 bp and 105458-110067 bp; GenBank: KU199713.1) and SSC (*ndhG*/*ndhI*; 119184-120988 bp and 120989-126029 bp; GenBank: KU199713.1) regions of the tobacco (*Nicotiana tabacum*) plastome were used. These sequences were synthesized by GeneArt (Thermo Fisher Scientific, Waltham, MA, USA) and cloned into the pMK vector. The chloroplast-specific dual selection cassette (*Prrn-SD::aadA::5'UTR::mGFP::3'UTR*) of the PLD-PTD-GFP plasmid (Kwon *et al.,* 2013) was PCR amplified and blunt-cloned into the *Pme*I site of either IR (*trnI/trnA*) or SSC (*ndhG/ndhI*) synthetic sequences generating the Gen1 and pSSC plasmids. The pair of primers 1Fw/Rv were used to amplify the selection cassette for Gen1 and pSSC plasmids. For construction of Gen2, the dual selection cassette of Gen1 was amplified using the pair of primers 2Fw/Rv adding a *Psi*I restriction sites at both 5' and 3' ends. Gen2 was then constructed by cloning the selection cassette into the *Psi*I site located in the backbone of Gen1.

For sequence analysis of integration, plasmids containing ~1.7 kb plastomic sequence located at the *trnI*/*trnA* integration site of wild-type potato and lines containing $_{e}$Gen1, $_{e}$Gen2$^{\Delta}$, $_{e}$Gen2 lines 1-3, were constructed by cloning a PCR

amplicon obtained using genomic DNA and 3Fw/Rv primers into *Kas*I/*Hind*III restriction site of the pUC plasmid (Norrander *et al.,* 1983). These plasmids were then transformed into *E. coli* and isolated for sequencing.

Plant growth *in vitro* conditions

Solanum tuberosum 'Desirée' (potato) plants were grown in sterile conditions in Magenta GA7 boxes containing MS Reg media (4.33 g/l Murashige and Skoog (MS) basal salt mixture; 25 g/l sucrose; 100 mg/l myo-inositol; 170 mg/l sodium phosphate monobasic monohydrate; 440 mg/l calcium chloride dihydrate; 0.9 mg/l thiamine-HCl; 2 mg/l glycine; 0.5 mg/l nicotinic acid; 0.5 mg/l pyridoxine-HCl; 1X MS vitamins; 3 g/l phytagel; pH 5.8). Transplastomic lines were grown in selective MS rooting media (4.33 g/l MS basal salt mixture; 1X Gamborg B5 vitamins; 30 g/l sucrose; 200 mg/l spectinomycin; 3 g/l phytagel; pH 5.8). All reagents for tissue culture were purchased from Phyto Technology Laboratories (Lenexa, KS, USA), except spectinomycin dihydrochloride pentahydrate, which was purchased from Millipore (Billerica, MA, USA). The purity of spectinomycin is critical to ensuring efficient selection of transplastomic lines. Both wild-type and transgenic plants were kept in a controlled environmental room at 16 hours of light and 8 hours of dark. The temperature was kept at 22-24°C during all light/dark cycle. Tissue culture/selection/regeneration steps for generation of transplastomic lines were performed in the same controlled environment.

Production of transplastomic lines

The gene gun PDS-1000/He delivery system (Bio-Rad, Hercules, CA, USA) was used to transform chloroplasts (Occhialini *et al.,* 2019). Transplastomic plants were obtained from transformed leaf material by applying a tissue culture/selection/regeneration protocol as described previously (Valkov *et al.,* 2011). Approximately 6 cm^2 of leaf tissue collected from one month-old potato plants grown in sterile conditions was placed in the center of a Petri dish containing M6M media (4.33 g/l MS basal salt mixture; 1X Gamborg B5 vitamins; 30 g/l sucrose; 18.2 g/l mannitol; 18.2 g/l sorbitol; 0.8 mg/l zeatin riboside (ZR); 2 mg/l 2,4-dichlorophenoxyacetic acid (2,4-D); 3 g/l phytagel; pH 5.8). The tissue was kept overnight in the dark at room temperature before transformation. For transformation, 0.3 mg of 0.6 µm gold-particles were used to bind 1 µg of plasmid following the manufacturer's protocol (Seashell Technology, La Jolla, CA, USA). Tissue was bombarded at 6 cm-distance using 1,100 psi rupture disks. After two days of incubation in the dark at room temperature, leaf material was cut in small pieces (5 mm^2) and placed in selective M6 media (4.33 g/l MS basal salt mixture; 1X Gamborg B5 vitamins; 30 g/l sucrose; 0.8 mg/l zeatin riboside (ZR); 2 mg/l 2,4-dichlorophenoxyacetic acid (2,4-D); 400 mg/l spectinomycin; 3 g/l phytagel; pH 5.8) at the growth condition described previously. After one month-incubation the tissue was transferred to selective Ti media (4.33 g/l MS basal salt mixture; 1X Gamborg B5 vitamins; 16 g/l glucose; 3 mg/l zeatin riboside (ZR); 2 mg/l indole acetic acid (IAA); 1 mg/l gibberellic acid (GA_3); 400 mg/l spectinomycin; 3 g/l phytagel; pH 5.8). 4-8 weeks later, transplastomic green callus was obtained

from transformed leaves. Green callus was transferred to selective DH media (2.16 g/l MS basal salt mixture $NH_4NO_3^-$ free; 268 mg/l NH_4Cl; 1X Nitsch vitamin mixture; 2.5 g/l sucrose; 36.4 g/l mannitol; 100 mg/l casein hydrolysate; 80 mg adenine hemisulfate; 2.5 mg/l zeatin riboside (ZR); 0.1 mg/l indole acetic acid (IAA); 400 mg/l spectinomycin; 3 g/l phytagel; pH 5.8) for another month of growth, and after that placed on MON media (4.33 g/l MS basal salt mixture; 1x Gamborg B5 vitamins; 30 g/l sucrose; 0.1 mg/l naphthaleneacetic acid (NAA); 5 mg/l zeatin riboside (ZR); 400 mg/l spectinomycin; 3 g/l phytagel; pH 5.8) for regeneration of shoots. Primary transplastomic shoots were transferred in Magenta boxes containing selective MS rooting media for roots regeneration. For the second and third round of transplastomic plants the same protocol of selection/regeneration was performed.

Plant phenotypic analysis

To synchronize plant growth, apical shoots were collected from plants of the same age and then transferred in Magenta boxes containing MS Reg media without selection. After 2 weeks of growth *in vitro*, wild-type and transgenic plantlets with roots were transferred in 3.8 l pots containing soil Pro-Mix BK25 (Griffin Greenhouse Supplies, Inc. Tewksbury, MA, USA) and kept growing for 10 weeks until anthesis in a controlled environment (16 and 8 hours of light/dark cycle at the temperature of 22-24°C). At this developmental stage, plant height (cm), number of nodes, foliar and total plant fresh and dry weight (g) were collected. Before collecting the dry weight data, plant tissue was dried for 1 week at 50°C. ImageJ 1.41o software (National Institute of Health, Bethesda, MD, USA) was used to calculate the leaf area by image analysis of leaf scans (Occhialini *et al.,* 2016).

The LI-6800 portable photosynthesis system (LI–COR Biosciences, Lincoln, NE, USA) was used to determine the CO_2 assimilation (A) values per unit of leaf area (µmol m^{-2} s^{-1}) of both wild-type and transplastomic plants. The leaf gas exchange measurements were performed at a temperature of 23°C, 400 µmol mol air^{-1} of CO_2, 1000 µmol photons m^{-2} s^{-1} of irradiance 0.8-1.2 kPa of vapor-pressure deficit (VPD leaf) and a 200 µmol s^{-1} of flow rate (Occhialini *et al.,* 2020). The portable CCM-200 plus chlorophyll content meter (OPTI-SCIENCES Inc., Hudson, NH, USA) was used to determine the leaf chlorophyll content index (CCI). The IBM SPSS software was used to perform statistical analysis and means were compared using ANOVAs with post-hoc Tukey ($p<0.05$).

Total DNA extraction and PCR analysis

For extraction of total genomic DNA from leaves a CTAB-based procedure was used (Occhialini *et al.,* 2020). Approximately 50 mg of leaf tissue frozen in liquid nitrogen was finely ground in an Eppendorf tube. The macerated leaf material was resuspended in 500 µl of CTAB extraction buffer (2% hexadecyltrimethyl ammonium bromide; 1% (w/v) polyvinyl pyrrolidone; 100 mM Tris-HCl; 1.4 M NaCl; 20 mM EDTA; 0.1 mg/ml RNaseA), thoroughly vortexed and incubated for

10 minutes at room temperature. Samples were further incubated for 30 minutes at 60°C, and then the cellular debris was eliminated by centrifugation at 15,000 × g for 5 minutes. An equal volume of a solution containing chloroform/isoamyl alcohol (24:1) was added to the clarified supernatant. The sample was vortexed for 5 seconds and centrifuged at 4°C for 1 minute at 15,000 × g. The upper aqueous phase was transferred in a new tube, and the DNA was precipitated by adding an equal volume of ice-cold dry isopropanol. The samples were incubated for 30 minutes on ice and then centrifuged at 4°C for 30 minutes at 15,000 × g. The DNA precipitated in the tube was washed in 500 µl of ice-cold 75% (v/v) ethanol. The air-dried pellet was resuspended in 50-100 µl of sterile H_2O and quantified using a Nanodrop spectrophotometer.

Two pairs of primers 4Fw/Rv and 5Fw/Rv were used to check the integration in the IR (*trnI/trnA*) and SSC (*ndhG/ndhI*) sites of the plastome. The two pair of primers 6Fw/Rv and 7Fw/Rv were used to check for the presence of full-length s*mGFP* (encoding soluble monomeric green fluorescent protein, NCBI ID: AEX93343.1) and *aadA* (encoding the streptomycin 3''-adenylyltransferase, NCBI ID: AAR14532.1) genes, respectively. The pairs of primers 8Fw/Rv and 9Fw/Rv were used to check the presence of *KanR* (encoding aminoglycoside 3'-phosphotransferase, GenBank: APB62235.1) and *SpcR* (spectinomycin adenyltransferase, GenBank: AAA72848.1) selective genes of the backbone, respectively. The primers 10Fw/Rv were used to amplify an internal fragment of *rbcL* (*Solanum tuberosum* ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit, NCBI ID: 4099985) used as loading control.

E. coli back-transformation with episomal vectors extracted from leaf tissue

A total volume of 25 µl of chemically competent *E. coli* TOP10 (Thermo Fisher Scientific, Waltham, MA, USA) were transformed using 500 ng of genomic DNA from episomal lines using the heat-shock method. Transformed cells were grown in Luria-Bertani (LB) agar media (10 g/l of bacto-tryptone; 5 g/l of yeast extract; 10 g/l NaCl; 15 g/l bacto agar; pH 7) containing 50 µg/ml kanamycin. Pure preparations of episomal DNA ($_e$Gen1, $_e$Gen2 and $_e$Gen2$^{\Delta}$) were extracted from bacterial cells using QIAprep Spin Miniprep Kit (QIAGEN, Valencia, CA, USA). The presence of left and right homologous arms, along with the internal cassette, were tested by PCR using the pair of primers 11Fw/Rv, 12Fw/Rv and 13Fw/Rv, respectively. The nucleotide sequence of extra-plastomic DNA was determined by Sanger DNA sequencing (Massachusetts General Hospital MGH, Center for Computational & Integrative Biology CCIB, DNA Core, Boston, MA, USA).

Total RNA extraction and qRT- PCR

Total RNA preparations were obtained from leaf tissue collected from both transgenic lines and wild-type controls using the Tri-Reagent (Molecular Research Center, Inc, Cincinnati, OH, USA) according to manufacturer's protocol. For each RNA extraction ~50 mg of fresh leaf tissue was used.

Thereafter, RNA preparations were subjected to DNase treatment followed by cleaning using the RNA Clean & Concentrator Kit (Zymogen, Irvine, CA, USA) according to manufacturer's instruction. The Super Script III Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA, USA) was used to synthesize the first cDNA strand following the manufacturer's instruction. PCRs using several cDNA dilutions were performed to test the presence of full-length sequence of the two transgenes of the cassette, *aadA* (aminoglycoside-3″-adenylyltransferase, GenBank: ARK38551.1) and s*mGFP* (soluble monomeric green fluorescent protein, GenBank: AEX93343.1). The *rbcL* (Solanum tuberosum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit, NCBI ID: 4099985) and *ef1α* (Solanum tuberosum elongation factor 1-alpha; XM_006343390.2) genes were used as internal reference for plastome and nuclear genome, respectively.

qPCR for copy number determination

qPCR was performed in a total volume of 15 µl in 1X PowerUp™ SYBR™ Green Master Mix (Thermo Fisher Scientific, Waltham, MA, USA), using 5 ng of pure genomic DNA and 0.5 µM of both forward and reverse primers. The pairs of primers 16Fw/Rv and 17Fw/Rv were used to detect the backbone gene *KanR* of all episomal plasmids (encoding aminoglycoside 3'-phosphotransferase, GenBank: APB62235.1) and a unique sequence of $_e$Gen2 backbone, respectively. The plastome internal control *rbcL* (*Solanum tuberosum* plastome, NCBI Gene ID: 4099985) and the nuclear gene *actin PoAc58* (X55749.1) were detected using the pair of primers 14Fw/Rv and 18Fw/Rv. For episomal vs plastome copy number determination 0.1, 1, 20, 40, 60, 80, 100, 120, 140, 160, 180 and 250 pg of purified episomal plasmid (either $_e$Gen1 or $_e$Gen2) were used as standards of copy number. Per each standard, the DNA copy number was calculated using the equation: (ng DNA × 6.022×10^{23}) / (DNA length bp × 1×10^{9} × 650). Linear regression graphs plotting Ct values of DNA standards (Y axis) vs log10 of copy number of standards (X axis) were used to calculate the copy number of both episomal plasmid and plastome in genomic DNA samples. Wild-type plants and blanks were used as negative controls. Microsoft Excel software was used to process data and for their graphical representation.

qRT-PCR was also performed in a total volume of 15 µl in 1X PowerUp™ SYBR™ Green Master Mix (Thermo Fisher Scientific, Waltham, MA, USA), using cDNA diluted 1 to 3 and 0.5 µM of each primer. The pair of primers 19Fw/Rv and 20Fw/Rv were used to amplify the *aadA* (aminoglycoside-3″-adenylyltransferase, GenBank: ARK38551.1) and s*mGFP* (soluble monomeric green fluorescent protein, GenBank: AEX93343.1) gene of the selection cassette, respectively, whereas 15Fw/Rv were used to amplify the internal reference gene *ef1α* (*S. tuberosum* elongation factor 1-alpha; XM_006343390.2). The relative gene expression data were represented using the $2^{-\Delta CT}$ method.

qPCR and qRT-PCR primers were designed to amplify a fragment of ~100 bp at a compatible annealing temperature of ~57ºC using the online software Primer3 input v. 0.4.0 (Howard Hughes Medical Institute and by the National Institutes of Health) (Koressaar and Remm, 2007; Untergasser *et al.,* 2012). qPCR and qRT-PCR were performed using a QuantStudio™ 6 Flex Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA), whereas data were acquired using the QuantStudio™ Real-Time PCR Software v1.1 (Thermo Fisher Scientific, Waltham, MA, USA). The data are expressed as mean ± standard deviation (sd) of the indicated biological and technical replicates. The IBM SPSS software was used to perform statistical analysis and means separation was evaluated using ANOVA Tukey HSD ($p<0.05$).

Southern blot analysis

The inverted repeat region (IR) of the chloroplast genome of potato (*Solanum tuberosum*; GenBank: NC_008096.2; from 104457 to 104978 bp) was used to design a DNA probe for detection of Gen1 vector integration. The *KanR* gene sequence from the vector backbone pMK GeneArt (Thermo Fisher Scientific, Waltham, MA, USA) was used to design the probe to detect extra-plastomic DNAs ($_e$Gen1 or $_e$Gen2). The *aadA* gene (GenBank: ARK38551.1) was used to design to probe to detect the selection cassette. The IR, *KanR* and *aadA* DNA-probes labelled with digoxigenin-(DIG)-sUTP were synthesized using the PCR DIG Probe Synthesis Kit (Roche, Indianapolis, IN, USA) and the pair of primers 21Fw/Rv, 22Fw/Rv and 23Fw/Rv, respectively. 1 µg of total genomic DNA from leaf tissue was digested using *Kas*I/*Hind*III or *Fsp*I/*Fse*I restriction enzymes, for detection of IR and *aadA* or *KanR* fragments, respectively. DNA samples were separated on 0.9% agarose gel, depurinated, denatured and then transferred on Hybond-N+ nylon membrane (GE Healthcare, Life Sciences, Marlborough, MA, USA) (Lin *et al.,* 2014). The anti-digoxigenin-AP Fab fragments detection kit (Roche, Indianapolis, IN, USA) was used for detection of the DIG-labelled probe signal.

Confocal microscopy

Healthy leaves from 3-4-week-old wild-type potato plants and transplastomic lines grown in sterile conditions were imaged using an Olympus Fv1000 confocal microscope (Olympus, Center Valley, PA, USA) equipped with both traditional (argon and HeNe) and diodes (405, 440, 473, 559 and 635-nm) lasers. The green fluorescent protein (smGFP) was excited at 488 nm and detected at 509 nm of wavelength emission. The chlorophyll auto-fluorescence was excited using at 543 nm and detected at 667 nm of wavelength emission. Digital images were acquired using Olympus FV10-ASW Viewer software Ver.4.2a (Olympus, Center Valley, PA, USA). Confocal images were processed using ImageJ 1.41o (National Institute of Health, Bethesda, MD, USA).

Results and discussion

Construction of Gen1 vectors with long-homologous synthetic arms as a first step in the design-build-test-cycle of the mini-synplastome

To understand the necessary architecture for development of a mini-synplastome for potato plastid engineering, a plastid engineering vector was designed with homology to the most common transgene integration site of tobacco, *trnI*/*trnA* (Daniell *et al.,* 2016). In order to encompass the most likely components for episomal plasmid replication from tobacco (NICE1 and *ori A2* and *A1*) (Krishnan and Rao, 2009; Staub and Maliga, 1994), long homology arms (~2.8 and ~4.6 kb) were utilized (Figure 3.1a). The use of tobacco homology arms for engineering potato plastids was intentional with the goal to prevent direct sequence homology to enable tracking of recombination. It is known that intergenic spacer regions among *Solanaceae* chloroplast genomes are not highly conserved (Daniell *et al.,* 2006). It should also be noted that the homologous NICE1 sequence in potato contains an insertion of 102 bp not found in the NICE1 sequence from tobacco, as indicated in Figure 3.1A (position 14, Figure 3.1a;). Further, a series of single nucleotide polymorphisms (SNPs) and a multi-cloning site (MCS) were introduced into the tobacco homology arms to provide more markers to identify common sites of recombination and enable alternative cloning strategies (Figure 3.1a). In total, the synthetic generation 1 (Gen1) vector contained ~7.4 kb of native tobacco plastid sequence flanking the *trnI*/*trnA* integration site with 35 separate SNP and indel mutations (Figure 3.1a). For selection of engineered lines, a common dual selection (*aadA* and *smGFP*) cassette (Kwon *et al.,* 2013) was cloned into the *trnI/trnA* site in Gen1. As a control for HR-mediated chloroplast biotechnology, a vector for potato plastid transformation with homology arms of ~1.8 and ~5 kb targeting the *ndhG/ndhI* integration site within the small single copy region (SSC), and the same dual selection cassette, was used (Occhialini *et al.,* 2020). It should be noted that the Gen1 construct was designed to recombine with the native plastome along the ~7.4 kb homology arms, and integrate the transgene cassette into the commonly used *trnI*/*trnA* integration site on the native plastome. However, we hypothesized that persistence of episomal DNA would be possible due to the selective pressure within the compartment (plastid). This was anticipated to reach a steady state in the plastid if there was constant recombination back-and-forth between an episomal plasmid and the native plastome.

Screening for transgenic lines harboring plastid-replicating episomal units

To test our hypothesis, the Gen1 construct was transformed into potato leaf discs using biolistics and selected on spectinomycin. Selection yielded numerous transplastomic shoots. Molecular analysis confirmed that with the Gen1 construct 67% of the transplastomic shoots had the transgene cassette integrated into the *trnI*/*trnA* integration site of the native plastome (Figure 3.1b-c). Of the remaining shoots, the *KanR* gene could be amplified by PCR in 2 of the 5 Gen1-containing plants hinting at the presence of an episomally-replicating plasmid

(Supplementary Figure 3.1). The 3 remaining lines were hypothesized to result from recombination into an undetected site or represented escapes from selection. For the control vector, pSSC, the transgene cassette was correctly integrated into the *ndhG/ndhI* integration site in all transplastomic lines with no episomal plasmids detected (Figure 3.1b-c; Supplementary Figure 3.1). In all transplastomic lines obtained, plants grew vigorously *in vitro* and high GFP fluorescence was observed in chloroplasts (Supplementary Figure 3.2).
To confirm that Gen1-containing plants with undetectable transgene integration and amplification of the *KanR* gene, contained the episomally-replicating Gen1 backbone, total leaf genomic DNA was analyzed by Southern blot and PCR (Figure 3.2 and Supplementary Figure 3.3, respectively). Southern blots performed using two different probes (IR and *aadA*) confirmed no detectable transgene integration into the *trnI/trnA* site of the native plastome in Gen1 episome-containing plants, as compared to plants with the transgenes from Gen1 integrated into the native plastome (~3.8 kb bands in both *IR* and *aadA* DNA blots). Southern blots using a *KanR* probe similarly displayed the full-length backbone in the Gen1 episome-containing plants (Figure 3.2a-b).

Characterization of the first generation of episomal units ($_e$Gen1) extracted-back from leaf tissue

Episomes contained in chloroplasts of selected transgenic lines were isolated by back-transformation in *E. coli*. Unsurprisingly, after bacteria transformation using total genomic DNA isolated from plants containing the episomal Gen1 backbone, colonies were recovered that contained plasmids (Figure 3.2c), indicating that the Gen1 episomal backbone persisted in plastids as replicating extraplastomic DNA. PCR and sequence analysis of entire plasmids extracted from *E. coli* demonstrated the presence of the Gen1 backbone along with full-length homologous arms devoid of the transgene cassettes (Supplementary Figure 3.4). Since earlier molecular analysis (Figure 3.2; Supplementary Figure 3.3) confirmed that the transgene cassette was not present in the expected integration site, it was hypothesized that the transgene cassette recombined in an undetermined location, which is not uncommon in chloroplast engineering (Gray *et al.,* 2009) (Figure 3.2a). The presence of identical profiles of multiple DNA bands at higher molecular weight (>3.8 kb, with a main ~6 kb band) obtained with both *IR* and *aadA* probes further support this hypothesis (Figure 3.2b).

Using the sequence variability described earlier, it was possible to track recombination events within the homology arms of Gen 1 by comparing the sequence of the initial Gen1 construct with the sequence of the plasmid isolated from whole plant genomic DNA of stable transplastomics ($_e$Gen1). Sequence analysis of $_e$Gen1 demonstrated numerous recombination events between the homology arms and the endogenous plastome. The homologous arms of $_e$Gen1 recombined with the native potato plastome replacing 17 of the 35 mutations present in the bombarded version of the vector (M2, M10, M12-14, M18-29).

Based on the pattern of mutations that were converted from the synthetic version of the homology arms to potato-specific sequences it was hypothesized that there were 4 recombination events. Unlike standard chloroplast biotechnology, in this instance we were concerned with HR of native plastome sequence into the episomal construct, as opposed to HR from a vector into and between the native plastome. In previous work using transformation vectors with homology arms from distinct species, 100% correction of the heterologous sequences was observed (Ruhlman *et al.,* 2010).

The $_e$Gen1 episome is maintained as extra-plastomic DNA throughout multiple plant developmental stages

Despite only ~50% conversion of the homology arms of $_e$Gen1 to potato-specific sequence, $_e$Gen1 was the only plasmid recovered during all 3 rounds of tissue culture, and isolated from plants grown on potting mix including tubers. Since the sequence of $_e$Gen1 was invariant across multiple generations, we hypothesized that the sequence of $_e$Gen1 was imposing limits to recombination across the other heterologous sites. Further, no mutations were detected in the backbone of the episomal plasmid, suggesting that the backbone was devoid of significant homology with the native plastome. The lack of mutations within the backbone was crucial to understanding how to better design episomally replicating vectors without integration of foreign DNA. Surprisingly, the average copy number of $_e$Gen1 was ~1 copy per endogenous plastome in plants maintained in tissue culture (Figure 3.2d) and persisted in 3 successive tissue culture generations (Supplementary Figure 3.5). In addition, $_e$Gen1 could be isolated from all life cycle stages through anthesis on plants grown in pots without selective pressure (Figure 3.2d; Supplementary Figure 3.6). Over time, the average copy number of $_e$Gen1 per plastome shifted from 2.04 ± 0.56 at 2 weeks to 3.91 ± 1.44 at 4 weeks in pots before stabilizing at ~2.5 copies per plastome throughout the remaining life cycle of the plant (Figure 3.2d). In fact, in plants grown from tubers harvested from $_e$Gen1-containing plants, $_e$Gen1 could still be recovered with no changes to the sequence (Supplementary Figure 3.7). While plants were not sexually reproduced, it is anticipated, based on earlier work with NICE1 in tobacco, that $_e$Gen1 would be present at minimal to no copy number in potato seed (Staub and Maliga, 1994). While researchers have demonstrated that highly inefficient paternal transmission of plastid DNA is possible (Ruf *et al.,* 2007; Svab and Maliga, 2007), it is generally accepted that plastid transformation is an effective method of bioconfinement of GM crops. Based on these findings, we speculate that mini-synplastome technology would provide a similar high degree of bioconfinement.

Design of the mini-synplastome (Gen2) for plastid genetic engineering

As discussed previously, sequence analysis of $_e$Gen1 recovered after subsequent rounds of tissue culture (1-3) and at different plant developmental stages in potted plants indicated no variability in sequence, confirming a stabilized episomal architecture. Thus, the plants from this first design-build-test

cycle taught us how to design the mini-synplastome. However, to achieve the ultimate goal of chloroplast engineering without the integration of foreign DNA, we modified $_e$Gen1 to express transgenes directly from the episomal construct by inserting the dual selection cassette flanked by the vector backbone instead of the homology arms. This modification led to the design and assembly of the Gen2 version of the construct (Figure 3.3). It should be noted that the sequence for the ~7.4 kb homology arms flanking the *trnI/trnA* integration site present in Gen1 remained in Gen2, although the transgene cassettes were no longer inserted into the integration site, as described above (Figure 3.3a). In shuttle vectors previously engineered using NICE1 (Staub and Maliga, 1994) and multiple dinoflagellate *ori* (Min *et al.,* 2015a), transgene integration with the endogenous plastome was observed, thus it was necessary to eliminate these recombination events to meet the requirement of the mini-synplastome engineering platform. After assembly, Gen2 was transformed into potato leaf discs by biolistics and selected against spectinomycin. In this round, however, the spectinomycin resistance gene, *aadA*, was expressed from the episomal construct, not from a transformed plastome. The spectinomycin resistant shoots harbored no cassette integration into the plastome, as determined by Southern blot analysis; however, plants were positive for Gen2 (Figure 3.3b; Supplementary Figure 3.8).

Synplastomic lines harbor the $_e$Gen2 plasmid as non-integrating episomal units with intact transgene backbone structure

The back-transformation to *E. coli* of total leaf genomic DNA preparations from Gen2 lines followed by sequence analysis of the entire plasmid recovered ($_e$Gen2) confirmed the presence of an episomally replicating construct (Supplementary Figure 3.9). Compared with $_e$Gen1, the homologous arms of $_e$Gen2 were comprised entirely of potato-specific sequences with the exception of M1 and M35. Since M1 and M35 were at the end of the homology arms, the lack of correction of these mutations was expected since the flanking sequence on one side of the mutations were 98 bp and 7 bp, respectively. In fact, M1 and M35 were the only mutations that were not corrected in both $_e$Gen1 and $_e$Gen2. The correction of all other mutations in $_e$Gen2, as opposed to $_e$Gen1, indicated that insertion of the transgene cassettes in between the homology arms of $_e$Gen1 causes disruption of HR leading to the lack of near complete correction. Whole plasmids recovered from *E. coli* revealed two distinct versions of $_e$Gen2, full-length $_e$Gen2 and $_e$Gen2$^{\Delta}$ (Figure 3.3a). Full-length $_e$Gen2 had complete homology arms, whereas $_e$Gen2$^{\Delta}$ had a recombination event between the *Prrn* promoter sequence on the homology arm with the *Prrn* promoter on the transgene cassette, eliminating ~7.9 kb of sequence (Figure 3.3a). Sequence similarity between the *Prrn* promoter on the homology arm and transgene cassette were known; however, the homologous flanking regions at the reconstitution sites were only 168 and 33 bp with the total recombination event reconstituting a mere 220 bp of sequence (Figure 3.3a). Conventional tobacco chloroplast engineering typically uses homologous arms of ~1 kb (much longer

for lettuce) (Daniell *et al.,* 2019b; Kumari *et al.,* 2019; Kwon *et al.,* 2016; Kwon *et al.,* 2018; Meeker *et al.,* 1988; Park *et al.,* 2020; Ruhlman *et al.,* 2010) although the use of smaller direct repeats of up to 174 bp in length has been used for direct-repeat mediated excision of marker genes (Iamtham and Day, 2000). In terms of efficiency, even with the goal of introduction of a single SNP into the native plastome, homology arms of ~ 600 bp are recommended (Martin Avila *et al.,* 2016). Similarly, as indicated in the generation of NICE1, imperfect direct repeats as small as 16 bp can facilitate effective recombination in plastids (Staub and Maliga, 1994). Evolutionary studies of plastid genomes have revealed similar small inversions indicating some mechanism for recombination of 10-20 bp direct or imperfect direct repeats (Day and Madesis, 2007). Thus, while not unprecedented, the efficiency of this within episome recombination was unexpected.

Southern blots performed on $_e$Gen2-containing plants displayed only the presence of the wild-type *trnI/trnA* fragment (~1.6 kb bands, Figure 3.3b) confirming that there was no integration into the native plastome at this site. Sequence analysis performed on the sequence flanking the integration site revealed a truncated *trnI/trnA* fragment resulting from a ~100 bp deletion corresponding to M14 (Figure 3.1a) that was present in the endogenous plastome in a single $_e$Gen2$^{\Delta}$ -containing plant. As shown in Figure 3.2b, in a control $_e$Gen1-containing plant, a higher molecular weight band (~ 6 kb) associated with the wild-type *trnI/trnA* fragment was observed (Figure 3.3b). As described previously, this fragment was likely the result of an unpredicted homologous recombination event between the transgene cassette and plastome while the backbone of $_e$Gen1 was still episomally maintained (Figure 3.2). A plant with integration of the transgene cassette from Gen1, along with wild-type plants, were used as positive (~3.8 kb band) and negative (~1.6 kb band) controls for cassette integration into the *trnI/trnA* site of the native plastome (Figure 3.3b). Southern blots probed with *KanR* DNA revealed the presence of the full-length backbone in $_e$Gen2-containing plant lines 1-3 (~4.3 kb Figure 3.3b). In $_e$Gen2$^{\Delta}$-containing plant lines and $_e$Gen2-containing line 1 a high molecular weight band (≥ 10 kb) was observed, which indicated remaining uncut $_e$Gen2 (Figure 3.3b). In total, these data confirmed that $_e$Gen2 persisted as a non-integrating episomal plasmid, indicating that the transgenes were effectively expressed from the episomes. Therefore, we concluded that $_e$Gen2 was the first mini-synplastome vector to be developed.

The $_e$Gen2 episome confers efficient transgene expression and it is present at high copy number throughout all plant developmental stages

RT-qPCR performed using cDNA preparations from mini-synplastomic lines indicated that $_e$Gen2 conferred high *smGFP* and *aadA* expression in chloroplasts ($_e$Gen2-containing lines 1-3 and $_e$Gen2$^{\Delta}$-containing lines) (Figure 3.3c and Supplementary Figure 3.10). However, compared to a homoplasmic plant line with integration of the Gen1 transgene cassette, *smGFP* expression was reduced

by ~ 90, 50 and 75% in the $_e$Gen2-containing lines 1-3. For *aadA* expression a reduction of ~85% to ~60% was observed in the same $_e$Gen2-containing lines 1-3. It should be noted that *smGFP* and *aadA* expression were under the control of the same promoter, but different 5'UTRs, leading to higher expression of *smGFP* relative to *aadA* in all lines. The level of transgene expression for $_e$Gen2$^{\Delta}$-containing lines were not significantly different from the control line with integration of the Gen1 transgene cassette. The Gen1 control line represents the typical HR strategy for chloroplast engineering and the expression of transgenes from $_e$Gen2$^{\Delta}$-containing lines achieved comparable expression to the current state-of-the-art. Surprisingly, comparison of *smGFP* and *aadA* expression in a representative $_e$Gen1-containing line was 120 and 250% higher when compared to a plant line with integration of the Gen1 transgene cassette, but without persistent of the episome (Figure 3.3c). The extremely high-level expression of the $_e$Gen1-containing line indicated that, for the purposes of heterologous protein production, there may be benefit in using a combined integrating/episomal strategy. The presence of plastid-localized GFP signal was further confirmed through microscopic analysis of both $_e$Gen2- and $_e$Gen2$^{\Delta}$-containing lines (Figure 3.3d).

In $_e$Gen2-containing lines 1-3 and $_e$Gen2$^{\Delta}$-containing lines, the episomes persisted; however, the copy number and ratio of $_e$Gen2 to $_e$Gen2$^{\Delta}$ varied by line. In $_e$Gen2-containing line 2 it was possible to recover both $_e$Gen2$^{\Delta}$ and $_e$Gen2, whereas the other lines consisted exclusively of $_e$Gen2 or $_e$Gen2$^{\Delta}$ (Figure 3.3e, graph all forms vs $_e$Gen2 only; Supplementary Figure 3.11). Further, the lines containing only $_e$Gen2 ($_e$Gen2-containing lines 1 and 3) were present at lower average copy number compared to the native plastome, ~0.65 and 0.9 copies per plastome, whereas lines containing only $_e$Gen2$^{\Delta}$ or a mix of $_e$Gen2 and $_e$Gen2$^{\Delta}$ had an average copy number exceeding that of the native plastome, 2.63 ± 1.15 and 1.76 ± 1.13, respectively (Figure 3.3e, graph all forms). In order to test the stability of $_e$Gen2 throughout the complete life cycle of potato, a growth experiment using three independent $_e$Gen2-containing lines was initiated. For comparison, an $_e$Gen1-containing line, an integrating Gen1 line, and wild-type plants were used as controls. For each line, 5 plants were used as biological replicates, as it was anticipated that there may be plant to plant variation due to the constant potential for recombination within the homology arms. Recombination leading to the emergence of $_e$Gen2$^{\Delta}$ from $_e$Gen2 and the persistence of these two forms were investigated at different developmental stages until anthesis (Figure 3.4; Supplementary Figure 3.12). After two weeks in tissue culture (without selection) and 1 week in pots, all analyzed plants carried only $_e$Gen2 with the exception of one plant from line 2 which had $_e$Gen2$^{\Delta}$ (Supplementary Figure 3.12). After 4 weeks on potting mix, all $_e$Gen2-containing line 1 plants and 4 out of 5 $_e$Gen2-containing line 3 plants had undetectable levels of either episomal construct, whereas all $_e$Gen2-containing line 2 plants maintained episomal constructs (two carrying $_e$Gen2 and three only $_e$Gen2$^{\Delta}$). At 7 and 10 weeks (anthesis) all $_e$Gen2-containing line 2 plants were stably harboring

only $_e$Gen2$^{\Delta}$ (Supplementary Figure 3.12). For all $_e$Gen2-containing plants where the episomal construct could no longer be detected by PCR of the vector backbone, the transgenes were also not detected, confirming the complete elimination of recombination with the native plastome and restoration of the wild-type genotype (Supplementary Figure 3.12). For all $_e$Gen2-containing line 2 plants, episomal constructs could be recovered throughout the entire life cycle through anthesis, with an average of ~1-1.5 copies per plastome maintained from 1-10 weeks on potting mix without selection (Figure 3.4a). In $_e$Gen2-containing line 2 plants, the large standard deviations of copy number of the episomal construct reflect the high variability observed in different plants growing in tissue culture and on potting mix (Figure 3.3e and 3.4a).

In order to determine if there was pleiotropic or genetic load effects that were caused by the persistence of the episomal plasmid, the phenotype of $_e$Gen2-containing line 2 plants along with wild-type and transplastomic ($_e$Gen1-containing and integrating Gen1) control plants grown for 10 weeks in pots (Figure 3.4b), was analyzed. Plant height and number of nodes as well as total fresh and dry weight were not significantly different between $_e$Gen2-containing line 2 plants and wild-type control plants (Figure 3.4c-f). Considering leaf biomass per unit of foliar area, the $_e$Gen2-containing line 2 plants accumulated congruent fresh biomass (Figure 3.4g) but ~20% less dry biomass than wild-type (Figure 3.4h). However, there was no significant difference between $_e$Gen2-containing line 2 plants and the transplastomic controls (Figure 3.4h). No difference was observed in both chlorophyll content index and leaf CO_2 assimilation between $_e$Gen2-containing line 2 and wild-type plants, indicating there was congruent chloroplast function among lines (Figure 3.4i-j).

Conclusions

The complete design-build-test cycle used to develop the mini-synplastome platform is illustrated in Figure 3.5. Based on the results obtained from this work, mini-synplastomes provide significant promise as an alternative to traditional HR-based plastid engineering. The ability to express transgenes without the requirement for integration of foreign DNA into the host genome represents a paradigm shift in plastid engineering. With regards to the regulatory landscape, chloroplast biotechnology has long remained in the “gray area” since HR can be achieved without the use of any plant pathogens or pathogenic sequences, thus resulting in exemption from USDA-APHIS regulation 7 CFR part 340 (Daniell *et al.,* 2019b). However, in order for commercialization of chloroplast engineered plants, it is desirable for selection genes to be removed prior to approval. To-date, marker-free transplastomics have been generated through the use of recombinases (Corneille *et al.,* 2001; Hajdukiewicz *et al.,* 2001) , loop-out recombination (Ruf *et al.,* 2007), transient cointegration of two chloroplast vectors (Klaus *et al.,* 2004), or direct repeat recombination of transgenes integrated into the large single copy region of soybean (Dufourmantel *et al.,* 2007) or the

inverted repeat region of lettuce (Daniell *et al.,* 2020; Daniell *et al.,* 2019b; Kumari *et al.,* 2019; Park *et al.,* 2020). As demonstrated in this work, it is possible to design integrating episomal constructs, where the desired transgenes are incorporated into the native plastome through HR, while the selection genes are expressed on the episomal construct. In this scenario, once the selective pressure is removed, the episomal construct can be eliminated if it is designed for that function. In this case, the transgenes of interest will be retained. Facile methods for marker-free chloroplast engineering will be advantageous in a germplasm development pipeline.

In this work, engineered episomal constructs were demonstrated to persist throughout the life cycle of potato plants, regardless of whether *or* not transgenes were expressed from the episomal construct, or integrated into the native plastome. Compared to the current state-of-the-art in plastid engineering, the integration of a transgene into the *trnI/trnA* site by HR, $_{e}$Gen1-containing lines, with a persisting episome and integration of the transgene cassette into the *trnI/trnA* site, had up to a 250% increase in transgene expression. It is known that copy number has a direct impact on the expression level of foreign genes integrated into the chloroplast genome; integration of GFP into the trnI/trnA region of the Inverted Repeat region resulted in 25-fold higher expression than integration into the large single copy region (rbcL/accD) (Krichevsky *et al.,* 2010). One explanation for the increase may be the potential for the transgene cassette to constantly recombine between the episomal construct and the plastome, leading to increased copy numbers of the transgene cassette. Indeed, including the chloroplast origin of replication within the chloroplast DNA flanking sequence used for transgene integration (Lugo *et al.,* 2004; Nielsen *et al.,* 1993) facilitated achieving homoplasmy even after the first round of selection (Daniell *et al.,* 1998; Guda *et al.,* 2000). While demonstrated here in potato, episomal plastid engineering is not relegated to land plants, and represents a strategy that may be utilized for engineering plastids of other species, such as algae, or other organellar genomes, such as mitochondria. The simplicity of the approach, and the ability to express transgenes at levels similar to HR-based plastid engineering validates its use for production of heterologous proteins for use in a variety of fields, such as biopharming, enzyme production, and crop improvement. Further refinement of the episomal strategy using chloroplast-specific *ori* from different organisms, including higher plants and algae (Carrillo and Bogorad, 1988; Chiu and Sears, 1992; Daniell *et al.,* 1990; Karas *et al.,* 2015; Kunnimalaiyaan and Nielsen, 1997; Meeker *et al.,* 1988; Nisbet *et al.,* 2004; Waddell *et al.,* 1984; Wang *et al.,* 1984), and the use of single or multiple *ori* with different activities could enable precise tuning of the copy number of the episomal constructs, adding another layer for control of gene expression in complex pathways. In essence, this work demonstrates a novel technology with significant improvements on the current state of plastid engineering, which should enable synthetic biology in plants.

References

Barbrook, A.C., Dorrell, R.G., Burrows, J., Plenderleith, L.J., Nisbet, R.E. and Howe, C.J. (2012) Polyuridylylation and processing of transcripts from multiple gene minicircles in chloroplasts of the dinoflagellate *Amphidinium carterae*. *Plant Mol. Biol* **79**, 347-357.

Barbrook, A.C. and Howe, C.J. (2000) Minicircular plastid DNA in the dinoflagellate *Amphidinium operculatum*. *Mol. Gen. Genet* **263**, 152-158.

Carrillo, N. and Bogorad, L. (1988) Chloroplast DNA replication *in vitro*: site-specific initiation from preferred templates. *Nucleic Acids Res.* **16**, 5603-5620.

Chan, H.T., Xiao, Y., Weldon, W.C., Oberste, S.M., Chumakov, K. and Daniell, H. (2016) Cold chain and virus-free chloroplast-made booster vaccine to confer immunity against different poliovirus serotypes. *Plant Biotechnol. J.* **14**, 2190-2200.

Chiu, W.L. and Sears, B.B. (1992) Electron microscopic localization of replication origins in *Oenothera* chloroplast DNA. *Mol. Gen. Genet.* **232**, 33-39.

Corneille, S., Lutz, K., Svab, Z. and Maliga, P. (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-lox site-specific recombination system. *Plant J.* **27**, 171-178.

Daniell, H., Datta, R., Varma, S., Gray, S. and Lee, S.B. (1998) Containment of herbicide resistance through genetic engineering of the chloroplast genome. *Nat. Biotechnol.* **16**, 345-348.

Daniell, H., Lee, S.B., Grevich, J., Saski, C., Quesada-Vargas, T., Guda, C., Tomkins, J. and Jansen, R.K. (2006) Complete chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other *Solanaceae* genomes. *Theor. Appl. Genet.* **112**, 1503-1518.

Daniell, H., Lin, C.S., Yu, M. and Chang, W.J. (2016) Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol.* **17**, 134.

Daniell, H., Mangu, V., Yakubov, B., Park, J., Habibi, P., Shi, Y., Gonnella, P.A., Fisher, A., Cook, T., Zeng, L., Kawut, S.M. and Lahm, T. (2020) Investigational new drug enabling angiotensin oral-delivery studies to attenuate pulmonary hypertension. *Biomaterials* **233**, 119750.

Daniell, H., Rai, V. and Xiao, Y. (2019a) Cold chain and virus-free oral polio booster vaccine made in lettuce chloroplasts confers protection against all three poliovirus serotypes. *Plant Biotechnol. J.* **17**, 1357-1368.

Daniell, H., Ribeiro, T., Lin, S., Saha, P., McMichael, C., Chowdhary, R. and Agarwal, A. (2019b) Validation of leaf and microbial pectinases: commercial launching of a new platform technology. *Plant Biotechnol. J.* **17**, 1154-1166.

Daniell, H., Vivekananda, J., Nielsen, B.L., Ye, G.N., Tewari, K.K. and Sanford, J.C. (1990) Transient foreign gene expression in chloroplasts of cultured tobacco cells after biolistic delivery of chloroplast vectors. *Proc. Natl. Acad. Sci. USA* **87**, 88-92.

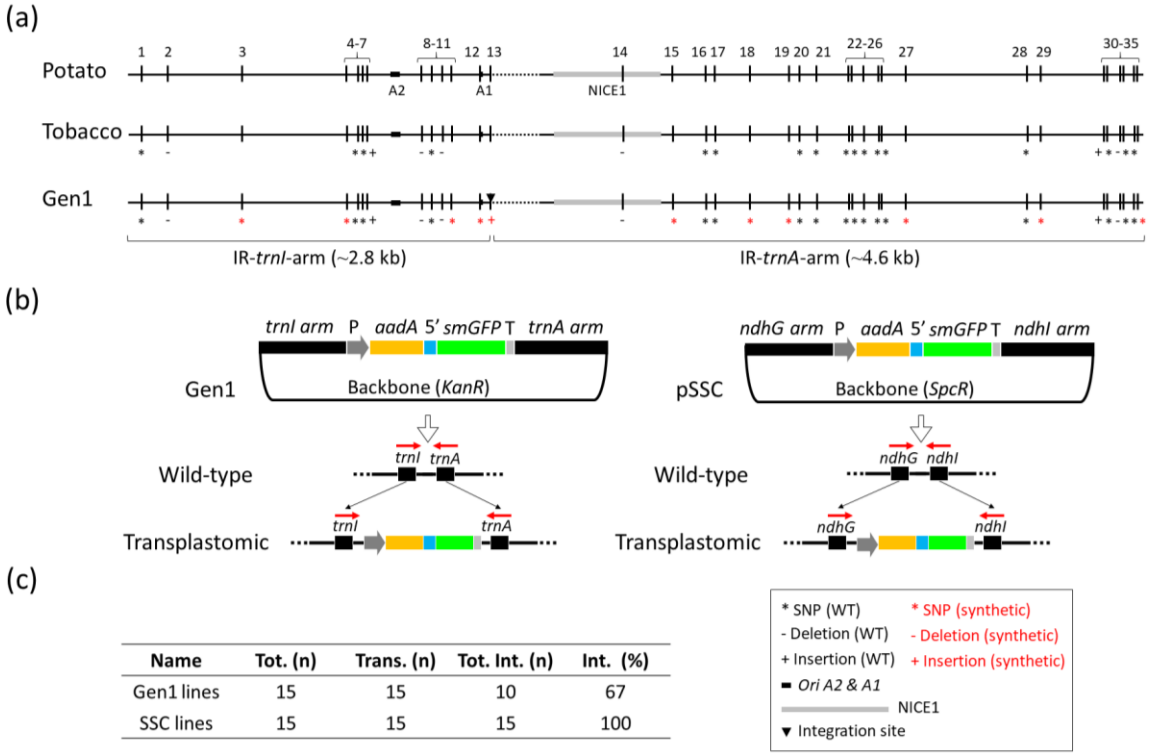
Day, A. and Madesis, P. (2007) DNA replication, recombination, and repair in plastids. In: *Cell and Molecular Biology of Plastids* (Bock, R. ed) pp. 65-119. Berlin, Heidelberg: Springer Berlin Heidelberg.
De Cosa, B., Moar, W., Lee, S.B., Miller, M. and Daniell, H. (2001) Overexpression of the *Bt cry2Aa2* operon in chloroplasts leads to formation of insecticidal crystals. *Nat. Biotechnol.* **19**, 71-74.
Dufourmantel, N., Dubald, M., Matringe, M., Canard, H., Garcon, F., Job, C., Kay, E., Wisniewski, J.P., Ferullo, J.M., Pelissier, B., Sailland, A. and Tissot, G. (2007) Generation and characterization of soybean and marker-free tobacco plastid transformants over-expressing a bacterial 4-hydroxyphenylpyruvate dioxygenase which provides strong herbicide tolerance. *Plant Biotechnol. J.* **5**, 118-133.
Fox, J.L. (2012) First plant-made biologic approved. *Nat. Biotechnol.* **30**, 472-472.
Gray, B.N., Ahner, B.A. and Hanson, M.R. (2009) Extensive homologous recombination between introduced and native regulatory plastid DNA elements in transplastomic plants. *Transgenic Res.* **18**, 559-572.
Guda, C., Lee, S.B. and Daniell, H. (2000) Stable expression of a biodegradable protein-based polymer in tobacco chloroplasts. *Plant Cell Rep.* **19**, 257-262.
Hajdukiewicz, P.T., Gilbertson, L. and Staub, J.M. (2001) Multiple pathways for Cre/lox-mediated recombination in plastids. *Plant J.* **27**, 161-170.
Howe, C.J., Barbrook, A.C., Koumandou, V.L., Nisbet, R.E., Symington, H.A. and Wightman, T.F. (2003) Evolution of the chloroplast genome. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **358**, 99-106; discussion 106-107.
Howe, C.J., Nisbet, R.E. and Barbrook, A.C. (2008) The remarkable chloroplast genome of dinoflagellates. *J. Exp. Bot.* **59**, 1035-1045.
Iamtham, S. and Day, A. (2000) Removal of antibiotic resistance genes from transgenic tobacco plastids. *Nat. Biotechnol.* **18**, 1172-1176.
Karas, B.J., Diner, R.E., Lefebvre, S.C., McQuaid, J., Phillips, A.P., Noddings, C.M., Brunson, J.K., Valas, R.E., Deerinck, T.J., Jablanovic, J., Gillard, J.T., Beeri, K., Ellisman, M.H., Glass, J.I., Hutchison, C.A., 3rd, Smith, H.O., Venter, J.C., Allen, A.E., Dupont, C.L. and Weyman, P.D. (2015) Designer diatom episomes delivered by bacterial conjugation. *Nat. Commun.* **6**, 6925.
Klaus, S.M., Huang, F.C., Golds, T.J. and Koop, H.U. (2004) Generation of marker-free plastid transformants using a transiently cointegrated selection gene. *Nat. Biotechnol.* **22**, 225-229.
Koressaar, T. and Remm, M. (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**, 1289-1291.
Kota, M., Daniell, H., Varma, S., Garczynski, S.F., Gould, F. and Moar, W.J. (1999) Overexpression of the *Bacillus thuringiensis* (Bt) Cry2Aa2 protein in chloroplasts confers resistance to plants against susceptible and Bt-resistant insects. *Proc. Natl. Acad. Sci. USA* **96**, 1840-1845.
Koumandou, V.L. and Howe, C.J. (2007) The copy number of chloroplast gene minicircles changes dramatically with growth phase in the dinoflagellate *Amphidinium operculatum*. *Protist* **158**, 89-103.

Krichevsky, A., Meyers, B., Vainstein, A., Maliga, P. and Citovsky, V. (2010) Autoluminescent plants. *PLoS One* **5**, e15461.
Krishnan, N.M. and Rao, B.J. (2009) A comparative approach to elucidate chloroplast genome replication. *BMC Genomics* **10**, 237.
Kumar, S., Hahn, F.M., Baidoo, E., Kahlon, T.S., Wood, D.F., McMahan, C.M., Cornish, K., Keasling, J.D., Daniell, H. and Whalen, M.C. (2012) Remodeling the isoprenoid pathway in tobacco by expressing the cytoplasmic mevalonate pathway in chloroplasts. *Metab. Eng.* **14**, 19-28.
Kumari, U., Singh, R., Ray, T., Rana, S., Saha, P., Malhotra, K. and Daniell, H. (2019) Validation of leaf enzymes in the detergent and textile industries: launching of a new platform technology. *Plant Biotechnol. J.* **17**, 1167-1182.
Kunnimalaiyaan, M. and Nielsen, B.L. (1997) Fine mapping of replication origins (*oriA* and *oriB*) in *Nicotiana tabacum* chloroplast DNA. *Nucleic Acids Res.* **25**, 3681-3686.
Kwon, K.C., Chan, H.T., León, I.R., Williams-Carrier, R., Barkan, A. and Daniell, H. (2016) Codon optimization to enhance expression yields insights into chloroplast translation. *Plant Physiol.* **172**, 62-77.
Kwon, K.C., Sherman, A., Chang, W.J., Kamesh, A., Biswas, M., Herzog, R.W. and Daniell, H. (2018) Expression and assembly of largest foreign protein in chloroplasts: oral delivery of human FVIII made in lettuce chloroplasts robustly suppresses inhibitor formation in haemophilia A mice. *Plant Biotechnol. J.* **16**, 1148-1160.
Kwon, K.C., Verma, D., Jin, S., Singh, N.D. and Daniell, H. (2013) Release of proteins from intact chloroplasts induced by reactive oxygen species during biotic and abiotic stress. *PLoS One* **8**, e67106.
Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A. and Hanson, M.R. (2014) A faster Rubisco with potential to increase photosynthesis in crops. *Nature* **513**, 547-550.
Lu, Y., Rijzaani, H., Karcher, D., Ruf, S. and Bock, R. (2013) Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. *Proc. Natl. Acad. Sci. USA* **110**, E623-632.
Lugo, S.K., Kunnimalaiyaan, M., Singh, N.K. and Nielsen, B.L. (2004) Required sequence elements for chloroplast DNA replication activity *in vitro* and in electroporated chloroplasts. *Plant Sci.* **166**, 151-161.
Malhotra, K., Subramaniyan, M., Rawat, K., Kalamuddin, M., Qureshi, M.I., Malhotra, P., Mohmmed, A., Cornish, K., Daniell, H. and Kumar, S. (2016) Compartmentalized metabolic engineering for artemisinin biosynthesis and effective malaria treatment by oral delivery of plant cells. *Mol. Plant* **9**, 1464-1477.
Martin Avila, E., Gisby, M.F. and Day, A. (2016) Seamless editing of the chloroplast genome in plants. *BMC Plant Biol.* **16**, 168.
Meeker, R., Nielsen, B. and Tewari, K.K. (1988) Localization of replication origins in pea chloroplast DNA. *Mol. Cell. Biol.* **8**, 1216-1223.

Min, S.R., Davarpanah, S.J., Jung, S.H., Park, Y., Liu, J.R. and Jeong, W.J. (2015a) An episomal vector system for plastid transformation in higher plants. *Plant Biotechnol. Rep.* **9**, 443-449.
Min, S.R., Jung, S.H., Liu, J.R. and Jeong, W.J. (2015b) The fate of extrachromosomal DNAs in the progeny of plastid-transformed tobacco plants. *Plant Biotechnol. Rep.* **9**, 431-442.
Mühlbauer, S.K., Lössl, A., Tzekova, L., Zou, Z. and Koop, H.U. (2002) Functional analysis of plastid DNA replication origins in tobacco by targeted inactivation. *Plant J.* **32**, 175-184.
Nielsen, B.L., Lu, Z. and Tewari, K.K. (1993) Characterization of the pea chloroplast DNA *oriA* region. *Plasmid* **30**, 197-211.
Nielsen, B.L. and Tewari, K.K. (1988) Pea chloroplast topoisomerase I: purification, characterization, and role in replication. *Plant Mol. Biol.* **11**, 3-14.
Nisbet, R.E., Koumandou, L., Barbrook, A.C. and Howe, C.J. (2004) Novel plastid gene minicircles in the dinoflagellate *Amphidinium operculatum*. *Gene* **331**, 141-147.
Norrander, J., Kempe, T. and Messing, J. (1983) Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26**, 101-106.
Occhialini, A., Lin, M.T., Andralojc, P.J., Hanson, M.R. and Parry, M.A. (2016) Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO_2. *Plant J.* **85**, 148-160.
Occhialini, A., Pfotenhauer, A.C., Frazier, T.P., Li, L., Harbison, S.A., Lail, A.J., Mebane, Z., Piatek, A.A., Rigoulot, S.B., Daniell, H., Stewart, C.N., Jr. and Lenaghan, S.C. (2020) Generation, analysis, and transformation of macro-chloroplast potato (*Solanum tuberosum*) lines for chloroplast biotechnology. *Sci. Rep.* **10**, 21144.
Occhialini, A., Piatek, A.A., Pfotenhauer, A.C., Frazier, T.P., Stewart, C.N., Jr. and Lenaghan, S.C. (2019) MoChlo: a versatile modular cloning toolbox for chloroplast biotechnology. *Plant Physiol.* **179**, 943-957.
Park, J., Yan, G., Kwon, K.C., Liu, M., Gonnella, P.A., Yang, S. and Daniell, H. (2020) Oral delivery of novel human IGF-1 bioencapsulated in lettuce cells promotes musculoskeletal cell proliferation, differentiation and diabetic fracture healing. *Biomaterials* **233**, 119591.
Ruf, S., Karcher, D. and Bock, R. (2007) Determining the transgene containment level provided by chloroplast transformation. *Proc. Natl. Acad. Sci. USA* **104**, 6998-7002.
Ruhlman, T., Verma, D., Samson, N. and Daniell, H. (2010) The role of heterologous chloroplast sequence elements in transgene integration and expression. *Plant Physiol.* **152**, 2088-2104.
Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and

Sugiura, M. (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J.* **5**, 2043-2049.
Staub, J.M. and Maliga, P. (1994) Extrachromosomal elements in tobacco plastids. *Proc. Natl. Acad. Sci. USA* **91**, 7468-7472.
Staub, J.M. and Maliga, P. (1995) Marker rescue from the *Nicotiana tabacum* plastid genome using a plastid/*Escherichia coli* shuttle vector. *Mol. Gen. Genet.* **249**, 37-42.
Svab, Z., Hajdukiewicz, P. and Maliga, P. (1990) Stable transformation of plastids in higher plants. *Proc. Natl. Acad. Sci. USA* **87**, 8526-8530.
Svab, Z. and Maliga, P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proc. Natl. Acad. Sci. USA* **90**, 913-917.
Svab, Z. and Maliga, P. (2007) Exceptional transmission of plastids and mitochondria from the transplastomic pollen parent and its impact on transgene containment. *Proc. Natl. Acad. Sci. USA* **104**, 7003-7008.
Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012) Primer3--new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115.
Valkov, V.T., Gargano, D., Manna, C., Formisano, G., Dix, P.J., Gray, J.C., Scotti, N. and Cardi, T. (2011) High efficiency plastid transformation in potato and regulation of transgene expression in leaves and tubers by alternative 5' and 3' regulatory sequences. *Transgenic Res.* **20**, 137-151.
Vickery, B.P., Vereda, A., Casale, T.B., Beyer, K., du Toit, G., Hourihane, J.O., Jones, S.M., Shreffler, W.G., Marcantonio, A., Zawadzki, R., Sher, L., Carr, W.W., Fineman, S., Greos, L., Rachid, R., Ibáñez, M.D., Tilles, S., Assa'ad, A.H., Nilsson, C., Rupp, N., Welch, M.J., Sussman, G., Chinthrajah, S., Blumchen, K., Sher, E., Spergel, J.M., Leickly, F.E., Zielen, S., Wang, J., Sanders, G.M., Wood, R.A., Cheema, A., Bindslev-Jensen, C., Leonard, S., Kachru, R., Johnston, D.T., Hampel, F.C., Jr., Kim, E.H., Anagnostou, A., Pongracic, J.A., Ben-Shoshan, M., Sharma, H.P., Stillerman, A., Windom, H.H., Yang, W.H., Muraro, A., Zubeldia, J.M., Sharma, V., Dorsey, M.J., Chong, H.J., Ohayon, J., Bird, J.A., Carr, T.F., Siri, D., Fernández-Rivas, M., Jeong, D.K., Fleischer, D.M., Lieberman, J.A., Dubois, A.E.J., Tsoumani, M., Ciaccio, C.E., Portnoy, J.M., Mansfield, L.E., Fritz, S.B., Lanser, B.J., Matz, J., Oude Elberink, H.N.G., Varshney, P., Dilly, S.G., Adelman, D.C. and Burks, A.W. (2018) AR101 oral immunotherapy for peanut allergy. *N. Engl. J. Med.* **379**, 1991-2001.
Waddell, J., Wang, X.M. and Wu, M. (1984) Electron microscopic localization of the chloroplast DNA replicative origins in *Chlamydomonas reinhardii*. *Nucleic Acids Res.* **12**, 3843-3856.
Wang, X.M., Chang, C.H., Waddell, J. and Wu, M. (1984) Cloning and delimiting one chloroplast DNA replicative origin of *Chlamydomonas*. *Nucleic Acids Res.* **12**, 3857-3872.
Xiao, Y. and Daniell, H. (2017) Long-term evaluation of mucosal and systemic immunity and protection conferred by different polio booster vaccines. *Vaccine* **35**, 5418-5425.

Zhang, J., Khan, S.A., Hasse, C., Ruf, S., Heckel, D.G. and Bock, R. (2015) Pest control. Full crop protection from an insect pest by expression of long double-stranded RNAs in plastids. *Science* **347**, 991-994.



Name	Tot. (n)	Trans. (n)	Tot. Int. (n)	Int. (%)
Gen1 lines	15	15	10	67
SSC lines	15	15	15	100

Figure 3.1. Vector design. (a) Design of *trnI*/*trnA* homologous arms (~2.8 and ~4.6 kb) used in Gen1 vector construction. An alignment of homologous regions in the potato and tobacco plastome along with the two synthetic sequences used in the design of the Gen1 vector are indicated. A total of 35 regions of sequence difference are located within homologous regions. A vertical bar (|) without any symbols indicates sequence identity compared to the potato plastome, whereas an asterisk (*), minus (-) or plus (+) indicate single nucleotide polymorphism (SNP), sequence deletion or addition, respectively. Black symbols indicate wild-type mutations between potato and tobacco, whereas red symbols indicated synthetic mutations introduced in Gen1 sequences. *Ori*A2 and A1 (black bars) and NICE1 (gray bar) are also indicated. (b) Schematic representation of two DNA constructs (Gen1 and pSSC) used in this work and their integration site in the native plastome. IR *trnI/trnA* homologous arms of the Gen1 vector (~2.8 and ~4.6 kb respectively, including *oriA2* and *A1*) and *ndhG/ndhI* homologous arms of the pSSC (~1.8 and ~5 kb, respectively) are indicated. A selection cassette located between arms is indicated: *Prom-SD* (P): *rrn* promoter along with a Shine-Dalgarno sequence (gray); *aadA*: spectinomycin resistance gene (yellow); *5'UTR*: 5' untranslated region (blue); *smGFP*: gene encoding the soluble monomeric green fluorescent protein (green); and *3'UTR* (T): 3' untranslated region (light gray). Backbone vectors containing the kanamycin (*KanR*) or spectinomycin (*SpcR*) resistance gene are indicated in Gen1 and pSSC constructs. Gen1 and pSSC integration into *trnI/trnA* and *ndhG/ndhI* sites of the native plastome and the location of primers used to check integration (red

arrows) are indicated. (c) Table indicating the percentage of vector integration in transplastomic lines. Tot. (n): total number of lines analyzed; Trans. (n): total number of positive lines for the presence of *aadA* and *smGFP* genes; Tot. int. (n): total number of lines with vector integration; Int. (%): percentage of plants with vector integration.

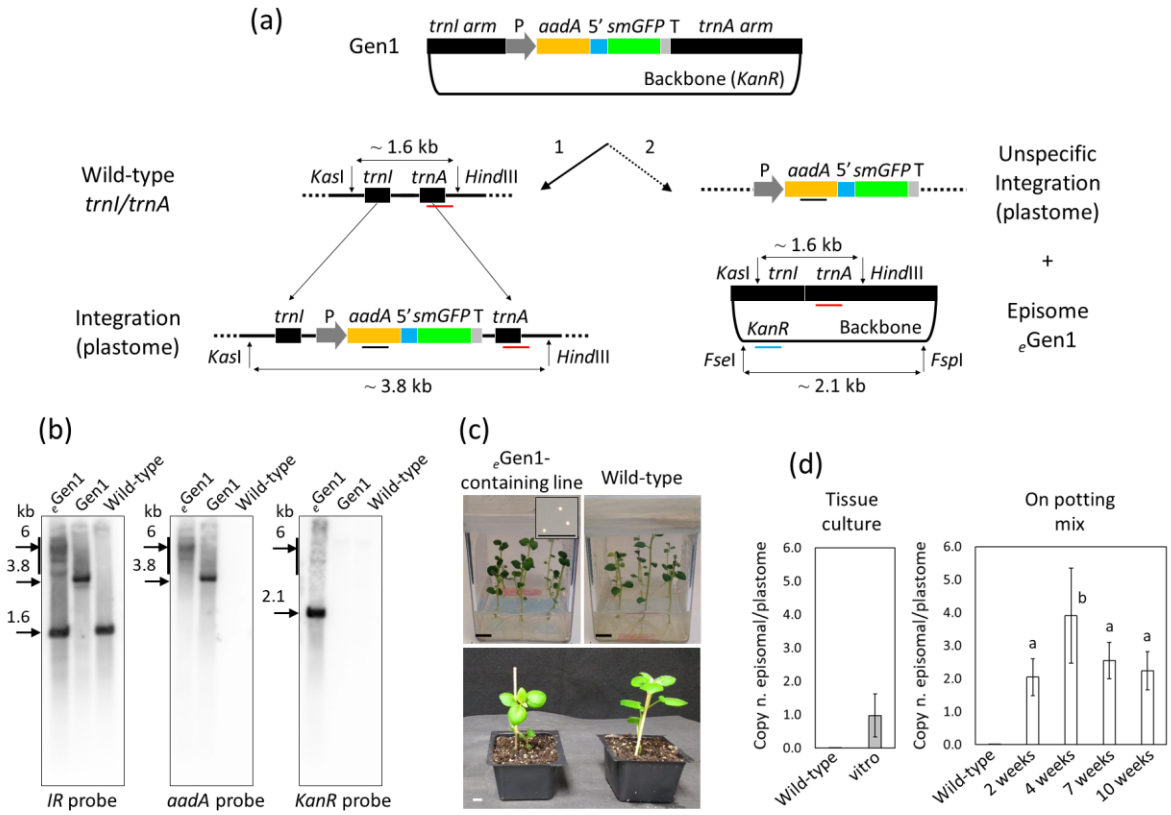


Figure 3.2. Characterization of $_e$Gen1-containing lines. (a) Schematic representation of the possible fates of the Gen1 vector in plants are shown: (1) integration of the transgene cassette into the *trnI/trnA* site of the potato plastome; (2) nonspecific integration of the transgene cassette into an unknown site, while an episome ($_e$Gen1) is formed from the vector backbone and the remaining homologous sequence. The *rrn* promoter and a Shine-Dalgarno sequence (P; gray); *aadA* gene (yellow); 5' untranslated region (5'; blue); *smGFP* gene (green); 3' untranslated region (T; light gray); the *trnI/trnA* arms; and the backbone vectors containing *KanR* gene are indicated in the figure. The location of *IR*, *aadA* and *KanR* probes are indicated with red, dark and blue bars, respectively. The molecular weights of DNA fragments obtained using *Kas*I/*Hind*III (for *IR* and *aadA* probes) and *Fse*I/*Fsp*I (for the *KanR* probe) restriction enzymes are also indicated. (b) Southern blot analysis performed using an *IR, aadA* or a *KanR* probe and total leaf genomic-DNA preparations from an $_e$Gen1-containing line, a Gen1-integrating line, and a wild-type control is shown. Molecular weight of DNA fragments (kb) are indicated in the blots. (c) Wild-type potato plants along with the $_e$Gen1-containing line grown *in vitro* and for 2 weeks in potting mix are shown. Insert shows bacterial colonies transformed with $_e$Gen1 extracted from leaf tissue (scale bar: 10 mm). (d) Graph summarizing the ratio of copy number of episomal DNA vs the copy number of plastome (copy n. episomal/plastome) in genomic DNA preparations of $_e$Gen1-containing plants by qPCR. Wild-type potato plants were used as negative control. Two graphs representing *in vitro*

plants at the second round of tissue culture and plants grown on potting mix without selection (2, 4, 7 and 10 weeks) are shown. Results are expressed as mean ± standard deviation. For plants in tissue culture, 5 biological and 9 technical replicates per biological replicate were used. For plants on potting mix, 3 biological and 4 technical replicates per each biological replicate were used. Means separation was evaluated using ANOVA Tukey HSD ($p<0.05$). Statistical significance is indicated by different letters.

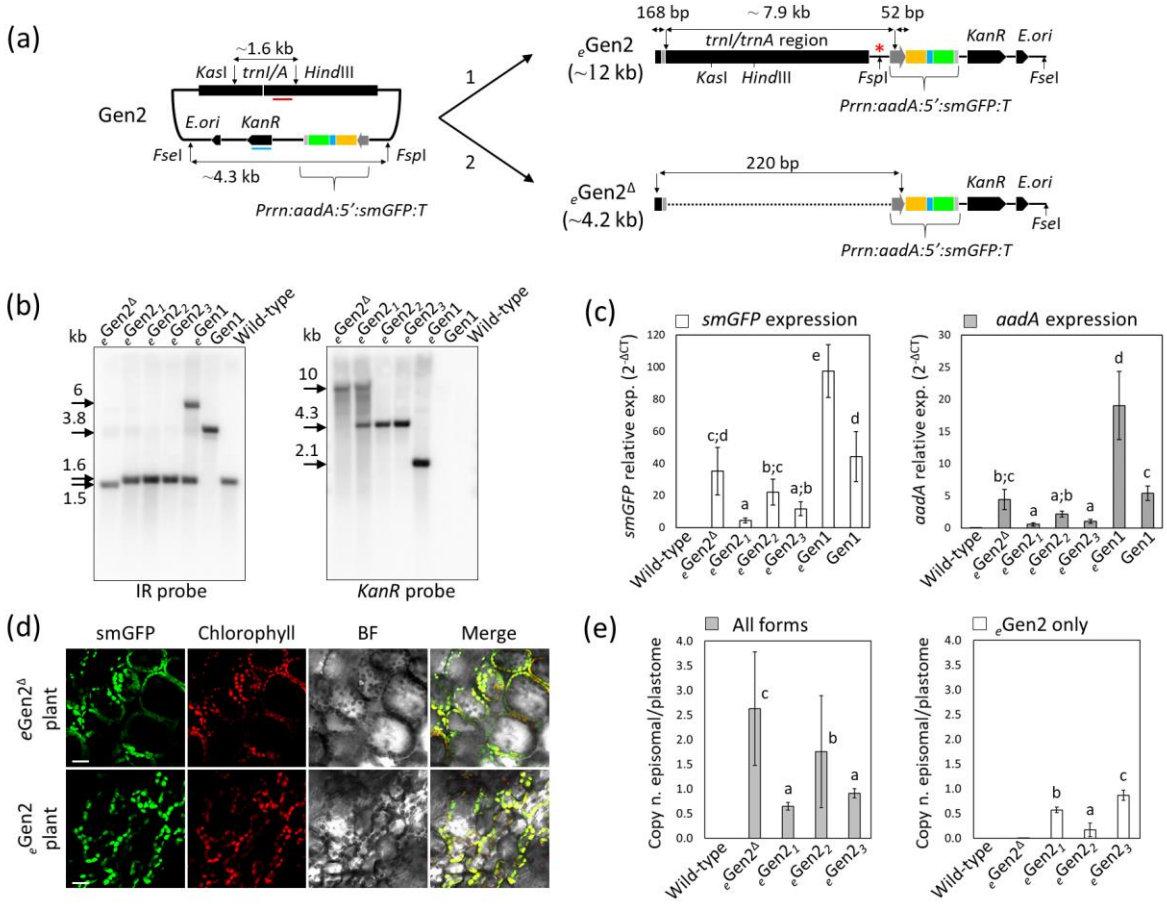


Figure 3.3. Characterization of synplastomic $_e$Gen2-containing lines. (a) Schematic representation of the possible fates of the Gen2 vector in plants are shown: 1) full-length $_e$Gen2 and/or $_e$Gen2$^{\Delta}$. For simplicity of sequence comparison, these two plasmids, $_e$Gen2 and $_e$Gen2$^{\Delta}$, are represented in linear form. In $_e$Gen2$^{\Delta}$, ~7.9 kb was excised by loop-out recombination reconstituting *Prrn*. Perpendicular black arrows indicate deletion sites, while parallel double arrows indicate molecular weight of fragments. The *trnI/trnA* homologous region and the backbone vector containing the *KanR* gene, an *E. coli* origin of replication (*E.ori*) and the dual selection cassette are indicated. The *Prrn* promoter fused to a ribosome binding site (P, deep gray); *aadA* gene (yellow); 5'UTR (5', blue); *smGFP* gene (green); and the 3'UTR/terminator (T, gray) are indicated in the selection cassette. Restriction enzyme combinations used for Southern blots, *Kas*I/*Hind*III and *Fse*I/*Fsp*I, together with the sizes of predicted DNA fragments are indicated. The *Kas*I/*Hind*III and *Fse*I/*Fsp*I fragments were detected by a ~0.5 kb probe designed on *trnI/trnA* (red bar) or *KanR* (blue bar), respectively. (b) Southern blot analysis performed using either an *IR* or *KanR* probe and leaf total DNA preparations from $_e$Gen2-containing lines 1-3 and a line harboring $_e$Gen2$^{\Delta}$. DNA samples from $_e$Gen1-containing and Gen1-integrating lines along with wild-type plants were used as a comparison. Molecular weight of DNA fragments (kb) are indicated in the blots. (c) qRT-PCRs using cDNA

preparations from $_e$Gen2-containing lines 1-3 and $_e$Gen2$^{\Delta}$ lines. $_e$Gen1-containing and Gen1-integrating lines along with wild-type controls have been included. Graphs showing the relative expression of *smGFP* (white bars) and *aadA* (gray bars) compared to the internal reference gene *ef1* (y axis: $2^{-\Delta CT}$) are shown. The results are expressed as mean ± sd (standard deviation) of 3 biological and 3 technical replicates. (d) Confocal image showing smGFP localization to the chloroplast in both $_e$Gen2- and $_e$Gen2$^{\Delta}$-containing lines. smGFP (green), chlorophyll (red), bright field (BF) and merged images are shown. Scale bars: 20 µm. (e) Graph summarizing the ratio of episomal plasmid copy number to the copy number of the plastome (copy n. episomal/plastome) in genomic DNA preparations of $_e$Gen2- and $_e$Gen2$^{\Delta}$-containing lines determined by qPCR. Wild-type plants were used as a negative control. PCR analyses were performed to calculate the ratio of copy number for all forms ($_e$Gen2 and $_e$Gen2$^{\Delta}$; gray bars) and only the full-length $_e$Gen2 (white bars). Results are expressed as mean ± standard deviation of 5 biological and 3 technical replicates per each biological replicate. In graphs C and E, mean separation was evaluated using ANOVA Tukey HSD ($p<0.05$), and statistical significance is indicated by different letters.

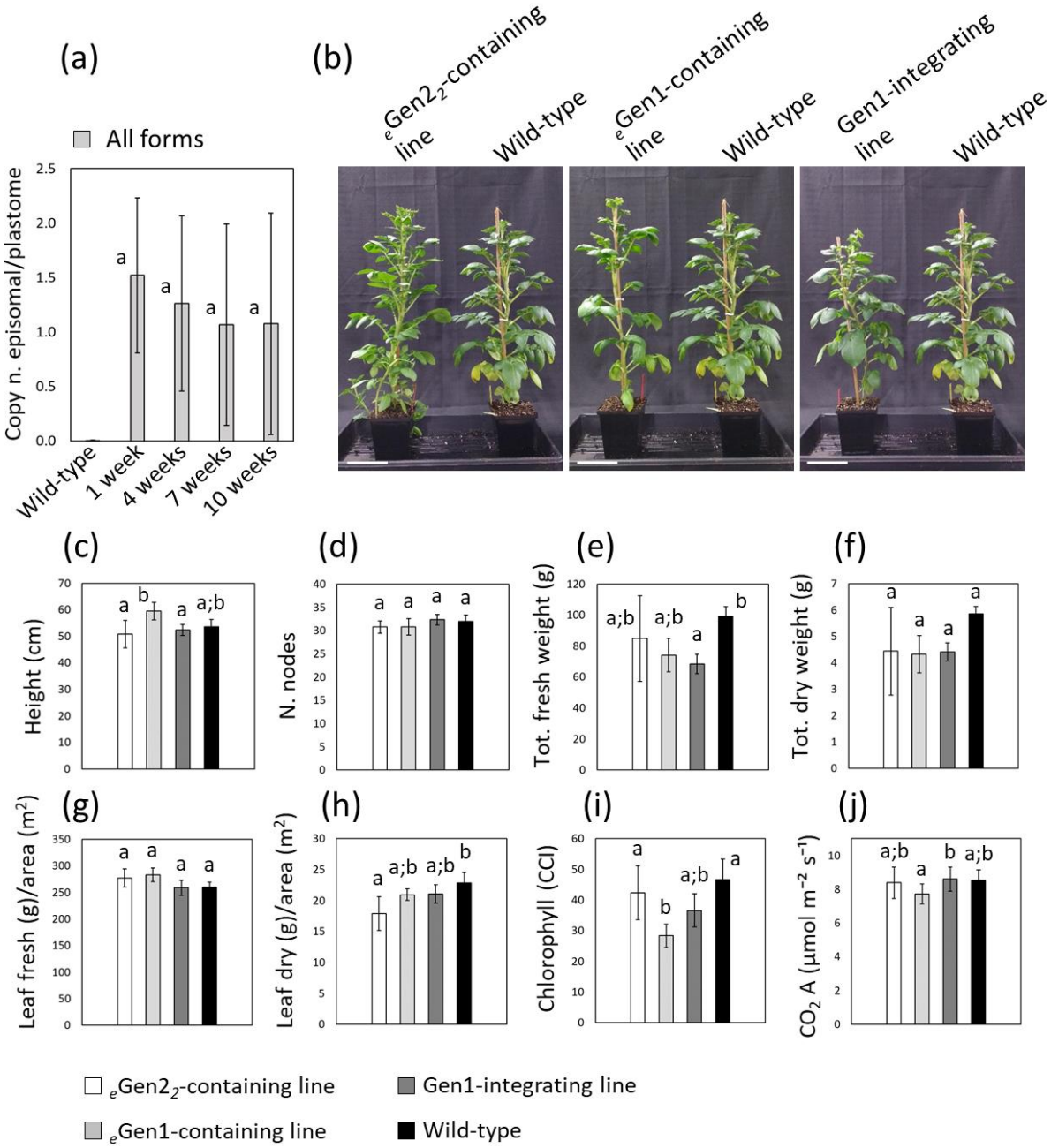


Figure 3.4. Growth characteristics of synplastomic plants. (a) Graph summarizing the ratio of episomal plasmid copy number vs the copy number of plastome (copy n. episomal/plastome) in genomic DNA preparations of $_e$Gen2-containing line 2 at different plant developmental stages (1, 4, 7 and 10 weeks in pots) determined by qPCR. Wild-type plants were used as negative controls. PCRs were performed to determine the copy number for all plasmid forms ($_e$Gen2 and $_e$Gen2$^{\Delta}$). The results shown are mean ± standard deviation of 5 biological replicates (plants 1-5) and 3 technical replicates per biological replicate. (b) Images showing 10-week-old $_e$Gen2-containing line 2, an $_e$Gen1-

containing line and a Gen1-integrating transplastomic line compared to wild-type control plants. Scale bar: 10 cm. Graphs represent various plant characteristics: (c) Plant height; (d) number (N.) of nodes; (e) total fresh weight; (f) total dry weight; (g) ratio of leaf fresh weight (g) to foliar area (m^2); (h) ratio of leaf dry weight (g) to foliar area (m^2); (i) chlorophyll content index (CCI); (j) leaf CO_2 assimilation (µmol m^{-2} s^{-1}). The results are expressed as mean ± standard deviation of 5 plants per each transgenic line and wild-type control. For all graphs means separation was evaluated using ANOVA Tukey HSD ($p<0.05$). Statistical significance is indicated by different letters.

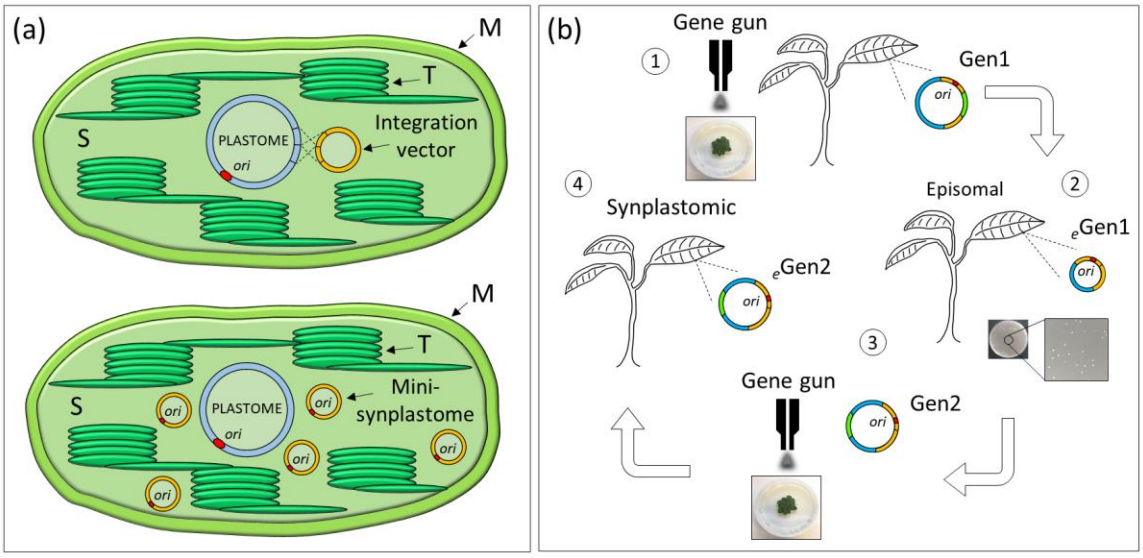
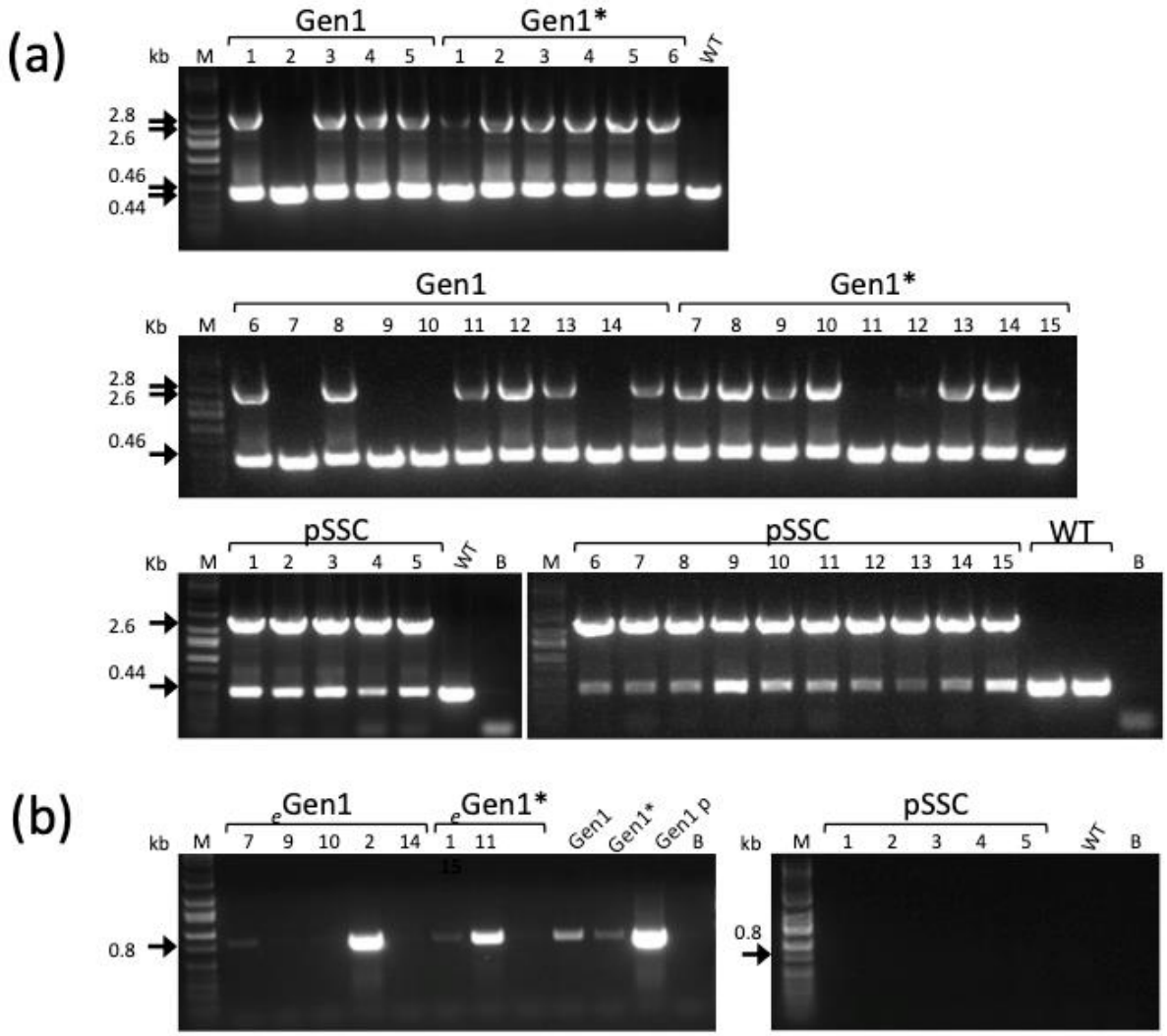
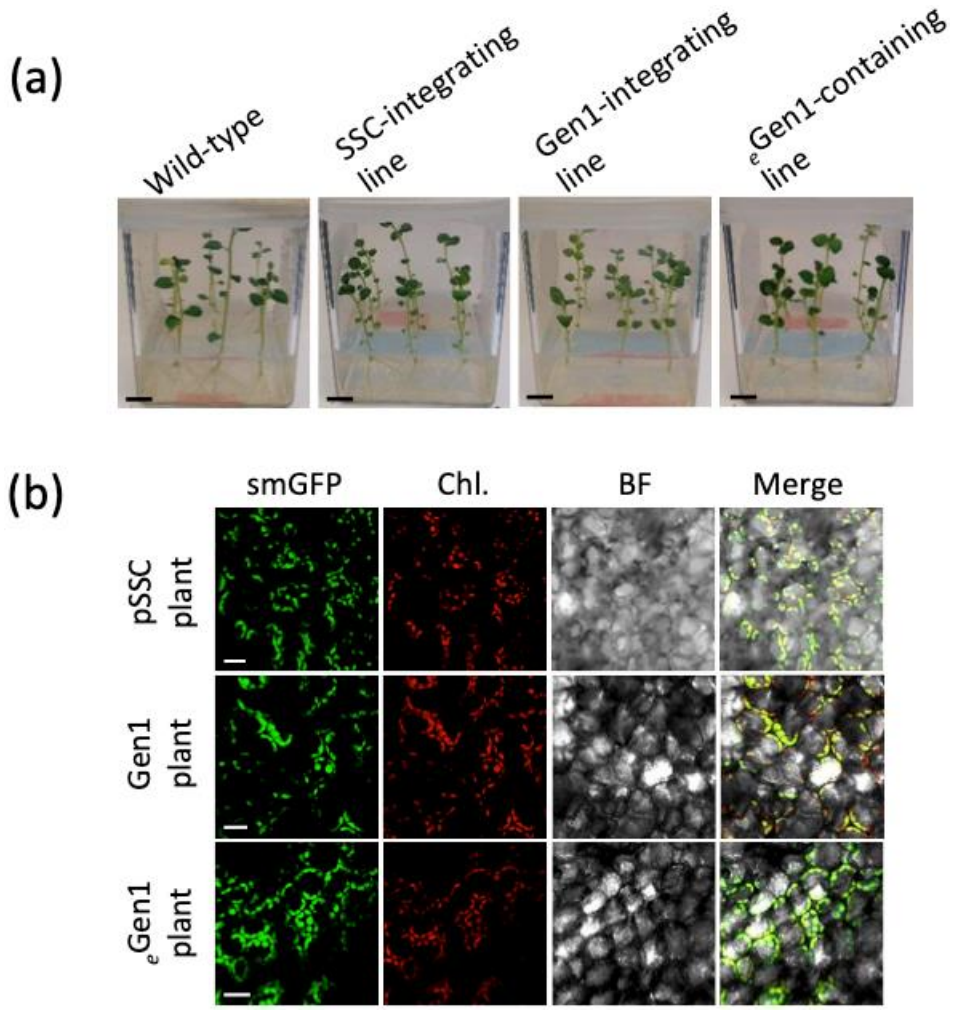


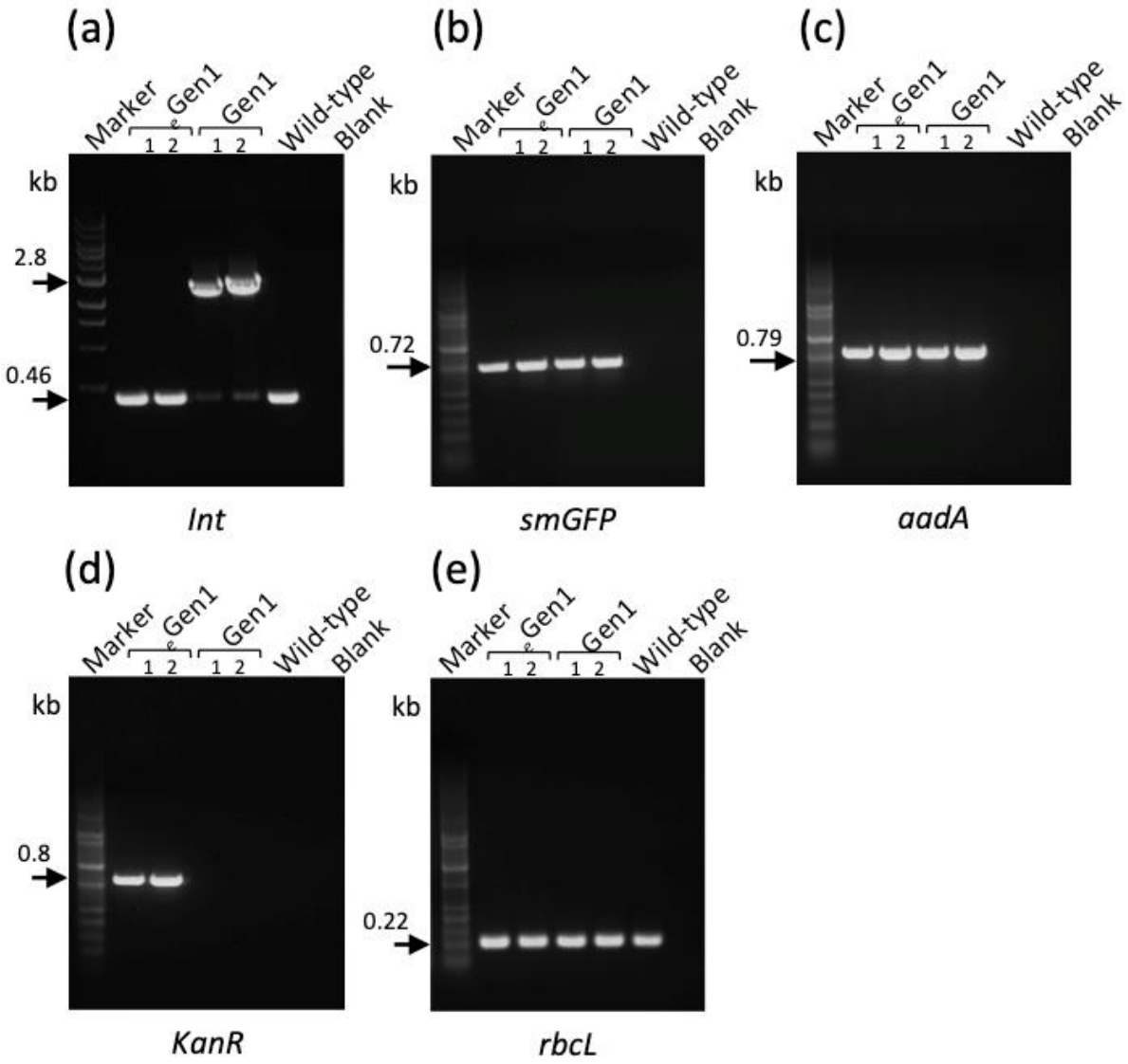
Figure 3.5. Plastid genetic engineering using the mini-synplastome. (a) Different methods of chloroplast transformation. The traditional method of chloroplast transformation is based on the utilization of vectors able to integrate into the plastome by homologous recombination (top image). A new method of chloroplast transformation based on the utilization of non-integrating episomal vectors able to persist as small independent synthetic plastomes, mini-synplastomes (bottom image). The chloroplast genome (plastome), integration vector and mini-synplastomes are indicated in the chloroplast stroma (S). The chloroplast origin of replication *(ori,* in red), thylakoids (T) and chloroplast membranes (M) are also indicated. (b) The design-build-test cycle used to develop the mini-synplastome platform for chloroplast transformation: (1) Biolistic transformation of leaf tissue using chloroplast-specific integration vectors (Gen1) and selection for transplastomic lines without vector integration; (2) Screening for lines containing the episome, $_e$Gen1, and isolation of this plasmid by back-transformation into *E. coli*; (3) Based on sequence analysis informed from $_e$Gen1, Gen1 was used as backbone to synthesize the Gen2 plasmid which was used to transform chloroplasts; (4) Synplastomic lines containing the episome $_e$Gen2 were generated. Homologous arms (yellow), chloroplast *ori* (red), transgene cassette (green) are indicated in the plasmids.



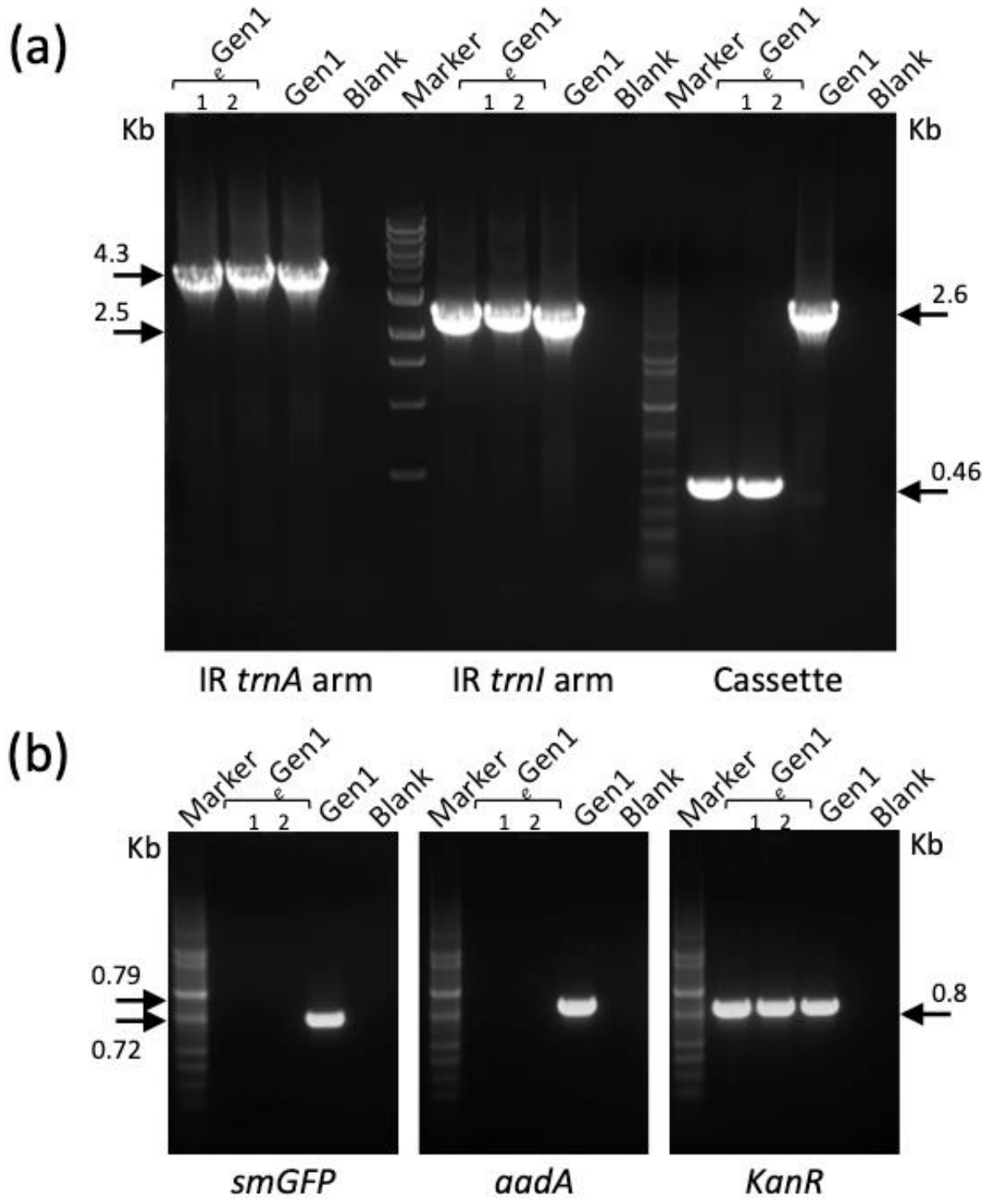
Supplementary Figure 3.1: Screening for putative Gen1 episome-containing plants. (a) Vector integration in the plastome of Gen1 and pSSC-integrating lines. PCRs for IR (*trnI/trnA*) or SSC (*ndhG/ndhI*) regions of the plastome were used to check vector integration in Gen1 and pSSC lines. Gen1-integrating lines originated with another version of the transgene cassette are indicated with asterisk (*). Per each construct, PCR samples from lines 1-15 are shown. DNA bands of 2.8 and 2.6 kb indicate integration of Gen1 and pSSC, respectively. DNA bands of 0.46 and 0.44 kb indicate wild-type (WT) IR and SSC regions of the plastome, respectively. Wild-type (WT) samples, blanks (B) and DNA markers (M; kb) are shown. (b) Presence of backbone vector in different transplastomic lines. PCRs using primers for *KanR* (0.8 kb) indicate the presence of the backbone vector in leaf samples of two putative Gen1 lines containing the episome ($_e$Gen1; line 2 and 7). PCRs using primers for *SpcR* confirm the absence of backbone in pSSC-integrating plants. Positive (Gen1 plasmid; p) and negative (Gen1-integrating lines) controls along with wild-type (WT) samples, blanks (B) and DNA markers (M; kb) are shown.



Supplementary Figure 3.2: Phenotype of transgenic lines and wild-type controls. (a) Images showing 2-week-old *in vitro* transgenic lines along with wild-type controls. pSSC and Gen1-integrating lines, along with a plant containing $_{e}$Gen1 are shown. (b) Confocal images showing smGFP localization into the chloroplast stroma of leaf mesophyll cells from transgenic lines in A. smGFP (green), chlorophyll (Chl., red), bright-field (BF; gray) and merged images are indicated. Scale bars: 10 mm (a); 20 µm (b).

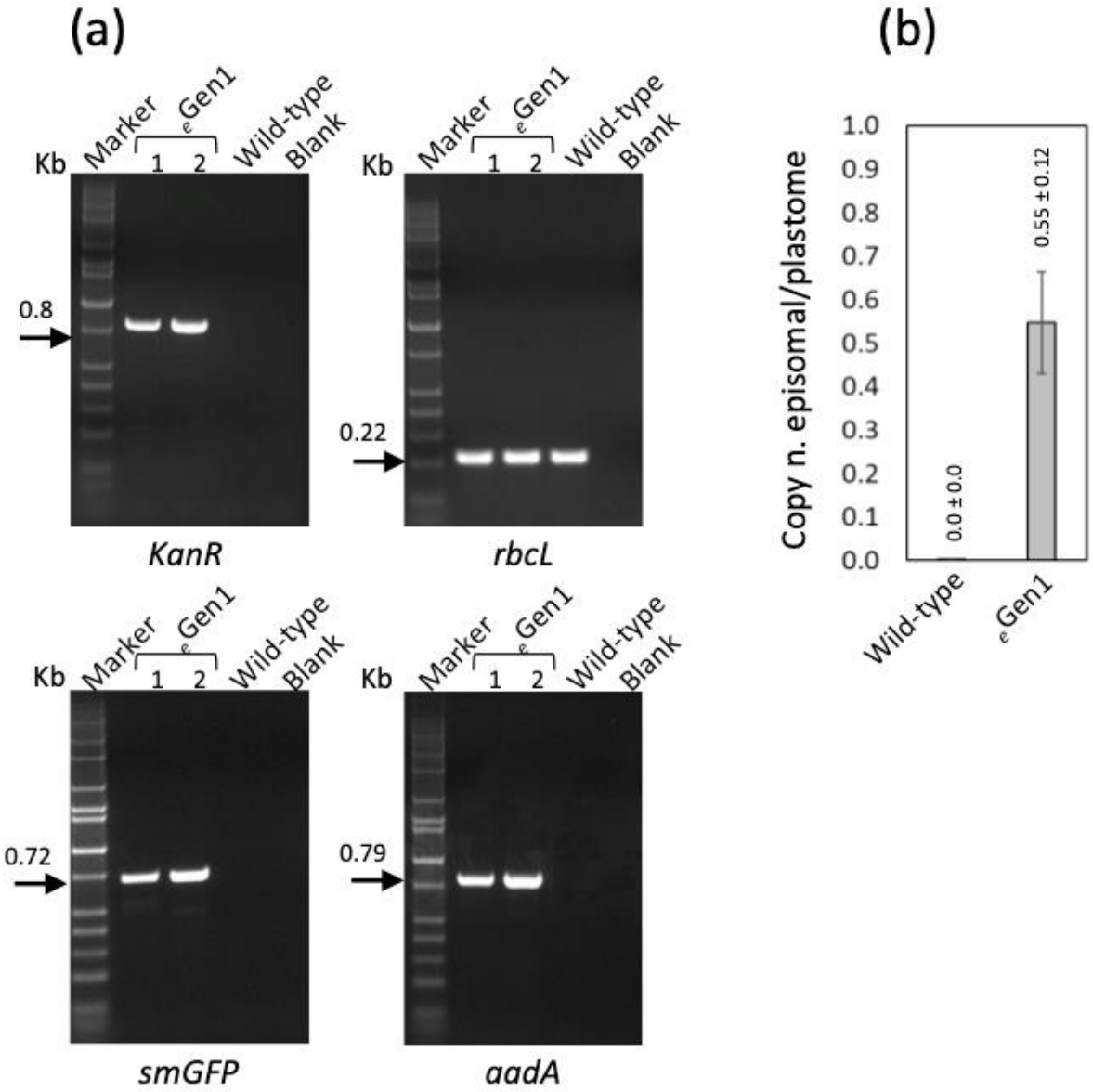


Supplementary Figure 3.3: Characterization of episomal lines containing $_e$Gen1. (a) PCRs to check vector integration (Int.) in the IR (*trnI/trnA*) region of $_e$Gen1-containing and Gen1-integrating lines, respectively. Two independent lines of both genotypes are indicated (1 and 2). Sample 2 corresponds to the $_e$Gen1-containing line 2 (Figure S1) used in this study. DNA-bands at 2.8 kb indicate transgenes integration into the IR (*trnI/trnA*) site, whereas the presence of 0.46 kb bands indicate wild-type IR regions. (b-e) PCRs for s*mGFP* (0.72 kb), *aadA* (0.79 kb), *KanR* (0.8 kb) and *rbcL* fragment (0.22 kb) are also indicated. Wild-type samples and blanks along with DNA molecular markers (kb) are shown.



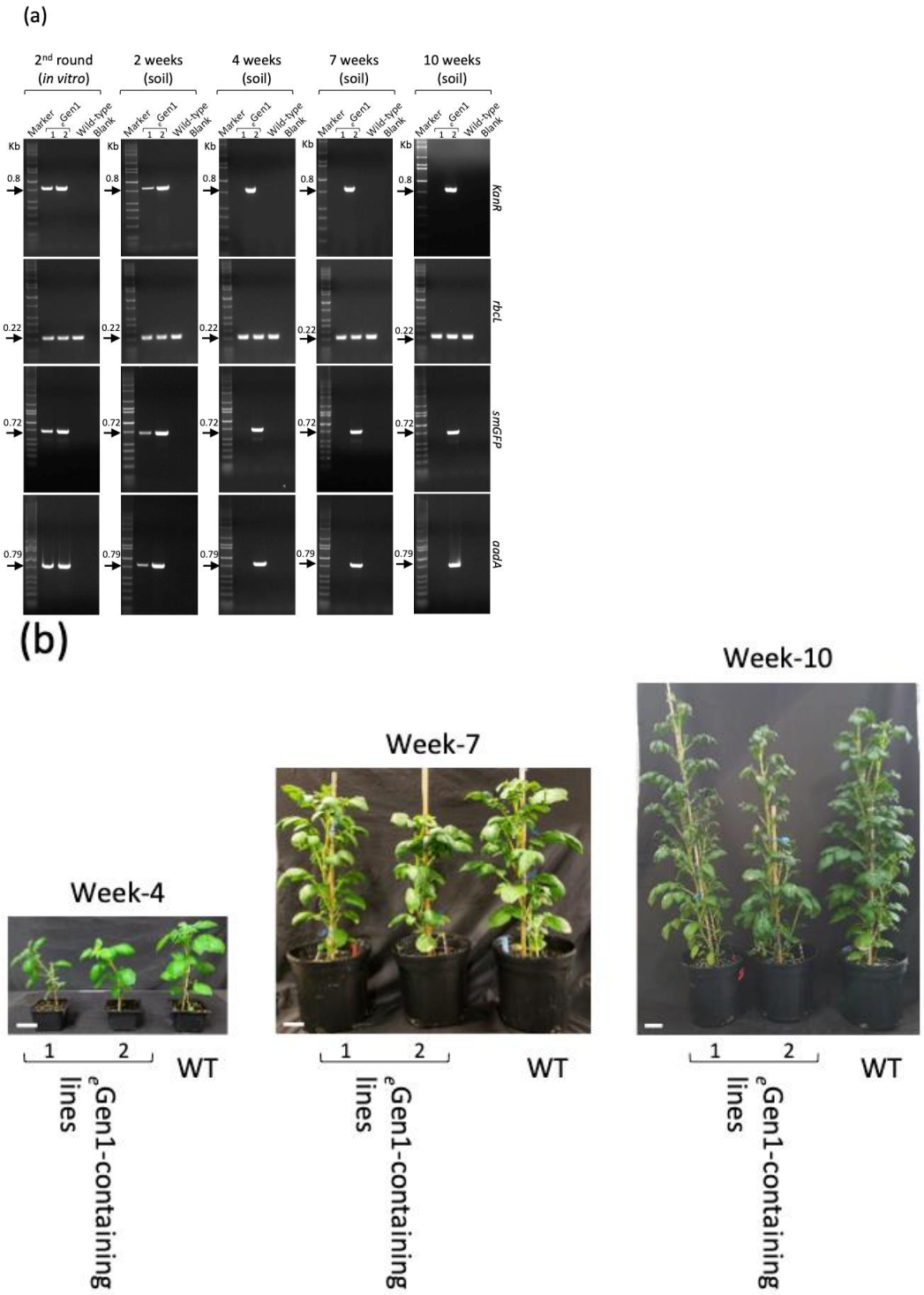
Supplementary Figure 3.4: PCR characterization of $_e$Gen1 plasmids extracted from leaf tissue of episome-containing lines. (a) PCRs using primers for the *trnA* and *trnI* homologous arms along with primers external of the dual-selection cassette were used to characterize the $_e$Gen1 plasmid extracted by back transformation into *E. coli*. Plasmids extracted from two independent

colonies (1 and 2) are indicated. DNA bands of 4.3 and 2.5 kb at the same molecular weight of the positive control, Gen1, indicate the presence of full-length homologous arms in both $_e$Gen1 plasmids. On the contrary, the presence of lower-molecular weight bands of 0.46 kb rather than 2.6 kb (Gen1) indicate removal of the dual selection cassette in $_e$Gen1. (b) PCRs using primers for s*mGFP* (0.72 kb), *aadA* (0.79 kb) and *KanR* gene (0.8 kb) confirmed the absence of the selection cassette and the presence of the backbone vector in $_e$Gen1. The Gen1, was used as positive control for comparison of the molecular weight of DNA bands. The negative controls (blanks) and DNA molecular markers (kb) are also indicate in the gels. These results have been confirmed by sequencing analysis of the entire $_e$Gen1 plasmid.



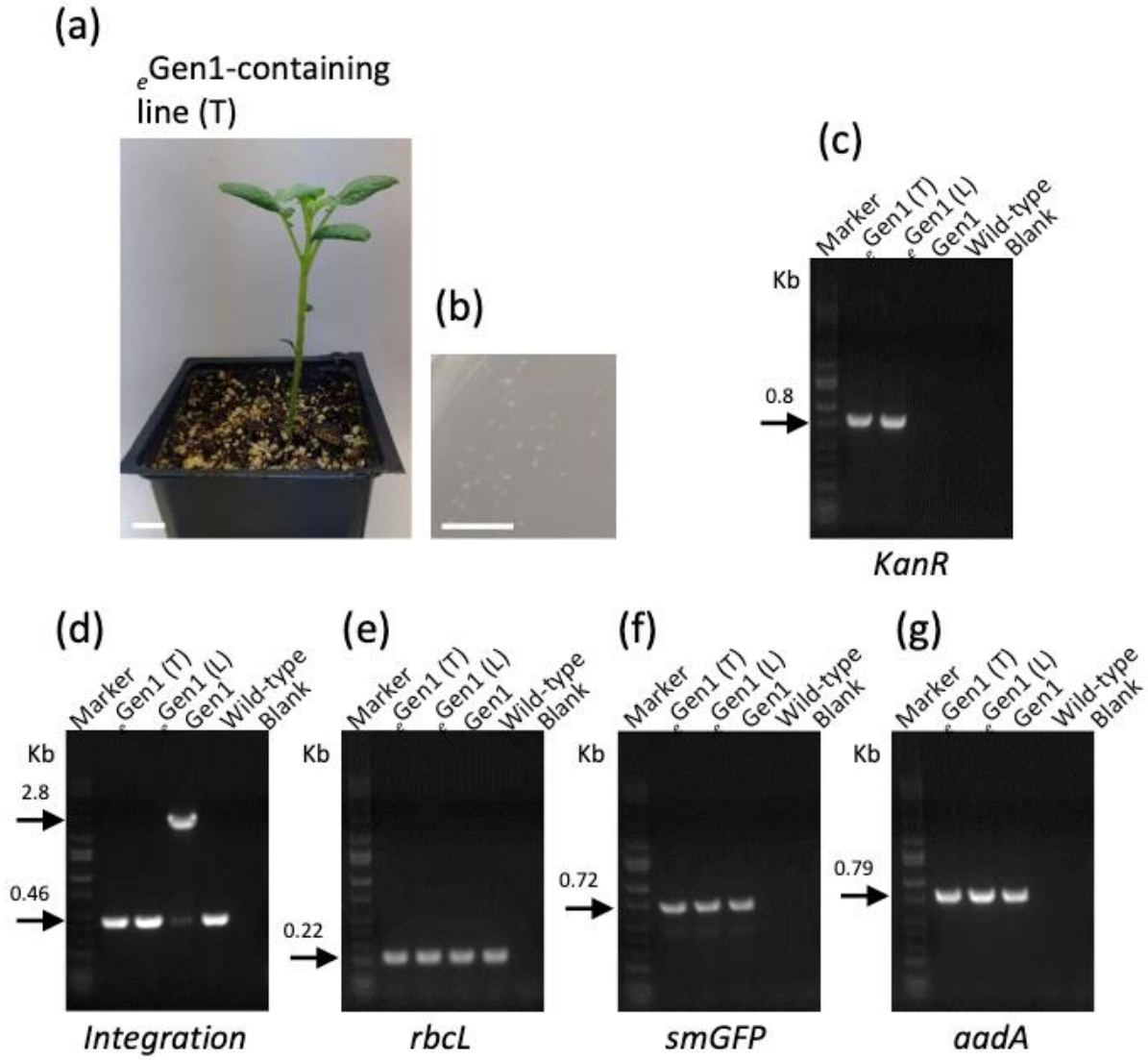
Supplementary Figure 3.5: Determination of the ratio of episome/plastome of $_e$Gen1-containing lines at the 3rd round of tissue culture. (a) PCRs for *KanR* (0.8 kb), *rbcL* fragment (0.22 kb), s*mGFP* (0.72 kb), and *aadA* (0.79 kb) using DNA samples extracted from two independent $_e$Gen1-containing lines (1 and 2) at the third round of tissue culture. The second $_e$Gen1-containing line 2 was the line selected for further study (Figure S1). The PCR profiles confirmed that $_e$Gen1 was stable at this developmental stage. Wild-type controls, blanks and molecular markers (kb) have been included. (b) Graph summarizing the ratio of copy number of episome vs plastome (copy n. episomal/plastome) in genomic DNA preparations of the selected $_e$Gen1-containing line determined by qPCR analysis from leaf tissue of transgenic plants and wild-type controls at the third round of tissue culture. Results are expressed as mean ± standard deviation

(SD) of a total of 3 biological and 4 technical replicates per each biological replicate.

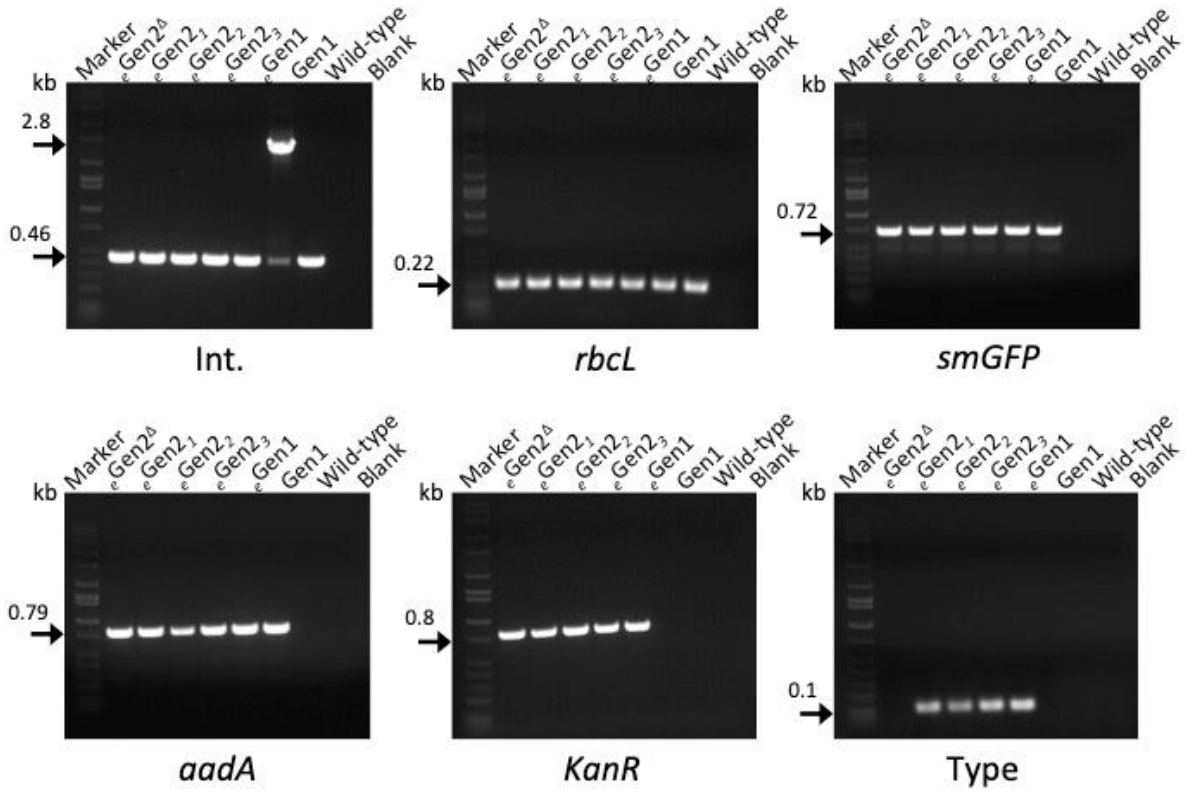


Supplementary Figure 3.6: Stability of $_e$Gen1 episome at different plant developmental stages. (a) PCRs for *KanR* (0.8 kb), *rbcL* fragment (0.22 kb), s*mGFP* (0.72 kb), and *aadA* (0.79 kb) using DNA samples extracted from $_e$Gen1-containing lines at the indicated developmental stages are shown. The PCR profiles confirmed that $_e$Gen1 (Figure S1) is stable throughout all plant

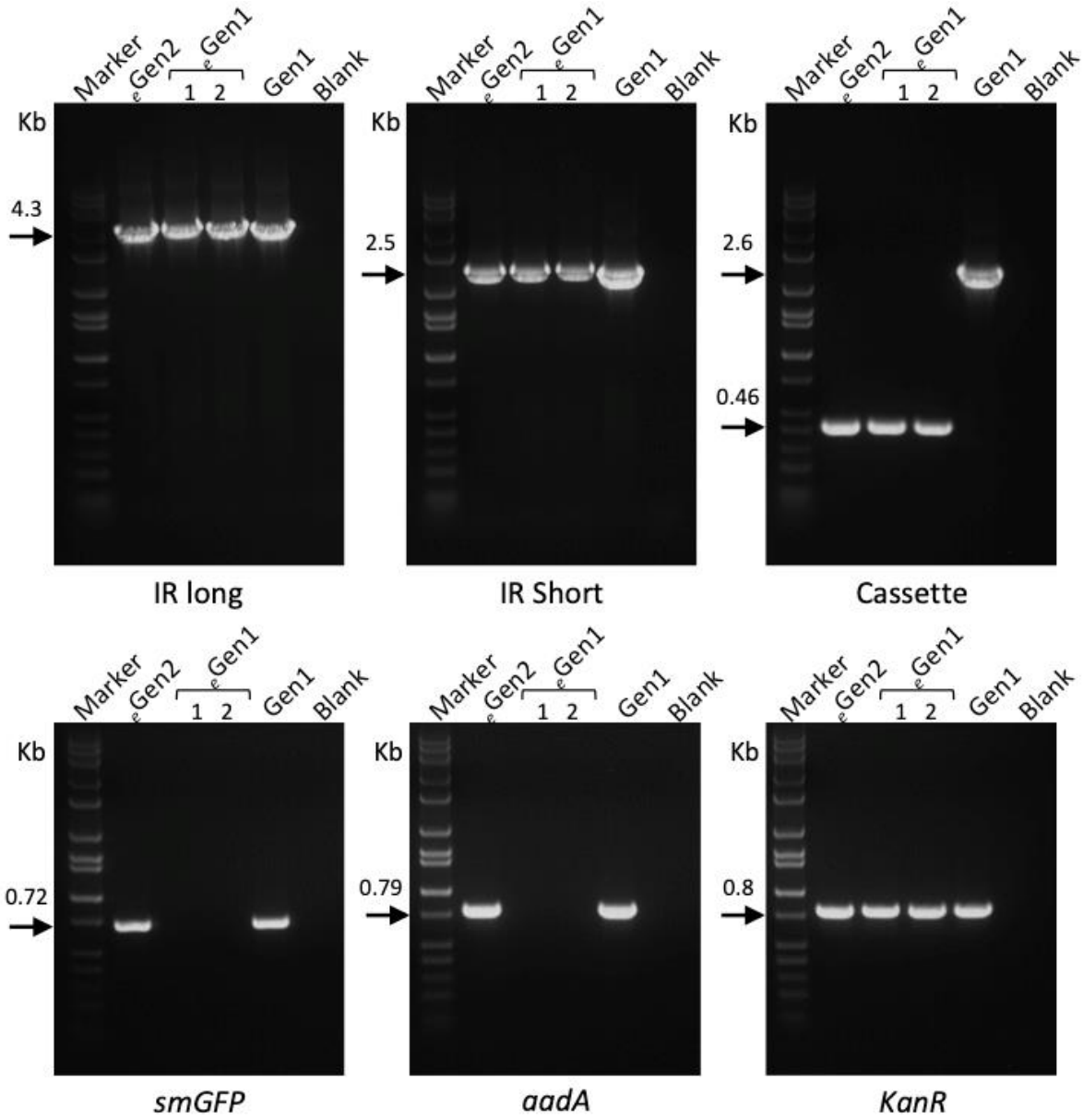
developmental stages. Wild-type samples (WT), blanks and molecular markers (kb) have been included. (b) Representative plants at the second round of tissue culture (2nd round *in vitro*) grown on potting mix for 4, 7 and 10 weeks (anthesis); $_{e}$Gen1-containing lines along with wild-type control plants (WT). Scale bars = 5 cm.



Supplementary Figure 3.7: Characterization of $_e$Gen1-containing lines originated from tubers. (a) A second generation $_e$Gen1-containing plant regenerated from tubers of $_e$Gen1-containing line 2 plants grown on potting mix (line 2, Figure S1). (b) Bacterial colonies transformed with $_e$Gen1 contained in leaf tissue is shown in A. (c-g) PCRs for detection of *KanR* (0.8 kb), integration (2.8 or 0.46 kb bands, for integration or not, respectively), *rbcL* fragment (0.22 kb), s*mGFP* (0.72 kb), and *aadA* (0.79 kb) using DNA samples extracted from $_e$Gen1-containing lines originated from tubers (T) or from tissue culture (L) are shown. A Gen1-integrating line along with wild-type controls, blanks and molecular markers (kb) are also shown. The PCR results confirmed the presence of $_e$Gen1 in plants originating from tubers. Scale bars = 10 mm (a-b).

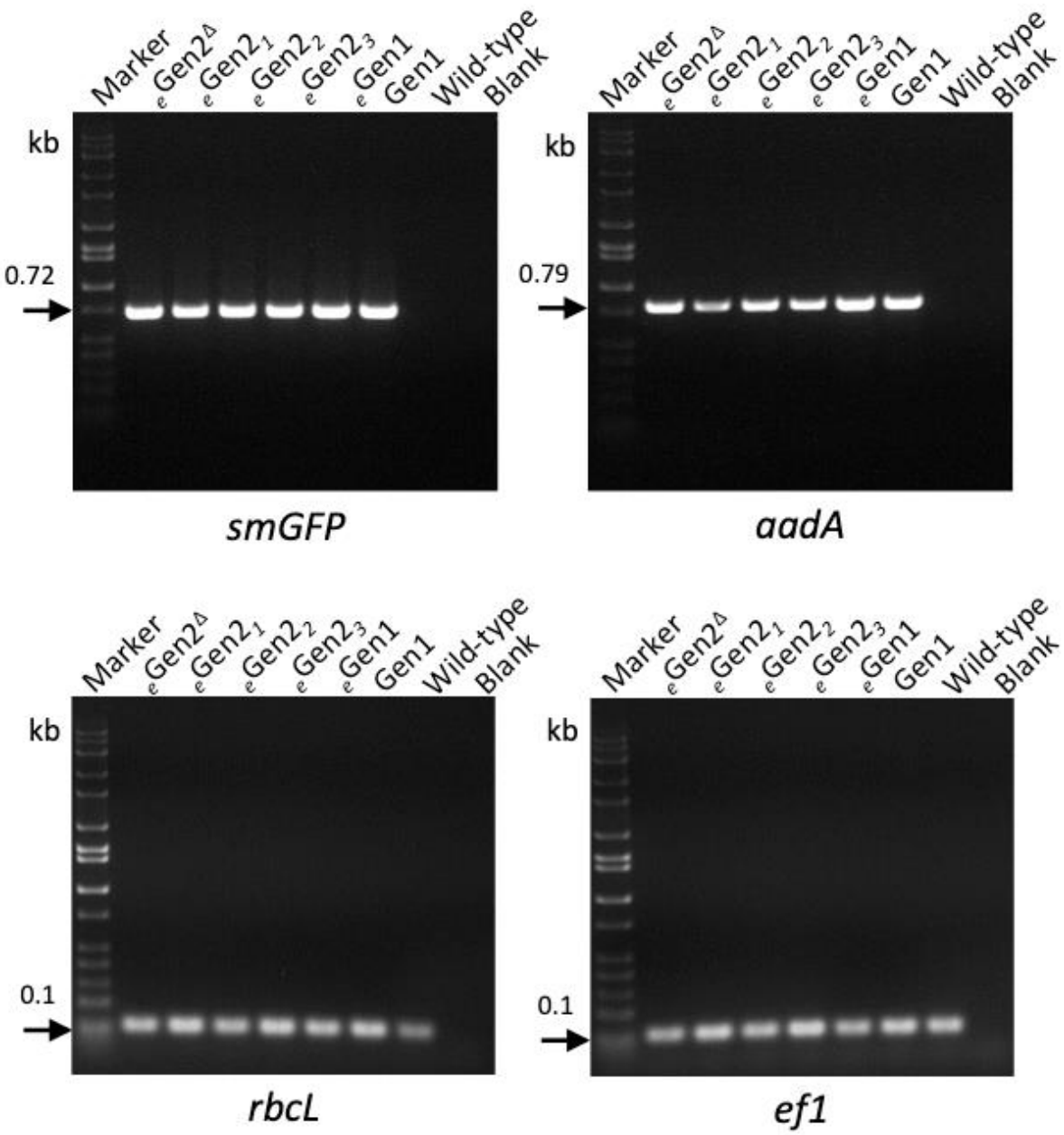


Supplementary Figure 3.8: Characterization of synplastomic $_e$Gen2-containing lines. Molecular characterization of synplastomic $_e$Gen2-containing lines at the first cycle of tissue culture. DNA-bands of 2.8 kb indicate correct integration into the *trnI/trnA* plastome site in a Gen1-integrating plant used as positive control (Int.). Lower-molecular weight bands of 0.46 kb indicate wild-type IR regions of the plastome in all $_e$Gen2-containing lines. PCRs for *rbcL* fragment (0.22 kb), s*mGFP* (0.72 kb), *aadA* (0.79 kb), *KanR* (0.8 kb) and backbone type (Type; 0.1 kb) to detect the presence of the full-length $_e$Gen2 are included. $_e$Gen2 ($_e$Gen2$_{1\text{-}3}$) or $_e$Gen2$^{\Delta}$-containing lines along with the two controls, $_e$Gen1-containing lines and Gen1-integrating lines are shown. Wild-type samples, blanks and molecular markers (kb) are also shown in the gels.

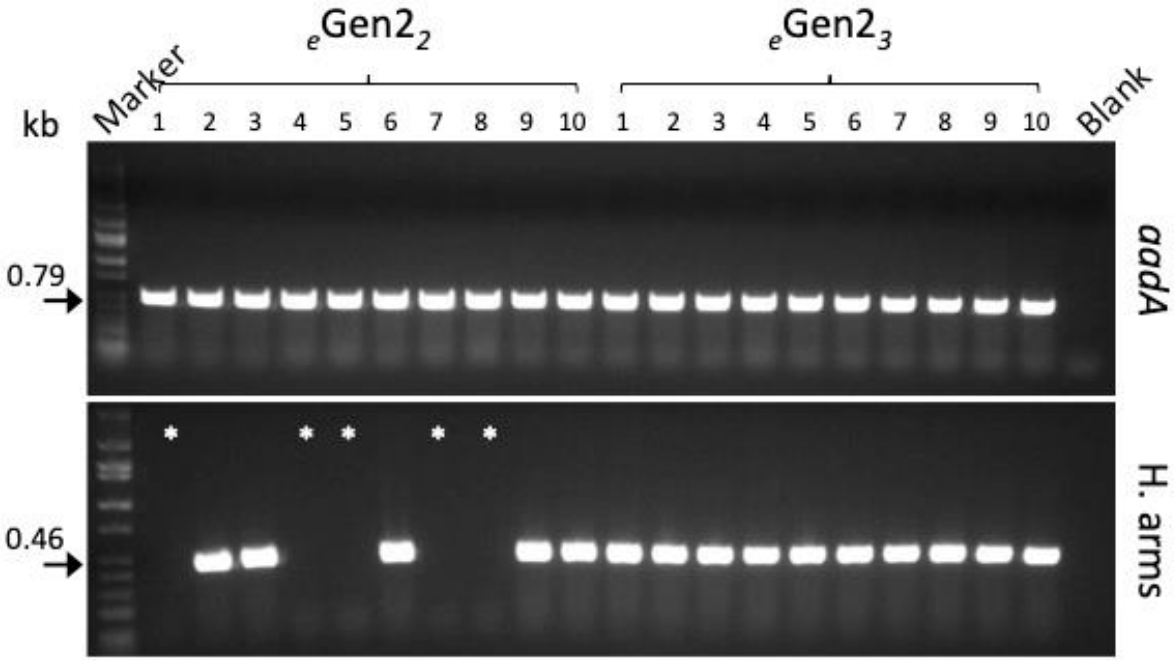


Supplementary Figure 3.9: PCR characterization of $_e$Gen2 extracted from synplastomic plants. PCRs for IR *trnI/trnA* homologous arms (short and long, respectively) along with the dual-selection cassette were used to characterize $_e$Gen2 extracted from synplastomic lines by back transformation into *E. coli*. PCR bands of 4.3 and 2.5 kb at the same molecular weight of the positive control, Gen1, indicate the presence of homologous arms. The presence of lower-molecular weight bands of 0.46 kb rather than 2.6 kb (Gen1) support correct transgene cassette location in the backbone. PCRs for s*mGFP* (0.72 kb), *aadA* (0.79 kb) and *KanR* gene (0.8 kb) confirmed the presence of both the backbone and transgenes. Two bacterial colonies containing $_e$Gen1 and Gen1 were isolated for comparison. The negative controls (blanks) and DNA molecular

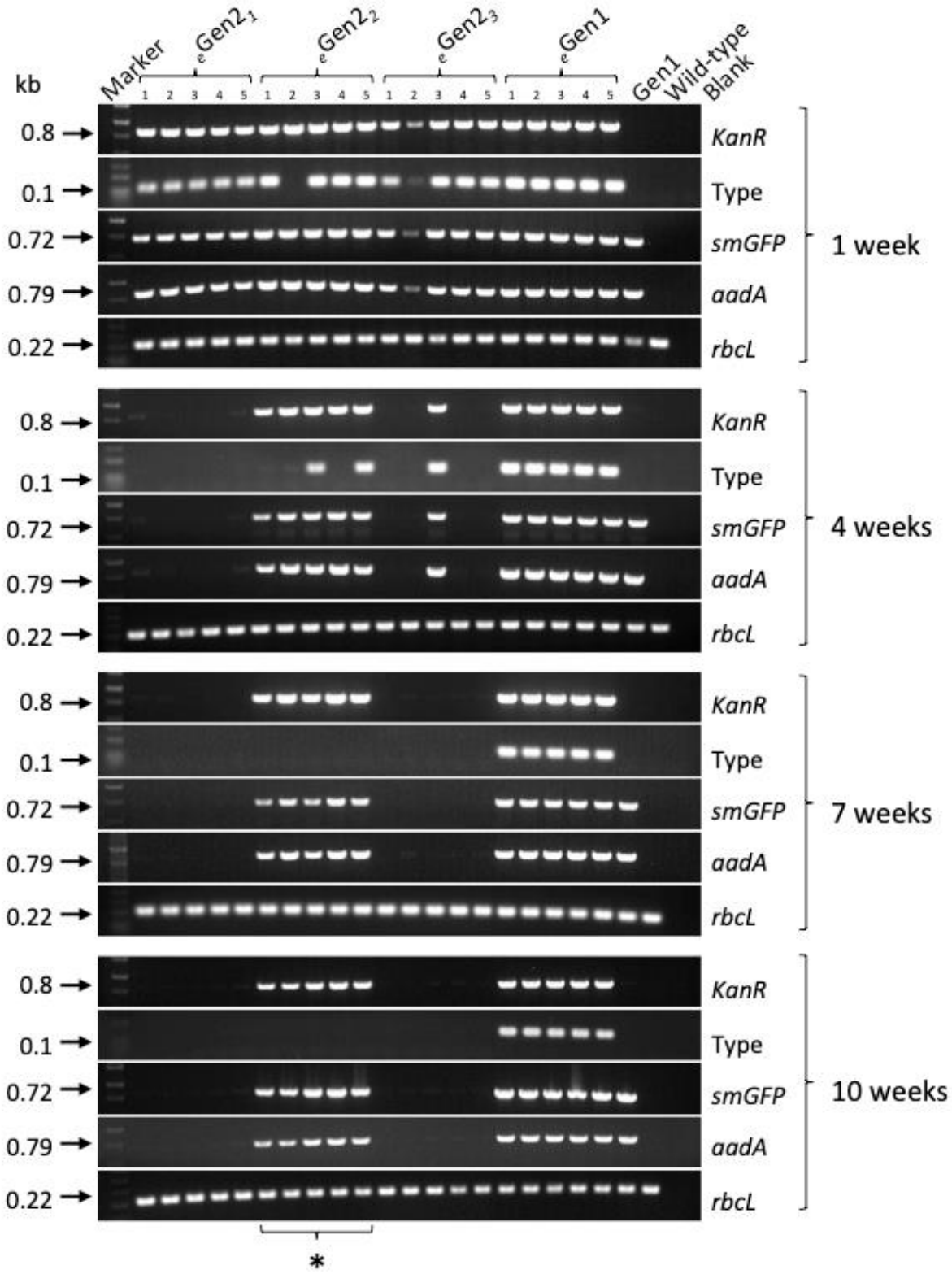
markers (kb) are also indicate in the gels. These PCR results were confirmed by sequence analysis of the entire *e*Gen2.



Supplementary Figure 3.10: Transgene expression in $_e$Gen2-containing lines. RT-PCRs using cDNA preparations from $_e$Gen2$_{1-3}$ and $_e$Gen2$^{\Delta}$-containing lines. $_e$Gen1-containing and a Gen1-integrating line along with wild-type controls and blanks have been included. A pair of primers specific for s*mGFP* and *aadA* were used to detect the full-length cDNAs (0.72 and 0.79 kb, respectively). PCRs for *rbcL* and *ef1* genes were used as loading controls (0.1 kb DNA bands) for the plastome and nuclear genome. Molecular weight markers (kb) are shown in the gel.



Supplementary Figure 3.11: PCRs on bacterial colonies transformed with $_e$Gen2 contained in leaf tissue. For each $_e$Gen2-containing line ($_e$Gen2$_2$ or $_e$Gen2$_3$), a total of ten *E. coli* colonies (1-10) were tested. The presence of the $_e$Gen2 backbone in all bacterial samples was confirmed by PCR-positive products for the *aadA* gene (0.79 kb). The presence of the homologous region of full-length $_e$Gen2 was detected by using primers for an internal region of 0.46 kb (H. arms). Asterisk (*) indicate bacterial samples from $_e$Gen2-containing line 2, which contained the $_e$Gen2$^{\Delta}$ form (~50% of total colonies). Blanks and molecular markers (kb) are shown in the gels.



Supplementary Figure 3.12: Stability of $_e$Gen2 at different plant developmental stages. Molecular characterization of $_e$Gen2-containing lines at

1, 4, 7 and 10 (anthesis) weeks on potting mix. $_{e}$Gen2-containing lines ($_{e}$Gen2$_{1\text{-}3}$), along with $_{e}$Gen1-containing and the Gen1-integrating control lines are shown. Five plants (1-5) per each genotype have been used. Wild-type samples, blanks and molecular markers (kb) are also shown in the gels. PCRs for *KanR* (0.8 kb), minisynplastome type (Type; 0.1 kb) to detect the presence of full-length $_{e}$Gen2 or $_{e}$Gen2$^{\Delta}$, s*mGFP* (0.72 kb), *aadA* (0.79 kb) and *rbcL* fragment (0.22 kb) are included. Asterisk (*) indicate the line able to harbor $_{e}$Gen2$^{\Delta}$ until anthesis (10 weeks on potting mix).

Research overview

In this work we describe several approaches to alter plastid morphology or methods of transgene expression. In the first chapter, we found that large starch granule potato plants could be generated through genome-editing of *FtsZ1*. These plants displayed increased viscosity parameters, though otherwise performed similarly to wild-type. Non-GMO methods of genome-editing are discussed, and these methods could be employed to create similar plants. In the second chapter, we describe the generation of similar macro-plastid plants that can be used for chloroplast transformation. While the efficiency did not increase from the use of these plants, they were amenable to transformation. The installation of large constructs such as synthetic genomes would likely need to be conducted through microinjection, and these plants may be useful for such purposes. The third chapter explores the generation of an alternative chloroplast transformation method. As opposed to transgene integration via homologous recombination, we describe methods to maintain transformation vectors within the organelle separate from the plastome. Plants were generated that expressed transgenes from these vectors, and were maintained once plants were grown outside of tissue culture and selection pressure. Together, these chapters incorporate many similar themes of chloroplast morphology and methods of gene expression. It is our hope that the knowledge gained from these efforts can further plant biotechnology and synthetic biology.

Conclusions

This section discusses conclusions for each research objective and how it addresses the identified gaps identified in the introduction. The objectives are listed once again for reference.

Objective 1. Knockout *FtsZ1* to produce large starch granule potatoes
This study successfully knocked out *FtsZ1* through genome-editing in order to produce large starch granule potatoes. Two of the lines contained starch granules that were up to 1.97-fold larger than wild-type. These two lines grew the same as wild-type after conducting an in depth growth study, and produced the same mass of tubers. Starch was extracted from these lines and characterized to determine viscosity parameters when using it as a thickener. One line showed a 2.07-fold increase in peak viscosity as compared to wild-type. These lines could be produced in an alternative manner where no foreign DNA was integrated into the potato genome. This would allow for rapid release on the market, where producers could theoretically use less of the starch generated from these lines for a variety of production purposes.

Objective 2. Generate macro-chloroplast potato plants for chloroplast transformation

This study successfully generated macro-chloroplast plants for chloroplast transformation. The macro-chloroplast lines did display some fitness penalties, as they were delayed in growth and did not reach the same height as wild-type. Regardless, they were amenable for transformation. While the level of transformation efficiency did not increase from the use of these, they nonetheless expressed transgenes at the same level as wild-type. Organelle specific expression of GFP was quantified and shown to reach the same level as wild-type. Though transformation efficiency from these lines did not increase, they still may be useful when transforming in alternative methods. The sight-directed method of microinjection may require larger plastids to effectively insert the needle into the organelle. Integrating substantial constructs such as synthetic genomes will potentially require microinjection in order to prevent DNA shearing that could occur from the use of a gene gun.

Objective 3. Develop an alternative approach to transform chloroplasts

This study successfully elucidated an alternative way to transform chloroplasts. By incorporating origins of replications on transformation vectors, plants were recovered that contained replicating plasmids within the chloroplast organelle. These plasmids were characterized to determine which regions combined with the native plastome, and which regions persisted independently. The design-build-test mantra of synthetic biology was employed to generate several versions of these persisting plasmids, termed mini-synplastomes. The second generation of design allowed for transgene expression from the mini-synplastomes. These persisted and expressed at various levels, some which were higher than the native plastome. Several advantages of using mini-synplastomes were discussed, one being the generation of marker-free plants to alleviate regulatory concern.

Vita

Alexander C. Pfotenhauer was born in Nashville, Tennessee. He graduated from the University of Tennessee, Knoxville in 2017 with a Bachelor of Science in Food Science. He then attended graduate school at the University of Tennessee, Knoxville for his Master of Science and Doctor of Philosophy in Food Science, concentrating mostly on agricultural synthetic biology.