Genetic analysis of *in vitro* regeneration and methods for enhancing regeneration and gene delivery in lettuce (*Lactuca sativa* L.)

By

TAWNI BULL

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Horticulture and Agronomy
with a Designated Emphasis in Biotechnology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Richard Michelmore, Chair

Luca Comai

Allen Van Deynze

Committee in Charge

2022

Genetic analysis of *in vitro* regeneration and methods for enhancing regeneration and gene delivery in lettuce (*Lactuca sativa* L.)

ABSTRACT

Gene delivery into plant cells has been studied for multiple decades and *Agrobacterium*-mediated transformation is still the most commonly used method. Mostly, his requires *in vitro* regeneration of whole plants; however, *in vitro* regeneration is genotype dependent and multiple elite crops, such as cotton, pepper, and sunflower, remain recalcitrant to regeneration and transformation. The molecular mechanisms of *in vitro* regeneration have been thoroughly investigated in model species such as *Arabidopsis*, and ectopic expression of characterized molecular players has been shown to increase regeneration rates in other species. Lettuce (*Lactuca sativa* L.) follows the *de novo* organogenesis pathway and little effort has been invested in identifying molecular mechanisms important to this mode of regeneration in lettuce. Previously, I have reviewed the molecular determinants of *in vitro* regeneration in model species as well as their potential use to enhance regeneration rates in lettuce. Genetic analysis of regeneration using a recombinant inbred line (RIL) mapping population identified eight quantitative trait loci (QTLs) associated with *de novo* shoot organogenesis in lettuce. Multiple candidate genes with known functions in *de novo* organogenesis and somatic embryogenesis were at or near the peaks of each QTL. Furthermore, *in vitro* regeneration rates in multiple diverse lettuce genotypes were boosted by the introduction of a *GROWTH REGULATING FACTOR* (*GRF*) and *GRF-INTERACTING FACTOR* (*GIF*) gene fusion (GRF-GIF). Co-transformation of a GRF-GIF construct with a gene of interest construct increased recovery of transgenic plants harboring the gene of interest. The use of magnetic nanoparticles as a DNA delivery agent was explored. Results from this work will be used to validate candidate genes, increase regeneration rates, and develop genotype-independent

regeneration for gene delivery in lettuce. This work can be extended to other recalcitrant crops, particularly in the Compositae, such as sunflower.

ACKNOWLEDGEMENTS

It is with immense pride that I am completing my dissertation. With many ups and downs, the journey has taught me a tremendous amount about research, independence, and resilience. Many people have helped me along the way and I owe them more than I can say in the acknowledgements, but I will start with a thank you.

Thank you to my major professor, Richard Michelmore, for the time and effort you spent helping me develop into a better research scientist and writer. I thank you for the amount of time and work you invested in helping me improve my dissertation. I thank you for answering all my questions and for listening to me when I had a nervous breakdown about failed experiments. I am a much better person and scientist after working with you.

Thank you to all the Michelmore lab members that helped me troubleshoot problems, aided with experiments, and cooked delicious food for all the potlucks and celebrations. Thank you to Ali Ahmed for being a mentor and teaching me some of your cloning skills, Rongkui Han for introducing me to the basics of bioinformatics and coding, and Amber Flores for always giving me someone to talk to when I was stressed out. I would like to specially thank all my undergraduate students, David Galkowski, Tran Nyugen, Allissa Tran, Benjamin Walsh, and Dylan Wong. You all helped me tremendously throughout the various experiments presented in my dissertation.

Thank you to my parents, Tom and Melissa Middleton, who have given me the encouragement, support, and opportunities to get this far in my education. I owe them my deepest thanks for always supporting my goals even when I am thousands of miles away from home. They gave me the love and passion I have for agriculture, they taught me the work ethic and

determination it takes to do scientific research, and they provided me with experiences that leave me grateful for where I am in my career and life.

Thank you to my sister, Kelsi Katcher, for always being a great role model, giving me footsteps to follow in, and encouraging me to follow my dreams. Growing up watching you succeed in all aspects of life encouraged me to work hard and invest in myself and in my passions.

Thank you to my husband, Justin Bull, for your constant support, understanding, and patience as I finished my research and dissertation. You got to see first-hand the ups and downs I experienced throughout my studies and supported me through them all. You always listened when I needed to talk about science with someone, even though I know it is your least favorite subject. Your love and encouragement kept me progressing in my research and helped me keep some of my sanity.

Last, thank you to my sweet dog, Kya, for your patience during my long days in the lab and for all the joy you bring into my life

Table of Contents

Chapter 1: Introduction

Genetic analysis of in vitro regeneration and methods for enhancing regeneration and gene delivery in lettuce (*Lactuca sativa* L.)

In vitro plant regeneration is the process of regenerating whole organs and plants through the dedifferentiation and genetic reprogramming of cells. This process can follow multiple pathways, that require intricate molecular, regulatory, and signaling pathways. In tissue culture, plants can regenerate following either the *de novo* organogenesis (e.g., lettuce, tomato, pepper) or somatic embryogenesis (e.g., cotton, rice, wheat) pathways (Heidmann et al., 2011; Leelavathi et al., 2004; Michelmore et al., 1987; Murthy et al., 1996; Ozias-akins and Vasil, 1982; Rueb et al., 1994; Sun et al., 2015). *De novo* organogenesis is the process of developing new, whole organs that did not previously exist. For example, *de novo* shoot organogenesis and *de novo* root organogenesis occur separately forming only shoots and roots, respectively (Bustillo-Avendaño et al., 2018; Shin et al., 2020). Alternatively, somatic embryogenesis occurs when differentiated somatic cells undergo genetic and epigenetic reprogramming to form a bipolar somatic embryo, of which shoot and root regeneration occur simultaneously (von Arnold et al., 2002). Both modes of regeneration can occur directly from existing plant tissues or indirectly from a pluripotent cell mass called a callus.

Lettuce (*Lactuca sativa*) is a dicotyledonous plant that can be regenerated by indirect *de novo* organogenesis and was a model for early studies of *in vitro* regeneration (reviewed in Michelmore and Eash, 1985). As with many species, lettuce regeneration is genotype specific with low regenerating genotypes (*L. serriola* Armenian 999) and high regenerating genotypes (*L. sativa* oil seed PI251246). Lettuce is also amenable to *Agrobacterium*-mediated transformation, for which *in vitro* regeneration is a requirement (Curtis et al., 1994; Kanamoto et al., 2006; Michelmore et al., 1987); therefore, identifying genetic loci important to lettuce regeneration is

desirable to fully benefit from biotechnological approaches. This would also allow for the development of protocols for high efficiency, genotype-independent regeneration of lettuce.

In the past decades, multiple studies have identified molecular determinants for both modes of *in vitro* plant regeneration, although most of this work has been conducted in model species (reviewed in Bull and Michelmore, 2022). By synthesizing and studying the intricate molecular mechanisms and crosstalk between auxin and cytokinin signaling required for *in vitro* regeneration of other species, we can start to untangle the underlying mechanism of *de novo* shoot organogenesis in species such as lettuce. Further understanding the known molecular players of other species gives us potential targets for improving regeneration rates of recalcitrant lettuce genotypes. **Chapter 2** of this dissertation, entitled "Molecular determinants of *in vitro* plant regeneration: prospects for enhanced manipulation of lettuce (*Lactuca sativa* L.)," gives a detailed synopsis of the key molecular determinants of indirect *de novo* organogenesis and somatic embryogenesis identified in model species. In addition, **Chapter 2** also discusses how we can apply the knowledge from model species for the improvement of indirect *de novo* organogenesis in lettuce and recalcitrant related species.

Genetic analyses of quantitative trait loci (QTLs) have identified loci and candidate genes involved during *de novo* organogenesis and somatic embryogenesis of several species including *Arabidopsis*, cucumber, tomato, soybean, wheat, rice, and maize (Kwon et al., 2002; Lall et al., 2004; Ma et al., 2016; Motte et al., 2014; Salvo et al., 2018; Trujillo-Moya et al., 2011; Wang et al., 2018; Yang et al., 2011). Mapping populations segregating for rates of regeneration in lettuce are available (Han, 2021a; Han, et al., 2021b). In addition, the development of high quality and high-resolution genetic maps allows us to identify smaller confidence intervals for genetic loci and helps narrow down candidate gene searches (Han et al., 2021b). In **Chapter 3**, "Genetic analysis

of indirect *de novo* organogenesis in lettuce (*L. sativa* L.)", I describe the varying regeneration rates of a recombinant inbred line (RIL) population derived from an interspecific cross between a low regenerating parent (*L. serriola* accession Armenian 999) and a high regenerating parent (*L. sativa* oil seed PI251246). In addition, quantitative trait loci (QTLs) mapping was used to identify significant QTLs associated with multiple regeneration traits in lettuce. Identification and validation of candidate genes will provide targets for the enrichment of *in vitro* regeneration in lettuce.

Manipulation of transcription factors and genes required for successful *in vitro* regeneration can improve regeneration rates and transformation efficiencies of recalcitrant species (Debernardi et al., 2020; Jones et al., 2019; Kong et al., 2020). Altered expression of key transcription factors required for proper meristem development, such as *BABY BOOM* (*BBM*), *WUSCHEL* (*WUS*), and *SHOOT MERISTEMLESS* (*STM*), has increased the rates of *de novo* shoot organogenesis and somatic embryogenesis in multiple species (Boutilier et al., 2002; Gallois et al., 2002; Heidmann et al., 2011; Mayer et al., 1998; Zhao et al., 2002). In addition, *Agrobacterium*-mediated transformation of *GROWTH REGULATING FACTORs* (*GRFs*) increased regeneration in sugar beet, soybean, and sunflower (Kong et al., 2020). Also, introduction of chimeric transgenes composed of *GRFs* fused with its cofactor *GRF-INTERACTING FACTOR* (*GIF*) (*GRF-GIF*) leads to higher rates of regeneration and transformation (Debernardi et al., 2020). In **Chapter 4**, "GRF-GIF chimeric proteins enhance *in vitro* regeneration and *Agrobacterium*-mediated transformation efficiencies of lettuce (*Lactuca* spp.)," we analyze the efficacy of GRF-GIF chimeric transgenes from multiple species for the improvement of regeneration and transformation efficiencies in lettuce. Furthermore, we present a strategy for increasing the rate of transformation

and recovery of transgenic plants harboring a gene of interest. The results from this chapter will continue to be used for the improvement of regeneration and transformation efficiencies in lettuce.

Although *in vitro* regeneration and *Agrobacterium*-mediated transformation is routinely used for the development of transgenic plants, tissue culture is time consuming and labor intensive. Developing methods of gene delivery directly into germline cells would allow for faster recovery of transgenic or genome edited plants without going through tissue culture systems. One potential approach for improving gene delivery is through the use of nanoparticles. Infiltration of carbon nanotubes loaded with plasmid DNA into leaves showed expression of fluorescent reporters without transgene integration (Demirer et al., 2019). Optimization of this method would allow for delivery of genes required for genome editing without stable incorporation of the transgene. In **Chapter 5** of this dissertation, "Fast fluorescent titration quantification of plasmid DNA with DNA attractive magnetic nanoparticles," we develop a new method for loading DNA of large plasmids onto magnetic nanoparticles (MNs). Optimization of infiltration of MNs for gene delivery into intact plants would allow rapid introduction of genes with less stringent cargo capacity limits.

The knowledge of genetic loci and genes affecting regeneration in lettuce gained from this research can enhance regeneration and transformation rates of other recalcitrant genotypes (e.g., Mainspring, Oak Leaf, Royal Oak Leaf, Sangria) through the overexpression or genome editing of these specific genes. In addition, ectopic expression of identified regeneration candidate genes could increase regeneration efficiency of other recalcitrant species of the Compositae family, such as sunflower. A better understanding of the genetic loci and molecular mechanisms that underlie *in vitro* regeneration will help form a more generalized view of this process, which can be applied to other important crops of the Compositae family.

REFERENCES

Boutilier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C. M., van Lammeren, A. A. M., Miki, B. L. A., Custers, J. B. M., and van Lookeren Campagne, M. M. (2002). Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth. *Plant Cell*, *14*(8), 1737–1749.

Bull, T., and Michelmore, R., (2022). Molecular Determinants of *in vitro* Plant Regeneration: Prospects for Enhanced Manipulation of Lettuce (*Lactuca sativa* L.). Front. Plant Sci. 13:888425.

Debernardi, J. M., Tricoli, D. M., Ercoli, M. F., Hayta, S., Ronald, P., Palatnik, J. F., and Dubcovsky, J. (2020). A GRF–GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nature Biotechnology 2020 38:11*, *38*(11), 1274–1279.

Demirer, G. S., Zhang, H., Goh, N. S., González-Grandío, E., and Landry, M. P. (2019). Carbon nanotube-mediated DNA delivery without transgene integration in intact plants. *Nature Protocols*, *14*, 2954–2971.

Gallois, J.-L., Woodward, C., Reddy, G. V., and Sablowski, R. (2002). Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in Arabidopsis. *Development*, *129*(13), 3207–3217.

Han, R., Lavelle, D., Truco, M. J., and Michelmore, R. (2021a). Quantitative Trait Loci and Candidate Genes Associated with Photoperiod Sensitivity in Lettuce (*Lactuca* spp.). *Theoretical and Applied Genetics*, *134*(10), 3473–3487.

Han, R., Wong, A. J. Y., Tang, Z., Truco, M. J., Lavelle, D. O., Kozik, A., Jin, Y., and Michelmore, R. W. (2021b). Drone phenotyping and machine learning enable discovery of loci regulating daily floral opening in lettuce. *Journal of Experimental Botany*, *72*(8), 2979–2994.

Heidmann, I., de Lange, B., Lambalk, J., Angenent, G. C., and Boutilier, K. (2011). Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. *Plant Cell Reports*, *30*(6), 1107–1115.

Jones, T., Lowe, K., Hoerster, G., Anand, A., Wu, E., Wang, N., Arling, M., Lenderts, B., and Gordon-Kamm, W. (2019). Maize Transformation Using the Morphogenic Genes . *Methods in Molecular Biology*, *1864*, 81–93.

Kong, J., Martin-Ortigosa, S., Finer, J., Orchard, N., Gunadi, A., Batts, L. A., Thakare, D., Rush, B., Schmitz, O., Stuiver, M., Olhoft, P., and Pacheco-Villalobos, D. (2020). Overexpression of the Transcription Factor GROWTH-REGULATING FACTOR5 Improves Transformation of Dicot and Monocot Species. *Frontiers in Plant Science*, *11*.

Kwon, Y. S., Kim, K. M., Eun, M. Y., and Sohn, J. K. (2002). QTL mapping and associated

marker selection for the efficacy of green plant regeneration in anther culture of rice. *Plant Breeding*, *121*(1), 10–16.

Lall, S., Nettleton, D., DeCook, R., Che, P., and Howell, S. H. (2004). Quantitative Trait Loci Associated With Adventitious Shoot Formation in Tissue Culture and the Program of Shoot Development in *Arabidopsis*. *Genetics*, *167*(4), 1883–1892.

Ma, J., Deng, M., Lv, S. Y., Yang, Q., Jiang, Q. T., Qi, P. F., Li, W., Chen, G. Y., Lan, X. J., and Wei, Y. M. (2016). Identification of QTLs associated with tissue culture response of mature wheat embryos. *SpringerPlus*, *5*(1), 1–7.

Mayer, K. F. X., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T. (1998). Role of WUSCHEL in Regulating Stem Cell Fate in the Arabidopsis Shoot Meristem. *Cell*, *95*(6), 805–815.

Michelmore, R. W. and Eash, J. A. (1985). "Lettuce," in the Handbook of Plant Cell Culture vol 4, ed. D.A. Evans, W.R. Sharp, and P.V. Ammirato, 512-550.

Motte, H., Vercauteren, A., Depuydt, S., Landschoot, S., Geelen, D., Werbrouck, S., Goormachtig, S., Vuylsteke, M., and Vereecke, D. (2014). Combining linkage and association mapping identifies *RECEPTOR-LIKE PROTEIN KINASE1* as an essential Arabidopsis shoot regeneration gene. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(22), 8305–8310.

Salvo, S., Cook, J., Carlson, A. R., Hirsch, C. N., Kaeppler, S. M., Kaeppler, H. F., Salvo, S., Carlson, A. R., Kaeppler, S. M., and Kaeppler, H. F. (2018). Genetic Fine-Mapping of a Quantitative Trait Locus (QTL) Associated with Embryogenic Tissue Culture Response and Plant Regeneration Ability in Maize (*Zea mays* L.). *The Plant Genome*, *11*(2), 170111.

Trujillo-Moya, C., Gisbert, C., Vilanova, S., and Nuez, F. (2011). Localization of QTLs for in vitro plant regeneration in tomato. *BMC Plant Biology*, *11*.

Wang, Y., Zhou, Q., Zhu, G., Wang, S., Ma, Y., Miao, H., Zhang, S., Huang, S., Zhang, Z., and Gu, X. (2018). Genetic analysis and identification of a candidate gene associated with in vitro regeneration ability of cucumber. *Theoretical and Applied Genetics*, *131*(12), 2663–2675.

Yang, C., Zhao, T., Deyue, Y., and Gai, J. (2011). Mapping QTLs for Tissue Culture Response in Soybean (*Glycine max* (L.) Merr.). *Molecules and Cells*, *32*, 337–342.

Zhao, Q.-H., Fisher, R., and Auer, C. (2002). Developmental phases and *STM* expression during Arabidopsis shoot organogenesis. *Plant Growth Regulation*, *37*, 223–231.

Chapter 2: Molecular determinants of *in vitro* plant regeneration: prospects for enhanced manipulation of lettuce (*Lactuca sativa* L.)

ACKNOWLEDGMENTS

This work was previously accepted for publication as a review article:

Bull, T., and Michelmore, R., (2022). Molecular Determinants of *in vitro* Plant Regeneration: Prospects for Enhanced Manipulation of Lettuce (*Lactuca sativa* L.). Front. Plant Sci. 13:888425.

ABSTRACT

In vitro plant regeneration involves dedifferentiation and molecular reprogramming of cells in order to regenerate whole organs. Plant regeneration can occur via two pathways, *de novo* organogenesis and somatic embryogenesis. Both pathways involve intricate molecular mechanisms and crosstalk between auxin and cytokinin signaling. Molecular determinants of both pathways have been studied in detail in model species, but little is known about the molecular mechanisms controlling *de novo* shoot organogenesis in lettuce. This review provides a synopsis of our current knowledge on molecular determinants of *de novo* organogenesis and somatic embryogenesis with an emphasis on the former as well as provides insights into applying this information for enhanced *in vitro* regeneration in non-model species such as lettuce (*Lactuca sativa* L.).

INTRODUCTION

Plants have evolved a remarkable ability to regenerate tissues from differentiated organs, which involves the conversion of one cell type to others. Such plasticity provides the ability to regenerate whole organs and plants via dedifferentiation of cells and reprogramming of cell fates. There are three main types of regeneration: 1. Tissue regeneration, 2. *de novo* organogenesis, and 3. somatic embryogenesis (Sugimoto et al., 2019; Xu and Huang, 2014). Bryophytes have high capacity for tissue regeneration; for example, *Marchantia* spp. are capable of regenerating new meristems within their thallus (Yasui et al. 2019). However, vascular plants follow different regeneration pathways, which include *de novo* organogenesis or somatic embryogenesis (Figure 1). *De novo* organogenesis involves the regeneration of whole organs that did not previously exist. There are two types of *de novo* organogenesis: direct and indirect regeneration. Direct regeneration involves the development of organs directly from explants, whereas indirect regeneration involves an intermediate undifferentiated callus phase. For example, some plants, such as *Jatropha curcas* and succulents of the Cactaceae and Crassulaceae families (Preece, 2003; Severino et al., 2011), are capable of direct regeneration of new roots and shoots from stem cuttings. In contrast, many plants, such as lettuce, exhibit indirect organogenesis and regenerate shoots from calli (Michelmore and Eash, 1985). Somatic embryogenesis involves the regeneration of embryo or embryo-like structures from somatic cells, which can develop into a whole plant. In all forms of regeneration, cells must undergo dedifferentiation or transdifferentiation (reprogramming) into a more totipotent cell, ultimately changing the fate of the progenitor cell.

Plant tissue culture and totipotency were first proposed by Haberlandt in 1902 (Krikorian and Berquam, 2008; Thorpe, 2007), who attempted to culture isolated photosynthetic leaf cells. Although this proved unsuccessful, it was the start of many decades of work on developing and

improving plant tissue culture methods for multiple plant species. These failed experiments led to the development of root cultures using root tip cells in tomato and bud cultures. In 1904, embryo culture was first successful when embryos of crucifers (Brassicaceae) were isolated aseptically and grown in culture (Norstag, 1979). The first "true" plant tissue cultures were obtained on Knop's medium from cambial tissues of sycamore maple (*Acer psuedoplatanus*) by Gautheret in 1934. This approach was optimized by additions of auxin, indole acetic acid (IAA), and B vitamins. This resulted in tissues that could be grown indefinitely in culture and the regeneration of both roots and shoots (Gautheret, 1934, 1935, 1939). The previous studies, however, used explant tissues that already contained meristematic cells. It was not until 1948 that methods were developed to induce roots and shoots from non-meristematic explants (Skoog and Tsui, 1948). This drastically increased the number of species that could be studied using *in vitro* culture systems (Miller et al., 1955; Skoog and Miller, 1957), and led to the recognition of the importance of exogenous ratios of cytokinin and auxin in culture medium. The differing ratios were shown to affect cell fate transition to either rooting or shooting from callus cells (Skoog and Miller, 1957), where high ratios of auxin to cytokinin promoted root regeneration, low ratios of auxin to cytokinin promoted shoot regeneration, and intermediate levels promoted proliferation of callus tissues. From the early to mid-1900s, research helped develop common plant tissue culture methods and media still used today (Gautheret, 1942, 1955; Nobe´court, 1955; van Overbeek et al., 1941). The earliest plant tissue culture media were based on nutrient necessities of whole plants, with the most common being Knop's solution (Loomis and Schull, 1937). Numerous studies were conducted to optimize culture medium and in 1962, Murashige and Skoog reported a medium (MS salts) containing salt concentrations 25 times higher than those in Knop's solution; in particular this resulted in much higher concentrations of NO3- and NH4+. The development of MS salts is still considered to be

a major breakthrough in tissue culture because MS salts are still commonly used in plant tissue culture. The combination of exogenous plant hormones and appropriate salts allowed the study of basic plant biology questions about cell behavior, genetic improvement, disease biology, germplasm conservation, and clonal propagation.

Plant tissue culture to achieve *in vitro* regeneration was originally used to answer fundamental questions in plant biology but has since evolved to be foundational for genetic improvement, micropropagation, genetic engineering, and biotechnology (Chokheli et al., 2020; Loberant and Altman, 2010; Michelmore et al., 1987; Xu and Huang, 2014; Zhang et al., 2006). However, *in vitro* regeneration is not possible for all plant species and regeneration is very genotype dependent. Therefore, studying the molecular determinants of plant regeneration and exploiting these signaling pathways for improved *in vitro* regeneration of those recalcitrant genotypes and species is important. This review provides a synopsis of our current understanding of the pathways involved in *de novo* organogenesis and somatic embryogenesis. We focus on what is known of the molecular determinants of indirect *de novo* shoot organogenesis, which is the mode of regeneration in lettuce (*Lactuca sativa* L.). Finally, we describe future directions for improvement of *in vitro* regeneration of lettuce and other Compositae species.

MOLECULAR DETERMINANTS OF REGENERATION

Recently, many advances have been made towards understanding the cellular and molecular mechanisms that underlie plant regeneration (Ikeuchi et al., 2016; Sugimoto et al., 2019; Xu and Huang, 2014). Each of the regeneration processes described above have been studied in detail in model plants such as *Arabidopsis thaliana*. Each process entails a complex of molecular

players involved in signaling and developmental pathways that regulate the dedifferentiation (somatic embryogenesis) or reprogramming (*de novo* organogenesis) of cells.

Organogenic callus formation

Callus formation is the first step in indirect organogenesis. Based on morphology, calli are thought to be the result of the dedifferentiation of cells to form totipotent cells. Callus can originate from the initiation of multiple pathways that contain some overlap in gene expression (Fehér, 2019)and can be auxin or wound induced (Fehér, 2019). In *Arabidopsis*, auxin induced calli resemble pluripotent cells similar to root tip cells at the molecular level and originate from pluripotent pericycle cells located adjacent to xylem poles (Atta et al., 2009; Fehér, 2019; Sugimoto et al., 2010). Root cell-like, auxin-induced callus follows a similar pathway as lateral root formation. In contrast, wound-induced callus does not involve players of lateral root formation, but rather occurs via upregulation of cytokinin signaling (Ikeuchi et al., 2017; Iwase, Ohme-Takagi, et al., 2011). Due to the similarity of gene expression patterns during callus formation with other developmental pathways some consider it a form of transdifferentiation rather than dedifferentiation (Fehér, 2019).

Many genes and transcription factors that are involved in lateral root development are also critical players in auxin-induced callus formation (Figure 2). For example, the *LATERAL ORGAN BOUNDARIES* (*LBD*) family of genes, such as *LBD16*, *17*, *18*, and *29*, are critical to both lateral root formation and callus production (Fan et al., 2012; Feng et al., 2012; Lee H.W. et al., 2019; Xu et al., 2012). Ectopic expression of *LBD* genes led to the spontaneous formation of callus without exogenous applications of auxin and cytokinin, and repression of *LBD16* showed inhibited callus formation (Fan et al., 2012). In lateral root formation, *LBD16* and *LBD29* are positively

regulated by AUXIN RESPONSE FACTOR7 (ARF7) and ARF19, which provides evidence that *ARFs* are also involved in callus formation (Okushima et al., 2007). Furthermore, JUMONJI C DOMAIN CONTAINING PROTEIN 30 (JMJ30) interacts with ARF7 and ARF19 and directly binds to cis elements of *LBD16* and *LBD29*, promoting their expression (Lee et al., 2018). Other key players in both lateral root and callus formation are *ABERRANT LATERAL ROOT FORMATION 4* (*ALF4*) and *SOLITARY ROOT/IAA14* (*SLR/IAA14*). *ALF4* is involved in the earliest divisions of pericycle cells during lateral root formation. In *alf4* mutants, callus-forming capability was lost in multiple tissues (DiDonato et al., 2004; Sugimoto et al., 2010). It was later shown that *ALF4* is targeted for downregulation by CALLUS FORMATION RELATED-1 (CRF-1), which encodes an enzyme involved in very long chain fatty acid (VLCFA) biosynthesis (Shang et al., 2016). Another molecule involved in VLCFA biosynthesis is the AP2 transcription factor, PUCHI, which is also a key regulator controlling cell proliferation in lateral root primordia; *puchi-1* mutants resulted in both defective and disorganized lateral root and callus formation further indicating a link between these pathways (Trinh et al., 2019). SLR is a member of the auxin signaling protein family Aux/IAA, and *slr-1* mutants in *A. thaliana* were defective in both lateral root and callus formation (Fukaki et al., 2002; Shang et al., 2016). The functions of these genes and transcription factors provides evidence that callus formation and lateral root development have similar underlying mechanisms.In addition, callus formation can be initiated via a wound-induced signaling pathway and activation of a cytokinin response. Transcription factors involved during wound-induced callus formation include APETALA2/Ethylene Responsive Factor (AP2/ERF)-type transcription factors, WOUND-INDUCED DEDIFFERENTIATION1 (WIND1), and homologs (Iwase et al., 2013; Iwase, Mitsuda, et al., 2011; Iwase, Ohme-Takagi, et al., 2011). In Arabidopsis, expression of *WIND1* and homologs are upregulated upon wounding

and promote pluripotent callus formation at cut sites (Iwase et al., 2011a; Iwase et al., 2011b). Expression of *Arabidopsis WIND1* was also shown to induce callus formation in other species such as rapeseed, tomato, and tobacco (Iwase et al., 2013). A transcriptome analysis showed WIND1 activates over 2,000 genes involved in multiple pathways including wound-induced cellular reprogramming and vascular formation (Iwase et al., 2021).

Among the genes upregulated by WIND1 are those encoding for other AP2/ERF-type transcription factors including *PLETHORA* (PLT) genes (Iwase et al., 2021, Kareem et al., 2015). *PLT* genes work through the auxin signaling pathway, are often transcribed in response to auxin accumulation, and are activated downstream of *ARF7* and *ARF19* (Aida et al., 2004; Hofhuis et al., 2013)..*PLT3*, PLT5, and PLT7 upregulate *PLT1* and *PLT2*, which are important players in the regulation of lateral root formation, root apical meristem maintenance (RAM), and callus pluripotency (Durgaprasad et al., 2019; Xu et al., 2006). In *Arabidopsis, PLT1* is also upregulated by JANUS through the recruitment of RNA Polymerase II to the root meristem (Xiong et al., 2020). In addition to root meristem maintenance, PLT proteins play important roles in conjunction with *BABYBOOM/PLT4* (*BBM/PLT4*) in early embryogenesis (described further in section "Somatic embryogenesis"), and activate regeneration responses in shoot organs by regulating the shoot promoting factors *CUPPED-SHAPED COTYLEDON1 (CUC1)* and *CUC2* (Radhakrishnan et al., 2020). PLT-CUC2 together work through the auxin biosynthesis pathway and are essential for proper distribution and repolarization of auxin through PIN-FORMED (PIN) proteins (described further in section "*De novo* root organogenesis") (Kareem et al., 2015; Radhakrishnan et al., 2020; Shimotohno et al., 2018).

Callus formation also involves epigenetic regulators. One regulator, HISTONE ACETYLTRANSFERASE OF THE GNAT/MYST SUPERFAMILY 1 (HAG1), also known as

A. thaliana GENERAL CONTROL NONREPRESSED 5 (AtGCN5),acts upstream of *PLT1* and *PLT2* (Kim et al., 2018; Kornet and Scheres, 2009). HAG1 also epigenetically upregulates root meristem genes *WUSHCEL RELATED HOMEOBOX 5* (*WOX5*), *WOX14*, and *SCARECROW* (*SCR)* by acetylating the N terminus of histone 3. HAG1 is further involved in determining the root–shoot axis in embryo development and is a regulator of floral meristem activity (Kim et al., 2018). The RAM gene, *ROOT CLAVATA-HOMOLOG 1* (*RCH1*), is also highly expressed in callus (Sugimoto et al., 2010), providing further evidence of homologies between lateral root development and callus formation. Although initiation of callus can follow multiple pathways, this provides further evidence that each pathway contains overlapping players.

De novo root organogenesis

De novo root organogenesis is the process by which adventitious roots are formed from detached plant tissues such as leaves and stems. Multiple studies have investigated the regeneration of the RAM in *A. thaliana* (Casamitjana-Martinez et al., 2003; de Smet et al., 2008; Galinha et al., 2007; Müller and Sheen, 2008; Perilli et al., 2012; Tian et al., 2002). The quiescent center (QC) is the site of stem cell maintenance of the RAM that is regenerated after QC ablation or entire removal of the root tip; polar transportation of auxin driven by PIN-FORMED (PIN) proteins results in auxin accumulation in cells adjacent to the damaged QC cells, which drives the reprogramming to new QC cells (van den Berg et al., 1997; Wildwater et al., 2005).

One of the key molecular players in root organogenesis is auxin. In *Arabidopsis,* auxin accumulates at cut sites, which induces expression of the homeobox transcription factors WOX11 and WOX12 (Liu et al., 2014). WOX11 and WOX12 directly upregulate *WOX5*, *LBD16*, and *LBD29*, marking the first step in cell differentiation and the formation of root meristems (Goh et

al., 2012; Hu and Xu, 2016; Liu et al., 2014). Auxin accumulation at wound sites in *Arabidopsis* drives the expression of *PLT* genes (as seen in callus formation), which will in turn upregulate *SHORT ROOT* (*SHR*) (Kareem et al., 2015). The SHR proteins will localize to the nucleus, inducing the expression of *SCR*; SHR and SCR are both involved in QC identity and radial patterning (van den Berg et al., 1997; Wildwater et al., 2005). SCR and PLT work together with plant-specific teosinte-branched cycloidea PNCP (TCP) in PLT-TCP-SCR complexes to promote the organization of PIN proteins and expression of *WOX5* in new meristem QC cells (Shimotohno et al., 2018; Xu et al., 2006). Root primordia formation is inhibited in *shr, plt1*, and *plt2* mutants, indicating that these genes play an important role during root formation from root founder cells (Bustillo-Avendaño et al., 2018).

De novo shoot organogenesis

Shoot organogenesis may occur with direct regeneration of a shoot from an explant or indirect regeneration from a callus (Figure 1). Because a callus seems to resemble root tip cells rather than shoot cells at the molecular level, callus cells must undergo changes in gene expression that push the cells toward shoot development rather than root development. Shoot regeneration has been studied extensively in plants; however, while many genes and hormones have been identified as important players in the process (Figure 2), the detailed molecular interactions and pathways are unclear (reviewed in Ikeuchi et al., 2016; Lardon and Geelen, 2020; Su and Zhang, 2014; Xu et al., 2006; Xu and Huang, 2014).

Regeneration of shoots from callus requires the formation of a primary meristem or a shoot apical meristem (SAM) (Figure 3). Similar to the RAM, the SAM contains a population of pluripotent stem cells that give rise to all aboveground organs of a plant. The undifferentiated state of the

organizing center (OC), which is similar to the RAM QC, and surrounding stem cells is maintained by a feedback mechanism between WUSHEL (WUS) and the signaling peptide CLAVATA3 (CLV3) (Sarkar et al., 2007). Leaves and other lateral organs arise from the peripheral regions of the SAM and the stem arises from the basal cells, called the rib zone. The SAM also contains the central zone, which consists of a stem cell pool that will replenish cells in the peripheral and rib zones that have further differentiated (Bowman and Eshed, 2000; Kwiakowska, 2004). Unlike auxin accumulation in the RAM, the SAM contains high levels of cytokinins. Organization of auxin and cytokinin in cells help promote differentiation of pluripotent cells to either shoot or root cells.

Shoot regeneration from callus occurs in four stages: 1. formation of a pluripotent callus, 2. shoot promeristem formation, 3. shoot progenitor development, and 4. shoot regeneration (Shin et al., 2020). The development of a pluripotent callus cell mass (described further in section "Organogenic callus formation") that highly expresses the No Apical Meristem/*A. thaliana* activating factor (NAC) transcription factor genes, *CUC1* and *CUC2* (Gordon et al., 2007), transitions into promeristems composed of a primary meristem of actively dividing cells. Within the callus *CUC2* expression marks pre-meristematic regions by promoting cell proliferation and leading to the localized upregulation of a key shoot meristem regulator, *SHOOT MERSITEMLESS* (STM), and *PIN1*. As seen in *de novo* root organogenesis, PIN1 proteins polarly localize, denoting areas of cellular reprogramming toward promeristematic cells (Gordon et al., 2007). Both STM and PIN1 aid in the development of radial patterning as STM marks the promeristem and PIN1 marks primordia (Gordon et al., 2007). Because PIN1 proteins are important players in both promeristem formation and root *de novo* organogenesis, this suggests that auxin transport is important for both shoot and root meristem patterning.

Proper regulation and distribution of *CUC1, CUC2,* and *WUS* are critical for shoot progenitor cells. These NAC transcription factors are subject to upstream regulation during shoot promeristem formation. AP2/ERF-type transcription factors, ENHANCER OF SHOOT REGENERATION 1 (ESR1)/DORNROSCHEN (DRN) and ESR2/DRN-LIKE (DRNL) participate in upstream regulation of *CUC* genes by actively binding to the promoter and inducing expression (Banno et al., 2001; Ikeda et al., 2006; Kirch et al., 2003). Mutants of *esr1, esr2,* and *esr1 esr2* show a reduction in shoot regeneration. This is likely due to improper regulation of *CUC1* and *CUC2* (Matsuo et al., 2011). WIND1 also upregulates *ESR1* by directly binding to the vascular-responsive motifs in the *ESR1* promoter (Iwase et al., 2017), suggesting that WIND1 is important in multiple plant regeneration processes. PLT5 and PLT7, which are induced during callus production, also influence the expression of *CUC1* and *CUC2* (Kareem et al., 2015). This further suggests that the molecular players and pathways involved in shoot regeneration are intertwined.

WUSCHEL (WUS) is a key regulator of the SAM and is upregulated during shoot regeneration. Expression of *WUS* is an important part of the conversion of a promeristem to a shoot progenitor as it represses cell division, cell elongation, and auxin-induced expression. This directs cell fate toward shoot development rather than root development. Ectopic expression of *AtWUS* results in *de novo* meristem formation and organogenesis in multiple plant species including *Arabidopsis* (Gallois et al., 2002; Negin et al., 2017), rice (Victorathosayam and Sridevi, 2020), and cotton (Bouchabké-Coussa et al., 2013). *WUS* expression is restricted to high cytokinin domains, while *CUC2* expression tends to be restricted to low cytokinin and high auxin domains. This is consistent with high expression of *CUC2* during induction of callus on media using higher concentrations of auxin (Daimon et al., 2003; Kareem et al., 2015). Regulation of *WUS* is subject

to epigenetic regulation. METHYLTRANSFERASE1 (MET1), KRYPTONITE (KYP), histone acetyl transferase1 (HAC1), and JMJ14 are all required for proper expression of *WUS*, SAM organization, and shoot development (Ishihara et al., 2019; Li et al., 2011). *MET1* is induced by the cytokinin-CYCD3-E2FA module, which represses *WUS* expression, allowing cells to retain callus identity rather than transitioning to shoot cells. However, in later stages of *de novo* shoot organogenesis, *MET1* is spatially regulated, allowing for an increase in *WUS* expression in the inner cell layers of the callus (Liu et al., 2018). Previously, it was thought that *WUS* expression in the inner callus cell layers is directly activated by the cytokinin-responsive Type B ARABIDOPSIS RESPONSE REGULATORS (ARRS), ARR1, ARR2, ARR10, and ARR12 (Dai et al., 2017). However, a recent study showed that ARR1 is a strong inhibitor of callus formation and shoot regeneration. This occurs through indirect repression of *CLV3* by competitive binding with ARR12 (Liu et al., 2020). ARR1 also indirectly represses *WUS* by inducing expression of the auxin response repressor gene *INDOLE-3-ACETIC ACID INDUCIBLE17* (*IAA17*) (Liu et al., 2020). In addition, Type-B ARRs negatively regulate the expression of the auxin biosynthetic genes *YUCCA1* (*YUC1*) and *YUC4* (Meng et al., 2017). This results in indirect upregulation of *WUS* expression. Although it has been known for decades that auxin and cytokinin signaling is important for plant regeneration, these findings further untangle the underlying mechanisms of the signaling pathways.

Eukaryotic stem cells tend to have open chromatin states, while differentiated cells tend to have closed chromatin states (Shchuka et al., 2015). Epigenetic controls include Trithorax group (trxG) and Polycomb Group (PcG) proteins. The *A. thaliana* trxG, ATXR2, interacts with ARR1 and methylates the Type A ARRs, *ARR5* and *ARR7*, marking them for increased transcription. This leads to a repression of cytokinin signaling and a reduction in *de novo* shoot organogenesis (Lee

et al., 2021). PcG protein complexes, specifically POLYCOMB REPRESSIVE COMPLEX1 (PRC1) and PRC2, are chromatin modifiers and bind to Polycomb Response Elements (PRE) to keep genes transcriptionally repressed in order to fine-tune the balance between cell proliferation and cell differentiation (Köhler and Hennig, 2010). PRC2 suppresses leaf identity via H3K27me3 of leaf identity genes. PRC2 is also involved in callus formation as PRC2 mutants *curly leaf swinger (clf swn)* and *embryonic flower2 (emf2)* are incapable of developing callus from leaf and cotyledon explants but retained the ability to form callus in root explants (He et al., 2012). This suggests PCR2 represses leaf identify genes, allowing for the transition to root-like callus cells. Other instances of epigenetic regulation during the early stages of regeneration include gene priming by LYSINE-SPECIFIC DEMETHYLASE 1-LIKE 3 (LDL3), which involves the elimination of methylation of lysine 4 on histone 3 (H3K4me2) during callus formation. This indirectly promotes the expression of genes that are involved in shoot progenitor development (Ishihara et al., 2019).

Regulatory microRNA, miR156, plays a role in activating cytokinin signaling by targeting *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*). *SPL* genes control transitions in shoot development—juvenile-to-adult and vegetative-to-reproductive—by binding to and regulating Type-B ARRs, decreasing shoot regenerative capacity with age (Xu et al., 2016; Xu et al., 2015). miRNA156 expression is higher in younger tissues, which partially explains why younger explant tissue (i.e., cotyledons) is more amenable to *in vitro* regeneration, when compared to more mature tissue types. Type B ARRs and WUS also regulate the Type A ARRs, *ARR7* and *ARR15*, which negatively regulate cytokinin signaling (Buechel et al., 2010).

After proper development of shoot progenitor cells, activation of leaf identity genes will lead to the development of leaf tissues and leaf emergence. Two important players involved in

shoot regeneration are miR165 and miR166, both of which target HD-ZIP III transcription factor genes *PHABULOSA* (*PHB*), *PHAVOLUTA* (*PHV*), *REVOLUTA* (*REV*), *KANADI* (*KAN*), and *ARABIDOPSIS THALIANA HOMEOBOX GENE 8* (*ATHB8*) (Shin et al., 2020). *PHB, PHV, REV,* and *KAN* function in radial leaf patterning (abaxial vs. adaxial), and *phb*, *phv*, *rev*, and *kan* mutants show a transition of abaxial leaf fates into adaxial leaf fates as well as altered auxin gradients (Emery et al., 2003; McConnell et al., 2001; Zhou et al., 2019). *ATHB8* and *SHR* expression activate simultaneously and lead to leaf vein precursor cells (Gardiner et al., 2011). An RNA-induced silencing complex, ARGONAUTE10 *(*AGO10), helps sequester and repress the activity of miR165 and miR166. This indirectly promotes the activity of these leaf identity genes. Interestingly, accumulation of miR165/166 in overexpressing *Arabidopsis* mutants resulting in less HD-ZIP III transcription factor activity in shoot progenitor cells, increased the overall shoot regeneration (Xue et al., 2017). This suggests that leaf identity genes work to suppress *in vitro* transition from meristematic cells into shoot cells. In addition, *AGO10* is repressed by LBD12, resulting in reduced apical meristem size (Ma et al., 2017). Leaf identity genes are also subject to epigenetic regulation. TrxG proteins, ATX1, ATX4, ULTRAPETALA1 (ULT1), and PICKLE (PKL), act as antagonists of PCR1 and PCR2 to activate transcription of leaf identity genes, which will aid in the development of leaves from shoot progenitor cells (Köhler and Hennig, 2010). In *A. thaliana*, ATX4 protein tri-methylates histone 3 (H3K4me3) to increase the expression of the shoot identity genes *ARABIDOPSIS THALIANA HOMEOBOX GENE 1* (*ATH1*), *KNOTTED1-LIKE HOMEOBOX (KNOX) GENE 4* (*KNAT4*), *SAWTOOTH 1* (*SAW1*), *SAW2, TCP FAMILY TRANSCRIPTION FACTOR 10* (*TCP10*), and *YABBY 5* (*YAB5*) (Lee K. et al., 2019).

As elaborated above, *de novo* shoot regeneration is controlled by a complex network of genetic and epigenetic factors. Although we are gaining a more detailed understanding of the

molecular players involved in this network via forward and reverse genetic approaches, there is clearly more information to discover involving interactions between these genetic, epigenetic, and hormone signaling pathways.

Embryogenic callus formation

Formation of embryogenic callus results from acquisition of a new cell fate through expression of embryonic regulators. Similar to organogenic calli, embryogenic calli have been observed to originate from cells surrounding vascular tissue (pre-procambial cells) (de Almeida et al., 2012). Endogenous application of plant growth regulators such as auxin and cytokinin have been shown to induce proliferation of embryonic tissues in some species, such as soybean and cotton (Elhiti and Stasolla, 2022; Raza et al., 2020). This is similar to auxin-induced callus formation suggesting upregulation of *ARFs* such as *ARF7* and *ARF19* are also requirements for the formation of embryonic callus. Furthermore, *LEAFY COTYLEDON1* (*LEC1*) and *LEC2* genes are major embryonic regulators that control embryo maturation and development (Gaj et al., 2005). *LEC1* overexpression induced embryogenic callus formation in *Arabidopsis*; however, *lec1* and *lec2* mutants resulted in the development of fewer somatic embryos via only indirect somatic embryogenesis (Gaj et al., 2005). This suggests that *LEC1* is sufficient, but not necessary to the formation of embryogenic callus. Overexpression of the MADS-box transcription factor, AGAMOUS-LIKE 15 (AGL15), induced embryogenic callus-like structures on SAMs and extended embryonic development from callus in *Arabidopsis* (Harding et al., 2003). Expression of specific genes and presence of proteins have been observed in embryogenic callus, but not observed in non-embryogenic callus. The MADS-box gene, *CUS1*, whose amino acid sequence is highly similar to Arabidopsis AGL1 and AGL5 amino acid sequences, was detected in embryogenic callus during somatic embryogenesis in cucumber (Filipecki et al., 1997).

Additionally, in sugar cane, unique proteins were identified during embryonic callus formation including proteins related to metabolic activity and stress (Schuabb Heringer et al., 2015). Induction of somatic embryogenesis and formation of proembyrogenic masses on calli (described further in section "Somatic embryogenesis") involves different molecular players than formation of promeristems during organogenesis, but differences between embryogenic calli and organogenic calli formation, if any, are still not well characterized.

Somatic embryogenesis

A second type of *in vitro* regeneration is somatic embryogenesis. Somatic embryogenesis results when a differentiated somatic cell undergoes molecular changes and genetic/epigenetic reprogramming resulting in the formation of a bipolar somatic embryo. In tissue culture, somatic embryogenesis can be induced in response to the addition of exogenous plant growth regulators or the introduction of stressful conditions. Similar to *de novo* organogenesis, somatic embryogenesis may originate directly at wound sites of explants or indirectly from embryogenic callus (Quiroz-Figueroa et al., 2006). Species tend to regenerate either through *de novo* organogenesis (e.g., tomato, lettuce, pepper) or somatic embryogenesis (e.g., cotton, wheat, rice,) but rarely both (e.g., chickpea, purple coneflower) (Choffe et al., 2000; Heidmann et al., 2011; Leelavathi et al., 2004; Michelmore et al., 1987; Murthy et al., 1996; Ozias-akins and Vasil, 1982; Rueb et al., 1994; Sun et al., 2015).

Regulators and genetic determinants of embryo initiation are not well understood, although auxin signaling and accumulation are thought to play a major role. In tissue culture, addition of auxin is used to induce somatic embryogenesis by exposure of explants to high levels of auxin immediately followed by a transfer to auxin-free medium (Méndez-Hernández *et al*., 2019). This

allows for the formation of auxin gradients within the developing embryos—areas of high auxin promote *WUS* expression, which denote areas of future SAM development as mentioned previously (Ikeuchi et al., 2016). In *Arabidopsis,* several *ARFs* are both up and downregulated during the first steps of somatic embryogenesis, and multiple *arf* mutants showed inhibited somatic embryogenesis (Wójcikowska and Gaj, 2017). *SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES* (SERKs), specifically *SERK1* in *Arabidopsis,* are upregulated in embryonic callus and are continually expressed from megasporogenesis until the heart stage of the embryonic development (Hecht et al., 2001). Other genes, such as auxin-responsive gene *EgIAA9* from *Elaeis guineensis*, have been shown to be upregulated during somatic embryogenesis initiation (Ooi et al., 2012).

The transcription factor BABY BOOM (BBM) and the LEC1-AB13-FUS3-LEC2 (LAFL) complex are master regulators of somatic embryogenesis (Horstman et al., 2017; Jones et al., 2019). *BBM* encodes an AINTEGUMENTA-LIKE (AIL) AP2/ERF and directly regulates all LAFL genes. *LAFL* genes are also regulated by a BBM-like protein, PLT2 (Horstman et al., 2017). The *LAFL* gene group consists of the *LEC* transcription factor genes, including *LEC1*, *LEC2*, and *FUSCA3* (*FUS3*), and the ABA signaling transcription factor, ABSCISIC ACID INSENSITIVE 3 (ABI3). Somatic embryogenesis events are shown to significantly decrease in *lec* mutants (Gaj et al., 2005), and the overexpression of *LEC2* led to an increase in the expression of auxin biosynthesis genes *IAA30*, *YUC2*, *YUC4*, and *YUC10* (Junker et al., 2012; Stone et al., 2008), suggesting that *LEC* genes and the LAFL complex help promote auxin activity. LEC2 also induces the expression of *LEC1*, *LEAFY COTYLEDON 1-LIKE (L1L)*, *ABI3*, and *FUS3.* Another transcription factor, AGL15, has been shown to directly regulate *LAFL* (Zheng et al., 2009) and promote the expression of the AP2/ERF gene *At5g61590* (Zheng et al., 2013). *At5g61590* is a

relative of the *Medicago truncatula SOMATIC EMBRYO-RELATED FACTOR 1* (*MtSERF1*), which is essential for somatic embryogenesis (Mantiri et al., 2008). Recently, another MADS-box transcription factor, AGL18, was identified as an active regulator in somatic embryogenesis in *Arabidopsis* (Paul et al., 2022). Overexpression of *AGL18* resulted in an increase in somatic embryogenesis, while a decrease was observed in *agl18* mutants; *agl15 agl18* double mutants resulted in even less frequent development of somatic embryos. While the functions of AGL15 and AGL18 transcription factors were relatively redundant, different gene targets for each transcripation factor were present and an AGL15/AGL18 regulatory loop was identified. This provides evidence that AGL18 may act in conjunction with AGL15 during somatic embryogenesis. Along with *BBM*, *LAFL*, and *AGL15*, the ectopic expression of *WUS*, *PLT4*/*BBM*, *PLT5*/*EMBRYMAKER*, *MYB118*, and *RWP-RK DOMAIN-CONTAINING4* (*RKD4*)/*GROUNDED* (*GRD*) leads to the induction of somatic embryogenesis in *Arabidopsis* (Boutilier et al., 2002; Gallois et al., 2004; Harding et al., 2003; Lotan et al., 1998; Waki et al., 2011) .

The master regulators work with other transcription factors to balance auxin, gibberellin (GA), and abscisic acid (ABA) signaling. In particular, the balance of GA and ABA has a major role in controlling cell identity in the developing embryo. Embryonic cells have been shown to have a higher ratio of GA to ABA than somatic cells (Hu et al., 2008; Mitchum et al., 2006; Yamaguchi et al., 2001). The LAFL transcription factors, LEC1, LEC2, FUS3, and AGL15, downregulate GA biosynthesis genes (Curaba et al., 2004; Zheng et al., 2009), while FUS3 positively regulates the ABA pathway (Gazzarrini et al., 2004). LEC1 and LEC2 promote the expression of auxin biosynthesis genes (Braybrook et al., 2006; Junker et al., 2012), and AGL15 negatively regulates the auxin response genes, *ARF6*, *ARF8*, and *TRANSPORT INHIBITOR RESPONSE1* (*TIR1*) (Zheng et al., 2016). LEC1 and AGL15 positively regulate *ABI3*. Generally,

these transcription factors work to negatively regulate GA biosynthesis and positively regulate ABA and auxin biosynthesis, transitioning cells from embryonic cells (high GA/ABA ratios) into differentiated somatic cells (low GA/ABA ratios). MYB-family transcription factors, MYB118 and MYB115, also play important roles in somatic embryogenesis. These transcription factors promote the expression of *LEC1*; overexpression of both resulted in the formation of somatic embryos on root explants (Wang et al., 2008). The micro RNA miR396 is associated with somatic embryogenesis induction and helps control *PLT1* and *PLT2* (Szczygieł-Sommer and Gaj, 2019). Other evidence suggests that AGL15 forms protein complexes with SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES (SERKs), which are induced in response to auxin (Zheng et al., 2009). Ethylene has also been shown to impact somatic embryogenesis because interactions between ETHYLENE RESPONSE FACTOR 002 (ERF022) and LEC2, and the involvement of other AP2/ERF transcription factors have been observed (Horstman et al., 2017; Xu and Huang, 2014; Zheng et al., 2013). Reprogramming of somatic cells to form embryos and then back to differentiated somatic cells requires multiple hormone signaling pathways to work together.

Genomic DNA methylation patterns change during development. In mature tissues, *LEC1*, *LEC2*, and *AGL15* are hypermethylated in somatic cells, while hypomethylation has been seen of similar genes (e.g., *SERKs*, *LEC2*, *WUS)* in embryonic calli (Fraga et al., 2012). This suggests that somatic embryogenesis and genes involved in embryonic cell to somatic cell transition is subject to epigenetic regulation as the repression of embryonic genes leads to the development of mature and differentiated tissues. There is conflicting evidence for the role of DNA methylation in somatic embryogenesis. In some studies, the demethylation agent 5-azacitidine strongly inhibited embryogenesis in *Medicago truncatula* and *Arabidopsis* (Grzybkowska et al., 2018; Santos and

Fevereiro, 2002), while in other plants, such as in *Coco nucifera* and *Acca sellowiana*, its application increased embryogenesis (Fraga et al., 2012; Osorio-Montalvo et al., 2020). This suggests that differential DNA methylation is required for successful somatic embryogenesis but its effects are highly genotype and species dependent.

Two critical regulatory epigenetic pathways, PcG and PKL, are involved in the epigenetic regulation of genes during somatic embryogenesis. As in shoot organogenesis, the PRC2-mediated H3K27 methylation, part of the PcG pathway, is involved in the repression of *LEC1*, *LEC2*, and *FUS3*, aiding in the transition from embryonic to somatic cells (Makarevich et al., 2006). The Repressive LEC2 Element (RLE) in the *LEC2* promoter recruits PCR2 for methylation and repression of *LEC2* in somatic cells (Berger et al., 2011). Evidence supporting this includes an increase in somatic embryogenesis of *Arabidopsis* in vegetative tissue depleted of PRC2 (Mozgová et al., 2017). PRC2 has also been shown to negatively regulate other important regulators of plant regeneration including WOX5, WOX11, WUS, and STM. PKL encodes for a chromatin remodeling factor, which serves to rearrange nucleosome positions in order to regulate gene expression. Multiple studies have demonstrated that *pkl* mutants show an increase in the ectopic expression of *LEC1*, *LEC2*, and *FUS3*, resulting in embryonic traits in somatic tissues (Henderson et al., 2004; Ogas et al., 1997). This suggests that PKL plays a role in negatively regulating embryonic genes in somatic tissues. However, the specific molecular mechanism by which PKL works is still unclear.

Small signaling peptides in plant regeneration

Signaling peptides are important players in plant development. One family of signaling peptides, CLAVATA/ENDOSPERM SURROUNDING REGION (CLE), has central roles in

modulating stem cell differentiation during plant growth and development (Katsir et al., 2011). These peptides are post-translationally processed and contain a signal peptide targeting the protein for secretion, where it is used for cell-to-cell communication (Yamaguchi et al., 2016). In *A. thaliana*, there are 32 CLE peptides including CLV1, CLV2, and CLV3. CLV3 is secreted from cells and interacts with CLV1, a leucine-rich repeat receptor kinase, to maintain stem cell populations in the apical meristem (Clark et al., 1995; Hirakawa et al., 2008). In *clv1* and *clv3* mutants, plants develop enlarged shoot and floral apical meristems, suggesting improper signaling disruption to maintenance of stem cell populations (Clark et al., 1995). *WUS* promotes cell proliferation and division and upregulates *CLV1-CLV3*. This results in the downregulation of *WUS* by CLV1-CLV3 in a negative feedback loop. This feedback mechanism produces and maintains a constant stem cell pool (Brand et al., 2000; Mayer et al., 1998). Manipulating either WUS, CLV1, and/or CLV3 could lead to larger stem cell pools and greater potential for cell division. This in conjunction with downstream molecular players, such as *CUC* genes, *PLT* genes, or *SPL*, and could potentially lead to more and faster plant regeneration. However, this would require careful orchestration of the key regulators.

Growth regulating factors as agents for increased regeneration

Growth Regulating Factors (GRF) are a transcription factor family that regulates many aspects of plant growth and development including leaf, stem, root, seed development, flowering, regulation of stress, and plant longevity. The first GRF, Os-*GRF1*, was identified two decades ago during a differential expression study of responses of deep-water rice to GA (van der Knaap et al., 2000). GRFs have now been identified in many plant species, where typically eight to 20 different GRF genes are present in each genome (Omidbakhshfard et al., 2015). GRFs form complexes with their cofactor, GRF-interacting Factors (GIF), and will bind to *cis*-regulatory elements of different

developmental genes in plants (Kim, 2019). For example, *AtGRF7* binds to the promoter of the AP2/ERF gene *Dehydration responsive element binding protein2A* (*DREB2A*) and represses gene expression in leaf veins (Kim et al., 2012). In *Arabidopsis, GRFs* have been shown to be expressed in leaf and root tissue where prolific cell growth is occurring and tend to decrease with plant age (Hewezi et al., 2012; Kim et al., 2003; Lee et al., 2009; Szczygieł-Sommer and Gaj, 2019).

GRF proteins are post-transcriptionally regulated by miR396 throughout the course of plant development; miR396 recognizes and binds to GRF, resulting in degradation or translational arrest. Expression of miR396 occurs at low levels in leaf primordia that gradually increase throughout organ development and maturity (Rodriguez et al., 2010). Expression of *AtGRF2* is restricted to specific portions of the leaf during development through antagonistic expression of *miR396* (Rodriguez et al., 2010). In rice, *miR396* mutants resulted in an upregulation of multiple *GRF* genes, in particular *GRF3*. These mutants also produced plants with longer leaves and shorter internodes (Miao et al., 2020). Because of their involvement in organ development, GRF and miR396 are potential targets for increasing *in vitro* regeneration.

GRFs regulate players important for *in vitro* regeneration. GRF proteins from rice, OsGRF3 and OSGRF10 repress promoter activity of a KNOX gene, *Oskn2* (Kuijt et al., 2014). In the same study, barley GRF, BGRF1, repressed *Hooded/Bkn3*, a barley KNOX gene, and overexpression of *OsGRF10*, *AtGRF4*, *AtGRF5*, and *AtGRF6* repressed activity of *KNAT2* in *Arabidopsis*. In addition, overexpression of *OsGRF3* and *OsGRF10* induced root and shoot formation on primary tillers of rice (Kuijt et al., 2014).Because regulation of *KNOX* genes is necessary for cell identity transitions from meristem cells to mature organ cells (Hake et al., 1995, 2004; Tsuda et al., 2011), the reported functions of these GRFs demonstrate the potential importance of GRFs in both organogenesis and somatic embryogenesis. An RNA-seq study in rice

showed upregulation of *OsGRF6* resulted in an increase in expression of the auxin biosynthesis gene, *OsYUCCA-like,* and signaling genes, *OsARF2, OsARF7, OsARF11* (Gao et al., 2015). In addition, altered expression of *GRF* and *GIF* affect root growth through regulation of *PLT1*, *PLT2*, and *SCR* (Ercoli et al., 2018). In Arabidopsis, the double mutant *gif1/an3 gif2* and the triple mutant *gif1/an3 gif2 gif3* both showed the formation of a disorganized QC and larger RAM, while overexpression of *GRF3* with a mutated miRNA binding site (*rGRF3*) resulted in smaller meristems (Ercoli et al., 2018). These studies provide evidence that GRFs and GIFs are upstream regulators of molecular determinants involved in callus formation and shoot meristem identity, giving altered expression of GRFs and GIFs the potential to increase regeneration in plants.

GRFs and GIFs have now been shown to enhance regeneration capacity and rates in plants. Ectopic expression of *AtGRF5* and orthologs increased callus production in canola and shoot organogenesis in sugar beet, soybean, and sunflower; also, ectopic expression of the maize *GRF5* ortholog increased formation of embryogenic calli indicating that GRFs regulate multiple *in vitro* regeneration pathways (Kong et al., 2020). In addition, transformation with a chimeric *GRF-GIF* gene fusion can increase the rate and number of regenerates in wheat, rice, and citrus (Debernardi et al., 2020). Independent transformations and co-transformations of multiple wheat *GRF*s fused with *GIF*s were studied including *GRF4*, *GRF5*, *GIF1*, *GIF2*, and *GIF3*; the chimeric transgene composed of a fusion between GRF4 and GIF1 (GRF4-GIF1) resulted in the highest frequency of regeneration in wheat among all combinations of GRFs and GIFs tested. In addition to increased regeneration, shoot regeneration and transgenesis in wheat was successful without the use of cytokinins in the culture medium. Furthermore, regeneration could be induced from leaf explants rather than immature embryos. The efficacy of chimeric transgene was also tested in the dicotyledonous species, *Citrus*, using the *Citrus* and *Vitis* GRF4 and GIF1 homologs (Debernardi

et al., 2020). Furthermore, the use of the microRNA insensitive *rGRF4-GIF* resulted in greater stimulation of regeneration in wheat, rice, and *Citrus*. This is a major breakthrough and will be exploited for the regeneration of recalcitrant species and cultivars, leading to a likelihood of higher transformation rates.

PROSPECTS FOR ENHANCED REGENERATION IN LETTUCE

Synopsis of studies on the regeneration of lettuce

Lettuce, *Lactuca sativa* L. (Compositae), is a dicotyledonous plant that can be regenerated by indirect *de novo* shoot organogenesis (Figure 4) and was a model for early studies of regeneration (reviewed in Michelmore and Eash, 1985). Some genotypes regenerate readily on a variety of media formulations and growth regulators; however, some lettuce genotypes are recalcitrant to regeneration. Lettuce is also amenable to *Agrobacterium*-mediated transformation (Michelmore et al., 1987). Protocols for high efficiency, genotype-independent regeneration of lettuce are required in order to fully benefit from biotechnological approaches, including genome editing, for crop improvement. Given differences in regeneration rates of different genotypes and the wealth of knowledge from model species described above, top-down and bottom-up approaches to the molecular basis of regeneration in lettuce could lead to protocols for enhanced regeneration of multiple genotypes.

Lettuce regeneration has been studied for many decades. Lettuce was among the first plants to be tested for regeneration. The first studies on *in vitro* regeneration of lettuce failed to produce shoots from leaves of *L. sativa* and *L. canadensis* (LaRue, 1936; LaRue, 1933). Later, regeneration of lettuce shoots was successful with the addition of adenine and kinetin to the growth medium (Doerschug and Miller, 1967). In this study, the regenerative capability of hypocotyl, cotyledon,

and mature leaf explants was tested on the same base medium with different combinations of IAA, kinetin, and adenine, and cotyledons were shown to be the most effective explant source for shoot regeneration. In the same study, kinetin was effective at promoting the transition from callus formation to shoot regeneration (Doerschug and Miller, 1967). This suggested that in lettuce, as shown in other plant species, high levels of cytokinin promotes the formation of shoot meristems that results from the transition of cell fate from root-like callus cells to shoot cells. Later studies focused on the optimization of factors influencing lettuce regeneration, including media formulations, plant growth regulator use, light requirements, temperature, explant type, and genotype (Doerschug and Miller, 1967; Sasaki, 1979; Kadkade and Siebert; 1977, Koevary; 1978, Sasaki; 1982, Alconero, 1983; Webb, 1984; Michelmore and Eash, 1985). Light intensity and photoperiod were shown to be also important for lettuce regeneration; cotyledon explants developed well-formed shoots with a 16-hour photoperiod but significantly fewer shoots formed in the dark; additionally, the presence of red light doubled the number of buds and shoots (Kadkade and Seibert, 1977). In aggregate, callus formation occurred on all lettuce cultivars studied when using both auxins and cytokinins in the culture medium, although there were differences between genotypes. Shoot regeneration was elicited when the medium contained cytokinins with little or no auxins. Although mature leaves and hypocotyls showed regenerative capabilities, cotyledons were the most amenable explant source for regeneration.

Indirect *de novo* shoot organogenesis in lettuce involves cell divisions of spongy, palisade, and epidermal cells. A cytohistological study of adventitious bud formation from cotyledon explants revealed initial divisions of spongy and palisade cells followed by divisions of epidermal cells to form tetrads (Ronchi and Gregorini, 1970). Callus was formed from the division of mesophyll cells and inward proliferation of epidermal cells. Subsequently, adventitious buds arose

from one or two epidermal cells, which led to the formation and organization of shoot apical meristems. This study provided the timeline and steps that occur during organogenesis; however, the tools were not available to study the underlying genetic and molecular constituents responsible for the changes in cell anatomy and transition of cell fate, particularly epidermal cells to meristematic centers.

Like most plant species, regenerative capacity is highly dependent on genotype and there is considerable variation in regenerative capacity among lettuce cultivars (Ampomah-Dwamena et al., 1997; Curtis et al., 1994; Michelmore et al., 1987; Mohebodini et al., 2011). There is no significant correlation to regeneration efficiency and morphological group (i.e., crisphead, butterhead, cos, and leaf). In a side-by-side study, highly regenerating genotypes included Bambino (crisphead), Iceberg (crisphead), Cobham Green (butterhead), Sweet Butter (butterhead), Simpson Elite (leaf), Rosalita (cos), and Paris White (cos); recalcitrant genotypes included Oak Leaf (leaf), Royal Oak Leaf (leaf), Sangria (crisphead), and Mainspring (butterhead) (Ampomah-Dwamena et al., 1997). Generation of stable transgenics of lettuce relies on *Agrobacterium*-mediated transformation and *in vitro* regeneration. Therefore, it is important to understand and identify the genetic and molecular players to increase regeneration in order to manipulate recalcitrant lettuce varieties.

Known molecular determinants for regeneration in lettuce

There have been few studies on the molecular determinants of regeneration in lettuce. A dominant mutation of the ethylene receptor ETR1-1 was shown to inhibit shoot regeneration in lettuce (Kim and Botella, 2004). Lettuce cultivars LEI26 and Seagreen were transformed using *Agrobacterium*-mediated transformation for the introduction of GUS under the control of the

CaMV 35S constitutive promoter and the mutated ethylene receptor *etr1-1* under the control of a leaf senescence-specific promoter, *sag12*. Transformations with 35S:GUS showed high regenerative potential with 85% of explants developing shoots, while the introduction of *sag12:etr1-1* significantly reduced regenerative potential with only 2.86% of explants producing shoots. Explants transformed with *sag:etr1-1* also stimulated root formation directly from cotyledon explants without the formation of callus (Kim and Botella, 2004). This suggests that ethylene responses are important in *in vitro* lettuce regeneration in which inhibiting ethylene receptors promotes root formation and inhibits callus and shoot formation. This is consistent with observations of other ethylene response factors during *in vitro* regeneration, such as the early expression of AP2/ERF transcription factors during callus formation and the involvement of ERF022 activity during somatic embryogenesis (Horstman et al., 2017; Iwase et al., 2011a; Iwase et al., 2011b; Xu and Huang, 2014; Zheng et al., 2013)

Data is limited for lettuce on the effects of the pathways and molecular determinants described in other species. A recent study examined the chronological expression of homeobox genes during *in vitro* regeneration of lettuce (Farina et al., 2021). Gene expression profiles of lettuce homologs to the homeobox WOX family transcription factor genes *WUS* (*LsWUS1L* and *LsWUS2L*) and the KNOTTED1-LIKE homeobox family transcription factor gene *ST-M* (*LsSTM*), were examined in cotyledon explants over 12 days on inductive medium. A time course analysis showed a steady increase of expression of *LsWUS1*; in early days of culture, increased expression of *LsWUS2L* correlated with the formation of poorly vacuolated cells with large nuclei in the explants. Expression of *LsSTM1L* also drastically increased in early days of culture, followed by a later decrease, suggesting that it helps recruit proteins and regulates expression of genes needed for the initiation of regeneration in lettuce (Farina et al., 2021). This parallels patterns of *WUS* and

STM expression observed early in plant regeneration, specifically during the formation of shoot promeristems and meristematic centers from callus in *Arabidopsis* (Daimon et al., 2003; Zhang et al., 2017). This is also consistent with the essential role WUS plays in maintaining the stem cell pool that is critical for proper SAM function (Sarkar et al., 2007). The CCAAT-binding transcription factors, *LEC1* and *LEC2*, play a major role in development and maturation of embryos (described further in sections "Embryogenic callus formation" and "Somatic embryogenesis"). Nothing has been reported for homologs of *LEC1* and *LEC2* in lettuce. It would be interesting to overexpress homologs of these transcription factors in lettuce to determine if this results in enhanced regeneration as in *Arabidopsis*, tobacco, and cassava (Gaj et al., 2005; Guo et al., 2013; Brand et al., 2019). Similarly, over-expression of *CUC1* and *CUC2* as well as *PLT* genes (described further in sections "Organogenic callus formation", "*De novo* root organogenesis", and "*De novo* shoot organogenesis") may also result in enhanced regeneration of lettuce as in Arabidopsis (Ikeda et al., 2006; Kareem et al., 2015; Matsuo et al., 2009).

MADS-box genes in lettuce

MADS-box transcription factors, particularly AGL15 and AGL18, are major molecular players involved in *in vitro* regeneration (described further in sections "Embryogenic callus formation" and "Somatic embryogenesis"). There are at least 82 MADS-box encoding genes in lettuce (Ning et al., 2019), most of which have been studied in relation to flowering time and floral development (reviewed in Han et al, 2021). Of these 82 genes, 23 encoded for M-type genes of the type 1 lineage and 59 floral genes of the type II lineage containing a MIKC domain. Within the type II MADs-box genes, ten belonged to the *AGL15* subfamily which contained homologs of Arabidopsis and tomato *AGL15* genes. Currently, no work has been reported on the role of lettuce MADs-box genes during *in vitro* regeneration. The ten genes identified in the *AGL15* subfamily

should be characterized for their roles in regeneration in lettuce; it should be tested whether over expression of ALG15 results in increased somatic embryogenesis as in *Arabidopsis* (Paul et al., 2022).

Growth regulating factors in lettuce

There are 15 *GRF* genes in lettuce and their chromosomal locations, gene structure, conserved motifs, and expression patterns have been characterized (Zhang et al., 2021). One *GRF* gene was studied in detail. *LsaGRF5* showed low expression in leaves and roots with high expression in reproductive buds, suggesting an important function in flower development. The GRF regulator, *miR396a*, had high expression in mature flowers and stems and low expression in reproductive buds. These data suggest that high levels of *LsaGRF5* expression in young tissues is coincident with actively dividing cells; as the cells and tissues mature, *LsaGRF5* becomes downregulated by *miR396a*; this is similar to what is observed in other species (described further in section "Growth regulating factors as agents for increased regeneration"). Furthermore, overexpression of *LsaGRF5* resulted in larger leaf size, while overexpression of miR396a resulted in smaller leaf size (Zhang et al., 2021). However, none of the 15 *GRF* genes have been characterized for their effects on regeneration in lettuce. Given the success of GRF or GRF-GIF fusions with other species (described further in section "Growth regulating factors as agents for increased regeneration"), it is likely that similar enhanced rates of regeneration and transformation will be reported soon.

CONCLUSIONS AND FUTURE PERSPECTIVES

The underlying processes of plant regeneration all involve cell fate transition by reprogramming gene expression. The several pathways involved in plant development and

regeneration are complex. Although each pathway has unique molecular players, many of the key regulators overlap and have important functions in each. Auxin and cytokinin signaling pathways play a major role in regulating multiple regenerative pathways and accompany the genome-wide switch in gene expression profile during the early stages of regeneration. Other phytohormones, such as GA, ABA, and ethylene, also contribute to plant regeneration and cell fate transition.

Many of the players and regulators important for *in vitro* regeneration have been studied in model species, such as *Arabidopsis,* but have not been functionally characterized in non-model species such as lettuce. The complete genome sequence of *L. sativa* (Reyes-Chin-Wo et al., 2017) has provided useful genic targets for modification by genome editing. Currently, genome editing of lettuce requires *Agrobacterium*-mediated transformation, which requires *in vitro* regeneration; therefore, studying molecular determinants and understanding pathways controlling regeneration in lettuce has great value. Identifying orthologs of genes discussed in this review and then characterizing them in other systems, such as lettuce, will help form a more generalized understanding of *in vitro* regeneration in plants. Further studies on identification of recalcitrant varieties, quantitative trait locus analyses on varieties with varying regenerative capabilities, and expression profiles during *in vitro* regeneration could provide insight into other genes regulated during *in vitro* regeneration of lettuce. Understanding these pathways in lettuce will allow for a better understanding of the pathways in other important crops, particularly within the Compositae family such as sunflower, artichoke, safflower, and many ornamentals.

FIGURES

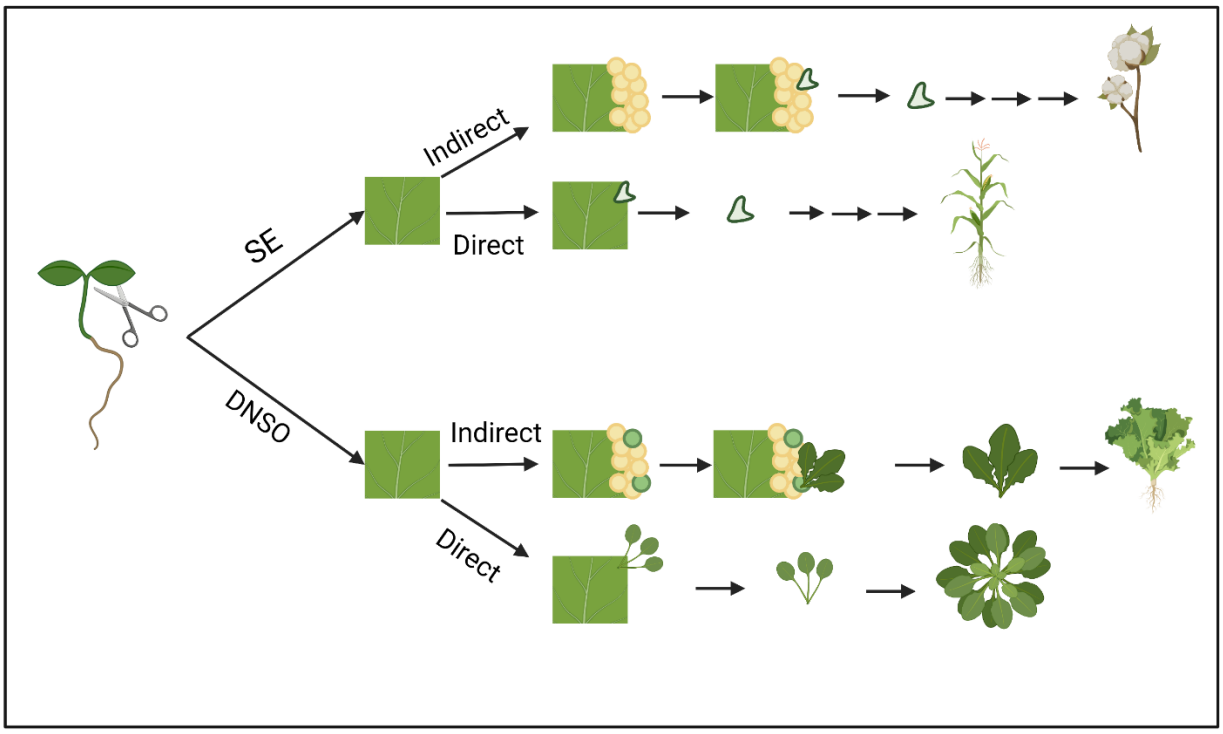


Figure 2.1. Pathways of in vitro regeneration of vascular plants. Somatic embryogenesis (SE) and *de novo* shoot organogenesis (DNSO) can occur directly on the explant or indirectly with the formation of pluripotent callus as an intermediate step. Species that are capable of regeneration for each pathway are represented from top to bottom: cotton, maize, *Arabidopsis*, and lettuce. Figure created using BioRender (https://biorender.com/).

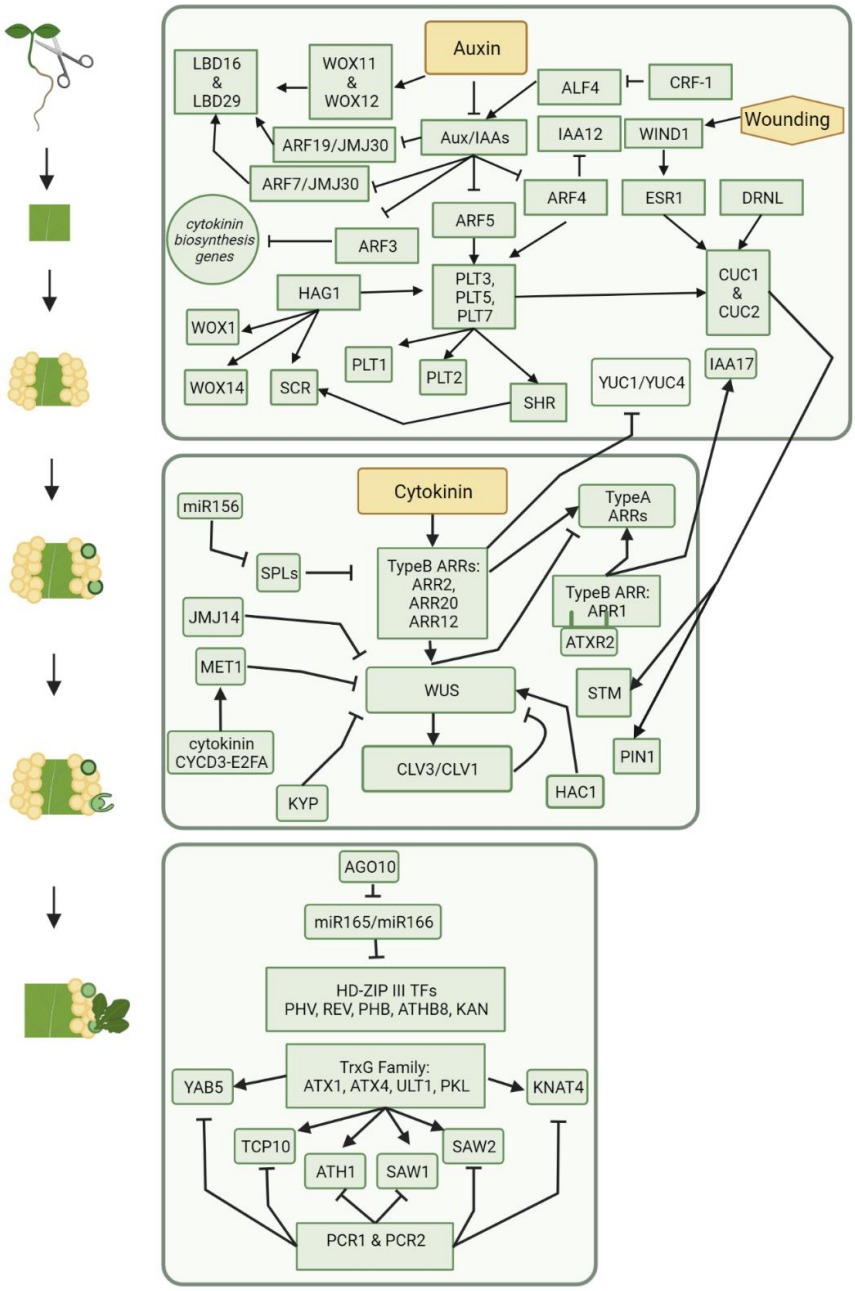


Figure 2.2. The progression of molecular players during indirect de novo shoot organogenesis. Callus is formed on auxin rich medium and includes signaling pathways represented in box one. Shoot promersitems and meristematic centers are formed on cytokinin rich medium and include signaling pathways represented in box two. Shoot regeneration follows meristem formation and is represented by the signaling pathways included in box three. Figure created using BioRender (https://biorender.com/).

.

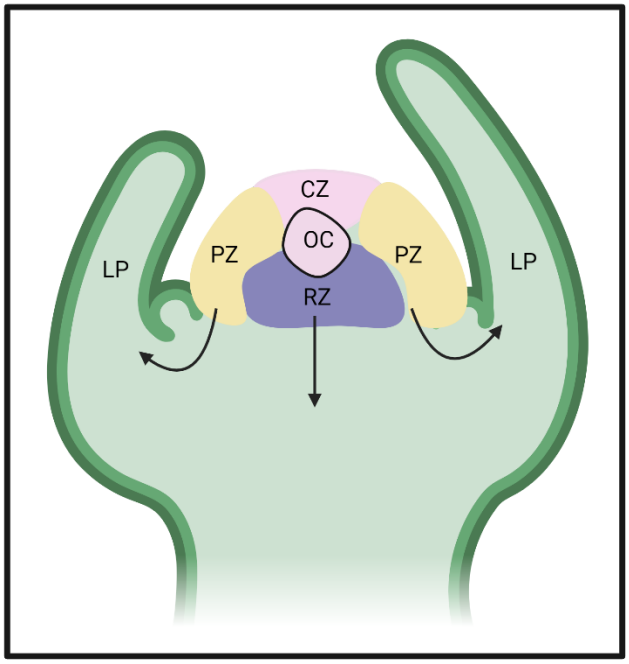


Figure 2.3. Functional domains of the shoot apical meristem (SAM). The organizing center (OC) is part of the central zone (CZ), which consists of a stem cell pool that replenishes cells to the peripheral zone (PZ) and rib zone (RB). The black arrows represent the direction of differentiating cells from the PZ to form leaf primordia (LP) and the RZ to form the stem. WUS expression is high in the OC and is regulated by CLV3/CLV1 from the CZ in a negative feedback loop. Figure created using BioRender (https://biorender.com/).

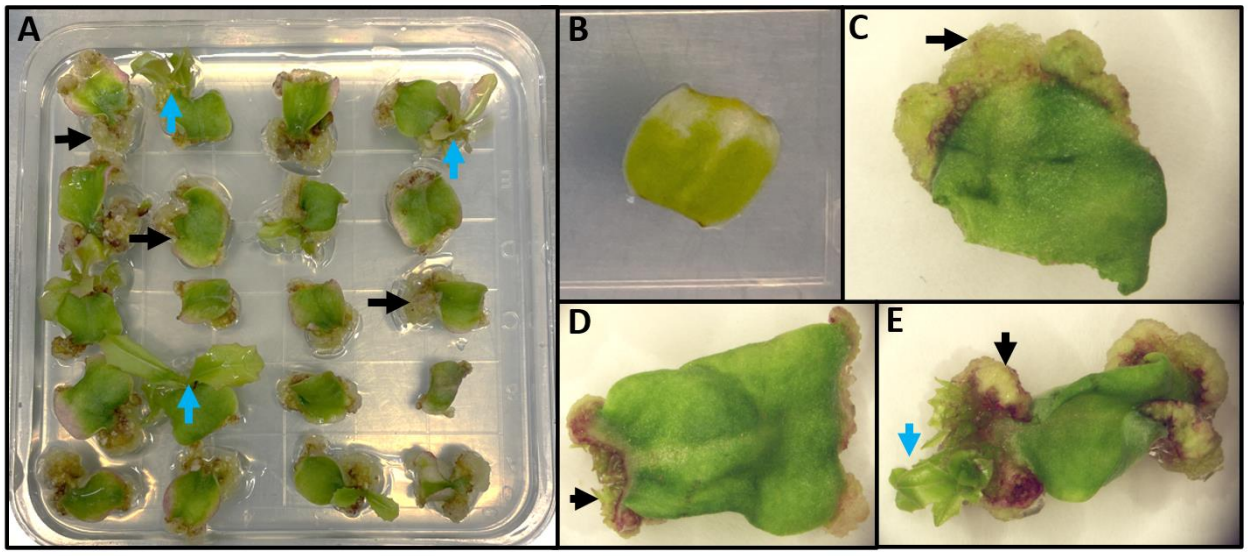


Figure 2.4. Indirect de novo shoot organogenesis in lettuce. **A.** A plate of 20 explants undergoing indirect de novo shoot regeneration. Black arrows represent friable callus formation at the wounded end of explants; blue arrows represent shoot regeneration from calli. **B.** An explant before callus formation. **C.** An explant during callus formation (black arrow). **D.** First organized growth from callus (black arrow) **E.** Indirect shoot regeneration (blue arrow) from callus (black arrow).

REFERENCES

Aida, M., Beis, D., Heidstra, R., Willemsen, V., Blilou, I., Galinha, C., Nussaume, L., Noh, Y. S., Amasino, R., and Scheres, B. (2004). The *PLETHORA* genes mediate patterning of the *Arabidopsis* root stem cell niche. *Cell*, *119*(1), 109–120.

Ampomah-Dwamena, C., Conner, A. J., and Fautrier, A. G. (1997). Genotypic response of lettuce cotyledons to regeneration in vitro. *Scientia Horticulturae*, *71*(3–4), 137–145.

Atta, R., Laurens, L., Boucheron-Dubuisson, E., Guivarc'h, A., Carnero, E., Giraudat-Pautot, V., Rech, P., and Chriqui, D. (2009). Pluripotency of *Arabidopsis* xylem pericycle underlies shoot regeneration from root and hypocotyl explants grown *in vitro*. *The Plant Journal*, *57*(4), 626–644.

Banno, H., Ikeda, Y., Niu, Q.-W., and Chua, N.-H. (2001). Overexpression of Arabidopsis ESR1 Induces Initiation of Shoot Regeneration. *The Plant Cell*, *13*(12), 2609.

Berger, N., Dubreucq, B., Roudier, F., Dubos, C., and Lepiniec, L. (2011). Transcriptional Regulation of Arabidopsis LEAFY COTYLEDON2 Involves RLE, a cis-Element That Regulates Trimethylation of Histone H3 at Lysine-27. *The Plant Cell*, *23*(11), 4065.

Bouchabké-Coussa, O., Obellianne, M., Linderme, D., Montes, E., Maia-Grondard, A., Vilaine, F., and Pannetier, C. (2013). Wuschel overexpression promotes somatic embryogenesis and induces organogenesis in cotton (*Gossypium hirsutum* L.) tissues cultured in vitro. *Plant Cell Reports 2013 32:5*, *32*(5), 675–686.

Boutilier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C. M., van Lammeren, A. A. M., Miki, B. L. A., Custers, J. B. M., and van Lookeren Campagne, M. M. (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell*, *14*(8), 1737–1749.

Bowman, J. L., and Eshed, Y. (2000). Formation and maintenance of the shoot apical meristem. *Trends in Plant Science*, *5*(3), 110–115.

Brand, A., Quimbaya, M., Tohme, J., and Chavarriaga-Aguirre, P. (2019). Arabidopsis LEC1 and LEC2 Orthologous Genes Are Key Regulators of Somatic Embryogenesis in Cassava. *Frontiers in Plant Science*, *10*.

Brand, U., Fletcher, J. C., Hobe, M., Meyerowitz, E. M., and Simon, R. (2000). Dependence of Stem Cell Fate in Arabidopsis on a Feedback Loop Regulated by CLV3 Activity. *Science*, *289*(5479), 617–619.

Braybrook, S. A., Stone, S. L., Park, S., Bui, A. Q., Le, B. H., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (2006). Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(9), 3468–3473.

Buechel, S., Leibfried, A., To, J. P. C., Zhao, Z., Andersen, S. U., Kieber, J. J., and Lohmann, J. U. (2010). Role of A-type ARABIDOPSIS RESPONSE REGULATORS in meristem maintenance and regeneration. *European Journal of Cell Biology*, *89*(2–3), 279–284.

Bustillo-Avendaño, E., Ibáñez, S., Sanz, O., Sousa Barros, J. A., Gude, I., Perianez-Rodriguez, J., Micol, J. L., del Pozo, J. C., Moreno-Risueno, M. A., and Pérez-Pérez, J. M. (2018). Regulation of Hormonal Control, Cell Reprogramming, and Patterning during De Novo Root Organogenesis. *Plant Physiology*, *176*(2), 1709–1727.

Casamitjana-Martínez, E., Hofhuis, H. F., Xu, J., Liu, C. M., Heidstra, R., and Scheres, B. (2003). Root-specific CLE19 overexpression and the so1l/2 suppressors implicate a CLV-like pathway in the control of Arabidopsis root meristem maintenance. *Current Biology*, *13*(16), 1435–1441.

Choffe, K. L., Victor, J. M. R., Murch, S. J., and Saxena, P. K. (2000). In vitro regeneration of *Echinacea purpurea* L.: Direct somatic embryogenesis and indirect shoot organogenesis in petiole culture. *In Vitro Cell. Dev. Biol.-Plant*, *36*, 30–36.

Chokheli, V. A., Dmitriev, P. A., Rajput, V. D., Bakulin, S. D., Azarov, A. S., Varduni, T. V., Stepanenko, V. V., Tarigholizadeh, S., Singh, R. K., Verma, K. K., and Minkina, T. M. (2020). Recent Development in Micropropagation Techniques for Rare Plant Species. *Plants 2020, Vol. 9, Page 1733*, *9*(12), 1733.

Clark, S. E., Running, M. P., and Meyerowitz, E. M. (1995). CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same processes as CLAVATA1. *Development*, *121*(7), 2057–2067.

Curaba, J., Moritz, T., Blervaque, R., Parcy, F., Raz, V., Herzog, M., and Vachon, G. (2004). AtGA3ox2, a Key Gene Responsible for Bioactive Gibberellin Biosynthesis, Is Regulated during Embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis. *Plant Physiology*, *136*(3), 3660.

Curtis, I. S., Power, J. B., Blackhall, N. W., de Laat, A. M. M., and Davey, M. R. (1994). Genotype-independent transformation of lettuce using *Agrobacterium tumefaciens*. *Journal of Experimental Botany*, *45*(279), 1441–1449.

Dai, X., Liu, Z., Qiao, M., Li, J., Li, S., and Xiang, F. (2017). ARR12 promotes de novo shoot regeneration in *Arabidopsis thaliana* via activation of WUSCHEL expression. *Journal of Integrative Plant Biology*, *59*(10), 747–758.

Daimon, Y., Takabe, K., and Tasaka, M. (2003). The CUP-SHAPED COTYLEDON Genes Promote Adventitious Shoot Formation on Calli. *Plant and Cell Physiology*, *44*(2), 113–121.

de Almeida, M., Vieira De Almeida, C., Mendes, E., Gilvano, G. •, Brondani, E., and Fiori De Abreu-Tarazi, M. (2012). Pre-procambial cells are niches for pluripotent and totipotent stem-like cells for organogenesis and somatic embryogenesis in the peach palm: a histological study. *Plant Cell Reports*, 31, 1495–1515.

de Smet, I., Vassileva, V., de Rybel, B., Levesque, M. P., Grunewald, W., van Damme, D., van Noorden, G., Naudts, M., van Isterdael, G., de Clercq, R., Wang, J. Y., Meuli, N., Vanneste, S., Friml, J., Hilson, P., Jürgens, G., Ingram, G. C., Inzé, D., Benfey, P. N., and Beeckman, T. (2008). Receptor-like kinase ACR4 restricts formative cell divisions in the Arabidopsis root. *Science*, *322*(5901), 594–597.

Debernardi, J. M., Mecchia, M. A., Vercruyssen, L., Smaczniak, C., Kaufmann, K., Inze, D., Rodriguez, R. E., and Palatnik, J. F. (2014). Post-transcriptional control of GRF transcription factors by microRNA miR396 and GIF co-activator affects leaf size and longevity. *The Plant Journal*, *79*(3), 413–426.

Debernardi, J. M., Tricoli, D. M., Ercoli, M. F., Hayta, S., Ronald, P., Palatnik, J. F., and Dubcovsky, J. (2020). A GRF–GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nature Biotechnology 2020 38:11*, *38*(11), 1274–1279.

DiDonato, R. J., Arbuckle, E., Buker, S., Sheets, J., Tobar, J., Totong, R., Grisafi, P., Fink, G. R., and Celenza, J. L. (2004). Arabidopsis ALF4 encodes a nuclear-localized protein required for lateral root formation. *Plant Journal*, *37*(3), 340–353.

Doerschug, M. R., and Miller, C. O. (1967). Chemical Control of Adventitious Organ Formation in *Lactuca sativa* Explants. In *Source: American Journal of Botany*, *54*(4).

Durgaprasad, K., Roy, M. v., Venugopal M., A., Kareem, A., Raj, K., Willemsen, V., Mähönen, A. P., Scheres, B., and Prasad, K. (2019). Gradient Expression of Transcription Factor Imposes a Boundary on Organ Regeneration Potential in Plants. *Cell Reports*, *29*(2), 453-463.

Elhiti, M. and Stasolla, C. (2022). Transduction of Signals during Somatic Embryogenesis. *Plants*, *11*(2), 178.

Emery, J. F., Floyd, S. K., Alvarez, J., Eshed, Y., Hawker, N. P., Izhaki, A., Baum, S. F., and Bowman, J. L. (2003). Radial Patterning of Arabidopsis Shoots by Class III HD-ZIP and KANADI Genes. *Current Biology*, *13*(20), 1768–1774.

Fan, M., Xu, C., Xu, K., and Hu, Y. (2012). LATERAL ORGAN BOUNDARIES DOMAIN transcription factors direct callus formation in Arabidopsis regeneration. *Cell Research*, *22*(7), 1169–1180.

Farina, P., Fambrini, M., Pugliesi, C., and Viviani, A. (2021). Expression of homeobox genes during in vitro culture of *Lactuca sativa*. *Plant Biosystems*, *155*(3), 609–621.

Fehér, A. (2019). Callus, Dedifferentiation, Totipotency, Somatic Embryogenesis: What These Terms Mean in the Era of Molecular Plant Biology? *Frontiers in Plant Science.*

Feng, Z., Zhu, J., Du, X., and Cui, X. (2012). Effects of three auxin-inducible LBD members on lateral root formation in *Arabidopsis thaliana*. *Planta*, *236*(4), 1227–1237.

Filipecki, M. K., Sommer, H., and Malepszy, S. (1997). The MADS-box gene *CUS1* is expressed during cucumber somatic embryogenesis. *Plant Science*, *125*(1), 63–74.

Fraga, H. P. F., Vieira, L. N., Caprestano, C. A., Steinmacher, D. A., Micke, G. A., Spudeit, D. A., Pescador, R., and Guerra, M. P. (2012). 5-Azacytidine combined with 2,4-D improves somatic embryogenesis of *Acca sellowiana* (O. Berg) Burret by means of changes in global DNA methylation levels. *Plant Cell Reports*, *31*(12), 2165–2176.

Fukaki, H., Tameda, S., Masuda, H., and Tasaka, M. (2002). Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of Arabidopsis. *The Plant Journal*, *29*(2), 153–168.

Gaj, M. D., Zhang, S., Harada, J. J., and Lemaux, P. G. (2005). Leafy cotyledon genes are essential for induction of somatic embryogenesis of Arabidopsis. *Planta*, *222*(6), 977–988.

Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R., and Scheres, B. (2007). PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. *Nature*, *449*(7165), 1053–1057.

Gallois, J.-L., Nora, F. R., Mizukami, Y., and Sablowski, R. (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes and Development*, *18*(4), 375–380.

Gallois, J.-L., Woodward, C., Reddy, G. V., and Sablowski, R. (2002). Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in Arabidopsis. *Development*, *129*(13), 3207–3217.

Gardiner, J., Donner, T. J., and Scarpella, E. (2010). Simultaneous Activation of SHR and ATHB8 Expression Defines Switch to Preprocambial Cell State in Arabidopsis Leaf Development. *Developmental Dynamics*, *240*(1), 261–270.

Gao, F., Wang, K., Liu, Y., Chen, Y., Chen, P., Shi, Z., Luo, J., Jiang, D., Fan, F., Zhu, Y., and Li, S. (2015). Blocking miR396 increases rice yield by shaping inflorescence architecture. *Nature Plants*, *2*(1), 1–9.

Gautheret, R. J. (1934). Culture du tissus cambial. *Comptes Rendus Hebdomadaires Des Se´ances de l'Acade´mie Des Sciences*, *15*, 2195–2196.

Gautheret, R. J. (1935). *Recherches sur la culture des tissus végétaux*. Thèse Sc. Paris pp. 279.

Gautheret, R. J. (1939). Sur la possibilite´ de re´aliser la culture inde´finie des tissus de tubercules de carotte. *Comptes Rendus Hebdomadaires Des Se´ances de l'Acade´mie Des Sciences*, *208*, 118–120.

Gautheret, R. J. (1942). He´te´ro-auxines et cultures de tissus ve´ge´taux. *Bulletin de La Socie´te´de Chimie Biologique*, *24*, 13–21.

Gautheret, R. J. (1955). Sur la variabilite´ des proprie´te´s physiologiques des cultures de tissues ve´ge´taux. *Revista General de Bota^nica*, *65*, 5–112.

Gazzarrini, S., Tsuchiya, Y., Lumba, S., Okamoto, M., and McCourt, P. (2004). The Transcription Factor FUSCA3 Controls Developmental Timing in Arabidopsis through the Hormones Gibberellin and Abscisic Acid. *Developmental Cell*, *7*(3), 373–385.

Goh, T., Joi, S., Mimura, T., and Fukaki, H. (2012). The establishment of asymmetry in Arabidopsis lateral root founder cells is regulated by LBD16/ASL18 and related LBD/ASL proteins. *Development*, *139*(5), 883–893.

Gordon, S. P., Heisler, M. G., Reddy, G. V., Ohno, C., Das, P., and Meyerowitz, E. M. (2007). Pattern formation during de novo assembly of the Arabidopsis shoot meristem. *Development*, *134*(19), 3539–3548.

Grzybkowska, D., Morończyk, J., Wójcikowska, B., and Gaj, M. D. (2018). Azacitidine (5-AzaC)-treatment and mutations in DNA methylase genes affect embryogenic response and expression of the genes that are involved in somatic embryogenesis in Arabidopsis. *Plant Growth Regulation*, *85*(2), 243–256.

Guo, F., Liu, C., Xia, H., Bi, Y., and Zhao, C. (2013). Induced Expression of AtLEC1 and AtLEC2 Differentially Promotes Somatic Embryogenesis in Transgenic Tobacco Plants. *PLoS ONE*, *8*(8).

Hake, S., Char, B. R., Chuck, G., Foster, T., Long, J., and Jackson, D. (1995). Homeobox genes in the functioning of plant meristems. *Philosophical Transactions of the Royal Society B, 350* (1331), 45-51.

Hake, S., Smith, H. M. S., Holtan, H., Magnani, E., Mele, G., and Ramirez, J. (2004). The Role of KNOX Genes in Plant Development. *Annual Review of Cell and Developmental Biology*, *20*, 125–151.

Han, R., Truco, M. J., Lavelle, D. O., and Michelmore, R. W. (2021). A Composite Analysis of Flowering Time Regulation in Lettuce. *Frontiers in Plant Science*.

Harding, E. W., Tang, W., Nichols, K. W., Fernandez, D. E., and Perry, S. E. (2003). Expression and Maintenance of Embryogenic Potential Is Enhanced through Constitutive Expression of AGAMOUS-Like 15. *Plant Physiology*, *133*(2), 653.

He, C., Chen, X., Huang, H., and Xu, L. (2012). Reprogramming of H3K27me3 Is Critical for Acquisition of Pluripotency from Cultured Arabidopsis Tissues. *PLOS Genetics*, *8*(8), e1002911.

Hecht, V., Vielle-Calzada, J.-P., Hartog, M. v, Schmidt, E. D. L., Boutilier, K., Grossniklaus, U., and de Vries, S. C. (2001). The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 Gene Is Expressed in Developing Ovules and Embryos and Enhances Embryogenic Competence in Culture 1. *Plant Physiology* , *127*(3), 803–816.

Heidmann, I., de Lange, B., Lambalk, J., Angenent, G. C., and Boutilier, K. (2011). Efficient sweet pepper transformation mediated by the BABY BOOM transcription factor. *Plant Cell Reports*, *30*(6), 1107–1115.

Henderson, J. T., Li, H. C., Rider, S. D., Mordhorst, A. P., Romero-Severson, J., Cheng, J. C., Robey, J., Sung, Z. R., de Vries, S. C., and Ogas, J. (2004). PICKLE Acts throughout the Plant to Repress Expression of Embryonic Traits and May Play a Role in Gibberellin-Dependent Responses. *Plant Physiology*, *134*(3), 995–1005.

Hewezi, T., Maier, T. R., Nettleton, D., and Baum, T. J. (2012). The Arabidopsis MicroRNA396-GRF1/GRF3 Regulatory Module Acts as a Developmental Regulator in the Reprogramming of Root Cells during Cyst Nematode Infection. *Plant Physiology*, *159*(1), 321–335.

Hirakawa, Y., Shinohara, H., Kondo, Y., Inoue, A., Nakanomyo, I., Ogawa, M., Sawa, S., Ohashi-Ito, K., Matsubayashi, Y., and Fukuda, H. (2008). Non-cell-autonomous control of vascular stem cell fate by a CLE peptide/receptor system. *Proceedings of the National Academy of Sciences*, *105*(39), 15208–15213.

Hofhuis, H., Laskowski, M., Du, Y., Prasad, K., Grigg, S., Pinon, V., and Scheres, B. (2013). Phyllotaxis and Rhizotaxis in Arabidopsis Are Modified by Three PLETHORA Transcription Factors. *Current Biology*, *23*(11), 956–962.

Horstman, A., Li, M., Heidmann, I., Weemen, M., Chen, B., Muino, J. M., Angenent, G. C., and Boutiliera, K. (2017). The BABY BOOM Transcription Factor Activates the LEC1-ABI3-FUS3-LEC2 Network to Induce Somatic Embryogenesis. *Plant Physiology*, *175*(2), 848–857.

Hu, J., Mitchum, M. G., Barnaby, N., Ayele, B. T., Ogawa, M., Nam, E., Lai, W. C., Hanada, A., Alonso, J. M., Ecker, J. R., Swain, S. M., Yamaguchi, S., Kamiya, Y., and Suna, T. P. (2008). Potential sites of bioactive gibberellin production during reproductive growth in Arabidopsis. *Plant Cell*, *20*(2), 320–336.

Hu, X., and Xu, L. (2016). Transcription Factors WOX11/12 Directly Activate WOX5/7 to Promote Root Primordia Initiation and Organogenesis 1. *Plant Physiology*, *172*(4), 2363–2373.

Ikeda, Y., Banno, H., Niu, Q.-W., Howell, S. H., and Chua, N.-H. (2006). The ENHANCER OF SHOOT REGENERATION 2 gene in Arabidopsis Regulates CUP-SHAPED COTYLEDON 1 at the Transcriptional Level and Controls Cotyledon Development. *Plant and Cell Physiology*, *47*(11), 1443–1456.

Ikeuchi, M., Iwase, A., Rymen, B., Lambolez, A., Kojima, M., Takebayashi, Y., Heyman, J., Watanabe, S., Seo, M., de Veylder, L., Sakakibara, H., and Sugimoto, K. (2017). Wounding Triggers Callus Formation via Dynamic Hormonal and Transcriptional Changes. *Plant Physiology*, *175*(3), 1158–1174.

Ikeuchi, M., Ogawa, Y., Iwase, A., and Sugimoto, K. (2016). Plant regeneration: Cellular origins and molecular mechanisms. *Development*, 143(9), 1442–1451.

Ishihara, H., Sugimoto, K., Tarr, P. T., Temman, H., Kadokura, S., Inui, Y., Sakamoto, T., Sasaki, T., Aida, M., Suzuki, T., Inagaki, S., Morohashi, K., Seki, M., Kakutani, T., Meyerowitz, E. M., and Matsunaga, S. (2019). Primed histone demethylation regulates shoot regenerative competency. *Nature Communications 2019 10:1*, *10*(1), 1–15.

Iwase, A., Harashima, H., Ikeuchi, M., Rymen, B., Ohnuma, M., Komaki, S., Morohashi, K.,

Kurata, T., Nakata, M., Ohme-Takagi, M., Grotewold, E., and Sugimoto, K. (2017). WIND1 Promotes Shoot Regeneration through Transcriptional Activation of ENHANCER OF SHOOT REGENERATION1 in Arabidopsis. *The Plant Cell*, *29*(1), 54–69.

Iwase, A., Kondo, Y., Laohavisit, A., Takebayashi, A., Ikeuchi, M., Matsuoka, K., Asahina, M., Mitsuda, N., Shirasu, K., Fukuda, H., and Sugimoto, K. (2021). WIND transcription factors orchestrate wound-induced callus formation, vascular reconnection and defense response in Arabidopsis. *New Phytologist*, *232*(2), 734–752.

Iwase, A., Mitsuda, N., Ikeuchi, M., Ohnuma, M., Koizuka, C., Kawamoto, K., Imamura, J., Ezura, H., and Sugimoto, K. (2013). Arabidopsis WIND1 induces callus formation in rapeseed, tomato, and tobacco. *Plant Signaling and Behavior*, *8*(12).

Iwase, A., Mitsuda, N., Koyama, T., Hiratsu, K., Kojima, M., Arai, T., Inoue, Y., Seki, M., Sakakibara, H., Sugimoto, K., and Ohme-Takagi, M. (2011a). The AP2/ERF transcription factor WIND1 controls cell dedifferentiation in Arabidopsis. *Current Biology*, *21*(6), 508–514.

Iwase, A., Ohme-Takagi, M., and Sugimoto, K. (2011b). WIND1: A key molecular switch for plant cell dedifferentiation. *Plant Signaling and Behavior*, *6*(12), 1943–1945.

Jones, T., Lowe, K., Hoerster, G., Anand, A., Wu, E., Wang, N., Arling, M., Lenderts, B., and Gordon-Kamm, W. (2019). Maize Transformation Using the Morphogenic Genes. *Methods in Molecular Biology*, *1864*, 81–93.

Junker, A., Mönke, G., Rutten, T., Keilwagen, J., Seifert, M., Thi, T. M. N., Renou, J. P., Balzergue, S., Viehöver, P., Hähnel, U., Ludwig-Müller, J., Altschmied, L., Conrad, U., Weisshaar, B., and Bäumlein, H. (2012). Elongation-related functions of LEAFY COTYLEDON1 during the development of *Arabidopsis thaliana*. *The Plant Journal*, *71*(3), 427–442.

Kadkade, P., and Seibert, M. (1977). Phytochrome-regulated organogenesis in lettuce tissue culture. *Nature*, *270*(2), 49–50.

Kareem, A., Durgaprasad, K., Sugimoto, K., Du, Y., Pulianmackal, A. J., Trivedi, Z. B., Abhayadev, P. v, Pinon, V., Meyerowitz, E. M., Scheres, B., and Prasad, K. (2015). *PLETHORA* Genes Control Regeneration by a Two-step Mechanism. *Current Biology*, *25*(8), 1017–1030.

Katsir, L., Davies, K. A., Bergmann, D. C., and Laux, T. (2011). Peptide Signaling in Plant Development. *Current Biology*, *21*(9), R356–R364.

Kim, J. H. (2019). Biological roles and an evolutionary sketch of the GRF-GIF transcriptional complex in plants. *BMB Reports*, *52*(4), 227.

Kim, J. H., and Botella, J. R. (2004). Etr1-1 gene expression alters regeneration patterns in transgenic lettuce stimulating root formation. *Plant Cell, 78,* 69-73.

Kim, J. H., Choi, D., and Kende, H. (2003). The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. *The Plant Journal*, *36*(1), 94–104.

Kim, J., Yang, W., Forner, J., Lohmann, J. U., Noh, B., and Noh, Y. (2018). Epigenetic reprogramming by histone acetyltransferase HAG1/AtGCN5 is required for pluripotency acquisition in Arabidopsis. *The EMBO Journal*, *37*(20).

Kim, J.-S., Mizoi, J., Kidokoro, S., Maruyama, K., Nakajima, J., Nakashima, K., Mitsuda, N., Takiguchi, Y., Ohme-Takagi, M., Kondou, Y., Yoshizumi, T., Matsui, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2012). Arabidopsis GROWTH-REGULATING FACTOR7 Functions as a Transcriptional Repressor of Abscisic Acid-and Osmotic Stress-Responsive Genes, Including DREB2A W. *The Plant Cell*, *24*(8), 3393–3405.

Kirch, T., Simon, R., Grünewald, M., and Werr, W. (2003). The DORNRÖSCHEN/ENHANCER OF SHOOT REGENERATION1 Gene of Arabidopsis Acts in the Control of Meristem Cell Fate and Lateral Organ Development. *The Plant Cell*, *15*(3), 694.

Koevary, K. Tissue culture propagation of head lettuce. HortScience, 12, 459-460.

Alconero, R. Regeneration of plants from cell suspensions of *Lactuca saligna*, *Lactuca sativa*, and *Lactuca serriola*. *HortScience*, *18*(3), 305-307.

Köhler, C., and Hennig, L. (2010). Regulation of cell identity by plant polycomb and trithorax group proteins. *Current Opinion in Genetics and Development*, *20*(5), 541–547.

Kong, J., Martin-Ortigosa, S., Finer, J., Orchard, N., Gunadi, A., Batts, L. A., Thakare, D., Rush, B., Schmitz, O., Stuiver, M., Olhoft, P., and Pacheco-Villalobos, D. (2020). Overexpression of the Transcription Factor GROWTH-REGULATING FACTOR5 Improves Transformation of Dicot and Monocot Species. *Frontiers in Plant Science, 11*.

Kornet, N., and Scheres, B. (2009). Members of the GCN5 histone acetyltransferase complex regulate PLETHORA-mediated root stem cell niche maintenance and transit amplifying cell proliferation in Arabidopsis. *Plant Cell*, *21*(4), 1070–1079.

Krikorian, A. D., and Berquam, D. L. (1969). Plant cell and tissue cultures: the role of HABERLANDT 1. *The Botanical Review*, *35*, 59–67.

Kuijt, S. J. H., Greco, R., Agalou, A., Shao, J., ’t Hoen, C. C. J., Övernäs, E., Osnato, M., Curiale, S., Meynard, D., van Gulik, R., Maraschin, S. de F., Atallah, M., de Kam, R. J., Lamers, G. E. M., Guiderdoni, E., Rossini, L., Meijer, A. H., and Ouwerkerk, P. B. F. (2014). Interaction between the GROWTH-REGULATING FACTOR and KNOTTED1-LIKE HOMEOBOX Families of Transcription Factors. *Plant Physiology*, *164*(4), 1952–1966.

Kwiatkowska, D. (2004). Structural integration at the shoot apical meristem: models, measurements, and experiments. *American Journal of Botany*, *91*(9), 1277–1293.

Lardon, R., and Geelen, D. (2020). Natural variation in plant pluripotency and regeneration. *Plants*, 9(10), 1–28.

LaRue, C. D. (1933). Regeneration in Mutilated Seedlings. *Proceedings of the National Academy of Sciences of the United States of America*, 53–63.

LaRue, C. D. (1936). Tissue Culture of Spermatophytes. *Proceedings of the National Academy of the United States of America*, *22*(4), 201–209.

Lee, B. H., Ko, J. H., Lee, S., Lee, Y., Pak, J. H., and Kim, J. H. (2009). The Arabidopsis GRF-INTERACTING FACTOR Gene Family Performs an Overlapping Function in Determining Organ Size as Well as Multiple Developmental Properties. *Plant Physiology*, *151*(2).

Lee, H. W., Cho, C., Pandey, S. K., Park, Y., Kim, M. J., and Kim, J. (2019). LBD16 and LBD18 acting downstream of ARF7 and ARF19 are involved in adventitious root formation in Arabidopsis. *BMC Plant Biology*, *19*(1).

Lee, K., Park, O. S., Choi, C. Y., and Seo, P. J. (2019). ARABIDOPSIS TRITHORAX 4 Facilitates Shoot Identity Establishment during the Plant Regeneration Process. *Plant and Cell Physiology*, *60*(4), 826–834.

Lee, K., Park, O. S., Go, J. Y., Yu, J., Han, J. H., Kim, J., Bae, S., Jung, Y. J., and Seo, P. J. (2021). Arabidopsis ATXR2 represses de novo shoot organogenesis in the transition from callus to shoot formation. *Cell Reports*, *37*(6), 109980.

Lee, K., Park, O. S., and Seo, P. J. (2018). JMJ30-mediated demethylation of H3K9me3 drives tissue identity changes to promote callus formation in Arabidopsis. *The Plant Journal*, *95*(6), 961–975.

Leelavathi, S., Sunnichan, V. G., Kumira, R., Vijaykanth, G. P., Bhatnagar, R. K., and Reddy, V. S. (2004). A simple and rapid Agrobacterium-mediated transformation protocol for cotton (*Gossypium hirsutum* L.): Embryogenic calli as a source to generate large numbers of transgenic plants. *Plant Cell Reports*, *22*, 465-47.

Li, W., Liu, H., Cheng, Z. J., Su, Y. H., Han, H. N., Zhang, Y., and Zhang, X. S. (2011). DNA Methylation and Histone Modifications Regulate De Novo Shoot Regeneration in Arabidopsis by Modulating WUSCHEL Expression and Auxin Signaling. *PLOS Genetics*, *7*(8), e1002243.

Liu, H., Zhang, H., Dong, Y. X., Hao, Y. J., and Zhang, X. S. (2018). DNA METHYLTRANSFERASE1-mediated shoot regeneration is regulated by cytokinin-induced cell cycle in Arabidopsis. *New Phytologist*, *217*(1), 219–232.

Liu, J., Sheng, L., Xu, Y., Li, J., Yang, Z., Huang, H., and Xu, L. (2014). WOX11 and 12 Are Involved in the First-Step Cell Fate Transition during de Novo Root Organogenesis in Arabidopsis. *The Plant Cell*, *26*(3), 1081–1093.

Liu, Z., Dai, X., Li, J., Liu, N., Liu, X., Li, S., and Xiang, F. (2020). The Type-B Cytokinin Response Regulator ARR1 Inhibits Shoot Regeneration in an ARR12-Dependent Manner in Arabidopsis. *The Plant Cell*, *32*(7), 2271–2291.

Loberant, B., and Altman, A. (2010). Micropropagation of Plants. *Encyclopedia of Industrial Biotechnology: Bioprocess, Bioseparation, and Cell Technology,* ed. Michael C. Flickinger. John Wiley and Sons, Inc., 1-17.

Loomis, W.G. and Schull, C.A. (1937). Methods in Plant Physiology, p.58. McGraw-Hull, New York.

Lotan, T., Ohto, M. A., Matsudaira Yee, K., West, M. A. L., Lo, R., Kwong, R. W., Yamagishi, K., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (1998). Arabidopsis LEAFY COTYLEDON1 Is Sufficient to Induce Embryo Development in Vegetative Cells. *Cell*, *93*(7), 1195–1205.

Ma, W., Wu, F., Sheng, P., Wang, X., Zhang, Z., Zhou, K., Zhang, H., Hu, J., Lin, Q., Cheng, Z., Wang, J., Zhu, S., Zhang, X., Guo, X., Wang, H., Wu, C., Zhai, H., and Wan, J. (2017). The LBD12-1 Transcription Factor Suppresses Apical Meristem Size by Repressing Argonaute 10 Expression. *Plant Physiology*, *173*(1), 801.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and Köhler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Reports*, *7*(9), 947–952.

Mantiri, F. R., Kurdyukov, S., Lohar, D. P., Sharopova, N., Saeed, N. A., Wang, X. D., Vandenbosch, K. A., and Rose, R. J. (2008). The Transcription Factor MtSERF1 of the ERF Subfamily Identified by Transcriptional Profiling Is Required for Somatic Embryogenesis Induced by Auxin Plus Cytokinin in *Medicago truncatula*. *Plant Physiology*, *146*(4), 1622–1636.

Matsuo, N., Makino, M., and Banno, H. (2011). Arabidopsis ENHANCER OF SHOOT REGENERATION (ESR)1 and ESR2 regulate in vitro shoot regeneration and their expressions are differentially regulated. *Plant Science*, *181*(1), 39–46.

Matsuo, N., Mase, H., Makino, M., Takahashi, H., and Banno, H. (2009). Identification of ENHANCER OF SHOOT REGENERATION 1-upregulated genes during in vitro shoot regeneration. *Plant Biotechnology*, *26*(4), 385–393.

Mayer, K. F. X., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T. (1998). Role of WUSCHEL in Regulating Stem Cell Fate in the Arabidopsis Shoot Meristem. *Cell*, *95*(6), 805–815.

McConnell, J. R., Emery, J., Eshed, Y., Bao, N., Bowman, J., and Barton, M. K. (2001). Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. *Nature 2001 411:6838*, *411*(6838), 709–713.

Méndez-Hernández, H. A., Ledezma-Rodríguez, M., Avilez-Montalvo, R. N., Juárez-Gómez, Y. L., Skeete, A., Avilez-Montalvo, J., De-La-Peña, C., and Loyola-Vargas, V. M. (2019). Signaling Overview of Plant Somatic Embryogenesis. *Frontiers in Plant Science*, *10*.

Meng, W. J., Cheng, Z. J., Sang, Y. L., Zhang, M. M., Rong, X. F., Wang, Z. W., Tang, Y. Y., and Zhang, X. S. (2017). Type-B ARABIDOPSIS RESPONSE REGULATORs Specify the Shoot Stem Cell Niche by Dual Regulation of WUSCHEL. *The Plant Cell*, *29*(6), 1357–1372.

Miao, C., Wang, D., He, R., Liu, S., and Zhu, J. K. (2020). Mutations in MIR396e and MIR396f increase grain size and modulate shoot architecture in rice. *Plant Biotechnology Journal*, *18*(2), 491–501.

Michelmore, R. W. and Eash, J. A. (1985). "Lettuce," in the Handbook of Plant Cell Culture vol 4, ed. D.A. Evans, W.R. Sharp, and P.V. Ammirato, 512-550.

Michelmore, R., Marsh, E., Seely, S., and Landry, B. (1987). Transformation of lettuce (*Lactuca sativa*) mediated by *Agrobacterium tumefaciens*. *Plant Cell Reports*, *6*, 439–442.

Miller, C., Skoog, F., von Saltza, M. H., and Strong, F. M. (1955). Kinetin, a cell division factor from desoxyribonucleic acid. *Journal of the American Chemical Society*, *77*, 1392.

Mitchum, M. G., Yamaguchi, S., Hanada, A., Kuwahara, A., Yoshioka, Y., Kato, T., Tabata, S., Kamiya, Y., and Sun, T. P. (2006). Distinct and overlapping roles of two gibberellin 3-oxidases in Arabidopsis development. *Plant Journal*, *45*(5), 804–818.

Mohebodini, M., Javaran, M. J., and Alizadeh, H. (2011). Effects of genotype, explant age and growth regulators on callus induction and direct shoot regeneration of Lettuce (*Lactuca sativa* L.). *Australian Journal of Crop Science*, 5(1), 92-95.

Mozgová, I., Muñoz-Viana, R., and Hennig, L. (2017). PRC2 Represses Hormone-Induced Somatic Embryogenesis in Vegetative Tissue of *Arabidopsis thaliana*. *PLOS Genetics*, *13*(1), e1006562.

Müller, B., and Sheen, J. (2008). Cytokinin and auxin interaction in root stem-cell specification during early embryogenesis. *Nature*, *453*(7198), 1094–1097.

Murthy, B. N. S., Victor, J., Singh, R. P., Fletcher, R. A., and Praveen, K. S. (1996). In vitro regeneration of chickpea (*Cicer arietinum* L.): Stimulation of direct organogenesis and somatic embryogenesis by thidiazuron. *Plant Growth Regulation*, *19*, 233–240.

Negin, B., Shemer, O., Sorek, Y., and Williams, L. E. (2017). Shoot stem cell specification in roots by the WUSCHEL transcription factor. *PLOS ONE*, *12*(4), e0176093.

Ning, K., Han, Y., Chen, Z., Luo, C., Wang, S., Zhang, W., Li, L., Zhang, X., Fan, S., and Wang, Q. (2019). Genome-wide analysis of MADS-box family genes during flower development in lettuce. *Plant Cell and Environment*, *42*(6), 1868–1881.

Nobe´court, P. (1955). Variations de la morphologie et de la structure de cultures de tissues ve´ge´taux. *Berichte Der Schweizerische Botanischen Gesellschaft*, *65*, 475–480.

Norstog, K. 1979. Embryo culture as a tool in the study of comparative and developmental morphology, p. 179–202. In: W.R. Sharp, P.O. Larsen, E.F. Paddock, and V. Raghavan. Plant cell and tissue culture. Ohio State Univ. Press, Columbus.

Nuti Ronchi, V., and Gregorini, G. (1970). Histological study of adventitious bud formation on *Lactuca sativa* cotyledons cultured in vitro. *Giornale Botanico Italiano*, *104*(6), 443–455.

Ogas, J., Cheng, J. C., Sung, Z. R., and Somerville, C. (1997). Cellular Differentiation Regulated by Gibberellin in the *Arabidopsis thaliana* pickle Mutant. *Science*, *277*(5322), 91–94.

Okushima, Y., Fukaki, H., Onoda, M., Theologis, A., and Tasaka, M. (2007). ARF7 and ARF19 Regulate Lateral Root Formation via Direct Activation of LBD/ASL Genes in Arabidopsis. *The Plant Cell*, *19*(1), 118–130.

Omidbakhshfard, M. A., Proost, S., Fujikura, U., and Mueller-Roeber, B. (2015). Growth-Regulating Factors (GRFs): A Small Transcription Factor Family with Important Functions in Plant Biology. *Molecular Plant*, *8*(7), 998–1010.

Ooi, S. E., Choo, C. N., Ishak, Z., and Ong-Abdullah, M. (2012). A candidate auxin-responsive expression marker gene, EgIAA9, for somatic embryogenesis in oil palm (*Elaeis guineensis* Jacq.). *Plant Cell, Tissue and Organ Culture*, *110*(2), 201–212.

Osorio-Montalvo, P., De-La-Peña, C., Carlos Oropeza, Geovanny Nic-Can, Córdova-Lara, I., Castillo-Castro, E., and Luis Sáenz-Carbonell. (2020). A peak in global DNA methylation is a key step to initiate the somatic embryogenesis of coconut palm (*Cocos nucifera* L). *Plant Cell Reports*, *39*, 1345–1357.

Ozias-akins, P., and Vasil, I. K. (1982). Plant Regeneration from Cultured Immature Embryos and Inflorescences *Triticum aestivum* L. (Wheat): Evidence for Somatic Embryogenesis. *Protoplasma*, 110, 95–105.

Paul, P., Joshi, S., Tian, R., Diogo Junior, R., Chakrabarti, § Manohar, and Perry, S. E. (2022). The MADS-domain factor AGAMOUS-Like18 promotes somatic embryogenesis. *Plant Physiology*, *188*, 1617–1631.

Perilli, S., di Mambro, R., and Sabatini, S. (2012). Growth and development of the root apical meristem. *Current Opinion in Plant Biology*, *15*(1), 17–23.

Preece, J. E. (2003). A Century of Progress with Vegetative Plant Propagation. *HORTSCIENCE*, 38(5), 1015-1025.

Quiroz-Figueroa, F. R., Rojas-Herrera, R., Galaz-Avalos, R. M., and Loyola-Vargas, V. M. (2006). Embryo production through somatic embryogenesis can be used to study cell differentiation in plants. *Plant Cell, Tissue and Organ Culture 2006 86:3*, *86*(3), 285–301.

Radhakrishnan, D., Shanmukhan, A. P., Kareem, A., Aiyaz, M., Varapparambathu, V., Toms, A., Kerstens, M., Valsakumar, D., Landge, A. N., Shaji, A., Mathew, M. K., Sawchuk, M. G., Scarpella, E., Krizek, B. A., Efroni, I., Mähönen, A. P., Willemsen, V., Scheres, B., and Prasad, K. (2020). A coherent feed-forward loop drives vascular regeneration in damaged aerial organs of plants growing in a normal developmental context. *Development (Cambridge)*, *147*(6).

Raza, G., Singh, M. B., and Bhalla, P. L. (2020). Somatic Embryogenesis and Plant Regeneration from Commercial Soybean Cultivars. *Plants*, *9*(1), 38.

Reyes-Chin-Wo, S., Wang, Z., Yang, X., Kozik, A., Arikit, S., Song, C., Xia, L., Froenicke, L., Lavelle, D. O., Truco, M. J., Xia, R., Zhu, S., Xu, C., Xu, H., Xu, X., Cox, K., Korf, I., Meyers, B. C., and Michelmore, R. W. (2017). Genome assembly with in vitro proximity ligation data and whole-genome triplication in lettuce. *Nature Communications 2017 8:1*, *8*(1), 1–11.

Rodriguez, R. E., Mecchia, M. A., Debernardi, J. M., Schommer, C., Weigel, D., and Palatnik, J. F. (2010). Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396. *Development*, *137*(1), 103–112.

Rueb, S., Leneman, M., Schilperoort, R. A., and Hensgens, L. A. M. (1994). Efficient plant regeneration through somatic embryogenesis from callus induced on mature rice embryos (*Oryza sativa* L.) *Plant Cell, Tissue and Organ Culture,* 36, 259-264.

Santos, D., and Fevereiro, P. (2002). Loss of DNA methylation affects somatic embryogenesis in *Medicago truncatula*. *Plant Cell, Tissue and Organ Culture*, *70*, 155–161.

Sarkar, A. K., Luijten, M., Miyashima, S., Lenhard, M., Hashimoto, T., Nakajima, K., Scheres, B., Heidstra, R., and Laux, T. (2007). Conserved factors regulate signaling in *Arabidopsis thaliana* shoot and root stem cell organizers. *Nature*, *446*(7137), 811–814.

Sasaki, H. (1982). Effect of Temperature and Light on Adventitious Bud Formation of Lettuce Hypocotyl Tissue Cultured in vitro. *Journal of the Japanese Society of Horticulture*, *51*(2), 187–194.

Sasaki, H. (1979). Physiological and Morphological Studies on Development of Vegetable Crops VI. Effect of Several Auxins, Cytokinins and Cytokinin-ribosides on the Adventitious Bud Formation of Lettuce Hypocotyl Tissue Cultured in vitro. *Physiological and Morphological Studies on Development of Vegetable Crops*, *48*(1), 67–72.

Schuabb Heringer, A., Barroso, T., Ferreira Macedo, A., Santa-Catarina, C., Martins, G. H., Souza, F., Iochevet, E., Floh, S., Apolinário De Souza-Filho, G., and Silveira, V. (2015). Label-Free Quantitative Proteomics of Embryogenic and Non-Embryogenic Callus during Sugarcane Somatic Embryogenesis. *PLOS One*, *10*(6).

Severino, L. S., Lima, R. L. S., Lucena, A. M. A., Freire, M. A. O., Sampaio, L. R., Veras, R. P., Medeiros, K. A. A. L., Sofiatti, V., and Arriel, N. H. C. (2011). Propagation by stem cuttings and root system structure of *Jatropha curcas*. *Biomass and Bioenergy*, *35*(7), 3160–3166.

Shang, B., Xu, C., Zhang, X., Cao, H., Xin, W., and Hu, Y. (2016). Very-long-chain fatty acids restrict regeneration capacity by confining pericycle competence for callus formation in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(18), 5101–5106.

Shchuka, V. M., Malek-Gilani, N., Singh, G., Langroudi, L., Dhaliwal, N. K., Moorthy, S. D., Davidson, S., Macpherson, N. N., and Mitchell, J. A. (2015). Chromatin Dynamics in Lineage Commitment and Cellular Reprogramming. *Genes 2015, Vol. 6, Pages 641-661*, *6*(3).

Shimotohno, A., Heidstra, R., Blilou, I., and Scheres, B. (2018). Root stem cell niche organizer specification by molecular convergence of PLETHORA and SCARECROW transcription factor modules. *Genes and Development*, *32*, 1085–1100.

Shin, J., Bae, S., and Seo, P. J. (2020). *De novo* shoot organogenesis during plant regeneration. *Journal of Experimental Botany*, *71*(1), 63–72.

Skoog, F., and Miller, C. O. (1957). Chemical regulation of growth and organ formation in plant tissue cultures in vitro. *Symposia of the Society for Experimental Biology*, *11*, 118–131.

Skoog, F., and Tsui, C. (1948). Chemical control of growth and bud formation in tobacco stem segments and callus cultured in vitro. *American Journal of Botany*, *35*, 782–787.

Stone, S. L., Braybrook, S. A., Paula, S. L., Kwong, L. W., Meuser, J., Pelletier, J., Hsieh, T. F., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (2008). Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. *Proceedings of the National Academy of Sciences*, *105*(8), 3151–3156.

Su, Y. H., and Zhang, X. S. (2014). The Hormonal Control of Regeneration in Plants. In *Current Topics in Developmental Biology*, 108, 35–69.

Sugimoto, K., Jiao, Y., and Meyerowitz, E. M. (2010). Arabidopsis Regeneration from Multiple Tissues Occurs via a Root Development Pathway. *Developmental Cell*, *18*(3), 463–471.

Sugimoto, K., Temman, H., Kadokura, S., and Matsunaga, S. (2019). To regenerate or not to regenerate: factors that drive plant regeneration. In *Current Opinion in Plant Biology*, 47, 138–150.

Sun, S., Kang, X. P., Xing, X. J., Xu, X. Y., Cheng, J., Zheng, S. W., and Xing, G. M. (2015). Agrobacterium-mediated transformation of tomato (*Lycopersicon esculentum* L. cv. Hezuo 908) with improved efficiency. *Biotechnology and Biotechnological Equipment*, *29*(5), 861–868.

Szczygieł-Sommer, A., and Gaj, M. D. (2019). The miR396–GRF Regulatory Module Controls the Embryogenic Response in Arabidopsis via an Auxin-Related Pathway. *International Journal of Molecular Sciences 2019, Vol. 20, Page 5221*, *20*(20), 5221.

Thorpe, T. A. (2007). History of plant tissue culture. In *Molecular Biotechnology*, 37(2), 169–180.

Tian, Q., Uhlir, N. J., and Reed, J. W. (2002). Arabidopsis SHY2/IAA3 inhibits auxin-regulated gene expression. *Plant Cell*, *14*(2), 301–319.

Trinh, D. C., Lavenus, J., Goh, T., Boutté, Y., Drogue, Q., Vaissayre, V., Tellier, F., Lucas, M., Voß, U., Gantet, P., Faure, J. D., Dussert, S., Fukaki, H., Bennett, M. J., Laplaze, L., and Guyomarc'h, S. (2019). PUCHI regulates very long chain fatty acid biosynthesis during lateral root and callus formation. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(28), 14325–14330.

Tsuda, K., Ito, Y., Sato, Y., and Kurata, N. (2011). Positive Autoregulation of a KNOX Gene Is Essential for Shoot Apical Meristem Maintenance in Rice. *The Plant Cell*, *23*(12), 4368–4381.

van den Berg, C., Willemsen, V., Hendriks, G., Weisbeek, P., and Scheres, B. (1997). Short-range control of cell differentiation in the Arabidopsis root meristem. *Nature*, *390*(6657), 287–289.

van der Knaap, E., Kim, J. H., and Kende, H. (2000). A Novel Gibberellin-Induced Gene from Rice and Its Potential Regulatory Role in Stem Growth. *Plant Physiology*, *122*(3), 695.

van Overbeek, J., Conklin, M. E., and Blakeslee, A. F. (1941).). Factors in coconut milk essential for growth and development of very young *Datura* embryos. *Science*, *94*(26), 350–351.

Victorathisayam, T., and Sridevi, G. (2020). Ectopic Expression of WUSCHEL (AtWUS) Gene Alters Plant Growth and Development in. *Rice. Plant*, *8*(3), 43–53.

Waki, T., Hiki, T., Watanabe, R., Hashimoto, T., and Nakajima, K. (2011). The Arabidopsis RWP-RK Protein RKD4 Triggers Gene Expression and Pattern Formation in Early Embryogenesis. *Current Biology*, *21*(15), 1277–1281.

Wang, X., Niu, Q. W., Teng, C., Li, C., Mu, J., Chua, N. H., and Zuo, J. (2008). Overexpression of PGA37/MYB118 and MYB115 promotes vegetative-to-embryonic transition in Arabidopsis. *Cell Research 2009 19:2*, *19*(2), 224–235.

Webb, D.T., Torres, L.D., Fobert, P. Interactions of growth regulators, explant age, and culture environment controlling organogenesis from lettuce cotyledons in vitro. *Canadian Journal of Botany*, *62*(3).

Wójcikowska, B., and Gaj, M. D. (2017). Expression profiling of AUXIN RESPONSE FACTOR genes during somatic embryogenesis induction in Arabidopsis. *Plant Cell Reports*, *36*(6), 843–858.

Xiong, F., Zhang, B.-K., Liu, H.-H., Wei, G., Wu, J.-H., Wu, Y.-N., Zhang, Y., and Li, S. (2020). Transcriptional Regulation of PLETHORA1 in the Root Meristem Through an Importin and Its Two Antagonistic Cargos. *The Plant Cell*, *32*, 3812–3824.

Xu, J., Hofhuis, H., Heidstra, R., Sauer, M., Friml, J., and Scheres, B. (2006). A Molecular Framework for Plant Regeneration. *Science*, *311*(5759), 385–388.

Xu, K., Liu, J., Fan, M., Xin, W., Hu, Y., and Xu, C. (2012). A genome-wide transcriptome profiling reveals the early molecular events during callus initiation in Arabidopsis multiple organs | Elsevier Enhanced Reader. *Genomics*, 100(2), 116-124.

Xu, L., and Huang, H. (2014). Genetic and Epigenetic Controls of Plant Regeneration. In *Current Topics in Developmental Biology*, 108, 1–33.

Xu, M., Hu, T., Zhao, J., Park, M.-Y., Earley, K. W., Wu, G., Yang, L., and Poethig, R. S. (2016). Developmental Functions of miR156-Regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Genes in *Arabidopsis thaliana*. *PLOS Genetics*, *12*(8), e1006263.

Xu, Z., Zhang, C., Ge, X., Wang, N., Zhou, K., Yang, X., Wu, Z., Zhang, X., Liu, C., Yang, Z., Li, C., Liu, K., Yang, Z., Qian, Y., and Li, F. (2015). Construction of a high-density linkage map and mapping quantitative trait loci for somatic embryogenesis using leaf petioles as explants in upland cotton (*Gossypium hirsutum* L.). *Plant Cell Reports*, *34*(7), 1177–1187.

Xue, T., Dai, X., Wang, R., Wang, J., Liu, Z., and Xiang, F. (2017). ARGONAUTE10 Inhibits In Vitro Shoot Regeneration Via Repression of miR165/166 in *Arabidopsis thaliana*. *Plant and Cell Physiology*, *58*(10), 1789–1800.

Yamaguchi, S., Kamiya, Y., and Sun, T. P. (2001). Distinct cell-specific expression patterns of early and late gibberellin biosynthetic genes during Arabidopsis seed germination. *Plant Journal*, *28*(4), 443–453.

Yamaguchi, Y. L., Ishida, T., and Sawa, S. (2016). CLE peptides and their signaling pathways in plant development. *Journal of Experimental Botany*, *67*(16), 4813–4826.

Yasui, Y., Tsukamoto, S., Sugaya, T., Nishihama, R., Wang, Q., Kato, H., Yamato, K.T., Fukaki, H., Mimura, T., Kubo, H., Theres, K., Kohchi, T., Ishizaki, K. (2019). GEMMA CUP_ASSOCIATED MYB1, an Ortholog of Axillary Meristem Regulators Is Essential in Vegetative Reproduction in *Marchantia polyphorma*. *Current Biology*, *29*(23), 3986-3995.

Zhang, B., Tong, Y., Luo, K., Zhai, Z., Liu, X., Shi, Z., Zhang, D., and Li, D. (2021). Identification of GROWTH-REGULATING FACTOR transcription factors in lettuce (*Lactuca sativa*) genome and functional analysis of LsaGRF5 in leaf size regulation. *BMC Plant Biology*, *21*(1), 485.

Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W., and Chua, N.-H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nature Protocols*, *1*, 641–646.

Zhang, Y., Jiao, Y., Jiao, H., Zhao, H., Zhu Y. Two-Step Functional Innovation of the Stem-Cell Factors WUS/WOX5 during Plant Evolution. *Molecular Biology and Evolution*, *34*(3), 640-653.

Zheng, Q., Zheng, Y., Ji, H., Burnie, W., and Perry, S. E. (2016). Gene Regulation by the AGL15 Transcription Factor Reveals Hormone Interactions in Somatic Embryogenesis. *Plant Physiology*, *172*(4), 2374–2387.

Zheng, Q., Zheng, Y., and Perry, S. E. (2013). AGAMOUS-Like15 Promotes Somatic Embryogenesis in Arabidopsis and Soybean in Part by the Control of Ethylene Biosynthesis and Response. *Plant Physiology*, *161*(4), 2113–2127.

Zheng, Y., Ren, N., Wang, H., Stromberg, A. J., and Perry, S. E. (2009). Global Identification of Targets of the Arabidopsis MADS Domain Protein AGAMOUS-Like15. *The Plant Cell*, *21*(9), 2563–2577.

Zhou, C., Han, L., Zhao, Y., Wang, H., Nakashima, J., Tong, J., Xiao, L., and Wang, Z. Y. (2019). Transforming compound leaf patterning by manipulating REVOLUTA in *Medicago truncatula*. *The Plant Journal*, *100*(3), 562–571.

Chapter 3: Genetic analysis of indirect *de novo* organogenesis in lettuce (*L. sativa* L.)

ACKNOWLEDGEMENTS

This chapter presents joint work with Rongkui Han and Allen Van Deynze.

I conducted the majority of the work including experimental design, initiating and maintaining tissue cultures, phenotyping, and data analysis. Rongkui Han constructed the genetic map and identified the variants within the candidate genes between the parents and the *L. sativa* cv. Salinas v11 reference gene annotation. Allen Van Deynze provided input and R scripts for data analysis, BLUP extractions, and multiple QTL mapping.

ABSTRACT

Plant *in vitro* regeneration is an important process for successful transgenesis and genome editing in plants. Multiple studies have focused on identifying molecular constituents of indirect *de novo* shoot organogenesis in other species, but little work has been conducted in non-model species such as lettuce (*Lactuca sativa* L.). The use of mapping populations and the implementation of genetic analyses of quantitative trait loci (QTLs) has allowed for the identification of loci and candidate genes involved during *in vitro* regeneration of several species. Like most species, *in vitro* regeneration is genotype dependent in lettuce, with the presence of high regenerating genotypes (*L. sativa* oil seed PI251246) and low regenerating genotypes (*L. serriola* accession Armenian 999). In this study, we used an f_8 population of 236 recombinant inbred lines (RILs) developed between an interspecific cross between parents, Armenian 999 and PI251246, with differing regeneration capacities. Multiple regeneration traits were observed over a period of 45 days for a total of 116 RILs in this population. Multiple QTL (MQTL) mapping was performed to identify significant QTLs associated with each trait and resulted in full QTL models explaining

17.1 to 59.7% of the phenotypic variation. In total, eight QTLs were identified with major QTLs revealed on Chromosomes 3, 8, and 9, and overlapping QTLs between regeneration traits. Preliminary analysis of candidate genes under each QTL revealed multiple developmental and transcription factor genes known to be involved in both *de novo* organogenesis and somatic embryogenesis in other species.

INTRODUCTION

Plant *in vitro* regeneration is an important process for successful transgenesis and genome editing in plants. Shoot regeneration can follow two pathways, *de novo* shoot organogenesis or somatic embryogenesis. Both pathways can occur via direct or indirect regeneration, where indirect regeneration contains an intermediate step of pluripotent callus formation (Chapter 2). Currently, *Agrobacterium*-mediated transformation is the most widely used method for development of transgenic and genome edited crops, which requires *in vitro* plant regeneration. Lettuce (*Lactuca sativa* L.), the focus of this chapter, regenerates by indirect shoot organogenesis.

Multiple studies have focused on identifying molecular constituents of indirect *de novo* shoot organogenesis in other species (reviewed in Chapter 2). Callus formation is the first step in indirect shoot organogenesis and is induced upon placement of explants on auxin rich medium. Pathways involved in callus and lateral root formation involve similar players and ectopic expression of genes required for lateral root formation triggered callus formation without the addition of plant hormones (DiDonato et al., 2004; Fan et al., 2012; Feng et al., 2012; Shang et al., 2016; Sugimoto et al., 2010). Other molecular determinants involved during *in vitro* regeneration include APETALA/ETHYLENE-RESPONSE FACTOR (AP2/ERF) transcription factors (Iwase et al., 2011; Kareem et al., 2015). These transcription factors are induced upon wounding and promote formation of auxin gradients, which help regulate shoot-promoting transcription factors (Kareem et al., 2015; Shimotohno et al., 2018). The subsequent steps in *de novo* shoot regeneration include formation of meristematic centers and shoot regeneration. Formation and maintenance of meristems is controlled by a negative feedback mechanism between key meristematic regulators, *WUSCHEL* (*WUS*) and *CLAVATA3* (*CLV*) (Sarkar et al., 2007). The expression of leaf identity genes, HD-ZIP III transcription factor genes, which are epigenetically regulated by Trithorax

group (trxG) and Polycomb Group (PcG) proteins, follow the formation of meristematic centers (Köhler and Hennig, 2010; Shin et al., 2020). Although the molecular players involved *in vitro* regeneration of model species are increasingly understood, little work has been done to study these pathways in non-model species, such as lettuce.

Genetic analyses of quantitative trait loci (QTLs) have identified loci and candidate genes involved during *in vitro* regeneration of several species including *Arabidopsis* (Lall et al., 2004; Motte et al., 2014), rice (Kwon et al., 2002; Kwon et al., 2001), wheat (Ma et al., 2016), maize (Salvo et al., 2018), soybean (Yang, Zhao, Yu, et al., 2011), cucumber (Wang et al., 2018), and tomato (Trujillo-Moya et al., 2011). A variety of mapping populations have been used including recombinant inbred line (RIL), near isogenic line (NIL), and backcross (BC) populations derived from low and high regenerating parents. Combined, these studies have identified QTLs associated with regeneration traits such as callus productivity, somatic embryogenesis, and shoot organogenesis. In tomato, multiple traits involved during *in vitro* shoot organogenesis were assessed and QTLs associated with those traits were identified across multiple linkage groups accounting for approximately 17 to 52% of the phenotypic variance (Trujillo-Moya et al., 2011). The candidate genes *LESK1* and the acid invertase gene were identified in tomato but have not yet been validated. Fine mapping of QTLs associated with embryogenic tissue culture response in maize led to the identification of an approximately 3,000 kb region; however, candidate genes from this region again have not been validated (Salvo et al., 2018). In *Arabidopsis*, combination of linkage and association mapping led to the identification of *RECEPTOR-LIKE PROTEIN KINASE1* (*RPK1*) as an important determinant of *de novo* shoot organogenesis (Motte et al., 2014). With the increasing quality of genomic data, QTL mapping and association studies will continue to help reveal loci and candidate genes involved in *in vitro* regeneration.

In this chapter, we describe a genetic analysis of regeneration in lettuce using an F_8 RIL population derived from a cross between a high regenerating oil seed type of *L. sativa* and a low regenerating accession of *L. serriola*. This population was evaluated for *in vitro* regeneration efficiency using multiple traits. Traits were strongly correlated; in total there were eight significant QTLs for regeneration in lettuce with overlapping QTLs for several traits. Candidate genes underlying each QTL included determinants of *in vitro* regeneration previously identified in other species. Our findings contribute to the current knowledge of genetic determinants involved during indirect *de novo* shoot organogenesis and will help improve regeneration of recalcitrant lettuce genotypes. This knowledge can also be applied to other species of the Compositae family, such as sunflower, that are recalcitrant to regeneration.

METHODS

Plant material and experimental design

An F_8 population of 236 recombinant inbred lines (RILs) was produced from single seed descent from a cross between *L. serriola* accession Armenian 999 and *L. sativa* oil seed line PI251246 (Han et al., 2021). RILs and parents were scored using a randomized block design. Initially, 120 RILS were randomly selected from the full population. These 120 RILS were randomly assigned to four batches of 30 RILs each. The same 120 RILs were randomly assigned twice more to four batches, for a total of three replication per RIL. Each batch also included both parents as controls.

After the QTL analysis had been completed, RILs with recombination breakpoints within the interval of major QTL on Chromosome 3 were selected for further phenotyping to fine map the region. Recombinants within a 1.5 LOD confidence interval were selected by examination of

the genotypes of all RILs across the interval using Excel version 2202. Additional replications were added for all recombinant RILs, resulting in six total replications for each recombinant line.

Preparation and maintenance of tissue cultures

Seeds of RILs and both parents were surface sterilized using 20% bleach and 5 µL of Tween20. The seeds were soaked for 20 minutes with constant agitation (250 rpm) and rinsed once with 100 mL of sterile distilled water. Twenty seeds were sown on 1/2x Hoagland's medium (0.815 g/L Hoagland's salt mixture, 8 g/L Phytoagar™) in 60 mm x 15 mm Fisherbrand® Petri dishes and incubated at 21°C under a 9/15 hr light/dark cycle with LED (Fluence Bioengineering, INC) lights providing approximately 13,000 lux.

After four days, explants were prepared by cutting off apical and basal tips of each cotyledon and were immediately placed adaxial side down on SHI medium (3.2 g/L SH salts, 30 g/L sucrose, 2 mL/L 500x MS vitamins, 8 g/L Phytoagar™, 0.1mg/L 6-BAP, 0.1mg/L 1-NAA) in Petri dishes. Twenty explants per RIL, per plate were prepared and incubated at 21°C under a 9/15 hr light/dark cycle with LED (Fluence Bioengineering, INC) lights providing approximately 13,000 lux. After 14 days in culture, explants were transferred to fresh SHI medium and returned to the growth chamber. At day 35, up to 12 shoots from each RIL were transferred to rooting medium (4.33 g/L MS salts, 30 g/L sucrose, 2 mL/L 500x MS vitamins, 8g/L Phytoagar™) in sterile Magenta boxes and returned to the growth chamber. On day 45, cultures were removed from the growth chamber and final evaluation of traits was conducted.

Evaluation of regeneration

Evaluation of cultures was conducted every third and fourth day over a period of 45 days. Each day of observation consisted of identifying the number of explants exhibiting the following traits: callus formation (CF), organized growth (OG), leaf emergence (LE), and multiple leaves (ML) per culture. These data were normalized by the total number of explants in culture. On day 35, all cultures were scored for the total number of individual shoots. The average number of shoots per explant was calculated (AS). The occurrence of rooting (OR) was scored on days 38, 42, and 45. After 45 days, growth curves for each culture were plotted and the area under the curve (AUC) was calculated for CF, OG, LE, ML, and OR. AUC was calculated in Microsoft Excel Version 2202 using the following formula:

$$AUC = \sum_{i=1}^{n-1} \frac{(y_i + y_{i+1}) \times (t_{i+1} - t_i)}{2}$$

Where y_i is the normalized phenotypic score of each RIL at the ith day of observation, t is the number of days at the ith observation, and n is the number of observations. The mean AUC of CP, OG, LE, ML, and RC and the mean AS for each RIL after three replications was calculated.

Data analysis

Descriptive statistics were calculated for the population. An analysis of variance (one-way ANOVA, $\alpha = 0.05$) was performed to evaluate the effect of genotype on regeneration rate as well as identify the presence of a batch effect.

Data was manually corrected in two ways. The first was corrected based on batch averages using the following formula:

$$y_{iBC} = y_i + (\mu - \bar{x})$$

Where y_{iBC} is the batch corrected phenotypic value, y_i is the raw phenotypic value, μ is the population mean, and $\bar{x}$ is the batch mean. Second, best linearized unbiased predictions (BLUPs) were extracted utilizing the lme4 package in R version 3.6.2 (2019-12-12) using the following model:

$$y_i \sim (1|Gt) + (1|Rep[Batch])$$

Where y_i is the raw phenotypic value, $(1|Gt)$ is referring to the RIL as a random effect, and $(1|Rep[Batch])$ is referring to the batch effect nested in replication as a random effect. BLUP adjusted ($BLUP_{adj}$) means were calculated by adding the grand mean to the extracted BLUP value. Distribution and correlations among raw means, batch corrected means, and BLUP adjusted means were analyzed and compared. Correlation between phenotypes and datasets was calculated using Pearson's correlation coefficient. All three datasets were used for QTL mapping.

Genetic map and QTL analysis

A reference-free k-mer based genetic map was developed using the AFLAP method (Fletcher et al., 2021). Thirty-one-mer hashes were produced independently for the read sets of each parent and individual progeny. The parental (F_0) hashes were then inspected to identify single-

copy 31-mers. To reduce redundancy, 31-mers for each parent were assembled using ABySS v2.2.2 (Jackman et al., 2017). Assembled fragments equal to or greater than 61 bp were extracted and a single representative 31-mer equal to coordinates 10 to 41 was selected for each fragment. The two sets of 31-mer markers from parents were then combined and scored against every progeny hash to obtain progeny genotypes. The markers in the progeny genotype table were ordered into a genetic map using Lep-MAP3 (Rastas, 2017). Reads for parents and RILs of the *L. serriola* accession Armenian 999 x *L. sativa* oil seed PI251246 interspecific population are available on NCBI under BioProjects PRJNA642889, PRJNA510128, and PRJNA478460 (Han et al., 2021).

All QTL analyses were conducted using the package "rqtl" in R version 3.6.2 (2019-12-12). Simple interval mapping was conducted to identify putative QTLs from the initial sampling of the population using 1,000-iteration permutation testing to identify a genome-wide significance threshold at $\alpha = 0.05$. Multiple QTL mapping (MQTL) with data derived from 1,000 imputations was performed using the scantwo and stepwiseqtl functions in R, allowing for additive and interactive effects (modified from Chunthawodtiporn et al., 2019). QTL modeling implemented a forward selection of up to ten QTLs, followed by a backward elimination to the null model. The model with the highest log of odds (LOD) ratio statistic was chosen. Significance thresholds for MQTL were calculated using a 1,000-iteration permutation test, and a penalized LOD (pLOD) score threshold was calculated for each trait. Individual terms with a LOD score greater than the genome-wide pLOD significance threshold and 5% phenotypic variance (PVE) represented were selected for the final adjusted model. Locations of 1.5 LOD drop intervals were identified using the refineqtl function in R. In addition, QTLs with PVEs greater than 10% were considered major QTLs (Chu et al., 2019; Li et al., 2020). QTLs were named using "q" followed by the phenotype,

dataset, and Chromosome. For example, a QTL identified on Chromosome 3 for OG raw means would be named *qOGR3*.

Candidate gene selection

K-mer markers were mapped to the PI251246 v1, Armenian 999 v1, and *L. sativa* cv. Salinas v11 genome assemblies to identify physical positions. A physical region of the PI251246 v1 and *L. sativa* cv. Salinas v11 assembly was extracted using the markers nearest to the flanking ends of the 1.5 LOD or fine mapped interval. PI251246 liftoff and *L. sativa* cv. Salinas v11 *de novo* annotations were analyzed within the 1.5 LOD or fine mapped interval. Genes known to be important players of *in vitro* regeneration were selected as candidate genes and physical distances of genes from the peak marker of the QTL were calculated.

RESULTS

Genetic map construction

The genotypes of RILs at 8,227 female and male specific k-mer markers were used to construct a genetic map. The genetic map was filtered to contain one k-mer marker per genetic bin for a total of 2,517 bin markers. These markers were quality filtered by analyzing the collinearity of the genetic position and the physical position in the *L. sativa* cv. Salinas v11 reference and the PI251246 and Armenian 999 draft genome assemblies. A total of 109 genetic markers were not collinear with flanking markers in the genome assembly and were therefore removed from the genetic marker dataset **(Supplementary Figure S3.1)**. The quality filtered genetic map contained 2,412 k-mer markers, which totaled 1,724 cM across nine chromosomal linkage groups (**Figure 3.1**). This is nearly identical to the genetic map generated for the same population using 2,677 SNP

markers (1,883 cM; Han et al., 2021) and similar to the genetic size of the reference mapping population *L. sativa* cv. Salinas x *L. serriola* UC96US23 (1,254 cM; Argyris et al., 2005). The mean distance between adjacent markers was 0.72 cM, with the majority (83.1%) of adjacent markers having a distance of less than 1 cM. The two largest gaps in the genetic map were 4.44 and 4.32 cM and were located between 187.641 and 192.089 cM on Chromosome 5 and 33.165 and 37.487 cM on Chromosome 7, respectively.

Analysis of regeneration efficiency

A total of 116 RILs were phenotyped for six different traits in three replicates randomized across 12 batches (**Supplementary Figures S3.2 and S3.3**). The mean and standard deviations of all traits for PI251246, Armenian 999, and RILs are shown in **Table 3.1**. No significant differences in CF ($p = 0.131$) and OR ($p = 0.13$) means were detected between parents; however, significant differences between parental means was observed for OG ($p=6^{-9}$), LE ($p = 1.93^{-8}$), ML ($p = 4.81^{-9}$), and AS ($p = 1.17^{-9}$). PI251246, on average, produced 2.22, 2.27, 2.5, and 3.85 times more OG, LE, ML, and AS than Armenian 999 (**Figure 3.2**). RIL population means differed significantly for all traits phenotyped with p values ranging from 6^{-4} to 2^{-16} (**Figure 3.3**). The population had continuous distributions for all traits **(Figure 3.4**). An ANOVA revealed a significant difference between batch means for CF ($p = 3.52^{-8}$), OG ($p = 0.04$), AIS ($p = 0.002$), and OR ($p = 0.003$) indicating the presence of a batch effect. To adjust for the batch effect, data was manually corrected based on batch means. An ANOVA on batch corrected means resulted in no significant differences between all traits except CF ($p = 0.001$). The distributions for each trait for each dataset of raw means, batch corrected means, and BLUP adjusted means are presented in **Figure 3.4.**

Pearson's correlation coefficients were used to calculate relationships between traits and datasets. Phenotypic traits were highly correlated with an r ranging from 0.893 to 0.995, 0.897 to 0.996, and 0.900 to 0.996 for raw, batch corrected, and BLUP adjusted means, respectively. Of all correlations tested, OR had the lowest correlation with all traits. Raw means, batch corrected means, and BLUP adjusted means were strongly correlated with batch corrected and BLUP adjusted means having the strongest correlation ($r > 0.99$) (**Figure 3.5**).

Simple interval mapping

In order to identify QTLs, we first used simple interval mapping (scanone function in R) using raw, batch corrected, and BLUP adjusted means, which revealed one to three putative QTLs for all phenotypes with overlapping QTLs between phenotypes (**Table 3.2**; **Supplementary Tables S3.1 and S3.2**). For all datasets, a putative QTL was revealed on Chromosome 2 for CF and a major QTL was revealed on Chromosome 3 across all other traits, except OR. An additional putative QTL was identified on Chromosome 1 for OG, LE, ML, and AS across all datasets. Other putative QTLs were revealed on Chromosome 4 for all traits across all datasets, except AS for raw and BLUP adjusted means. Using BLUP adjusted and batch corrected means, a putative QTL was also revealed on Chromosome 8. All three datasets were used for further analysis using the MQTL method. No significant putative QTLs were identified for OR; therefore, no further analysis was conducted for this trait.

Multiple QTL mapping

The three datasets (raw, batch corrected, and BLUP adjusted) were used for MQTL across all traits, and the results were compared. For each dataset, a total of eight similar QTLs were revealed to be associated with indirect *in vitro* regeneration in lettuce (**Tables 3.3 and 3.4;**

Supplementary Tables S3.3 to S3.6). These QTLs were located on Chromosomes 1, 3, 4, 8, and 9. QTLs specifically associated with individual traits varied slightly between datasets, with one to six QTLs identified for each. BLUP adjusted means resulted in the simplest model and explained the highest phenotypic variance with the fewest model terms. Therefore, results with BLUP adjusted means are described in detail below. Then, we briefly discuss similarities with QTLs revealed using the other two datasets (more detail in **Supplementary Tables S3.3 to S3.6**).

QTLs identified using BLUP adjusted means

BLUP adjusted means were calculated and used for MQTL mapping. Penalized LOD thresholds were calculated using 1,000-iteration permutation tests for each trait and ranged from 2.73 to 2.86. In total, eight QTLs were identified across all traits with multiple overlapping QTLs between traits (**Figure 3.6**). MQTL revealed one major QTL that was associated with CF on Chromosome 2 (PVE = 17.1%) and one major QTL on Chromosome 3 that was present for all other traits analyzed (PVE = 14.8% to 27.3%). In addition, QTLs identified on Chromosomes 1, 4, 8, and 9 were present for multiple traits, and one additional unique QTL was revealed on Chromosome 3 For AS. For the complete model, LOD scores ranged from 4.7 to 22.9 and had PVEs ranging from 17.1% to 59.7%. A full summary of the QTL models for CF, OG, and LE is described in **Table 3.3** and for ML and AS is described in **Table 3.4.**

Callus Formation

MQTL revealed one significant QTL associated with CF located on Chromosome 2 (*qCFB2*) (**Table 3.3**). This QTL had a LOD score of 4.7 and explained 17.1% of the phenotypic variance observed in the population. The male marker 21425_65 (44.742 cM) was located at the

peak of *qCFB2*, and the 1.5-LOD interval spanned 38.8 cM, with 6497_62 and 14691_67 as the nearest flanking markers of the interval.

Organized Growth

MQTL modeling revealed four QTLs associated with OG on Chromosomes 1, 3, 4, and 8 (**Table 3.3**). These QTL were named *qOGB1, qOGB3, qOGB4,* and *qOGB8,* respectively. The full model was composed of only additive effects, which resulted in a LOD score of 17.4 and represented 49.9% of the phenotypic variance. The major QTL, *qOGB3,* had the highest maximum LOD score of 9.4 and a PVE of 22.8%. The female marker 8472_64 (175.871 cM) was located at the peak of *qOGB3* and the 1.5 LOD interval, composed of 10 markers, spanned 6.705 cM. The remaining three QTLs had LOD scores ranging from 3.28 to 3.83 and PVEs ranging from 6.9 to 8.2%. The range of 1.5 LOD intervals covered 20.7 cM to 141.8 cM.

Leaf Emergence

Initially MQTL mapping and modeling revealed 10 QTLs with three interactions (**Figure 3.6**); however, only three major QTLs remained after deeper analysis and selection of QTLs that contained LOD scores exceeding the pLOD threshold of 2.86 and explained greater than 5% PVE (**Table 3.3)**. These QTLs were located on Chromosomes 3 (*qLEB3*), 8 (*qLEB8*), and 9 (*qLEB9*). The adjusted full model had a LOD score of 16.0, explained 47.0% of the phenotypic variance, and included both additive and interactive effects. The major QTL revealed for OG, *qOGB3*, was also present for LE; *qLEB3* had the highest LOD score (10.461) and PVE (27.28%) of all terms in the LE model. The 1.5 LOD interval covered 12.029 cM with the peak marker being male marker 6867_73, which is located 2.562 cM away from female marker 8472_64, which is located at the peak of *qOGB3*. The two additional major QTLs, *qLEB8* and *qLEB9*, both had maximum LOD

scores of 5.5 and PVEs of 12.9% and 13.1%, respectively. The 1.5 LOD interval covered 36.556 cM for *qLEB8* and 20 cM for *qLEB9*; *qOGB8* and *qLEB8* had different peak markers separated by approximately 93 cM and were, therefore, considered independent QTLs. One interaction was observed between *qLEB8* and *qLEB9*, which had a maximum LOD score of 2.9 and a PVE of 6.5%.

Multiple Leaves

MQTL revealed five QTLs associated with regeneration of ML and were located on Chromosomes 1, 3, 8, and 9, with two QTLs identified on Chromosome 8 (**Figure 3.6**; **Table 3.4**). The full model had a LOD score of 22.9, explained 59.7% of the phenotypic variation, and included three major QTLs. The same major QTL on Chromosome 3, *qMLB3*, was present and had the highest maximum LOD score (12.6) and PVE (26.1%) of all QTLs revealed for ML. The interval spanned 8.298 cM, consisted of 14 individual markers, and, like LE, male marker 6867_73 was located at the peak of *qMLB3*. QTLs identified on Chromosomes 8 (*qMLB8.2*) and 9 (*qMLB9*) had the second and third highest maximum LOD scores of 9.0 and 8.1, respectively. These QTLs accounted for 17.4% and 15.3% of the phenotypic variance and had intervals spanning 10.305 and 12.201 cM, respectively. A major QTL-QTL interaction was present between *qMLB8.2* and *qMLB9*, which was responsible for 11.5% of the phenotypic variance and had a maximum LOD score of 6.3. The remaining two QTLs, *qMLB1* and *qMLB8.1* had maximum LOD scores of 3.9 and 4.1, and explained 6.7 and 7.1% of the phenotypic variance, respectively. These QTLs also spanned 20.819 and 24.58 cM intervals. All QTLs associated with ML were revealed for OG and LE. Although the peaks of some QTLs were located at different genetic positions, the intervals overlapped and were, therefore, not considered to be unique, independent QTLs.

Average Shoots

MQTL revealed four QTLs, three of which overlapped with previous traits and one additional QTL located on Chromosome 3 (**Figure 3.6**; **Table 3.4**). The full model had a LOD score of 17.9, explained 50.9% of the phenotypic variance, and was composed of only additive effects. As with previous traits, a major QTL on Chromosome 3 was revealed, *qASB3.2*, which had a maximum LOD score of 6.7 and explained 14.8% of the phenotypic variance. The 1.5-LOD interval was the largest when compared to the major QTL on Chromosome 3 of other traits and spanned 22.586 cM. Similar to *qOGB3*, female marker 8472_64 was located at the peak of *qISB3.2*. Other overlapping QTLs, including a major QTL on Chromosome 8, were identified, although with slightly different peak positions than previously revealed; *qASB8* had a maximum LOD score of 6.6 and explained 14.5% of the phenotypic variance, while *qASB4* had a maximum of LOD score of 3.7 and explained 7.7% of the phenotypic variance. The additional, unique QTL revealed for AS was located on Chromosome 3, *qASB3.1*. This QTL had a maximum LOD score of 4.3 and explained 9.1% of the phenotypic variance.

Multiple QTL mapping using raw means

The putative QTLs identified using raw means were further analyzed using MQTL. Similar to BLUP adjusted means, eight QTLs in total were identified across all traits (**Supplementary Tables S3.3 and S3.4**). Four major QTLs were identified in total with one major QTL revealed for CF on Chromosome 2, one major QTL revealed for the remaining traits on Chromosome 3, and major QTLs on Chromosomes 8 and 9 for LE and ML. For all traits except CF, five to six QTLs were identified, which included overlapping QTLs between each trait. The major QTL revealed for CF, *qCFR2*, had a LOD score of 4.7 and a PVE of 17.1%. The major QTL revealed

on Chromosome 3 had a LOD score ranging from 9.4 to 11.7 and a PVE ranging from 20.6% to 26.6%. The QTL peak for *qCFR2* shifted 1.713 cM to male marker 12737_66 and the QTL peak for *qMLR3* shifted 2.56 cM to female marker 8472_64. Other major QTLs were identified on Chromosomes 8 and 9 for LE and ML traits, with LOD scores ranging from 7.8 to 8.9 and PVE ranging from 12.9% to 17.4%, respectively; *qLER8.2* and *qMLR8.2* had the same peak position as *qLEB8*; *qLER9* and *qMLR9* had different peak positions that were separated by 4.727 cM. Both peaks differed from *qLEB9* and *qMLB9*. Additional QTLs were revealed on Chromosome 1 across all traits, except CF, and Chromosome 4 for OG and AS. Full QTL models for each trait had LOD scores ranging from 4.5 to 26.0 and PVE ranging from 16.4% to 64.7%, with ML having the largest for both. Both additive and interactive effects were observed for LE, ML, and AS. A detailed summary of QTL models for CF, OG, and LE is represented in **Supplementary Table S3.3** and traits ML and AS are represented in **Supplementary Table S3.4**.

Multiple QTL mapping using batch corrected means

As with BLUP adjusted and raw means, full QTL models using stepwise regression were analyzed for each trait of the batch corrected means dataset. Resulting models revealed eight similar QTLs identified for previous datasets with major QTLs present on Chromosomes 2, 3, 8, and 9 (**Supplementary Tables S3.5 and S3.6**); *qCFC2* had a maximum LOD score of 4.7, a PVE of 16.3%, and had a matching peak marker with *qCFR2*. The major QTL on Chromosome 3 had maximum LOD scores ranging from 9.0 to 13.4 and PVEs ranging from 20.0% to 23.9%. The marker 8472_64 was located at the peak of the major QTL on Chromosome 3 for all traits, matching previous models and datasets. Other major QTLs on Chromosome 8 and Chromosome 9 also overlapped with previous datasets with LOD scores ranging from 6.1 to 8.6 and PVEs ranging from 11.5% to 13.8%. The marker located at the peak of *qLEC8.2* and *qMLC8.2* was the

same marker reported for *qLEB8*, and the marker at the peak of *qMLC9* was the same marker as reported for *qMLR9*. Full models resulted in LOD scores ranging from 4.7 to 27.4 and PVEs ranging from 16.3% to 66.6% with ML representing the highest for each. Full models for OG and AS included only additive effects, while models for LE and ML included additive QTLs and one QTL–QTL interaction between QTLs revealed on Chromosomes 8 and 9. A detailed summary of each QTL model identified for CF, OG, and LE is shown in **Supplementary Table S3.5**, and a summary of QTL models for ML and AS is shown in **Supplementary Table S3.6**.

Estimated effects and transgressive segregation

We focused on BLUP adjusted means to extract estimated effects from the MQTL models. Estimated effects determine the phenotypic effect of each QTL and identify which parental allele contributed to increased regeneration rates. The details of the estimated effects are shown in **Figure 3.7** and **Table 3.5**.The major QTL revealed for CF, *qCFB2*, had an estimated effect of 0.47 AUC. The major QTL present on Chromosome 3 had an estimated effect of approximately 1.9 AUC across OG, LE, and ML, and 0.16 for AS (**Figure 3.7**). The major QTL on Chromosome 8 (*qLEB8* and *qMLB8.2*) had an estimated effect of approximately 1.2 AUC. The third major QTL identified for LE (*qLEB9*) and ML (*qMLB9*) had estimated effects of 1.03 and 0.75 AUC. The remaining QTL had estimated effects ranging from 0.97 to 1.33 AUC and 0.12 to 0.16 AS. A significant interaction was observed between the QTLs revealed on Chromosomes 8 and 9 for LE and ML, which had estimated effects of 0.97 and 1.33, respectively (**Figure 3.8**). The PI25126 allele increased regeneration rates for all QTLs, except for the QTLs identified on Chromosome 9 (*qLEB9* and *qMLB9*) and Chromosome 2 (*qCFB2*) where the Armenian 999 allele contributed to increased regeneration rates.

The appearance of transgressive segregation is observed when phenotypes of progeny of segregating populations fall beyond the parental phenotypes. Some Armenian 999 x PI251246 RILs exhibited higher or lower mean regeneration rates across all traits, except for CF when compared to both parents (**Figure 3.3 and 3.7; Supplementary Figure S3.2**). For OG, LE, ML, and AS, 12.4, 14.2, 13.3, and 11.5%, respectively, of the population had lower regeneration rates than Armenian 999. Alternatively, 17.7, 16.8, 14.2, 11.5% had higher regeneration rates than PI251246 for the same traits, respectively. This observation was further supported by the presence of QTLs with both parental alleles contributing to higher rates of regeneration (**Table 3.5**).

Fine mapping of the major QTL on Chromosome 3

To further characterize a QTL associated with regeneration in lettuce, we focused on the major QTL on Chromosome 3, which was detected for all phenotypes across all datasets. Eighteen RILs with recombination events between 172.868 and 180.586 cM on Chromosome 3 were selected from the total population of 236 RILs. Nine of these 18 RILS had been previously phenotyped and had OG, LE, ML, and AS ranging from 2.4 to 20.59, 1.78 to 18.75, 1.08 to 17.88, and 0.075 to 1.825, respectively. Data was collected on three additional replications of these nine RILs, as well as on six replications of the remaining nine RILs that had not been phenotyped in the primary screen. New means calculated after additional phenotyping were used to fine map this region. RIL11, RIL58, and RIL138 were particularly informative because they exhibited recombination breakpoints near the peak markers, 8472_64 and 6867_73 **(Figure 3.9).** RIL11 had a breakpoint directly left (marker 2298_61) and RIL58 had a breakpoint directly right (marker 16393_61) of marker 8472_64. RIL138 had a marker directly right (17930_61) of marker 6867_73. For all traits, RIL58 had a regeneration rate that was in the bottom half of all RILs used for fine mapping, while RIL11 had a regeneration rate that was located in the top half (**Figure 3.9**).

RIL138 means were located in the top half of data for all traits, except AS. These RILs led to the identification of a 1.72 cM region located between marker 17930_61 and marker 2298_61 that may contain an important organogenesis candidate gene in lettuce. The phenotypes of the majority of RILs were consistent with their genotypes across this interval; however, RILs 103 and 115 had incongruent phenotypes and genotypes across this interval.

Physical locations of intervals and candidate genes

To identify physical coordinates of each QTL, we mapped the k-mer markers to *L. sativa* cv. Salinas reference genome assembly version 11. The majority of the physical intervals for all QTLs ranged from 11.9 Mb to 71.3 Mb. One QTL (*qOGB8*) had a physical interval spanning 199.8 Mb (**Table 3.7**). Using d*e novo* annotations of *L. sativa* cv. Salinas reference genome assembly version 11, a preliminary search of genes related to regeneration near the peaks of each QTL was conducted. All QTLs detected contain genes known to affect plant development and *in vitro* regeneration in other species (**Table 3.6**). Some candidate genes include MADS-box transcription factors, AP2/ERF transcription factors, MYB transcription factors, hormone signaling transcription factors, and multiple epigenetic regulators. Nucleotide alignments of parental versions of selected candidate genes with the Salinas v11 reference genome assembly resulted in multiple single nucleotide polymorphisms (SNPs) (**Table 3.6**). Functionalization of these variants will identify nonsynonymous and/or frameshift mutations of each parental candidate gene, which may narrow down the number of candidate genes used for validation studies.

DISCUSSION

In vitro regeneration is an important foundational process in biotechnological approaches to crop improvement. Although, plant regeneration techniques are widely used in multiple species,

successful regeneration is still genotype and species dependent and multiple, high value crops, such as cotton, pepper, and sunflower, are recalcitrant to regeneration (Gammoudi et al., 2018; Radonic et al., 2016; J. Wu et al., 2004). Further understanding of the regulatory networks, molecular determinants, and genetic loci that affect regeneration in multiple species will allow us to improve regeneration of recalcitrant cultivars. We identified important QTLs associated with indirect *de novo* shoot organogenesis in lettuce using an interspecific RIL population derived from a cross between *L. serriola* accession Armenian 999 and *L. sativa* oil seed PI251246, which had different regeneration capacities.

Regeneration in lettuce is genotype dependent and is under the control of multiple loci. MQTL identified QTLs associated with six traits throughout different stages of regeneration. Significant QTLs were revealed for five of the six traits and, in total, eight QTLs were identified with overlapping QTLs between traits. One QTL, *qCFB2*, was revealed for CF that was not present for other traits, suggesting that, in lettuce, callus formation is controlled by a unique locus that is not significantly involved in the progression of other regeneration traits such as organized growth, leaf formation, and shoot regeneration; *qCFB2* accounted for approximately 17% of the phenotypic variation, which is similar to the explained phenotypic variance of callus induction frequency in soybean (Yang et al., 2011). For the other traits, the complete models, including both main effects and interactions, explained 47% to 59.7% of the phenotypic variation, which is similar to PVE reported in other species including rice, tomato, and cucumber (Trujillo-Moya et al., 2011; Tyagi et al., 2010; Y. Wang et al., 2018). One major QTL on Chromosome 3 was consistently revealed across OG (*qOGB3*), LE (*qLEB3*), ML (*qMLB3*), and AS (*qASB3.2*) and had the highest LOD score and PVE for each model of each trait. This is likely related to a specific step in organogenesis that is required for the progression of regeneration. This is similar to a previous

study identifying one major locus that is associated with regeneration efficiency across multiple traits in tomato (Trujillo-Moya et al., 2011). MQTL using different datasets (raw means, batch corrected, BLUP adjusted) led to identification of the same eight QTLs, although with different combinations of QTLs for each trait except CF.

Alleles from both recalcitrant and high regenerating parents contributed to increased regeneration in lettuce. Individual RILs showed more extreme phenotypes than observed in either of the parents, indicating the presence of transgressive segregation. Along with alleles from the high regenerating parent, alleles from the recalcitrant line located under *qLEB9* and *qMLB9* contributed to increased regeneration in this RIL population. This is consistent with observations of transgressive segregation observed in other mapping populations used for *in vitro* regeneration studies in barley, *Arabidopsis*, and tomato (Mano et al., 1996; Schiantarelli et al., 2001; Trujillo-Moya et al., 2011). This also provides further evidence that *in vitro* regeneration is under the control of multiple loci, with specific allelic combinations leading to higher rates of regeneration even if the allele comes from the low regenerating parent.

Regeneration frequency is under polygenic control with genetic effects coming from multiple developmental and hormone signaling pathways. Multiple studies have shown an overlap between genes and molecular players involved during *in vitro* regeneration and plant development (reviewed in Chapter 2). Two major players seen in both *in vitro* regeneration and plant development are the family of MADS-box transcription factor genes, *AGAMOUS*. The *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes (Becker and Theißen, 2003; Ding, Ruan, Yu, Li, Song, Yang, and Gai, 2020). AGAMOUS-LIKE15 (AGL15) are key transcription factors during somatic embryogenesis, and downregulation of the *AGL15* gene allows for the transition of seed development to seed germination and vegetative growth (Zheng et al.,

2013; Zheng et al., 2009). Overexpression of *AGL15* increased callus formation and development of somatic embryos in *Arabidopsis* (Harding et al., 2003). In addition, *AGL15* directly regulates the LAFL complex, a master regulator of somatic embryogenesis, composed of LEAFY COTYLEDON1 (LEC1), ABSCISIC ACID INSENSITIVE (ABI13), FUSCA3 (FUS3), AND LEC2 (Horstman et al., 2017). Furthermore, *AGL15* regulates the accumulation of regulatory microRNA 156 (miR156), which post-transcriptionally regulates multiple genes during plant development and *in vitro* regeneration, including *SPL* genes (Nowak et al., 2020). Temporal expression of *SPL* genes is regulated by miR156 to control vegetative and floral development of plants (Wu and Poethig, 2006). In *Arabidopsis*, vegetative phase changes are regulated by *SPL3*, *SPL4*, and *SPL5* (Wu and Poethig, 2006).

There were genes encoding MADs-box transcription factors and *SPL* genes at or near peaks of detected QTLs. One candidate gene located 0.08 Mb away from the peak of the major QTL on Chromosome 3 was the MADS-box gene *AGAMOUS-LIKE 104* (*AGL104*). In *Arabidopsis*, *AGL104* functions in pollen maturation and tube growth (Adamczyk and Fernandez, 2009). In lettuce, 82 MADS-box genes have been identified, 10 of which belong to the AGL15 subfamily, which also included Arabidopsis and tomato *AGL15* homologs (Ning et al., 2019). However, *LsAGL104* was not represented in any of the subfamilies in this study and has not yet been functionalized in lettuce. A BLAST search of *AGL15* from *Arabidopsis* against the *L. sativa* cv. Salinas version 11 of the reference genome assembly revealed no significant hits within the lettuce genome. Therefore, *LsAGL104* in lettuce could be functionalized to participate in earlier stages of development such as leaf growth and maturation. An additional candidate gene located approximately 1.4 Mb away from the peak of *qOGB4* and *qASB4* is *LsSPL6*, which has been shown to be expressed in leaf primordia and is a downstream target of miR156 in *Arabidopsis* (Ding et

al., 2020; Xu et al., 2016); it may be likely that *LsAGL104* and *LsSPL6* indirectly regulate each other through the regulation of miR156. In addition, miR156 regulatory binds sites of *LsAGL104* and *LsSPL6* candidate genes could be further analyzed to identify potential mutations that may cause regulatory dysfunction in either of the parents leading to altered regeneration efficiencies.

A *KNOTTED1-LIKE HOMEOBOX GENE 6* (*KNAT6*) homolog in lettuce (*LsKNAT6*) was located 1.2 Mb downstream of the peaks of the major QTLs *qLEB8* and *qMLB8.2. KNAT6* is a *KNOTTED*-like homeobox (*KNOX*) gene and a closely related homolog of *SHOOT MERISTEMLESS* (*STM*), both of which function in plant development in *Arabidopsis* (Gordon et al, 2007). *STM* is a key meristem regulator and its expression in conjunction with *PIN-FORMED* (*PIN*) works in radial patterning of the shoot apical meristem (Gordon et al., 2007). *STM* expression is increased in callus cells of *Arabidopsis* cotyledon explants, where it delimits areas of promeristem formation during *de novo* shoot organogenesis (Gordon et al., 2007; Zhao et al., 2002). *STM* works in conjunction with *KNAT6* to promote meristem development, while *STM* works as an antagonist to the signaling peptide CLV3 to maintain meristem size (Nidhi et al., 2021). Additionally, *KNAT6* delimits the boundary between the shoot apical meristems and the cotyledons during embryogenesis or the emerging primordia during seedling development (Belles-Boix et al., 2006). Therefore the presence of a *KNAT6* homolog under these two QTLs is consistent with other *KNOX* genes, such as *STM*, playing a key role in meristem development during *in vitro* regeneration of lettuce.

In vitro regeneration in plants is accompanied by epigenetic changes and chromatin remodeling, which regulate cell fate transitions. Dedifferentiation and reprogramming of somatic cells to pluripotent cells are often a result of changes in methylation patterns and gene expression. Histone deacetylases (HDA), methyltransferases, and chromatin remodelers change methylation

and chromatin patterns in order to promote or decrease expression of specific genes throughout development (Fehér, 2015). Multiple epigenetic regulators were located at or near the peaks of several QTLs revealed in this study. One gene, *MEAF6*, encoding a chromatin modification related protein, was located approximately 0.83 Mb downstream of the peak of the major QTL revealed on Chromosome 3 for traits OG, LE, ML, and AS. A MEAF6-like isoform was shown to accumulate in globular embryos during somatic embryogenesis of cotton, suggesting that MEAF6 could potentially function in the early stages of *in vitro* regeneration in lettuce (Guo et al., 2019). *HDA15* was located 0.69 Mb downstream of the peak of *qOGB4*. *HDA15* is light-regulated and has functions during photomorphogenesis by suppressing hypocotyl elongation during germination and plays roles in auxin signaling in *Arabidopsis* (Kumar et al., 2021). These are both particularly interesting because of the epigenetic reprogramming required for changes in gene expression during early stages of indirect *in vitro* regeneration and the promotion of cell fate changes.

Further investigation and validation of the various potential candidate genes discussed above are required to show which genes are determining the phenotypic segregation in this population. Fine mapping of each QTL will narrow down candidate gene searches. Candidate genes nearest to the peak of major QTLs will be cloned and transformed into Armenian 999 and PI251246 to test for altered regeneration frequencies. In addition to the candidate genes described above, multiple predicted, uncharacterized proteins were found at or near the peak of each QTL. Characterizing these predicted proteins could lead to the discovery of novel genes involved in *in vitro* regeneration in lettuce. To our knowledge, this is the first study investigating loci responsible for successful *in vitro* regeneration in a Compositae species. Ectopic expression of genes that

stimulate regeneration could be exploited to increase recovery of plants in other recalcitrant genotypes of lettuce and species of the Compositae family, such as sunflower.

FIGURES AND TABLES

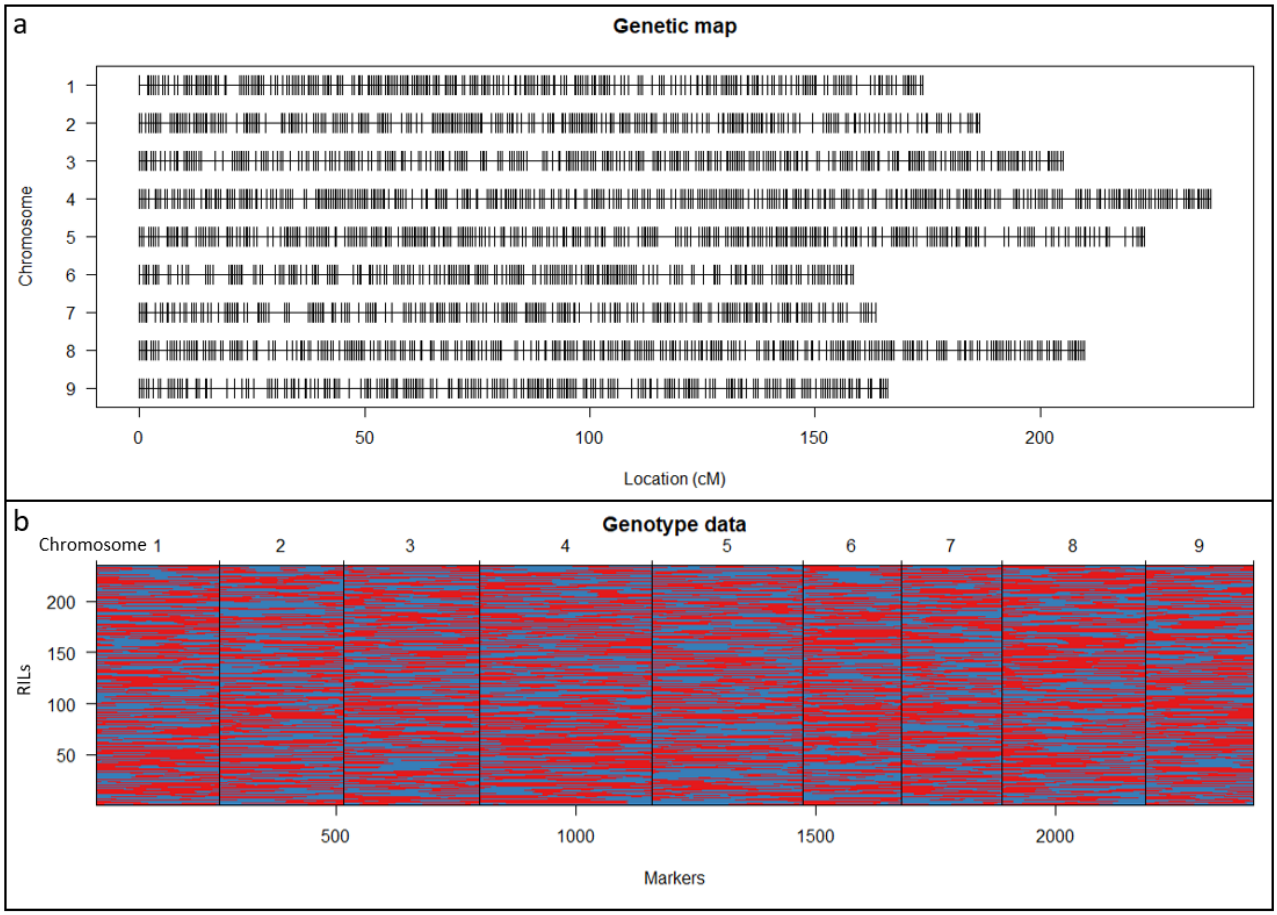


Figure 3.1. Genetic map and genotype data of the quality filtered k-mer markers used in this study. **a**) The genetic map representing 2,412 k-mer markers totaling 1,724 cM over 9 Chromosomes. **b**) The genotype data of the RIL population with blue representing the Armenian 999 allele and red representing the PI251246 allele.

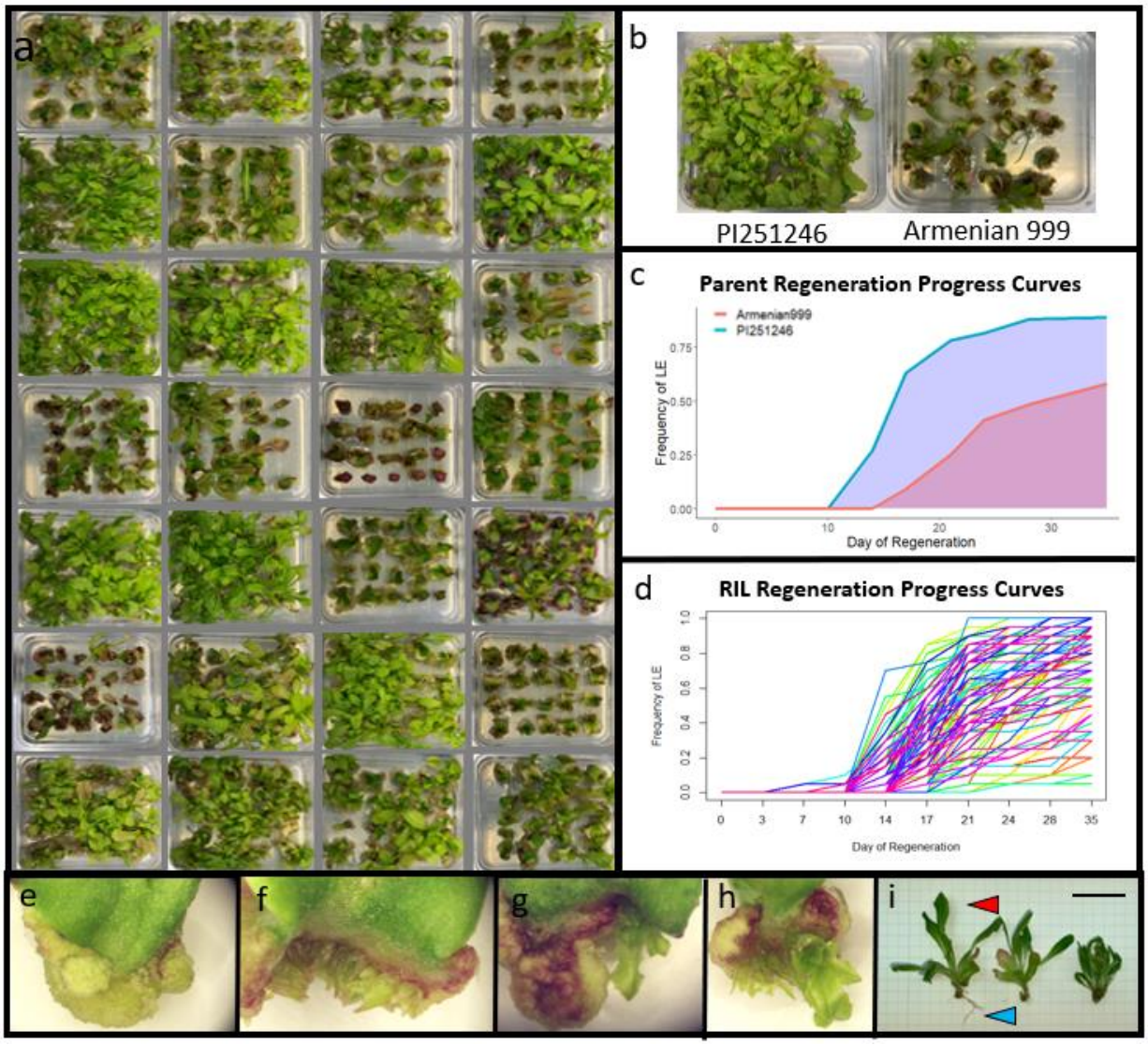


Figure 3.2. Variation in regeneration rates among the Armenian 999 x PI251246 F_8 RIL population. **a)** Regeneration of the RILs at day 35; **b)** Regeneration of the low regenerating (Armenian999) and the high regenerating (PI251246) parents at day 35; **c)** Mean regeneration progress curves for leaf emergence in the parents. The area under the curve (AUC) is represented by the blue (PI251246) and red (Armenian999) shaded regions; **d)** Mean regeneration progress curves for leaf emergence within 116 RILs. **e-i**) Visual representation of traits scored. **e**) callus formation; **f**) organized growth; **g**) leaf emergence; **h**) multiple leaves; **i**) shooting (red triangle) and occurrence of rooting (blue triangle); the black scale bar is representing 3 cm.

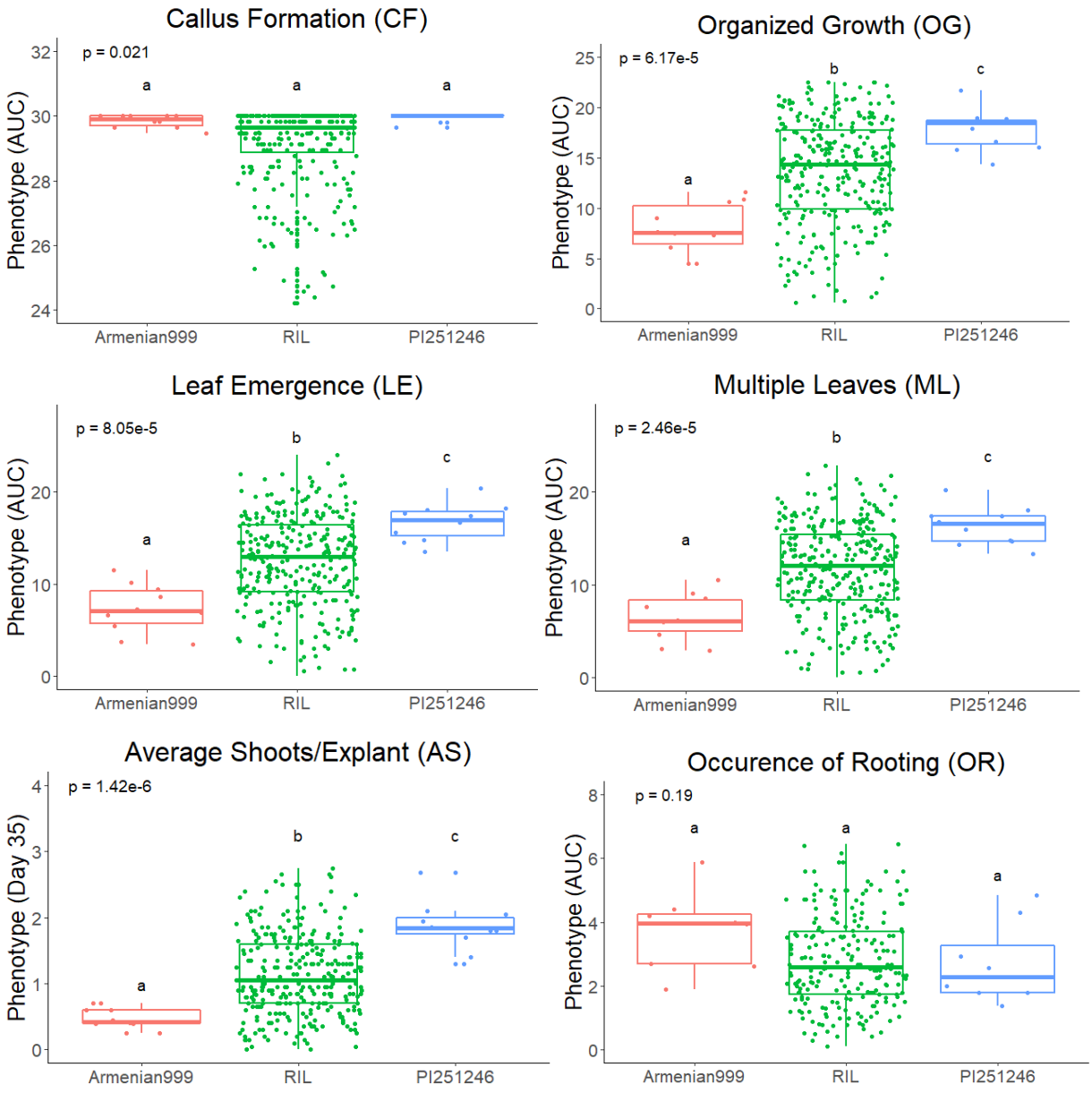


Figure 3.3. Boxplots representing the means and variation of CF, OG, LE, ML, AS, and OR in Armenian 999, PI251246, and the interspecific RIL population. P values are the result of a one-way ANOVA ($\alpha = 0.05$) comparing means of parents and individuals.

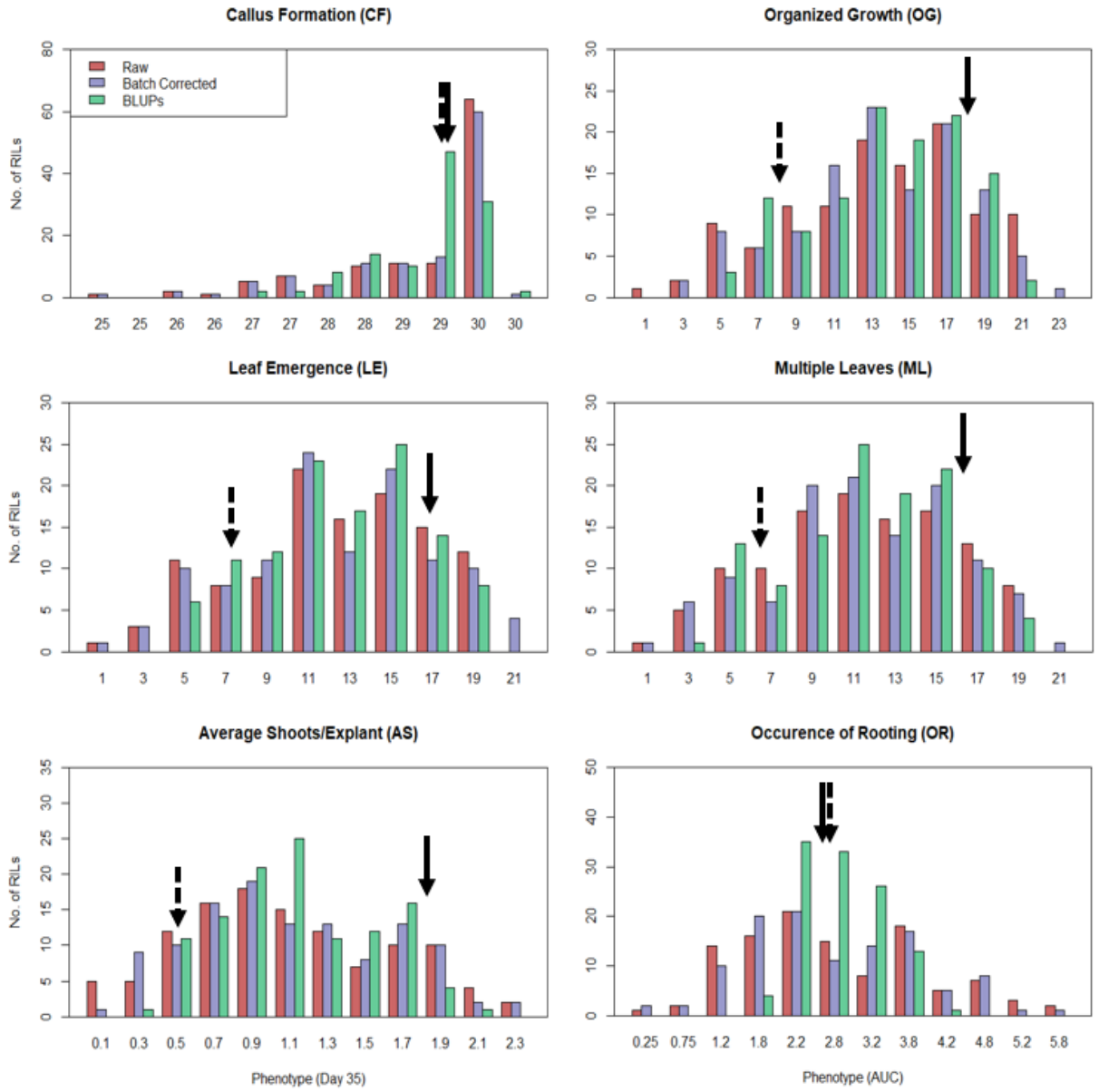


Figure 3.4. Histograms representing the distribution of regeneration traits OG, LE, ML, AS, and OR in the RIL population. Each distribution represents raw means (magenta), batch corrected means (purple), and BLUP adjusted means (teal) for each trait. Black arrows display the average phenotypes of Armenian999 (dashed) and PI251246 (solid).

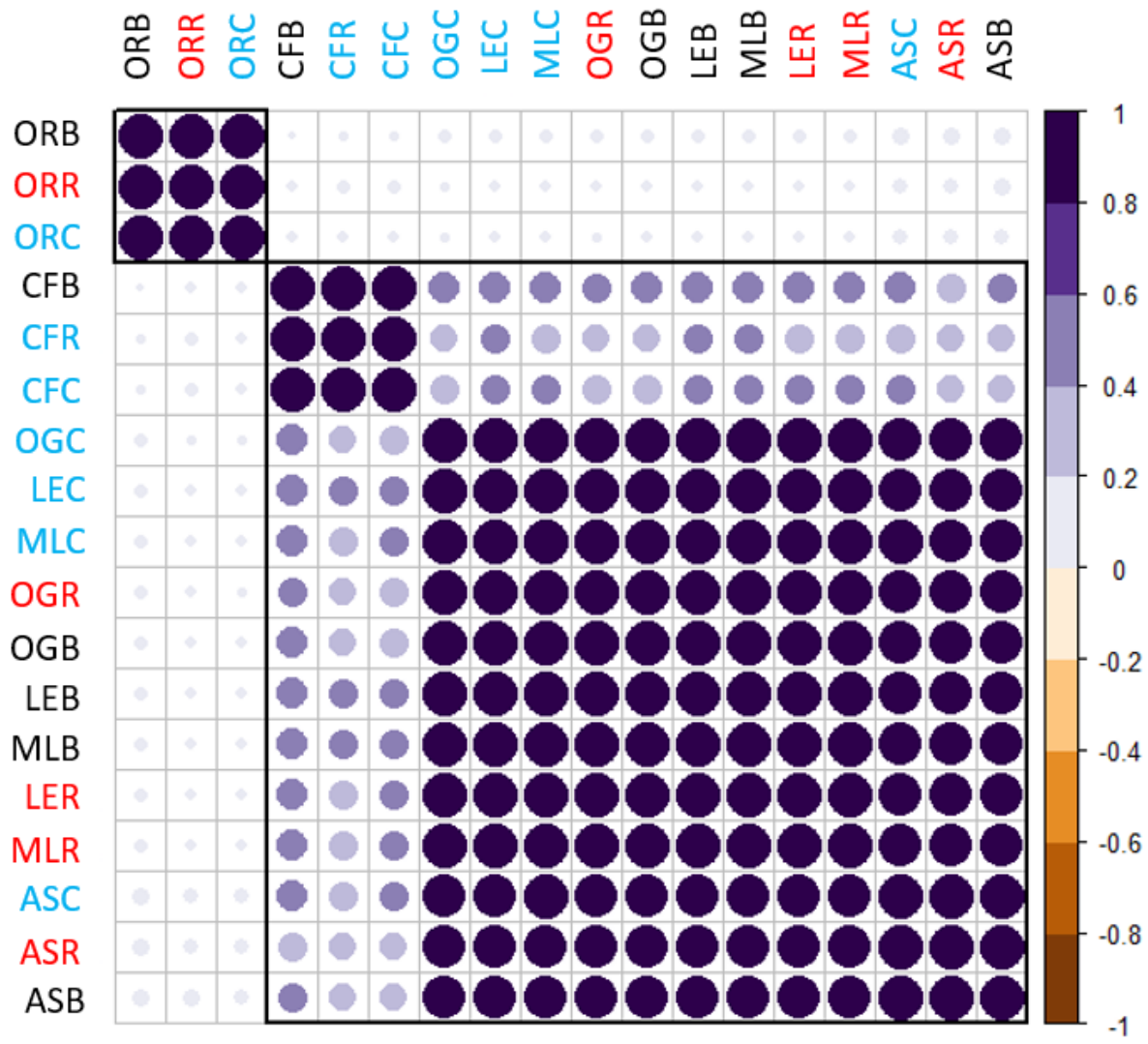


Figure 3.5. Pearson's correlations between callus formation (CF), organized growth (OG), leaf emergence (LE), multiples leaves (ML), average shoots (AS), and occurrence of rooting (OR). Correlations for each regeneration trait between raw means (red letters), BLUP adjusted means (black letters), and batch corrected means (blue letters) are also represented.

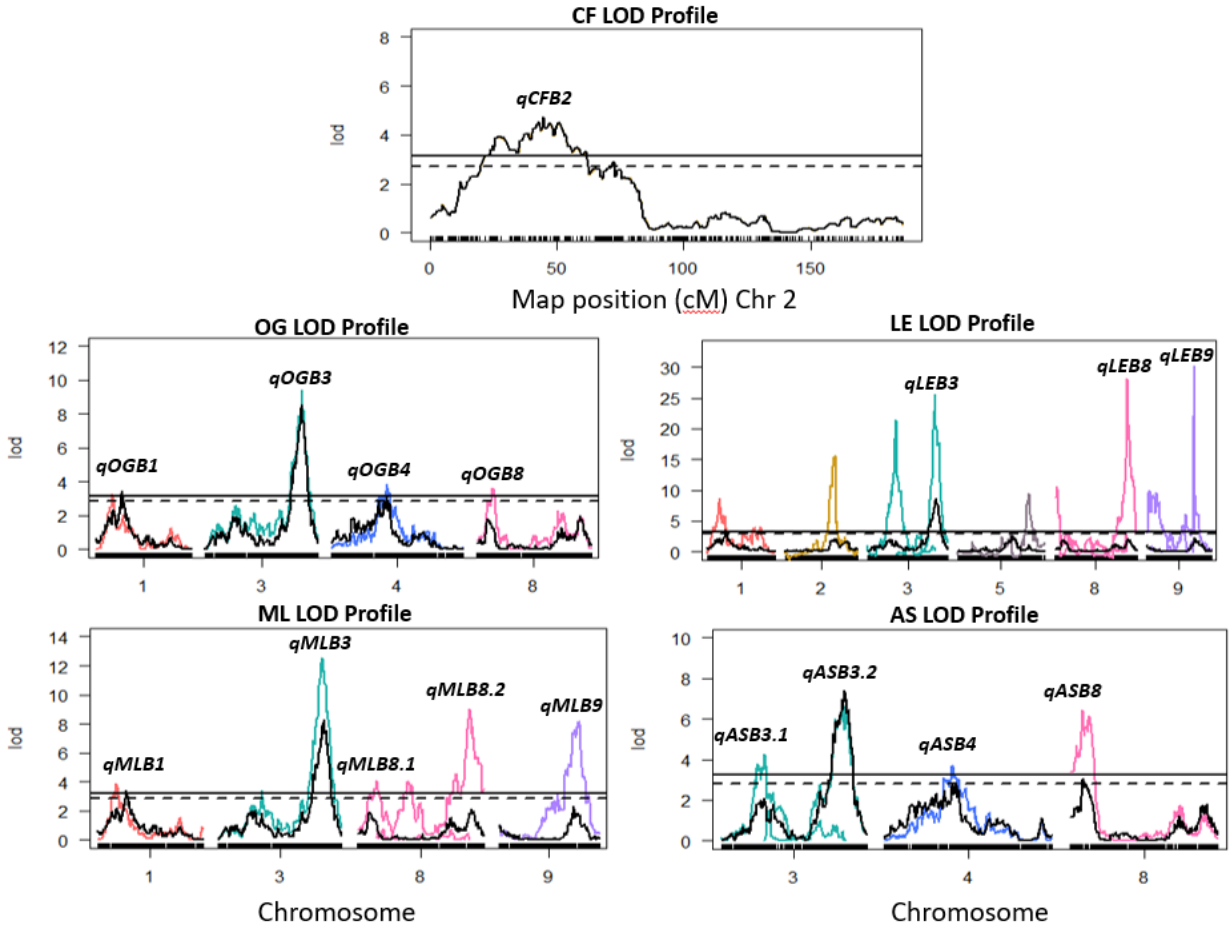


Figure 3.6. Significant associations between regeneration phenotypes and positions in the lettuce genome. The LOD profiles of callus formation (CF), organized growth (OG), leaf emergence (LE), multiple leaves (ML), and average shoots (AS) after simple interval mapping (black profiles) and multiple QTL mapping (colored profiles). Only chromosomes with significant associations are shown for each phenotype. Labeled QTLs were selected from the initial MQM models based on two criteria, 1) LOD scores exceeded the genome wide pLOD threshold and 2) QTLs represented greater than 5% phenotypic variance explained (PVE). Black horizontal lines represent the MQM genome wide pLOD significance threshold (solid) and simple interval mapping genome wide significance threshold (dashed) calculated using a 1000-iteration permutation test.

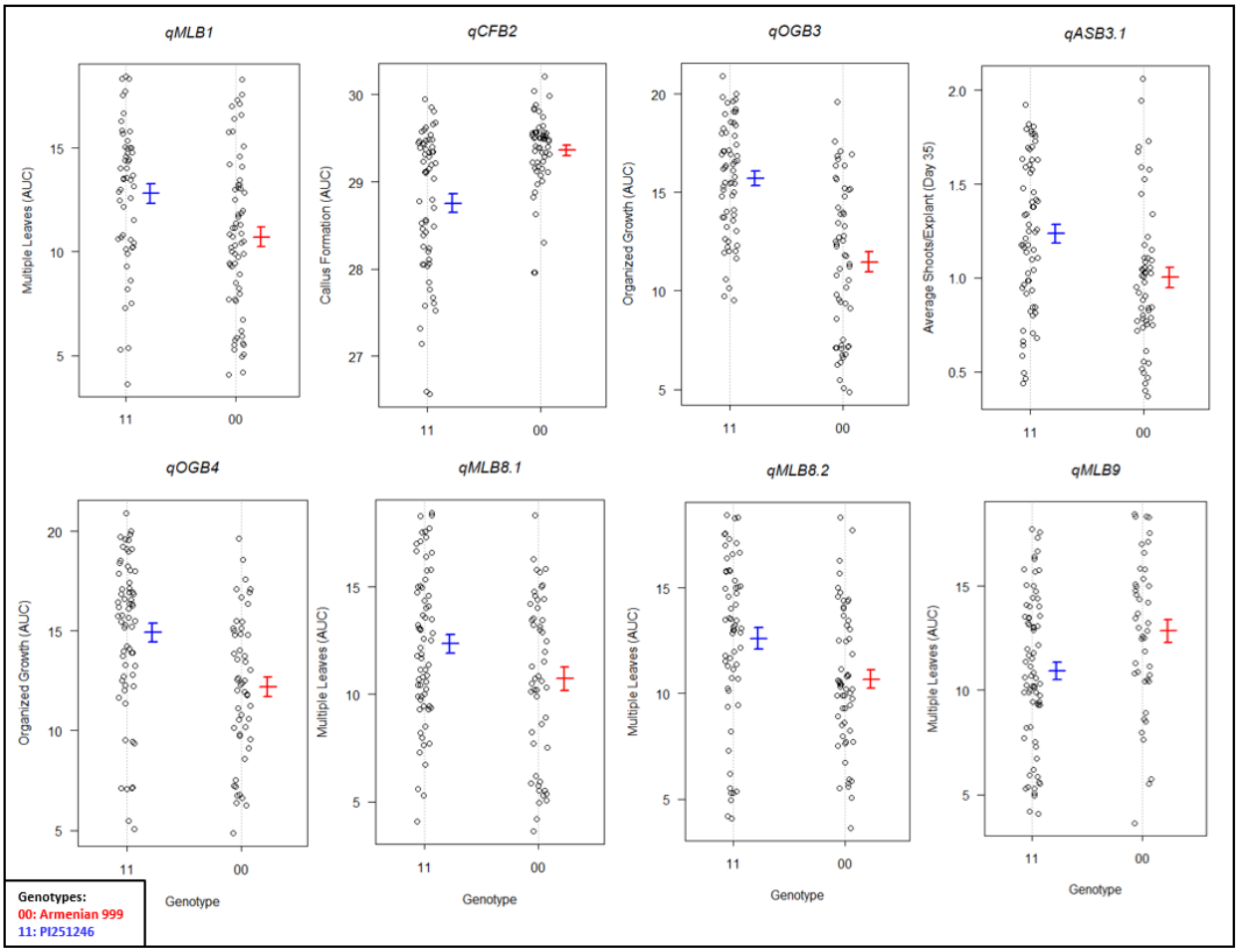


Figure 3.7. Estimated effects of QTLs using BLUP adjusted means. Each plot is representative of one of the eight total QTLs revealed for callus formation (CF), organized growth (OG), leaf emergence (LE), multiple leaves (ML), and average shoots (AS). The blue and red bars show the mean phenotypes of RILs with the PI251246 and Armenian 999 alleles, respectively, located at the peak of each QTL.

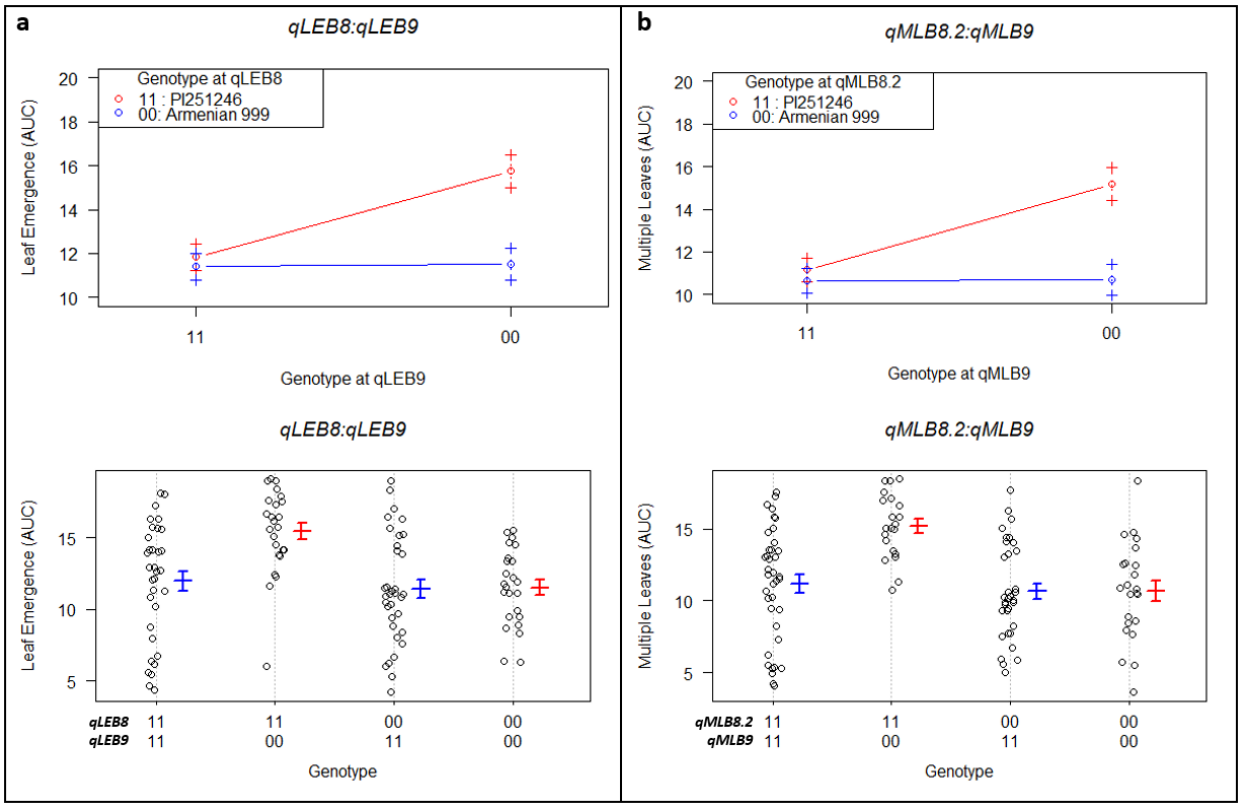


Figure 3.8. Interactions between QTLs identified on chromosomes 8 and 9 for leaf emergence and multiple leaves. PI251246 alleles are represented by "11" and Armenian 999 alleles are represented by "00." The blue and red bars show the mean phenotypes of RILs with the PI251246 and Armenian 999 alleles, respectively, located at the peak of each QTL. **a**) Interaction (top) and effect (bottom) plots between *qLEB8* and *qLEB9*. **b**) Interaction (top) and effect (bottom) plots between *qMLB8*.2 and *qMLB9*.

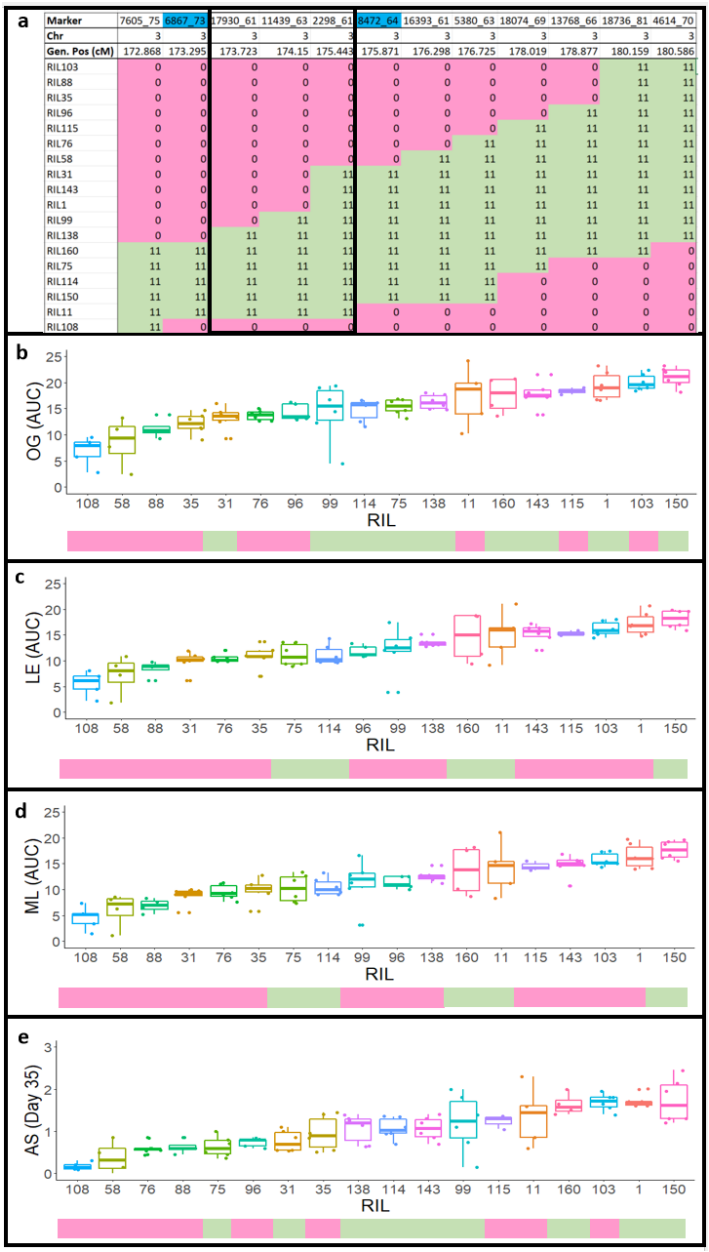


Figure 3.9. Phenotypes and genotypes of the eighteen RILs with recombination breakpoints within the interval of the major QTL on chromosome 3. **a)** Genotypes of eighteen RILs chosen for fine mapping. The alleles 0 and 11 are representative of the Armenian 999 and PI251246 parents, respectively. The blue highlighted markers represent the peaks identified for OG (right), LE (left), ML (left), and AS (right). The black thick-bordered is representative of the fine-mapped region based on the observed phenotypes of informative RILs. **b-e)** Boxplots representing the phenotypes of recombinant RILS for OG, LE, ML, and AS. Red (00) and green (11) bars under each boxplot displays the genotype at the QTL peak for each RIL.

Table 3.1. Means and standard deviations of PI251246, Armenian999, and the RIL population for callus formation (CF), organized growth (OG), leaf emergence (LE), multiple leaves (ML), average shoots (AS), and occurrence of rooting (OR).

Line	Statistic	CF*	OG*	LE*	ML*	AS+	OR*
PI251246	Mean	29.951	17.778	16.714	16.322	1.861	2.698
	Sd	0.115	1.987	1.989	1.983	0.366	1.256
Armenian999	Mean	29.883	8.000	7.367	6.522	0.483	3.709
	Sd	0.152	2.693	2.820	2.644	0.175	1.255
RILs	Mean	29.041	13.610	12.354	11.580	1.116	2.797
	Sd	1.400	5.295	5.094	4.012	0.614	1.407

* Area under the progress curve data
+ Time Point data

Table 3.2. Putative QTL (peak exceeds or is near the genome wide LOD significance threshold) identified using simple interval mapping and BLUP adjusted means.

Trait	LOD threshold†	Chr	LOD	Genetic Position (cM)	Marker Nearest Peak
CF	3.18	2	5.693	43.029	12737_66 ♂
OG	3.20	1	3.746	47.690	2944_61♀
		3	8.519	175.871	8472_64 ♀
		4	3.181	100.904	23694_81 ♂
LE	3.23	1	3.880	47.690	2944_61♀
		3	8.670	175.871	8472_64 ♀
		4	3.121	100.904	23694_81 ♂
ML	3.21	1	3.898	47.690	2944_61♀
		3	8.286	175.871	8472_64 ♀
		4	3.127	100.904	23694_81 ♂
AS	3.19	1	3.326	47.690	2944_61 ♀
		3	7.429	175.871	6867_73 ♂
		8	3.332	18.447	14938_66 ♀

† Genome wide LOD significance threshold ($\alpha = 0.05$), ♀ Female marker, ♂ Male marker

Table 3.3. LOD scores and PVE for the full QTL models and individual QTL terms for CF, OG, and LE using BLUP adjusted means. Markers, genetic (G. peak) positions, physical (P. peak) positions, and a 1.5 LOD interval is represented for each QTL.

Full Model				**QTL**							**1.5 LOD Interval**	
Trait	Model	LOD	PVE	QTL	Chr	LOD	PVE	Marker Nearest Peak	G.peak	P.peak	Start (cM)	End (cM)
CF	y ~ Q1	4.7	17.1	*qCFB2*	2	4.7	17.125	21425_65 ♀	44.742	60710106	23.200	62.000
OG	y ~ Q1 + Q2 + Q3 + Q4	17.4	49.9	*qOGB1*	1	3.3	6.9	855_63 ♀	30.098	202542423	20	50.696
				qOGCB3	3	9.4	22.8	8472_64 ♀	175.871	256626504 †	173.295	180
				qOGB4	4	3.8	8.2	7017_61 ♂	100.477	216241865	83.71	116.371
				qOGB8	8	3.6	7.7	13024_61 ♂	30.09	51843372	18.019	159.847
LE	y ~ Q4 + Q7 + Q10 + Q7:Q10	16.0	47.0	*qLEB3*	3	10.5	27.3	6867_73 ♂	173.295	255081484	168.557	180.586
				qLEB8	8	5.5	12.9	9475_61 ♀	186.103	291435679 †	158.993	195.549
				qLEB9	9	5.5	13.1	14019_74 ♂	123.508	177312968 †	116	136
				qLEB8: qLEB9	-	2.9	6.5	-	-	-	-	-

♀ Female marker, ♂ Male marker, * Multiple markers identified at peak, † Physical position of marker nearest peak that did map to v11

Table 3.4. LOD scores and PVE for the full QTL models and individual QTL terms for ML and AS using BLUP adjusted means. Markers, genetic (G. peak) positions, physical (P. peak) positions, and a 1.5 LOD interval is represented for each QTL.

Full Model				**QTL**							**1.5 LOD Interval**	
Trait	Model	LOD	PVE	QTL	Chr	LOD	%Var	Marker Nearest Peak	G.peak	P.peak	Start (cM)	End (cM)
ML	y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q4:Q5	22.9	59.7	*qMLB1*	1	3.9	6.7	9967_68 ♂	30.526	202338175	20	40.819
				qMLB3	3	12.6	26.1	6867_73 ♂	173.295	255081484	168	176.298
				qMLB8.1	8	4.1	7.1	5302_81 ♂	32.711	53273330	16.726	41.306
				qMB8.2	8	9.0	17.4	10352_61 ♀	186.53	291435679	182.238	192.543
				qMLB9	9	8.1	15.3	23467_66 ♂ *	130.84	181641895	121.799	136
				qMLC8.2: qMLC9	-	6.3	11.5	-	-	-	-	-
AS	y ~ Q1 + Q2 + Q3 + Q4	17.9	50.9	*qASB3.1*	3	4.3	9.1	1224_77 ♀	60.21	73668299	47.33	63.662
				qASB3.2	3	6.7	14.8	8472_64 ♀	175.871	256626504 †	158	180.586
				qASB4	4	3.7	7.7	13192_67 ♂	96.597	219585446	84	116.371
				qASB8	8	6.6	14.4	14938_66 ♀	18.447	38606210	16.299	32.711

♀ Female marker, ♂ Male marker, * Multiple markers identified at peak, † Physical position of marker nearest peak that did map to v11

Table 3.5. Estimated effects and parental contributions of each QTL revealed for BLUP adjusted means.

QTL	Chr	Estimated Effect (AUC)	Allele †
qCFB2	2	0.47	Armenian 999
qOGB1	1	1.0515	PI251246
qOGB3	3	1.9063	PI251246
qOGB4	4	1.163	PI251246
qOGB8	8	1.1209	PI251246
qLEB3	3	1.9696	PI251246
qLEB8	8	1.1471	PI251246
qLEB9	9	1.0247	Armenian 999
qLEB8:qLEB9	-	0.9687	-
qMLB1	1	0.9808	PI251246
qMLB3	3	1.9221	PI251246
qMLB8.1	8	1.0158	PI251246
qMLB8.2	8	1.2168	PI251246
qMLB9	9	0.7466	Armenian999
qMLB8.2:qMLB9	-	1.3256	-
qASB3.1	3	0.12772*	PI251246
qASB3.2	3	0.16223*	PI251246
qASB4	4	0.11632*	PI251246
qASB8	8	0.16167*	PI251246

† Parental allele contributing to higher regeneration rates
*Time point data (day 35) of average shoots per explant

Table 3.6. Candidate genes homologous to genes known to function in development or in vitro regeneration in other species located under each QTL identified using BLUP adjusted means, the gene coordinates from *L. sativa* cv. Salinas reference genome assembly version 11, and homologous genes in other species.

QTL	Chr	Lettuce Gene ID	Coordinates		Distance from peak (Mb)	Variants†	Ortholog	
							Species	Gene
qOGB1, qMLB1	1	LSAT1v11_C10004602	201507634	201516818	-1.034789	51^{P}, 0^{A}	*Arabidopsis*	*TONSOKU* (*TSK*)
qCFB2	2	LSAT1v11_C20006009	66296590	66298298	5.586484	1^{P}, 24^{A}	*Arabidopsis*	*LBD1* (LOB domain containing protein 1)
qASB3.1	3	LSAT1v11_C30012320	75536439	75538151	1.86814	16^{P}, 11^{A}	*Vitis vinifera*	*GAI1*
qASB3.1	3	LSAT1v11_C30012342	76429404	76433183	2.761105	7^{P}, 90^{A}	*Arabidopsis*	*PIN4*
qASB3.1	3	LSAT1v11_C30012189	70695699	70698656	-2.9726	3^{P}, 11^{A}	*Arabidopsis*	*MYB124*
qASB3.1	3	LSAT1v11_C30012355	76879418	76880710	3.211119	0^{P}, 5^{A}	*Arabidopsis*	*PILS2: PIN-LIKES 2*
qASB3.1	3	LSAT1v11_C30012400	78172317	78175482	4.504018	35^{P}, 25^{A}	*Arabidopsis*	*IAA32*
qOGB3, qASB3.2	3	LSAT1v11_C30014922	256708097	256710634	-0.081593	1^{P}, 18^{A}	*Arabidopsis*	*AGL104*
qOGB3, qASB3.2	3	LSAT1v11_C30014931	257126051	257128042	0.499547	3^{P}, 0^{A}	*Arabidopsis*	*MYB15*
qOGB3, qASB3.2	3	LSAT1v11_C30014934	257459200	257462059	0.832696	2^{P}, 32^{A}	*Mus musculus*	*MEAF6*
qLEB3, qMLB3	3	LSAT1v11_C30014872	254984336	254987602	-0.097148	0^{P}, 6^{A}	*Antirrhinum majus*	*G2/mitotic-specific cyclin-1*
qOGB4	4	LSAT1v11_C40020139	209529635	209530664	-6.71223	57^{P}, 1^{A}	*Arabidopsis*	*HDA15*
qOGB4	4	LSAT1v11_C40020114	207369655	207370948	-8.87221	0^{P}, 12^{A}	*Arabidopsis*	*SPL6*
qOGB4	4	LSAT1v11_C40020091	205657440	205661290	-10.584425	6^{P}, 48^{A}	*Arabidopsis*	*HDA5*
qASB4	4	LSAT1v11_C40020285	220625023	220625509	1.040063	0^{P}, 8^{A}	*Oryza sativa*	*ARF3*
qOGB8	8	LSAT1v11_C80040379	51599367	51602654	-0.244005	8^{P}, 55^{A}	*Arabidopsis*	*bHLH49*
qOGB8	8	LSAT1v11_C80040370	51078738	51081150	-0.764634	2^{P}, 22^{A}	*Arabidopsis*	*BIM2*
qOGB8	8	LSAT1v11_C80040366	51055521	51056601	-0.787851	0^{P}, 21^{A}	*Arabidopsis*	*BZIP44*
qMLB8.1	8	LSAT1v11_C80040442	54879341	54880363	1.606011	3^{P}, 1^{A}	*Arabidopsis*	*RAV1 AP2/ERF*
qMLB8.1	8	LSAT1v11_C80040485	57268233	57270714	3.994903	0^{P}, 28^{A}	*Arabidopsis*	*ERF118*
qMLB8.1	8	LSAT1v11_C80040488	57690886	57691257	4.417556	1^{P}, 5^{A}	*Arabidopsis*	*CRF4*

qASB8	8	LSAT1v11_C80040104	39768359	39771417	1.162149	4^{P}, 4^{A}	*Arabidopsis*	*CLV1*
qOGB8, qMLB8.2	8	LSAT1v11_C80045054	292298292	292302751	0.862613	2^{P}, 30^{A}	*Arabidopsis*	*MCM5*
qOGB8, qMLB8.2	8	LSAT1v11_C80045059	292649943	292650957	1.214264	0^{P}, 3^{A}	*Arabidopsis*	*KNAT6*
qMLB9	9	LSAT1v11_C90049620	180658692	180662807	-0.983203	136^{P}, 4^{A}	*Arabidopsis*	*MCM6*
qMLB9	9	LSAT1v11_C90049691	184247786	184248743	2.605891	16^{P}, 6^{A}	*Arabidopsis*	*MYB36*
qLEB9	9	LSAT1v11_C90049593	179346260	179346898	2.033292	8^{P}, 2^{A}	*Arabidopsis*	*TCP4*

† Number of variants between PI251246 ($\#^{P}$) or Armenian 999 ($\#^{A}$) and the Salinas v11 reference genome assembly

Table 3.7. Physical locations and flanking markers of each QTL interval revealed using BLUP adjusted means in the *L. sativa* cv. Salinas reference genome assembly version 11.

QTL				1.5 LOD Interval			
QTL	Chr	Marker Closest to Peak	P. peak	Start (bp)	End (bp)	Start Marker	End Marker
qCFB2	2	21425_65 ♀	60,710,106	34,186,052	77,355,398	6497_62	16030_61
qOGB1	1	855_63 ♀	202,542,423	233,960,631	162,665,600	3432_87	15275_74
qOGB3	3	8472_64 ♀	256,626,504 †	255,081,484	266,956,476	6867_73	15341_86
qOGB4	4	7017_61 ♂	216,241,865	240,003,690	189,504,059	8351_68	10829_63
qOGB8	8	13024_61 ♂	51,843,372	38,634,707	238,396,598	17040_68	18490_63
qLEB3	3	6867_73 ♂	255,081,484	254,469,695	266,956,476	1766_74	15341_86
qLEB8	8	9475_61 ♀	291,435,679 †	240,739,659	302,895,419	122_62	23456_68
qLEB9	9	14019_74 ♂	177,312,968 †	165,072,074	181,641,895	19973_75	23467_66
qMLB1	1	9967_68 ♂	202,338,175	233,960,631	170,844,556	3432_87	3906_82
qMLB3	3	6867_73 ♂	255,081,484	249,376,558	266,956,476	4093_65	15275_74
qMLB8.1	8	5302_81 ♂	53,273,330	38,634,707	63,459,900	17040_68	15563_75
qMB8.2	8	10352_61 ♀	291,435,679	288,544,336	302,851,329	15764_71	6807_75
qMLB9	9	23467_66 ♂ *	181,641,895	176,581,902	191,431,110	11326_81	5612_66
qASB3.1	3	1224_77 ♀	73,668,299	61,810,132	83,023,650	22193_71	3442_61
qASB3.2	3	8472_64 ♀	256,626,504 †	217,616,770	265,178,057	1631_76	4614_70
qASB4	4	13192_67 ♂	219,585,446	240,003,690	189,504,059	8351_68	10829_63
qASB8	8	14938_66 ♀	38,606,210	38,634,707	54,010,699	17040_68	6351_74

♀ Female marker, ♂ Male marker, * Multiple markers identified at peak, † Physical position of marker nearest peak that did map to v11

REFERENCES

Adamczyk, B. J., and Fernandez, D. E. (2009). MIKC* MADS Domain Heterodimers Are Required for Pollen Maturation and Tube Growth in Arabidopsis. *Plant Physiology*, *149*(4), 1713–1723.

Argyris, J., Mariá, A. E., Ae, J. J., Ochoa, O., Knapp, S. J., David, A. E., Still, W., Ger, A. E., Lenssen, M., Schut, J. W., Richard, A. E., Michelmore, W., and Bradford, K. J. (2005). Quantitative trait loci associated with seed and seedling traits in *Lactuca*. *Theoretical and Applied Genetics*, *111*, 1365–1376.

Becker, A., and Theißen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Molecular Phylogenetics and Evolution*, *29*(3), 464–489.

Belles-Boix, E., Hamant, O., Witiak, S. M., Morin, H., Traas, J., and Pautot, V. (2006). KNAT6: An *Arabidopsis* Homeobox Gene Involved in Meristem Activity and Organ Separation. *The Plant Cell*, *18*(8), 1900–1907.

Chu, Y., Chee, P., Culbreath, A., Isleib, T. G., Holbrook, C. C., and Ozias-Akins, P. (2019). Major QTLs for resistance to early and late leaf spot diseases are identified on chromosomes 3 and 5 in peanut (Arachis hypogaea). *Frontiers in Plant Science*, *10*, 883.

Chunthawodtiporn, J., Hill, T., Stoffel, K., and van Deynze, A. (2019). Genetic Analysis of Resistance to Multiple Isolates of *Phytophthora* capsici and Linkage to Horticultural Traits in Bell Pepper. *HortScience*, *54*(7), 1143–1148.

DiDonato, R. J., Arbuckle, E., Buker, S., Sheets, J., Tobar, J., Totong, R., Grisafi, P., Fink, G. R., and Celenza, J. L. (2004). *Arabidopsis ALF4* encodes a nuclear-localized protein required for lateral root formation. *Plant Journal*, *37*(3), 340–353.

Ding, X., Ruan, H., Yu, L., Li, Q., Song, Q., Yang, S., and Gai, J. (2020). miR156b from Soybean CMS Line Modulates Floral Organ Development. *Journal of Plant Biology*, *63*(2), 141–153.

Fan, M., Xu, C., Xu, K., and Hu, Y. (2012). LATERAL ORGAN BOUNDARIES DOMAIN transcription factors direct callus formation in Arabidopsis regeneration. *Cell Research*, *22*(7), 1169–1180.

Fehér, A. (2015). Somatic embryogenesis — Stress-induced remodeling of plant cell fate. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1849*(4), 385–402.

Feng, Z., Zhu, J., Du, X., and Cui, X. (2012). Effects of three auxin-inducible LBD members on lateral root formation in Arabidopsis thaliana. *Planta*, *236*(4), 1227–1237.

Fletcher, K., Zhang, L., Gil, J., Han, R., Cavanaugh, K., and Michelmore, R. (2021). AFLAP: assembly-free linkage analysis pipeline using k-mers from genome sequencing data. *Genome Biology*, *22*(1), 1–26.

Gammoudi, N., Pedro, T. S., Ferchichi, A., and Gisbert, C. (2018). Improvement of regeneration in pepper: a recalcitrant species. *In Vitro Cellular and Developmental Biology - Plant*, *54*(2), 145–153.

Gordon, S. P., Heisler, M. G., Reddy, G. V., Ohno, C., Das, P., and Meyerowitz, E. M. (2007). Pattern formation during de novo assembly of the Arabidopsis shoot meristem. *Development*, *134*(19), 3539–3548.

Guo, H., Guo, H., Zhang, L., Fan, Y., Fan, Y., Tang, Z., and Zeng, F. (2019). Dynamic TMT-Based Quantitative Proteomics Analysis of Critical Initiation Process of Totipotency during Cotton Somatic Embryogenesis Transdifferentiation. *International Journal of Molecular Sciences 2019, Vol. 20, Page 1691*, *20*(7), 1691.

Han, R., Wong, A. J. Y., Tang, Z., Truco, M. J., Lavelle, D. O., Kozik, A., Jin, Y., and Michelmore, R. W. (2021). Drone phenotyping and machine learning enable discovery of loci regulating daily floral opening in lettuce. *Journal of Experimental Botany*, *72*(8), 2979–2994.

Harding, E. W., Tang, W., Nichols, K. W., Fernandez, D. E., and Perry, S. E. (2003). Expression and Maintenance of Embryogenic Potential Is Enhanced through Constitutive Expression of AGAMOUS-Like 15. *Plant Physiology*, *133*(2), 653. He, C., Chen, X., Huang, H., and Xu, L. (2012). Reprogramming of H3K27me3 Is Critical for Acquisition of Pluripotency from Cultured Arabidopsis Tissues. *PLOS Genetics*, *8*(8), e1002911.

Horstman, A., Li, M., Heidmann, I., Weemen, M., Chen, B., Muino, J. M., Angenent, G. C., and Boutiliera, K. (2017). The BABY BOOM Transcription Factor Activates the LEC1-ABI3-FUS3-LEC2 Network to Induce Somatic Embryogenesis. *Plant Physiology*, *175*(2), 848–857.

Iwase, A., Ohme-Takagi, M., and Sugimoto, K. (2011). WIND1: A key molecular switch for plant cell dedifferentiation. *Plant Signaling and Behavior*, *6*(12), 1943–1945.

Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., Jahesh, G., Khan, H., Coombe, L., Warren, R. L., and Birol, I. (2017). ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. *Genome Research*, *27*(5), 768–777.

Kareem, A., Durgaprasad, K., Sugimoto, K., Du, Y., Pulianmackal, A. J., Trivedi, Z. B., Abhayadev, P. v, Pinon, V., Meyerowitz, E. M., Scheres, B., and Prasad, K. (2015). PLETHORA Genes Control Regeneration by a Two-step Mechanism. *Current Biology*.

Köhler, C., and Hennig, L. (2010). Regulation of cell identity by plant polycomb and trithorax group proteins. *Current Opinion in Genetics and Development*, *20*(5), 541–547.

Kumar, V., Thakur, J. K., and Prasad, M. (2021). Histone acetylation dynamics regulating plant development and stress responses. *Cellular and Molecular Life Sciences*, *78*, 4467–4486.

Kwon, Y. S., Kim, K. M., Eun, M. Y., and Sohn, J. K. (2002). QTL mapping and associated marker selection for the efficacy of green plant regeneration in anther culture of rice. *Plant Breeding*, *121*(1), 10–16.

Kwon, Y.-S., Kim, K.-M., Eun, M.-Y., and Sohn, J.-K. (2001). Quantitative Trait Loci Mapping Associated with Plant Regeneration Ability. *Molecules and Cells*, *11*(1), 64–67.

Lall, S., Nettleton, D., DeCook, R., Che, P., and Howell, S. H. (2004). Quantitative Trait Loci Associated With Adventitious Shoot Formation in Tissue Culture and the Program of Shoot Development in Arabidopsis. *Genetics*, *167*(4), 1883–1892.

Li, Z., Liu, P., Zhang, X., Zhang, Y., Ma, L., Liu, M., Guan, Z., Zhang, Y., Li, P., Zou, C., He, Y., Gao, S., Pan, G., and Shen, Y. (2020). Genome-wide association studies and QTL mapping uncover the genetic architecture of ear tip-barrenness in maize. *Physiologia Plantarum*, *170*(1), 27–39.

Ma, J., Deng, M., Lv, S. Y., Yang, Q., Jiang, Q. T., Qi, P. F., Li, W., Chen, G. Y., Lan, X. J., and Wei, Y. M. (2016). Identification of QTLs associated with tissue culture response of mature wheat embryos. *SpringerPlus*, *5*(1), 1–7.

Mano, Y., Takahashi, H., Sato, K., and Takeda, K. (1996). Mapping Genes for Callus Growth and Shoot Regeneration in Barley (*Hordeum vulgare* L.). *Breeding Science*, *46*, 137–142.

Motte, H., Vercauteren, A., Depuydt, S., Landschoot, S., Geelen, D., Werbrouck, S., Goormachtig, S., Vuylsteke, M., and Vereecke, D. (2014). Combining linkage and association mapping identifies RECEPTOR-LIKE PROTEIN KINASE1 as an essential Arabidopsis shoot regeneration gene. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(22), 8305–8310.

Nidhi, S., Preciado, J., and Liu Tie, ·. (2021). Knox homologs shoot meristemless (STM) and KNAT6 are epistatic to CLAVATA3 (CLV3) during shoot meristem development in *Arabidopsis thaliana*. *Molecular Biology Reports*, *48*, 6291–6302.

Ning, K., Han, Y., Chen, Z., Luo, C., Wang, S., Zhang, W., Li, L., Zhang, X., Fan, S., and Wang, Q. (2019). Genome-wide analysis of MADS-box family genes during flower development in lettuce. *Plant Cell and Environment*, *42*(6), 1868–1881.

Nowak, K., Moró, J., Wójcik, A., and Gaj, M. D. (2020). AGL15 Controls the Embryogenic Reprogramming of Somatic Cells in Arabidopsis through the Histone Acetylation-Mediated Repression of the miRNA Biogenesis Genes. *International Journal of Molecular Sciences*, *21*(18), 6733.

Radonic, L. M., Lewi, D. M., López, N. E., Esteban Hopp, H., Escandón, A. S., and Bilbao, M. L. (2016). Sunflower (*Helianthus annuus* L.). In K. Wang (Ed.), *Methods in Molecular Biology* (Vol. 1224). Springer, New York, NY.

Rastas, P. (2017). Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics*, *33*(23), 3726–3732.

Salvo, S., Cook, J., Carlson, A. R., Hirsch, C. N., Kaeppler, S. M., Kaeppler, H. F., Salvo, S., Carlson, A. R., Kaeppler, S. M., and Kaeppler, H. F. (2018a). Genetic Fine-Mapping of a Quantitative Trait Locus (QTL) Associated with Embryogenic Tissue Culture Response and Plant Regeneration Ability in Maize (*Zea mays* L.). *Plant Genome*, *11*(2), 170111.

Salvo, S., Cook, J., Carlson, A. R., Hirsch, C. N., Kaeppler, S. M., Kaeppler, H. F., Salvo, S., Carlson, A. R., Kaeppler, S. M., and Kaeppler, H. F. (2018b). Genetic Fine-Mapping of a Quantitative Trait Locus (QTL) Associated with Embryogenic Tissue Culture Response and Plant Regeneration Ability in Maize (*Zea mays* L.). *The Plant Genome*, *11*(2), 170111.

Sarkar, A. K., Luijten, M., Miyashima, S., Lenhard, M., Hashimoto, T., Nakajima, K., Scheres, B., Heidstra, R., and Laux, T. (2007). Conserved factors regulate signalling in Arabidopsis thaliana shoot and root stem cell organizers. *Nature*, *446*(7137), 811–814.

Schiantarelli, E., de la Peña, A., and Candela, M. (2001). Use of recombinant inbred lines (RILs) to identify, locate and map major genes and quantitative trait loci involved with in vitro regeneration ability in *Arabidopsis thaliana*. *Theoretical and Applied Genetics 2001 102:2*, *102*(2), 335–341.

Shang, B., Xu, C., Zhang, X., Cao, H., Xin, W., and Hu, Y. (2016). Very-long-chain fatty acids restrict regeneration capacity by confining pericycle competence for callus formation in arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(18), 5101–5106.

Shimotohno, A., Heidstra, R., Blilou, I., and Scheres, B. (2018). Root stem cell niche organizer specification by molecular convergence of PLETHORA and SCARECROW transcription factor modules. *Genes and Development*, *32*, 1085–1100.

Shin, J., Bae, S., and Seo, P. J. (2020). *De novo* shoot organogenesis during plant regeneration. *Journal of Experimental Botany*, *71*(1), 63–72.

Sugimoto, K., Jiao, Y., and Meyerowitz, E. M. (2010). Arabidopsis Regeneration from Multiple Tissues Occurs via a Root Development Pathway. *Developmental Cell*, *18*(3), 463–471.

Trujillo-Moya, C., Gisbert, C., Vilanova, S., and Nuez, F. (2011). Localization of QTLs for in vitro plant regeneration in tomato. *BMC Plant Biology*, *11*.

Tyagi, N., Dahleen, L. S., and Bregitzer, P. (2010). Candidate genes within tissue culture regeneration QTL revisited with a linkage map based on transcript-derived markers. *Crop Science*, *50*(5), 1697–1707.

Wang, Y., Zhou, Q., Zhu, G., Wang, S., Ma, Y., Miao, H., Zhang, S., Huang, S., Zhang, Z., and Gu, Xingfang. (2018). Genetic analysis and identification of a candidate gene associated with in vitro regeneration ability of cucumber. *Theoretical and Applied Genetic*, *131*, 2663–2675.

Wu, G., and Poethig, R. S. (2006). Temporal Regulation of Shoot Development in Arabidopsis Thaliana By Mir156 and Its Target SPL3. *Development (Cambridge, England)*, *133*(18), 3539.

Wu, J., Zhang, X., Nie, Y., Jin, S., and Liang, S. (2004). Factors Affecting Somatic Embryogenesis and Plant Regeneration From a Range of Recalcitrant Genotypes of Chinese Cottons (*Gossypium hirtsutum* L.). *In Vitro Cellular and Developmental Biology - Plants*, *40*, 371–375.

Xu, M., Hu, T., Zhao, J., Park, M.-Y., Earley, K. W., Wu, G., Yang, L., and Poethig, R. S. (2016). Developmental Functions of miR156-Regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Genes in *Arabidopsis thaliana*. *PLoS Genet*, *12*(8), 1006263.

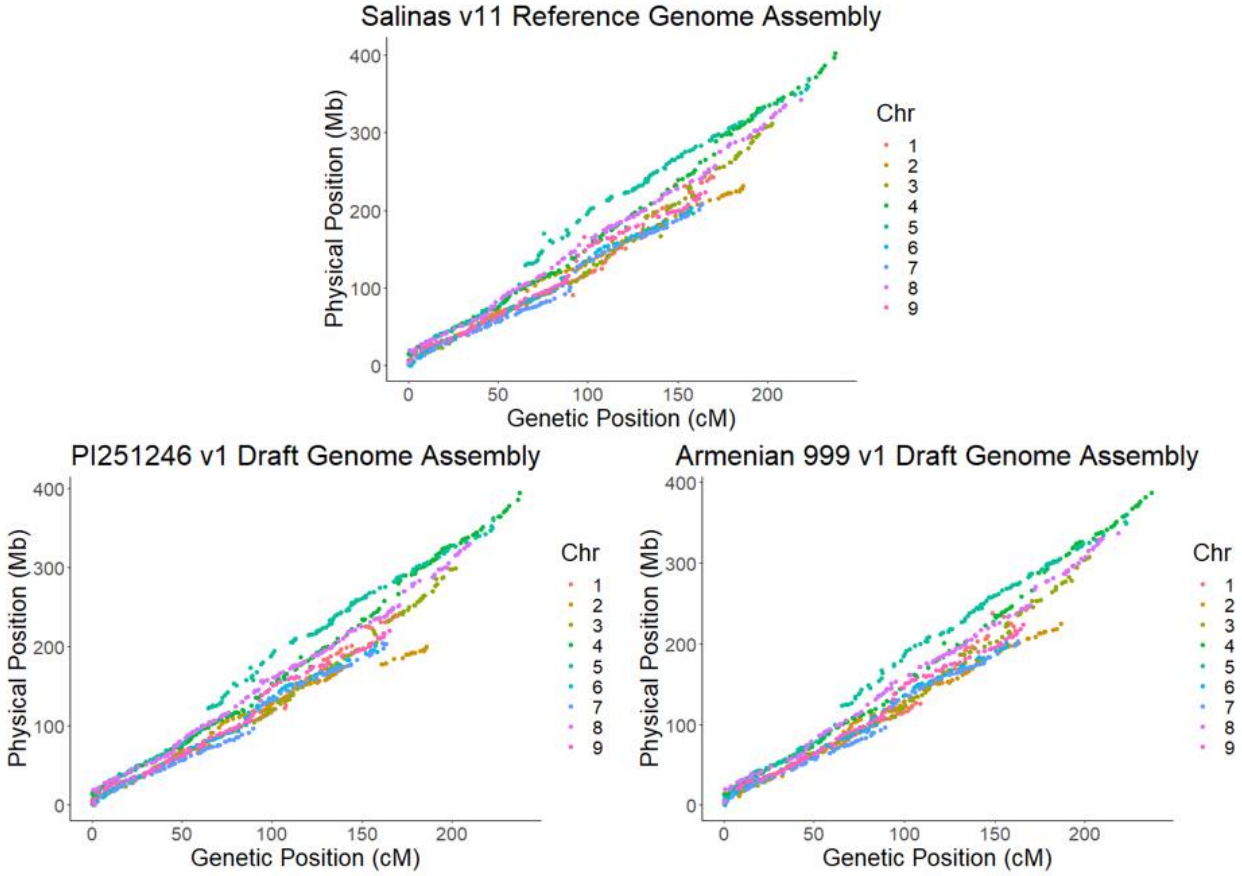
Yang, C., Zhao, T., Deyue, Y., and Gai, J. (2011). Mapping QTLs for Tissue Culture Response in Soybean (*Glycine max* (L.) Merr.). *Mol. Cells*, *32*, 337–342.

Zhao, Q.-H., Fisher, R., and Auer, C. (2002). Developmental phases and STM expression during Arabidopsis shoot organogenesis. *Plant Growth Regulation*, 37, 223-231.

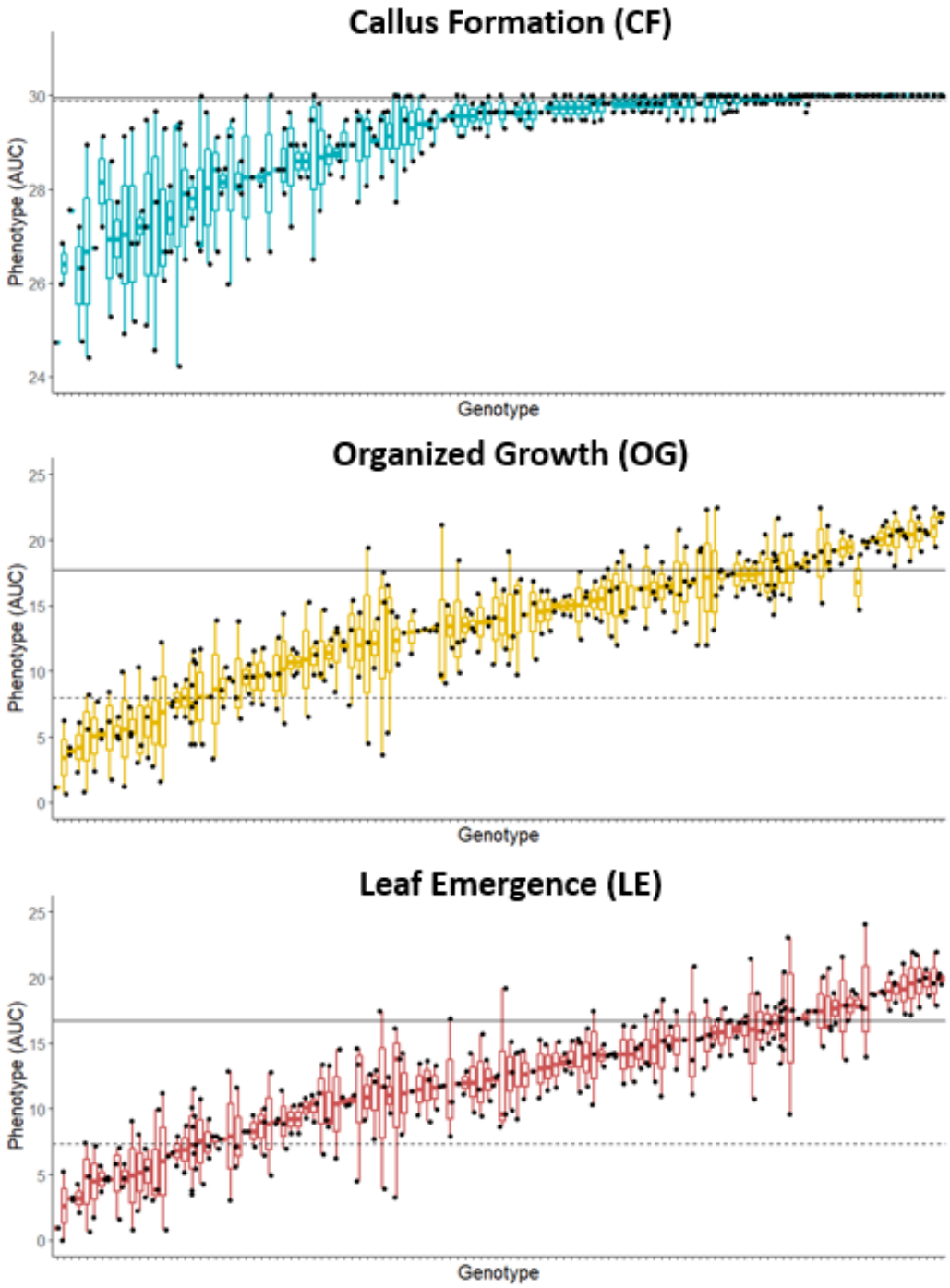
Zheng, Q., Zheng, Y., and Perry, S. E. (2013). AGAMOUS-Like15 Promotes Somatic Embryogenesis in Arabidopsis and Soybean in Part by the Control of Ethylene Biosynthesis and Response. *Plant Physiology*, *161*(4), 2113–2127.

Zheng, Y., Ren, N., Wang, H., Stromberg, A. J., and Perry, S. E. (2009). Global Identification of Targets of the Arabidopsis MADS Domain Protein AGAMOUS-Like15. *The Plant Cell*, *21*(9), 2563–2577.

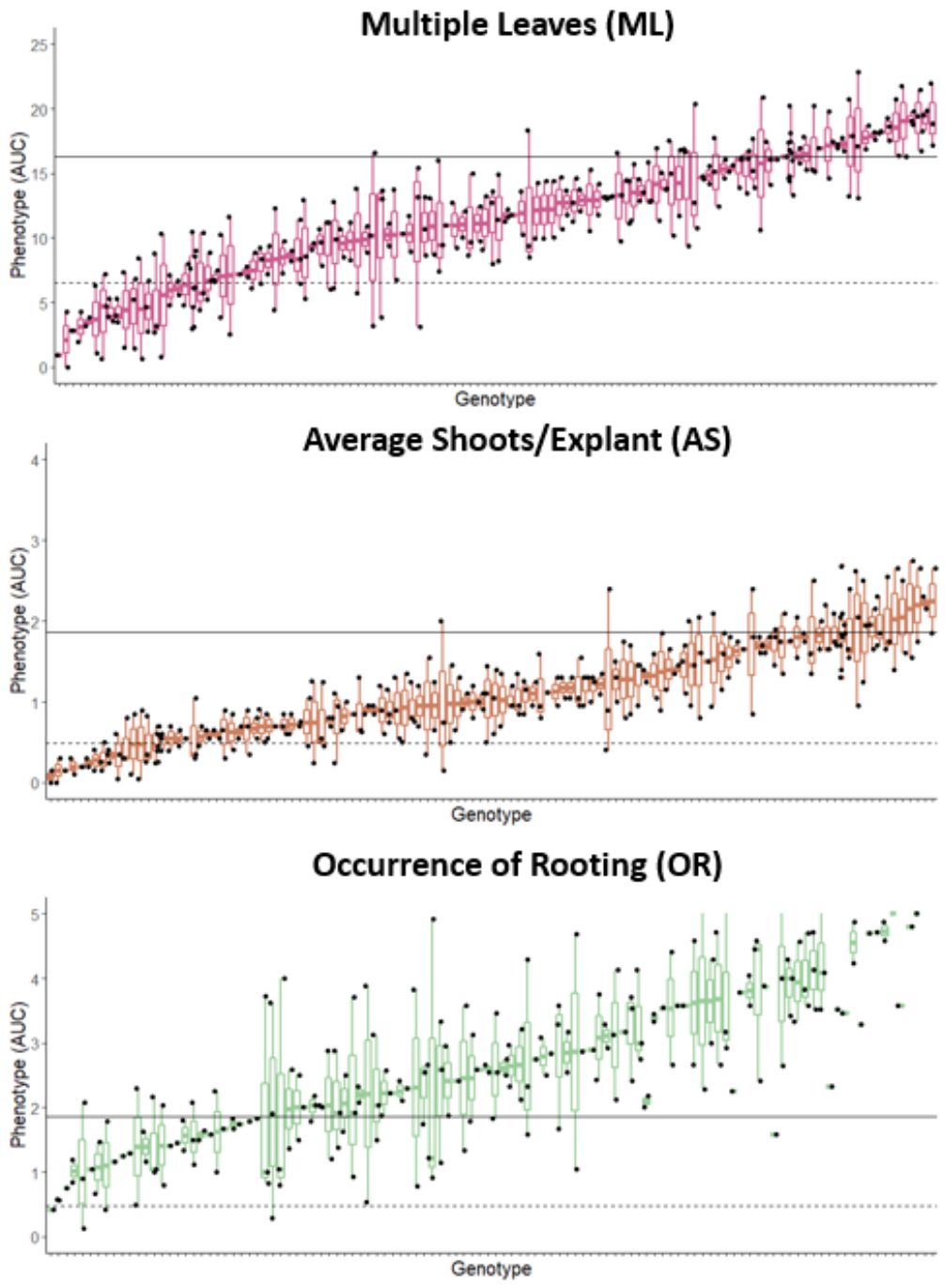
SUPPLEMENTARY FIGURES AND TABLES



Supplementary Figure S3.1. Collinearity of the k-mer markers after quality filtering the genetic map. Markers are mapped to the *L. sativa* cv Salinas version 11 reference, PI251246 v1 draft, and Armenian 999 v1 draft assemblies.



Supplementary Figure S3.2. Boxplots representing the means and variation of each RIL for CF ($p = 3.3^{-10}$), OG ($p < 2.0^{-16}$), and LE ($p < 2.0^{-16}$). P values are the result of a one-way ANOVA ($\alpha = 0.05$) comparing trait means of each RIL sampled from the population. Horizontal lines represent the means of Armenian 999 (dashed) and PI251246 (solid).



Supplementary Figure S3.3. Boxplots representing the means and variation of each RIL for ML ($p < 2.0^{-16}$), AS ($p < 2.0^{-16}$), and OR ($p < 2.0^{-16}$). P values are the result of a one-way ANOVA ($\alpha = 0.05$) comparing trait means of each RIL sampled from the population. Horizontal lines represent the means of Armenian 999 (dashed) and PI251246 (solid).

Supplementary Table S3.1. Putative QTL (peak exceeds or is near the genome wide LOD significance threshold) identified using simple interval mapping and raw means.

Trait	LOD threshold†	Chr	LOD	Genetic Position (cM)	Marker Nearest Peak
CF	3.12	2	5.335	43.029	12737_66 ♂
OG	3.32	1	3.694	47.690	2944_61♀
		3	8.021	175.871	8472_64 ♀
		4	3.253	100.904	23694_81 ♂
LE	3.26	1	3.85	47.690	2944_61♀
		3	8.28	175.871	8472_64 ♀
		4	3.21	100.904	23694_81 ♂
ML	3.26	1	3.841	47.690	2944_61♀
		3	7.869	175.871	8472_64 ♀
		4	3.241	100.904	23694_81 ♂
AS	3.23	1	3.139	47.690	2944_61 ♀
		3	7.270	175.871	8472_64 ♀

† Genome wide LOD significance threshold (α = 0.05), ♀ Female marker, ♂ Male marker

Supplementary Table S3.2. Putative QTL (peak exceeds or is near the genome wide LOD significance threshold) identified using simple interval mapping and batch corrected means.

Trait	LOD threshold†	Chr	LOD	Genetic Position (cM)	Marker Nearest Peak
CF	3.01	2	5.211	43.029	12737_66 ♂
OG	3.28	1	3.688	47.690	2944_61♀
		3	8.654	175.871	8472_64 ♀
		4	3.078	100.904	23694_81 ♂
LE	3.26	1	3.794	47.690	2944_61♀
		3	8.829	175.871	8472_64 ♀
		4	3.037	100.904	23694_81 ♂
ML	3.27	1	3.811	47.690	2944_61♀
		3	8.395	175.871	8472_64 ♀
		4	3.133	100.904	23694_81 ♂
AS	3.28	1	3.506	47.263	20467_61 ♂
		3	7.199	175.871	8472_64 ♀
		4	3.055	101.332	10909_69 ♂
		8	3.377	21.033	1653_61 ♀

† Genome wide LOD significance threshold ($\alpha = 0.05$), ♀ Female marker, ♂ Male marker

Supplementary Table S3.3. LOD scores and PVE for the full QTL models and individual QTL terms for CF, OG, and LE using raw means. Markers, genetic (G. peak) positions, physical (P. peak) positions, and a 1.5 LOD interval is represented for each QTL.

Full Model				**QTL**							**1.5 LOD Interval**	
Trait	Model	LOD	PVE	QTL	Chr	LOD	PVE	Marker Nearest Peak	P.Peak	G.peak	Start (cM)	End (cM)
CF	y ~ Q3	4.5	16.4	*qCFR2*	2	4.7	17.1	12737_66 ♂	43.029	59228938	28.000	59.357
OG	y ~ Q1 + Q2 + Q3 + Q4 + Q5	19.9	55.0	*qOGR1*	1	3.2	6.3	855_63 ♀	30.098	202542423	20	34.39
				qOGCR3	3	9.4	20.6	8472_64 ♀	175.871	256626504 †	174.15	180
				qOGR4	4	4.3	8.4	7017_61 ♂	100.477	216241865	84.568	106
				qOGR8	8	3.2	6.1	2797_65 ♂	28.804	51839580 †	16.726	189.548
				qOGR9	9	2.8	5.4	17180_86 ♂	123.081	177312968	106.268	130.84
LE	y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q4:Q5	22.9	60.1	*qLER1*	1	4.5	7.9	855_63 ♀	30.098	202542423	20	47.69
				qLER3	3	11.4	22.9	6867_73 ♂	173.295	255081484	172.013	180
				qLER8.1	8	4.2	7.3	5302_81 ♂	32.711	53273330	18.019	40.879
				qLER8.2	8	9.0	17.4	9475_61 ♀	186.103	291435679 †	182.238	188
				qLER9	9	8.6	16.5	20368_67 ♂	126.526	179084475	121.799	136
				qLER8.2: qLER9	-	6.8	12.6	-	-	-	-	-

♀ Female marker, ♂ Male marker, * Multiple markers identified at peak, † Physical position of marker nearest peak that did map to v11

Supplementary Table S3.4. LOD scores and PVE for the full QTL models and individual QTL terms for ML and AS using raw means. Markers, genetic (G. peak) positions, physical (P. peak) positions, and a 1.5 LOD interval is represented for each QTL.

Full Model				**QTL**							**1.5 LOD Interval**	
Trait	Model	LOD	PVE	QTL	Chr	LOD	PVE	Marker Nearest Peak	G.peak	P.peak	Start (cM)	End (cM)
ML	y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + Q5:Q6	26.0	64.7	qMLR1	1	5.4	8.5	9967_68 ♂	30.526	202338175	22.376	34.39
				qMLR3.1	3	3.7	5.7	1564_71 ♂	72.674	87423046 †	60.21	90.429
				qMLR3.2	3	11.7	21.2	8472_64 ♂	175.871	256626504 †	172.013	180
				qMR8.1	8	5.3	8.3	5302_81 ♂	32.711	53273330	20.606	41.306
				qMLR8.2	8	7.8	12.9	9475_61 ♀	186.103	291435679 †	183.524	188
				qMLR9	9	8.9	15.2	11326_81 ♂	121.799	176581902 †	121.799	136
				qMLC8.2: qMLC9	-	5.9	9.4	-	-	-	-	-
AS	y ~ Q1 + Q3 + Q4 + Q5 + Q6 + Q7 + Q3:Q4	19.8	54.8	qASR1	1	4.0	7.8	16768_65 ♂	27.508	205870117	17.584	31.819
				qASR3.1	3	4.4	8.8	12347_75 ♀	89.575	106029465	60	93.443
				qASR3.2	3	11.6	26.6	8472_64 ♀	175.871	256626504 †	175.443	176.298
				qASB4	4	4.8	9.5	10158_62 ♀	101.759	212831724 †	86.709	103.041
				qISB8	8	4.6	9.2	2797_65 ♂	28.804	51839580 †	25.756	34.862
				qISB3.1: qISR3.2	-	3.6	7.0	-	-		-	-

♀ Female marker, ♂ Male marker, * Multiple markers identified at peak, † Physical position of marker nearest peak that did map to v11

Supplementary Table S3.5. LOD scores and PVE for the full QTL models and individual QTL terms for CF, OG, and LE using batch corrected means. Markers, genetic (G. peak) positions, physical (P. peak) positions, and a 1.5 LOD interval is represented for each QTL.

Full Model				**QTL**							**1.5 LOD Interval**	
Trait	Model	LOD	PVE	QTL	Chr	LOD	PVE	Marker Nearest Peak	G.peak	P.peak	Start (cM)	End (cM)
CF	y ~ Q1	4.7	16.3	*qCFC2*	2	4.7	16.3	12737_66 ♂	43.029	60710106	25.337	54.000
OG	y ~ Q1 + Q2 + Q3 + Q4 + Q5	21.2	57.2	*qOGC1*	1	4.3	8.1	855_63 ♀	30.098	202542423	20	34.39
				qOGC3	3	10.8	23.2	8472_64 ♀	175.871	256626504 †	174	180
				qOGC4	4	4.9	9.2	7017_61 ♂*	100.477	216241865	84.568	106
				qOGC8.1	8	4.0	7.5	5302_81 ♂	32.711	53273330	18.019	36
				qOGC8.2	8	3.1	5.7	3416_79 ♂	149.119	228311632	140.962	194
LE	y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + Q5:Q6	27.1	66.2	*qLEC1*	1	5.8	8.8	9967_68 ♂	30.526	202338175	22.376	36
				qLEC3.1	3	4.1	6.0	1564_71 ♀	72.674	87423046 †	66.668	83.909
				qLEC3.2	3	13.4	23.9	8472_64 ♀	175.871	256626504 †	172.013	180.59
				qLEC8.1	8	5.8	8.9	5302_81 ♂	32.711	53273330	26.184	41.306
				qLEC8.2	8	7.3	11.5	9475_61 ♀	186.103	291435679 †	182.238	189.121
				qLEC9	9	8.1	13.0	11326_81 ♂	121.799	176581902 †	121.799	136
				qLEC8.2: qLEC9	-	5.4	8.2	-	-	-	-	-

♀ Female marker, ♂ Male marker, * Multiple markers identified at peak, † Physical position of marker nearest peak that did map to v11

Supplementary Table S3.6. LOD scores and PVE for the full QTL models and individual QTL terms for ML and AS using batch corrected means. Markers, genetic (G. peak) positions, physical (P. peak) positions, and a 1.5 LOD interval is represented for each QTL.

Full Model				**QTL**							**1.5 LOD Interval**	
Trait	Model	LOD	PVE	QTL	Chr	LOD	PVE	Marker Nearest Peak	G.peak	P.peak	Start (cM)	End (cM)
ML	y ~ Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + Q5:Q6	27.4	66.6	*qMLC1*	1	6.1	9.2	9967_68 ♂	30.526	202338175	24.086	34.39
				qMLC3.1	3	4.4	6.5	1564_71 ♀	72.674	87423046 †	68.816	83.909
				qMLC3.2	3	12.9	22.6	8472_64 ♀	175.871	256626504 †	172.013	180.159
				qMLC8.1	8	5.8	8.7	5302_81 ♂	32.711	53273330	26.184	41.306
				qMLC8.2	8	7.7	12.1	9475_61 ♀	186.103	291435679 †	183.524	189.121
				qMLC9	9	8.6	13.8	11326_81 ♂	121.799	176581902 †	121.799	136
				qMLC8.2: qMLC9	-	5.7	8.6	-	-	-	-	-
AS	y ~ Q1 + Q5 + Q6 + Q7	19.4	54.1	*qASC1*	1	4.8	9.8	855_63 ♀	30.098	202542423	19.305	30.098
				qASC3	3	9.0	20.0	8472_64 ♀	175.871	256626504 †	174.15	180.159
				qASC4	4	4.4	8.8	10158_62 ♀	101.759	212831724 †	84.568	108
				qASC8	8	6.1	12.8	2797_65 ♂	28.804	51839580 †	21.888	34.862

♀ Female marker, ♂ Male marker, * Multiple markers identified at peak, † Physical position of marker nearest peak that did map to v11

Chapter 4: GRF-GIF chimeric proteins enhance *in vitro* regeneration and *Agrobacterium*-mediated transformation efficiencies of lettuce (*Lactuca* spp.)

ACKNOWLEDGEMENTS

I conducted the majority of the work including initiation and maintenance of tissue cultures, *Agrobacterium*-mediated transformation experiments, screening of transgenics, data collection, and data analysis. Juan Debernardi developed tomato (JD746, JD747, JD749, and JD761), grape (JD638), and citrus (JD689), and empty vector (JD641) constructs. Theresa Hill (Allen Van Deynze lab) developed the pepper GRF-GIF construct (pTH1903).

ABSTRACT

The ability of plants to regenerate *in vitro* has been exploited for use in tissue culture systems for plant propagation, plant transformation, and genome editing. The success of *in vitro* regeneration is often genotype dependent and continues to be a bottleneck for *Agrobacterium*-mediated transformation and its implementation for improvement of some crop species. Manipulation of transcription factors that play key roles in plant development such as BABY BOOM (BBM), WUSCHEL (WUS), and GROWTH REGULATING FACTORs (GRFs) has improved regeneration and transformation efficiencies in several plant species. In this chapter, I demonstrate the use of *GRF-GIF* gene fusions from multiple species to boost regeneration efficiency and shooting frequency in four genotypes of wild and cultivated lettuce (*Lactuca spp.* L.). In addition, we show that GRF-GIFs with mutated miRNA 396 (miR396) binding sites increase regeneration efficiency and shooting frequency when compared to controls. I also present a co-transformation strategy for increased transformation efficiency and recovery of transgenic plants harboring a gene of interest. This work will be applied to other genotypes of lettuce to increase recovery of transgenic plants as well as other crops in the Compositae family.

INTRODUCTION

Plants have a remarkable capacity to regenerate whole organs from differentiated cells through dedifferentiation and reprogramming of cells toward new cell fates. This capability has been exploited for use in tissue culture systems for plant propagation, plant transformation, and genome editing (Altpeter et al., 2016). The success of *in vitro* regeneration is often genotype dependent and continues to act as a bottleneck for *Agrobacterium*-mediated transformation and its implementation for improvement of some crop species. Multiple species are capable of regenerating (e.g., lettuce, tomato, *Arabidopsis*) but many high value crops, such as cotton, sunflower, and pepper, are still recalcitrant to regeneration (Gammoudi et al., 2018; Wu et al., 2004). Although highly successful, classical plant breeding has germplasm and interspecific limitations. Employment of plant biotechnologies can increase available germplasm resources through use of *Agrobacterium*-mediated transformation and transgenesis to access sexually incompatible genepools.

Previously, research has been conducted to identify methods to improve somatic embryogenesis and *de novo* organogenesis for the development of transgenic lines (Debernardi et al., 2020; Elhiti et al., 2021; Jones et al., 2019; Kong et al., 2020; Stasolla and Yeung, 2003; Zheng et al., 2013). For example, incorporation of several genes encoding developmental regulators such as *LEAFY COTYLEDON1* (*LEC1*), *LEC2*, *WUSCHEL* (*WUS*), and *BABY BOOM* (*BBM*) increases somatic embryogenesis by promoting vegetative to embryonic cell transition (Jones et al., 2019). Although use of these genes increased regeneration rates, unregulated ectopic expression also induced pleiotropic phenotypes affecting cotyledon, hypocotyl, and shoot development (Boutilier et al., 2004). Pathways involved in *in vitro* regeneration include those for plant development;

therefore, the use of other developmental genes regulating cell proliferation and organ development should be explored for increasing regeneration of recalcitrant species.

GROWTH-REGULATING FACTOR (GRF) transcription factors are involved in regulating multiple stages of plant development including leaf, stem, root, seed, and flower development. GRFs tend to be associated with areas of prolific cell division during development (reviewed in Omidbakhshfard et al., 2015; Rodriguez et al., 2016). The first GRF, *OsGRF1*, was identified in rice and was observed to play a role in gibberellic acid induced stem elongation (van der Knaap et al., 2000). Now, multiple GRFs have been identified and studied in both monocotyledonous and dicotyledonous species including rice, barley, wheat, maize, *Arabidopsis*, *Brassica* spp., tomato, potato, and lettuce (Huang et al., 2021; Khatun et al., 2017; Li et al., 2016; Ma et al., 2017; Rosenquist et al., 2001; Wang et al., 2014; B. Zhang et al., 2021; D. F. Zhang et al., 2008). Individual species tend to have multiple GRFs each with different developmental functions and most with two conserved domains. One domain, the QLQ domain, likely functions in protein-protein interactions, and the second domain, the WRC domain, probably functions in nuclear targeting and DNA binding (Omidbakhshfard et al., 2015; van der Knaap et al., 2000). GRFs have a transcriptional co-factor, GRF-INTERACTING FACTOR (GIF), which interacts with the QLQ domain and enhances the function of GRFs. GRFs are also post-transcriptionally regulated by microRNA 396 (miR396), which recognizes and binds to the nucleotides of the WRC domain (Omidbakhshfard et al., 2015). These three together make up what is known as the miR396-GRF/GIF module in plant development.

Transformation using a *GRF* only or a *GRF-GIF* chimeric transgene increased regeneration efficiency in both monocotyledonous and dicotyledonous species (Debernardi et al., 2020; Kong et al., 2020). The ectopic expression of *AtGRF5* increased transgenic callus formation in canola

and shoot organogenesis in multiple varieties of sugar beet, soybean, and sunflower (Kong et al., 2020). In addition, overexpression of GRF orthologs in each species resulted in boosted shoot organogenesis and transformation efficiency. The overexpression of maize *AtGRF* orthologs were also shown to increase formation of embryogenic calli leading to higher levels of regeneration by somatic embryogenesis in maize (Kong et al., 2020). Furthermore, the use of GRFs in conjunction with GIFs also enhanced regeneration frequencies and rates in wheat, triticale, rice, and citrus (Debernardi et al., 2020). Transformations using species specific homologs of the *GRF4-GIF1* chimeric transgene resulted in better regeneration when compared to transformations of *GRF4* only, *GIF1* only, and a co-transformation of *GRF4* and *GIF1*. Transgenic wheat plants were also recovered in the absence of antibiotic-based selectable markers and hormones. In addition, citrus and grape *GRF4-GIF1* orthologs were tested in citrus and resulted in an approximately five-fold increase in regeneration (Debernardi et al., 2020). A resistant grape *GRF-GIF* (r*GRF-GIF)* gene fusion containing synonymous mutations in the miR396 binding site to prevent post-transcriptional degradation of the GRF mRNA was also transformed into citrus. This resulted in the highest regeneration efficiency observed throughout this study. Last, genome editing efficiency was increased when adding the *GRF4-GIF* fusion into an editing construct (Debernardi et al., 2020). Both studies observed an increase in regeneration of recalcitrant or low-transforming cultivars and resulted in developmentally normal plants.

In this chapter, I tested species specific versions of the *GRF-GIF* chimeric transgene and characterized their ability to increase regeneration and transformation efficiency in multiple lettuce (*Lactuca* spp.) genotypes. The aim of this study was to answer the following four questions: 1) *GRF-GIF* chimeric transgenes from which species most enhances regeneration and shooting frequency in lettuce, 2) does mutation of the miR396 binding site of the *GRF* fragment alter

regeneration and transformation efficiency when compared to the wildtype *GRF* coding sequence, 3) is enhancement of regeneration efficiency and shooting frequency using *GRF-GIF* gene fusions genotype-independent in lettuce, and 4) what plant selection and co-transformation strategy is most efficient when using *GRF-GIFs* for enhancing transformation rates with a gene of interest? The results of this study provide further evidence that the introduction of *GRF-GIF* chimeric transgenes can increase regeneration and transformation efficiency, even across plant families. In addition, a new strategy was developed for increasing the efficiency for introducing genes of interest into lettuce.

METHODS

Vectors and vector construction

The miRNA-resistant chimeric *GRF4-GIF1* coding sequences from grape and wildtype fusion from citrus as described previously (Debernardi et al., 2020) were used for this study. To identify the homologous *GRF* and *GIF* genes from tomato (Solanaceae), phylogenetic trees were generated using GRF and GIF protein sequences from wheat, rice, *Arabidopsis*, *Citrus*, grape (*Vitis vinifera*), and tomato (*Solanum lycopersicum*); the QLQ and WRC domains were used for GRF protein sequences and the SNH domain was used for GIF protein sequences (J. Debernardi, unpublished). The evolutionary history was inferred by using the Maximum Likelihood method, and the tree with the highest log-likelihood is shown in **Supplementary Figure S4.1.** A BLAST search with wheat and grape *GRF* and *GIF* genes was used to identify *GRF* and *GIF* genes in pepper (*Capsicum annuum*); pepper GRF protein sequences were aligned with GRF protein sequences from citrus, grape, *Arabidopsis*, and *Medicago truncatula*, and pepper GIF protein sequences were aligned with GIF protein sequences from several species (Debernardi et al., 2020;

T. Hill, unpublished). Both alignments were generated using T-Coffee (https://tcoffee.crg.eu/) with the M-Coffee aligner. Phylogenetic trees were generated using MEGA5 (**Supplementary Figure S4.2**). GRF and GIF protein sequences most closely related to the grape *GRF4* and *GIF1* sequences were used to synthesize *GRF-GIF* fusion constructs.

All constructs for transformation were generated using Gateway™ cloning (J. Debernardi and T. Hill, unpublished). Chimeric sequences were developed by fusing the *GRF* and *GIF* coding sequences with a four-alanine linker and synthesized using Genewiz (https://www.genewiz.com/en). All *GRF-GIF* sequences were first cloned into either pDONR™/Zeo or pDONR221. The pepper *GRF4-GIF1* chimeric coding sequence (pTH1903) was then subcloned into pEarlyGate100 (pEG100) using the L/R Gateway™ reaction. The wild-type tomato *GRF8-GIF4* was subcloned in pEG100 (JD761) and pGWB14 (JD746). Resistant tomato r*GRF8-GIF4* (JD747) and r*GRF12-GIF4* (JD749) were subcloned into pGWB14. Empty vectors of pGWB14 (JD641) (Debernardi et al., 2020) and pEG100 (pTB005) were used as controls for each transformation.

Preparation of bacterial cultures

All plasmids were transformed into *Agrobacterium tumefaciens* strain LBA4404. Initial bacterial cultures were prepared by inoculating 20 mL of MGL medium (5 g/L tryptone, 2.5 g/L yeast extract, 5 g/L NaCl, 5 g/L mannitol, 0.1 g/L $MgSO_4$, 0.25 g/L K_2HPO_4, 1.2 g/L glutamic acid, 15 g/L sucrose; pH 7.2) supplemented with rifampicin (50 mg/L) and kanamycin (50 mg/L) with one colony of *Agrobacterium tumefaciens* strain LBA4404 harboring GRF-GIF plasmids required for each experiment. Cultures were incubated overnight in an orbital shaker at 28 °C at 200 rpm. The following day, subcultures were prepared by inoculating 17 mL of TY medium (5 g/L tryptone, 3 g/L yeast extract; pH 7.2) supplemented with rifampicin (50 mg/L), kanamycin (50

mg/L), and acetosyringone (40 mg/L) with three mL of the previous overnight culture. Cultures were incubated overnight in an orbital shaker at 28 °C at 200 rpm. The following morning, cultures were diluted to an OD_{600} between 0.1 to 0.2. For co-transformation, diluted bacterial cultures were mixed at a 1:1 ratio, and the OD_{600} was remeasured to ensure correct bacterial density. Acetosyringone was added to the final diluted cultures prior to transformation at a final concentration of 200 uM.

Preparation of explants and transformations

Seeds were surface sterilized with 20% Clorox for 20 minutes with constant agitation at 250 rpm. Sterile seeds were rinsed three times with 100 mL of sterile distilled water and sown on 1/2x Hoagland's medium (0.815 g/L Hoagland Modified Basal Salt mixture [PhytoTech Labs Product ID# H353], 8g/L PhytoAgar™ [PlantMedia SKU# 40100072-1], pH 5.6-5.8). Seeds were incubated for four days in a 24 °C growth room under a 12/12hr light/dark cycle with Honeywell LED lights (Model #SH450505Q2004) providing approximately 8,700 lux. After four days, explants were prepared by cutting off apical and basal tips of cotyledons while submerged in 20 mL of respective *A. tumefaciens* suspension cultures. Cotyledon explants were immediately transferred to SH co-cultivation medium (3.2g/L SH Salts, 30g/L sucrose, 2ml/L 500x MS Vitamins, 8g/L PhytoAgar, pH 7.2) supplemented with acetosyringone (200 uM), 0.1 mg/L of 6-benzylaminopurine (6-BAP), and 0.1 mg/L of 1-Naphthalenacetic Acid (1-NAA) and incubated in the dark for three days at 24 °C. A total of 30 independent transformations were performed with approximately 80 cotyledon explants each (**Supplementary Table S4.1**). Each independent transformation was split into five replications of approximately 16 explants per plate. Independent transformations and experiments are described in more detail below.

Comparison of regeneration stimulated by GRF-GIF fusions from four plant species

A readily regenerating genotype, *L sativa* cv. Cobham Green, and an inconsistently regenerating genotype, *L. serriola* accession Armenian 999, were used for transformation. These genotypes were transformed using the wild-type tomato (JD761), citrus (JD689), pepper (pTH1903), miRNA-resistant grape (JD638) *GRF-GIF* fusions, and an empty vector control (pTB005). After incubation on co-cultivation, explants were transferred to SH Induction (SHI) medium (3.2 g/L Schenk and Hilderbrandt (SH) Basal Salt Mixture [PhytoTech Labs Product ID# S816], 30 g/L sucrose, 2 mL/L 500x Murashige and Skoog (MS) Vitamins [PhytoTech Labs Product ID# M533], 8 g/L PhytoAgar™; pH 5.6-5.8; add after autoclaving – 0.10 mg/L 6-BAP, 0.10 mg/L 1-NAA, 150 mg/L timentin, and 400 mg/L carbenicillin) supplemented with 10 mg/L Glufosinate-ammonium (BASTA) and incubated in a HiPoint growth chamber (model FH-1200 LED Z4) at 26 °C under a 12/12hr light/dark cycle for two weeks.

Regeneration rates were scored by observing traits following the progression of *de novo* shoot organogenesis (organized growth, leaf emergence, and shoot formation). After 14 days in culture, explants were transferred to fresh SHI and were scored for occurrence of regeneration by observing the presence of organized growth (presence of shoot meristems and/or leaf primordia), leaves (leaf emergence), and shoots, and then returned to the growth chamber. For the Armenian 999 transformations, the presence of organized growth, leaves, and shoots was observed and recorded on two additional days (days 20 and 30). After 24 and 35 days in culture for Cobham Green and Armenian 999, respectively, explants showing regeneration were transferred to SH elongation (SHE) medium (3.2 g/L SH Basal Salt mixture, 30 g/L sucrose, 2 mL/L 500x MS vitamins, 8 g/L PhytoAgar™, pH 5.6-5.8; add after autoclaving – 0.01 mg/L 6-BAP, 0.05 mg/L 1-NAA, 150 mg/L Timentin) supplemented with 10 mg/L BASTA and returned to the growth

chamber. After transfer, the total number of explants showing regeneration and the total number of shoots present was scored for each replication of each transformation. The final shooting frequency and regeneration efficiency was calculated after 35 and 45 days in culture for Cobham Green and Armenian 999, respectively. Up to ten shoots for each transformation were transferred to rooting medium (4.34 g/L MS Salts, 2 mL/L 500x MS Vitamins, 30 g/L sucrose, 8 g/L PhyoAgar™; pH 5.6 to 5.8; add after autoclaving – 150 mg/L timentin) supplemented with 10 mg/L BASTA. Rooted plants were transferred into soil, acclimated in a 24 °C growth room with Honeywell LED lights (Model #SH450505Q2004) providing approximately 8,700 lux for 10 days, and then transplanted in the greenhouse for observation of developmental patterns.

Analysis of regeneration stimulated by wildtype or miRNA-resistant GRF-GIFs

To test efficacy of regeneration using wildtype *GRF-GIF* and miRNA-resistant *GRF-GIF* (r*GRF-GIF*) fusions, Cobham Green and Armenian 999 were used for transformation. These two genotypes were transformed with the wildtype tomato *GRF8-GIF4* (JD746), the tomato r*GRF8-GIF4* (JD747), the tomato r*GRF12-GIF4* (JD749), and an empty vector control (JD641). After incubation on co-cultivation, explants were transferred to SHI medium supplemented with 100 mg/L kanamycin and incubated in a HiPoint growth chamber (model FH-1200 LED Z4) at 26 °C under a 12/12hr light/dark cycle for two weeks. After 14 days on SHI, each transformation was scored for the occurrence of regeneration by observing the number of explants exhibiting organized growth, leaves, and shoots. On day 20, Armenian 999 transformations were scored as above. After 21 days in culture for Cobham Green and 27 days in culture for Armenian 999, all explants exhibiting regeneration were transferred onto SHE medium and returned to the growth chamber. The number of explants with at least one shoot and the total number of shoots per replication was scored. On days 35 and 45 for Cobham Green and Armenian 999, respectively, the

final number of total shoots and explants showing regeneration were scored and the shooting frequencies and regeneration efficiencies were calculated. Up to ten shoots for each transformation were transferred to rooting medium (4.34 g/L MS Salts, 2 mL/L 500x MS Vitamins, 30 g/L sucrose, 8 g/L PhyoAgar™; pH 5.6 to 5.8) supplemented with 100 mg/L kanamycin. Rooted plants were transferred into soil, acclimated in a 24 °C growth room with Honeywell LED lights (Model #SH450505Q2004) providing approximately 8,700 lux for 10 days, and then transplanted in the greenhouse for observation of developmental patterns.

The effect of lettuce genotype on regeneration using rGRF-GIF

The lettuce cultivars, Cobham Green, Salinas, Valmaine, and the *L. serriola* accession, Armenian 999, were used for transformation. Each genotype was transformed with the best performing GRF-GIF, grape r*GRF4-GIF1*, and an empty vector control (pTB005). After co-cultivation, cultures were transferred to SHI supplemented with BASTA (10 mg/L) and incubated in a 24 °C growth room with Honeywell LED lights (Model #SH450505Q2004) providing approximately 8,700 lux and under a 12/12hr light/dark cycle for 20 days. After 20 days on induction medium, explants were transferred to new SHI supplemented with BASTA (10 mg/L) and scored for the presence of organized growth, leaves, and shoots, and returned to the growth room. Organized growth, leaf emergence, and shooting frequencies were also collected for all four genotypes after 30 days on induction medium. After 40 days on induction medium, all explants exhibiting regeneration in Cobham Green, Valmaine, Salinas, and Armenian 999 cultures were transferred to SHE and scored for the shooting frequency and regeneration efficiency and returned to the growth room.

Co-transformation strategies to generate transgenics with genes of interest using rGRF-GIF

The grape r*GRF4-GIF1* (JD638), was used to test for increased regeneration of transgenics with a gene of interest using *dsRED* expressed from the *L. sativa* ubiquitin promoter and terminator (*pLsUBI:dsRED:tLsUBI*) and carried in a separate strain of *A. tumefaciens*. *A. tumefaciens* cultures were prepared for transformation as described above. Co-transformation cultures were prepared by mixing diluted cultures of JD638 and *pLsUBI:dsRED:tLsUBI* in a 1:1 ratio to give a final OD_{600} ranging from 0.139 to 0.145. Explants were prepared and co-transformed using two mixtures: 1) JD638 + *pLsUBI:dsRED:tLsUBI* and 2) pTB005 + *pLsUBI:dsRED:tLsUBI* (**Table 4.3**). Four co-transformations of *L. sativa* cv. Cobham Green were performed as described above, with three transformations using mixture 1 and one transformation using mixture 2. After co-cultivation, explants were transferred to SHI medium consisting of different antibiotic-based selection regimes. The three co-transformations using mixture 1 were transferred to SHI medium supplemented with either 1) 100 mg/L kanamycin (further referred to as GRF-GIF CoTF Kan) to select for all cells successfully transformed with the gene of interest (*dsRED*), 2) 10 mg/L BASTA (further referred to as GRF-GIF CoTF BASTA) to select for all cells successfully transformed with the grape r*GRF-GIF*, or 3) 100 mg/L kanamycin and 10 mg/L BASTA (further referred to as GRF-GIF coTF Kan + BASTA) to select for events only transformed with both T-DNAs. The fourth transformation was selected on SHI supplemented with 100 mg/L kanamycin (further referred to GRF-GIF CoTF control) to act as a control for transformation rate in the absence of GRF-GIF.

Maintenance of tissue cultures was performed as described for previous transformation experiments. After incubation on co-cultivation, explants were transferred to SHI supplemented with the different antibiotics and incubated in a HiPoint growth chamber (model FH-1200 LED

Z4) at 26 °C under a 12/12hr light/dark cycle for two weeks. After 14 days in culture, explants were transferred to fresh SHI supplemented with respective antibiotics and scored for the occurrence of organized growth, leaves, and shoots, and returned to the growth chamber. After 20 days on SHI medium, explants showing regeneration were transferred to SHE medium supplemented with respective antibiotics, the frequency of organized growth was collected, and the cultures were returned to the growth chamber. After 30 days in culture, all regenerated shoots were transferred to rooting medium supplemented with respective antibiotics, the shooting frequency and regeneration efficiency data was collected, and the plants were returned to the growth chamber. On day 38, up to 10 rooted plants were transferred into soil, acclimated in a 24 °C growth room with Honeywell LED lights (Model #SH450505Q2004) providing approximately 8,700 lux for 10 days, and then transplanted in the greenhouse for observation of developmental patterns.

DNA extraction and screening of transgenics

DNA was extracted from leaf tissue (~10 to 30 mg) of the regenerated plants using the QIAGEN DNeasy® Plant Mini Kit (Cat. No. 69104). DNA was extracted from leaf tissue of 49 regenerated shoots from GRF-GIF coTF Kan to screen for the presence of the grape r*GRF4-GIF1* and *dsRED* transgenes. PCR amplification of each fragment was performed using Promega GoTaq® Green Master Mix. Primer sequences and PCR conditions for each reaction are described in **Supplementary Table S4.2.** In addition, This data was used to calculate co-transformation efficiency. Furthermore, for the co-transformation experiments, the expression of *dsRED* was observed using the Leica MZ16NF dissecting microscope and a Chroma® dsRED filter (Product # 49004; filter ET605/70m). Frequencies of *dsRED* expression were used to calculate transformation

efficiencies of GRF-GIF coTF Kan, GRF-GIF coTF Kan + BASTA, and GRF-GIF coTF control, and the co-transformation efficiency of GRF-GIF coTF BASTA.

Data Analysis

For every replication of each transformation, regeneration efficiencies were determined by dividing the total number of explants with at least one shoot by the number of inoculated explants and multiplied by 100. Transformation and co-transformation efficiencies were determined by dividing the total number of explants with at least one shoot containing one or both (co-transformation) transgenes by the total number of explants inoculated and multiplied by 100. Organized growth and leaf emergence efficiencies were calculated by dividing the total number of explants exhibiting each trait by the number of inoculated explants and multiplied by 100. Shooting frequencies were calculated by dividing the total number of shoots or transgenic shoots by the total number of inoculated explants and multiplied by 100. Analysis of variance (ANOVA, $\alpha = 0.05$) was used to observe significant differences between regeneration traits (organized growth, leaf emergence, and shoot frequency) and regeneration and transformation efficiencies of each transformation experiment. A Tukey's Honest Significant Difference (TukeyHSD, $\alpha = 0.05$) test was used to calculate all pairwise comparisons for each batch of transformations. For the genotype independence experiments, a Welch's t-test ($\alpha = 0.05$) was used to compare means between the empty vector control and the grape *GRF4-GIF1* for each genotype. Data analysis and visualization was performed using Microsoft Excel version 2202 and RStudio version 2021.09.1+372.

RESULTS

Comparison of regeneration stimulated by GRF-GIF fusions from four plant species

GRF-GIF chimeric transgenes from four dicotyledonous species were tested for their effect on regeneration efficiency and shooting frequency in two genotypes of lettuce. Analysis of *GRF-GIF* chimeric transgenes from tomato (*GRF8-GIF4*), pepper (*GRF4-GIF1*), citrus (*GRF4-GIF1*), and grape (r*GRF4-GIF1),* were tested on lettuce cultivar Cobham Green and *L. serriola* accession Armenian 999, which regenerate readily and inconsistently, respectively (**Table 4.1**). For all Cobham Green transformations with and without GRF-GIF, organized growth and leaves were observed as early as 14 days after explants were placed on induction medium (**Figure 4.1**). The pepper *GRF4-GIF1* and grape r*GRF4-GIF1* exhibited significantly more organized growth when compared to the empty vector control, with the grape rGRF4-GIF1 showing the highest efficiency of organized growth (75% of explants). Use of the citrus GRF-GIF also exhibited an increase in organized growth when compared to the empty vector control, although this difference was not significant (**Figure 4.1**). Use of the tomato GRF8-GIF4 resulted in the most similar organized growth rates to the control and was significantly lower than the pepper and grape GRF-GIFs. On day 14, leaf emergence was also present in all cultures with grape and pepper GRF-GIFs showing the highest frequency (24% and 28%, respectively), although no significant difference was detected between any treatments at day 14 (**Figure 4.1**).

Organized growth was also detected after 14 days on induction medium with Armenian 999 (**Supplementary Figure S4.3),** although at lower frequencies than with Cobham Green. The citrus (18.6%) and grape (17.3%) GRF-GIF fusions exhibited significantly higher frequencies of organized growth than the tomato GRF8-GIF4 (10.1%), pepper GRF4-GIF1 (4.8%), and the empty

vector control (3.9%). Because Armenian 999 regenerates slower than Cobham Green (**Table 4.2**), organized growth and leaf emergence was also recorded after 20 and 30 days on induction medium (**Figure 4.2**). At these time points, similar results were observed for Armenian 999 as for Cobham Green. On day 20, the grape construct exhibited a significantly higher frequency of organized growth (71.6%) compared to the empty vector control (34.8%) and the tomato *GRF8-GIF4* (45.5%) and was the highest among all of the GRF-GIFs tested. In addition, both pepper (63.0%) and citrus (60.5%) GRF-GIFs exhibited significantly higher organized growth when compared to the empty vector control. Leaf emergence was present for all transformations except the tomato GRF8-GIF4 and had average efficiencies ranging from 5 to 12% but these did not differ significantly. On day 30, all GRF-GIF treatments showed significantly higher organized growth than the control (39.8%) with mean efficiencies ranging from 75.7 to 93.9%. Significant differences in the frequency of leaf emergence were also detected on day 30 for the Armenian 999 cultures with the grape rGRF-GIF exhibiting 4.7-fold higher frequencies (54.3%) than the empty vector control (11.5%) (**Figure 4.2**). Furthermore, the citrus GRF-GIF showed 2.8-fold higher frequency of leaf emergence (32.7%) when compared to the control.

The shooting frequency (number of shoots per explant) and regeneration efficiency (fraction of explants with at least one shoot x 100) were calculated after 24 and 35 days on induction medium for Cobham Green and Armenian 999, respectively. For Armenian 999, a significant difference was detected between regeneration efficiencies when comparing all treatments (ANOVA; $\alpha = 0.05$; $p = 0.03$), but a TukeyHSD test did not detect significant pairwise differences between pairs of treatments (**Supplementary Figure S4.3)**. In Cobham Green, the *grape rGRF4-GIF1* exhibited boosted regeneration efficiency (51.8% of explants) and shooting frequency (0.91 shoots/explant) when compared to the empty vector control (35.0% of explants,

0.41 shoots/explant) (**Figure 4.1**). In addition, the grape rGRF4-GIF1 showed significantly increased shooting frequencies when compared to the citrus (0.64 shoots/explant) and tomato (0.43 shoots/explant) GRF-GIFs. The pepper GRF-GIF (0.79 shoots/explant) showed higher shooting frequency than both the control and the tomato GRF8-GIF4.

Final shooting frequency and regeneration efficiency data were collected for Cobham Green and Armenian 999 after 35 and 45 days in culture, respectively after being transferred to elongation medium. In Cobham Green, introduction of the grape r*GRF4-GIF1* fusion resulted in approximately a 2.1-fold increase in regeneration efficiency (75.1%) and a 4-fold increase in shooting frequency (2.32 shoots/explant) when compared to the control (36.3%, 0.59 shoots/explant) and the tomato GRF8-GIF4 (41.2%, 0.64 shoots/explant) (**Figure 4.1**). Furthermore, the citrus (1.51 shoots/explant) and pepper (1.40 shoots/explant) GRF-GIFs exhibited an increase in shooting frequencies when compared to both the control and tomato GRF8-GIF4. After 45 days on regeneration medium, introduction of the grape r*GRF4-GIF1* into Armenian 999 resulted in significantly higher shooting frequency (1.68 shoots/explant) when compared to all other transformations, and a significant increase in regeneration efficiency (53.0%) when compared to the pepper GRF4-GIF1 (14.7%), tomato GRF8-GIF4 (15.7%), and the control (11.4%) (**Figure 4.2**). This was approximately a 4.5 and 4.6-fold increase in shooting frequency and regeneration efficiency when compared to the empty vector control (0.37 shoots/explant, 11.4%). The citrus GRF4-GIF1 exhibited a 2.4-fold increase in shooting frequency (0.91 shoots/explant), although this difference was not statistically significant. In addition, the citrus GRF-GIF displayed a significant boost in regeneration efficiency (33.7%) when compared to the control. Overall, for both Cobham Green and Armenian 999, the introduction of a *GRF-GIF*

chimeric transgene resulted in an increase in shooting frequency and regeneration efficiency, with the grape r*GRF4-GIF1* resulting in the highest rates of both.

Analysis of regeneration stimulated by wildtype or miRNA-resistant GRF-GIFs

Chimeric GRF-GIF fusions from tomato were used to identify if synonymous mutations within the miR396 binding site (r*GRF-GIF*) would enhance regeneration when compared to the wildtype GRF-GIF sequences. Cobham Green and Armenian 999 were transformed with four different constructs containing either the wildtype tomato *GRF8-GIF4*, the tomato r*GRF8-GIF4*, the tomato r*GRF12-GIF4*, or an empty vector control (JD641). For both Cobham Green and Armenian 999, organized growth and leaf emergence was observed on day 14, although no significant differences were detected between the empty vector control and any of the GRF-GIF treatments (**Figure 4.4, Supplementary Figure S4.4**). Similarly, in Cobham Green no significant differences for shooting frequency or regeneration efficiency were observed between any treatments at 21 days (**Figure 4.4**); however, in Armenian 999, both resistant GRF-GIFs showed significantly higher organized growth and leaf emergence than the wildtype GRF-GIF on days 14 and 20 (**Figure 4.3, Supplementary Figure S4.4**).

Significant differences in shooting frequency and regeneration efficiency were observed after 30 and 45 days in culture for Armenian 999 (**Figure 4.3, Supplementary Figure S4.4**) and after 35 days in Cobham Green (**Figure 4.4, Table 4.1**). After 30 days in culture for Armenian 999, rGRF8-GIF4 regeneration efficiency (74.0%) and shooting frequency (1.29 shoots/explant) was significantly higher than both the wildtype GRF8-GIF4 (52.5%, 0.86 shoots/explant) and empty vector control (53.2%, 0.72 shoots/explant) (**Supplementary Figure S4.4**). In addition, the regeneration efficiency (70.8%) and shooting frequency (1.33 shoots/explant) of the rGRF12-GIF4

significantly increased compared to both the wildtype GRF8-GIF4 and the control. Furthermore, after 45 days in culture, both rGRF-GIFs had approximately 20% higher regeneration efficiencies and 1.5 and 2.4-fold higher shooting frequencies than the wildtype GRF-GIF and the empty vector control, respectively **(Figure 4.3).** In Cobham Green, mean regeneration efficiencies were 44.0% for both rGRF8-GIF4 and rGRF12-GIF4, which is approximately 4.7% and 5.7% higher than the regeneration efficiencies observed for the wildtype GRF8-GIF4 and the control, respectively (**Figure 4.4**). No difference in regeneration and shooting frequencies were observed between genotypes and no significant interactions were detected (**Table 4.2**). The rGRF8-GIF4 and rGRF12-GIF4 produced approximately 1.4-fold more shoots than both the wildtype GRF8-GIF4 and controls, which is similar to the increase in shooting frequency between miRNA396 resistant and wildtype GRF-GIFs as seen in Armenian 999.

Effect of lettuce genotype on regeneration using rGRF-GIF

To test the efficiency of the GRF-GIF system in lettuce, I transformed the highest performing *GRF-GIF* (grape r*GRF4-GIF1*) or an empty vector control (JD638) into four genotypes of lettuce: *L. sativa* cv. Cobham Green (a butterhead type), *L. sativa* cv. Salinas (a crisphead type), *L. sativa* cv. Valmaine (a romaine), and *L. serriola* Armenian 999 (wild accession). After 20 days on induction medium, the grape rGRF4-GIF1 exhibited significantly increased frequencies of organized growth in all genotypes when compared to the empty vector control **(Supplementary Figure S4.5**). Similar to before, the rGRF-GIF showed significantly higher frequencies of leaf emergence (50.0%) and shoots (0.21 shoots/explant) in Cobham Green, which was approximately a 2-fold increase when compared to the frequency of leaf emergence (25.0%) and shoots (0.12 shoots/explant) in the control. Furthermore, the grape rGRF4-GIF1 significantly increased both organized growth and leaf emergence frequencies in all genotypes

after 30 days on induction medium **(Supplementary Figure S4.6**). Transformation with the *GRF-GIF* displayed frequencies of organized growth ranging from 82% to 98% and leaf emergence frequencies ranging from 28% to 74%, while the empty vector control displayed organized growth and leaf emergence frequencies ranging from 24% to 69% and 0% to 41%, respectively.

After 40 days on induction medium, the grape rGRF4-GIF1 increased shooting frequency and regeneration efficiency of all genotypes when compared to the empty vector control (**Figure 4.5**). Introduction of the grape r*GRF4-GIF1* into Cobham Green resulted in a 2.1 and a 2.5-fold increase in shooting frequency and regeneration efficiency when compared to the empty vector control. In Armenian 999 cultures, transformation with the grape r*GRF4-GIF1* led to a 0.55 increase in shoots per explant and a 29.4% increase in regeneration efficiency when compared to the control. These values varied from values reported in the first experiment, which is most likely due to environmental difference (e.g. temperature, lights) between growth chambers. In Salinas, a significant increase in both shooting frequency (0.58 shoots/explant) and regeneration efficiency (36.3%) was observed when compared to the empty vector control (0.013 shoots/explant, 1.3%). In addition, the introduction of the r*GRF4-GIF1* into Valmaine significantly increased the shoot frequency (0.39 shoots/explant) and regeneration efficiency (26.8%) when compared to the control (0.02 shoots/explant, 2.4%). Therefore, grape rGRF4-GIF1 increased regeneration of all cultivars but this effect was greatest on the cultivars that regenerated poorly in its absence. The greatest enhancement was observed with regeneration efficiency of cv. Salinas from 1.3% to 36%.

Co-transformation strategies to generate transgenics with genes of interest using rGRF-GIF

To test the ability of a GRF-GIF fusion to increase the recovery rate of transgenic plants, I co-transformed the best performing GRF-GIF (grape rGRF4-GIF1) and a reporter construct

(*pLsUBI-dsRED-tLsUBI*) in a separate strain of *A. tumefaciens* into Cobham Green. In order to optimize a strategy to increase transformation efficiency of a gene of interest, I examined different antibiotic-selection treatments for each co-transformation experiment. Three co-transformations were performed of Cobham Green with the grape *rGRF4-GIF1* (BASTA resistance) and *pLsUBI-dsRED-tLsUBI* (kanamycin resistance), selecting on kanamycin (GRF-GIF coTF Kan), BASTA (GRF-GIF coTF BASTA), or kanamycin and BASTA (GRF-GIF coTF Kan + BASTA) (**Table 4.3**). As a control, a fourth co-transformation in Cobham Green was performed using an empty vector harboring BASTA resistance (pTB005) and *pLsUBI-dsRED-tLsUBI*. The control co-transformation selected for transformants on kanamycin (GRF-GIF coTF control).

Co-transformation with *GRF-GIF* boosted regeneration efficiency and shooting frequency **(Figure 4.6)** in lettuce. GRF-GIF coTF Kan showed the highest regeneration efficiency of all selection treatments with an average regeneration efficiency of 39.3%, which was significantly higher than the control (27.9%). No significant difference was detected between regeneration efficiencies GRF-GIF coTF BASTA (36.3%) and GRF-GIF coTF Kan + BASTA (37.3%). Multiple shoots were regenerated from both cut sites either side of individual explants; therefore, I calculated the shooting frequency of each co-transformation. Shooting frequencies differed significantly between treatments. GRF-GIF coTF BASTA (0.79 shoots per explant) and GRF-GIF coTF Kan (0.71 shoots per explant) had significantly increased frequencies of shooting, which were approximately two-fold higher than the control (0.38 shoots per explant). The GRF coTF Kan + BASTA (0.45) had a significantly lower shooting frequency when compared to the GRF-GIF coTF BASTA and GRF-GIF coTF kan. In addition, GRF-GIF coTF Kan + BASTA showed higher shooting frequency when compared to the control, although this difference was not significant. Therefore, co-cultivation with two strains carrying rGRF-GIF or the gene of interest

(*pLsUBI-dsRED-tLsUBI*) and selection only for the T-DNA with the gene of interest using kanamycin, led to a significant increase in both shooting frequency and regeneration efficiency.

To determine the proportion of transformants with only the gene of interest versus transformants with both transgenes, I checked *dsRED* expression in regenerated shoots to confirm the presence of the gene of interest for GRF-GIF coTF Kan and control experiments (**Figure 4.6**) and the presence of both T-DNAs by PCR of the coTF Kan treatment. The GRF-GIF coTF Kan treatment had a similar frequency of *dsRED* expressing plants (83%) as the control (89.2%), which did not differ significantly; however, the GRF-GIF coTF Kan produced approximately double the total amount of shoots when compared to the coTF control. In addition, I wanted to identify the frequency of regenerated shoots that showed *dsRED* expression for the other co-transformations (**Figure 4.6**). As expected, GRF-GIF coTF BASTA had a significantly lower number of shoots expressing *dsRED* when compared to other treatments (30.1%) and the GRF-GIF coTF Kan + BASTA resulted in similar frequencies of shoots expressing *dsRED* (90.1%) as the coTF control. A total of 49 shoots from the GRF-GIF coTF Kan treatment were screened of which 24 showed amplification for the *GRF-GIF* transgene (49%), resulting in a mean co-transformation efficiency of approximately 24% (**Figure 4.7**). This is similar to the dsRED expression-based estimate of co-transformation efficiency of the GRF-GIF coTF BASTA treatment. Therefore, co-transformation of an r*GRF-GIF* and a gene of interest in separate strains of *A. tumefaciens* increases the recovery of transgenic shoots harboring a gene of interest.

DISCUSSION

In vitro plant regeneration has been studied for decades; however, it remains a rate-limiting factor for *Agrobacterium*-mediated transformation for the development of transgenic plants and genome editing. Although considerable progress has been made in improving *in vitro* regeneration,

efficient transformation is still limited in some crop species, including sunflower, cotton, and pepper (Gammoudi et al., 2018; Wu et al., 2004). In addition, regeneration and transformation efficiencies in many species, including *Lactuca* spp., are genotype dependent. In this chapter, I tested *GRF-GIF* gene fusions from four species for their efficacy in increasing regeneration and transformation efficiency, and their ability to induce genotype-independent regeneration and transformation in lettuce.

My data showed that transformation using *GRF-GIF* gene fusions could increase regeneration efficiency and shooting frequency in multiple lettuce genotypes. In both Cobham Green and Armenian 999, the grape rGRF4-GIF1 resulted in the highest frequencies of shoots and increased regeneration efficiency. The wildtype pepper GRF4-GIF1 and citrus GRF4-GIF1 also increased shooting frequency in Cobham Green but did not significantly increase regeneration efficiency. Furthermore, the citrus GRF4-GIF1 increased regeneration efficiency, but did not significantly increase shoot frequency in Armenian 999. Increasing sample size might lead to a significant increase. Although both tomato and pepper are in the Solanaceae family, the wildtype tomato GRF8-GIF4 did not increase shooting frequency or regeneration efficiency in any transformation experiment, while the pepper GRF-GIF did. Therefore, taxonomic affinity between the source species for *GRF* and *GIF* genes and the target species may not be a predictor for the efficacy of *GRF-GIF* fusions. It would be interesting to compare to GRF and GIF protein sequences from lettuce to all the GRFs and GIFs used. In addition, this may allow for the developmental of predictor models and to identify potential relationships between source GRF-GIF fusions and target species.

MicroRNA resistant GRF-GIFs due to a mutated miR396 binding site boost regeneration when compared to the wildtype GRF-GIF. All constructs tested had four synonymous nucleotide

changes in the miRNA binding site. These mutations in the miR396 binding site prevent post-transcriptional down regulation by miR396. The grape rGRF4-GIF1 exhibited the highest shooting frequencies and regeneration efficiencies of all fusions tested. This is consistent with previous reports of the grape rGRF4-GIF1 providing the largest increase in regeneration in citrus compared to experiments testing a variety of GRF-GIFs in multiple species (Debernardi et al., 2020). Both of the miR396 resistant tomato GRF-GIFs, rGRF8-GIF4 and rGRF12-GIF4, enhanced shooting frequencies and regeneration efficiencies compared to the wildtype tomato GRF8-GIF4 and the empty vector control. This is consistent with previous work identifying regulatory patterns of miR396 and *GRF*s with mutated miR396 binding sites and the increased regeneration rates of other species when using resistant versions (Debernardi et al., 2014, 2020; Rodriguez et al., 2010). It would be interesting to perform transformations using resistant versions of the citrus and pepper *GRF*s to determine whether they boost regeneration efficiency and shooting frequency further compared to the wildtype versions and to levels higher than the tomato versions. Because the grape wildtype *GRF4-GIF1* was not available to us, I cannot conclude if the observed increase in regeneration was due to the source species or the presence of a mutated binding site. Repeating this experiment using the wildtype grape version would be informative because it would give us information about how distant evolutionarily, we can apply the GRF-GIF system.

The grape rGRF4-GIF1 increased regeneration in multiple diverse genotypes of lettuce exhibiting different tendencies to regenerate. This GRF-GIF increased regeneration efficiency and shooting frequency of four genotypes of lettuce, *L. serriola* acc. Armenian 999 (wild progenitor), and cultivars Cobham Green (butterhead), Salinas (crsiphead), and Valmaine (romaine). Crisphead (iceberg) and romaine type lettuce are the two most commonly consumed types of lettuce in the United States. The largest increase in regeneration was observed in genotypes that do not readily

regenerate in the absence of GRF-GIF. This indicates that the GRF-GIF system could provide successful, genotype-independent transformations of diverse lettuce cultivars. In addition, co-transformation using the grape r*GRF4-GIF1* and a construct harboring a gene of interest resulted in boosted transformation efficiency and higher recovery of transgenic plants. This may be a result from the accumulation of both the GRF-GIF induced regeneration from co-transformed cells in addition to the routinely recovered transgenics events seen with a single transformation. The increased transformation and regeneration efficiencies when *pLsUBI:dsRED:tLsUBI* was co-transformed with the grape r*GRF4-GIF1* in separate *A. tumefaciens* strains and regenerated on media selecting for only the gene of interest could be readily applied to enhance the generation of transgenic plants using extant constructs without modification. When there has been integration of both T-DNAs, the *GRF-GIF* will have to be segregated away in the next generation; this poses a constraint for clonally propagated crops.

Transformation of the *L. sativa* cultivars with r*GRF-GIF* did not result in an obvious change in phenotype. However, transformation of Armenian 999 with the grape r*GRF4-GIF1* resulted in abnormal development and maturation of vegetative shoots and leaves (**Supplementary Figure S4.7**). Armenian 999 has a different leaf morphology than the cultivars. Therefore, this abnormal phenotype may be due to the inability of miR396 to post-transcriptionally regulate the transgene to allow for proper development of leaves and shoots. However, this phenotype was not observed in Armenian 999 transformed with the two miRNA-resistant tomato fusions.

The experiments in this chapter complement previous reports and provide further evidence of the wide efficacy of the GRF-GIF system. Repeating these experiments using *GRF* and *GIF* genes derived from lettuce may boost regeneration efficiency even further. Fifteen *GRF* genes

have been identified in lettuce of which one has been shown to increase leaf size when ectopically expressed (Zhang et al., 2021). To my knowledge, no other studies have been conducted on identifying and characterizing *GIF* genes in lettuce. Identifying closest lettuce homologs of the *GRF*s and *GIF*s used in this study, particularly the grape r*GRF4-GIF1,* may lead to higher levels of regeneration when used as fusions in lettuce. In the future, this work will be extended to increase regeneration of other recalcitrant genotypes of lettuce, as well as to other recalcitrant crops of the Compositae family, such as sunflower.

FIGURES AND TABLES

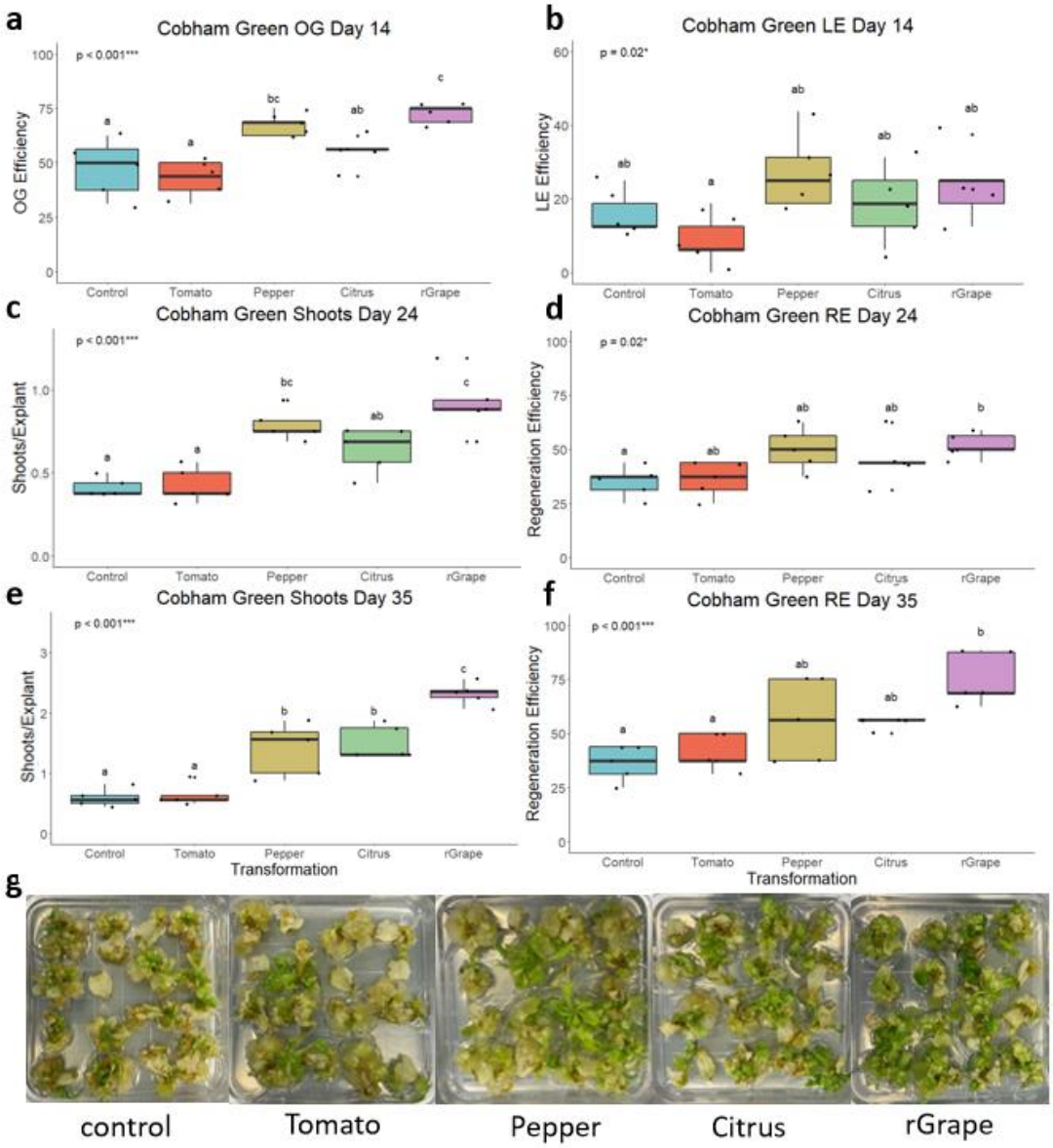


Figure 4.1. Regeneration rates of *L. sativa* Cobham Green after transformation with *GRF-GIF* gene fusions from tomato, pepper, citrus, and grape. **a-f)** Boxplots representing efficiencies of organized growth (a) and leaf emergence (b) at day 14, shoots per explant (c) and regeneration efficiency (d) at day 24, and shooting frequency (e) and regeneration efficiency (f) at day 35. Letters above each boxplot represent pairwise significance differences (TukeyHSD, $\alpha = 0.05$) and p-values were calculated using a one-way ANOVA ($\alpha = 0.05$). **g)** Tissue cultures of Cobham Green after 24 days on induction medium.

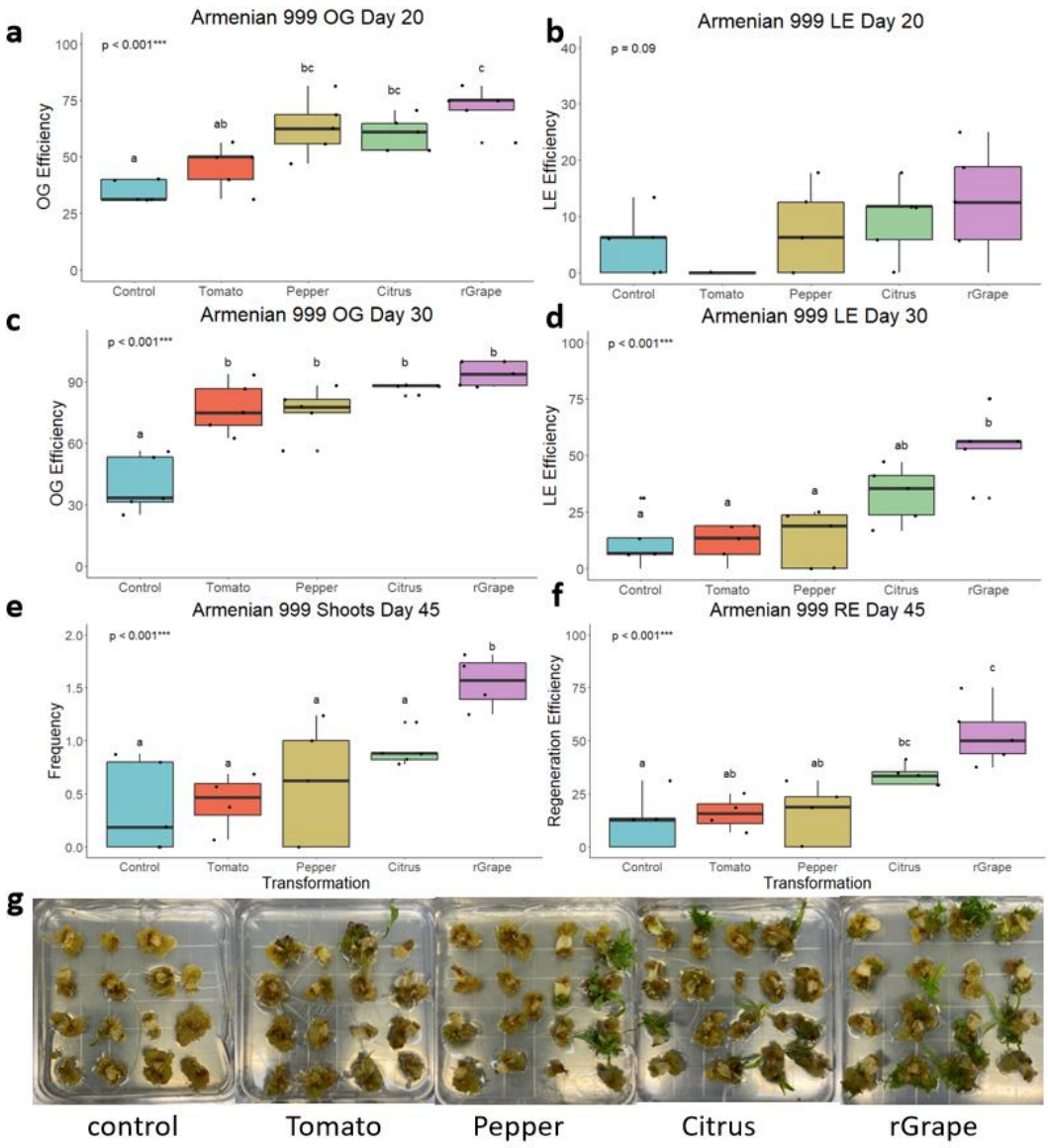


Figure 4.2. Regeneration rates of *L. serriola* accession Armenian 999 after transformation with *GRF-GIF* gene fusions from tomato, pepper, citrus, and grape. **a-f)** Boxplots representing efficiencies of organized growth (a) and leaf emergence (b) at day 20, organized growth (c) and leaf emergence (d) at day 30, and shooting frequency (e) and regeneration efficiency (f) at day 45. Letters above each boxplot represent pairwise significance differences (TukeyHSD, $\alpha = 0.05$) and p-values were calculated using a one-way ANOVA ($\alpha = 0.05$). **g)** Tissue cultures of Armenian 999 after 30 days on induction medium.

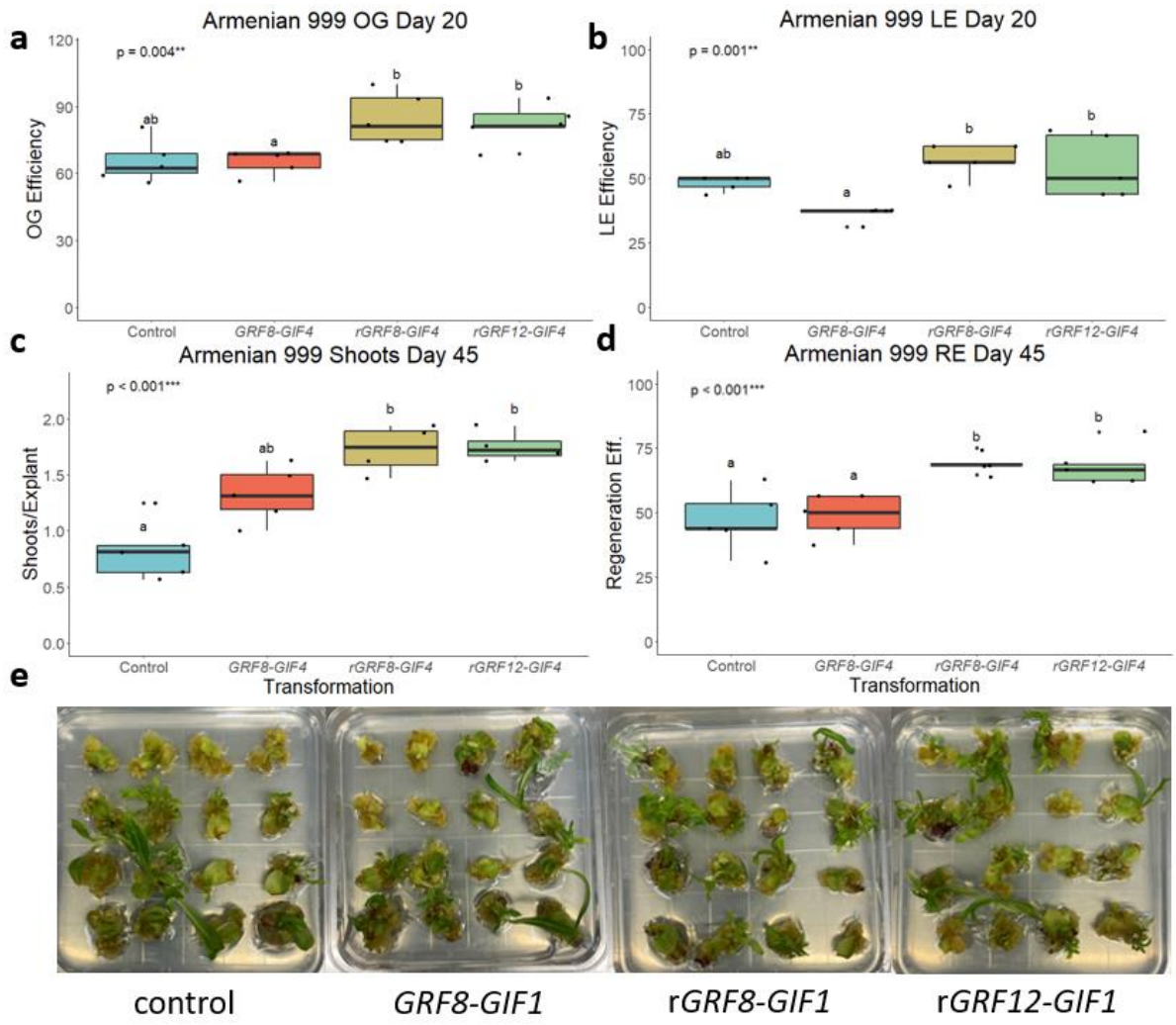


Figure 4.3. Regeneration rates of Armenian 999 after introduction of the wildtype and miR396 resistant tomato *GRF-GIF* fusions. **a-d)** Boxplots representing frequencies of organized growth (a) and leaf emergence (b) at day 20, and shooting frequency (c) and regeneration efficiency (d) at day 45. Letters above each boxplot represent pairwise significance differences (TukeyHSD, $\alpha = 0.05$) and p-values were calculated using a one-way ANOVA ($\alpha = 0.05$). **e)** Tissue cultures of Armenian 999 after 30 days on induction medium.

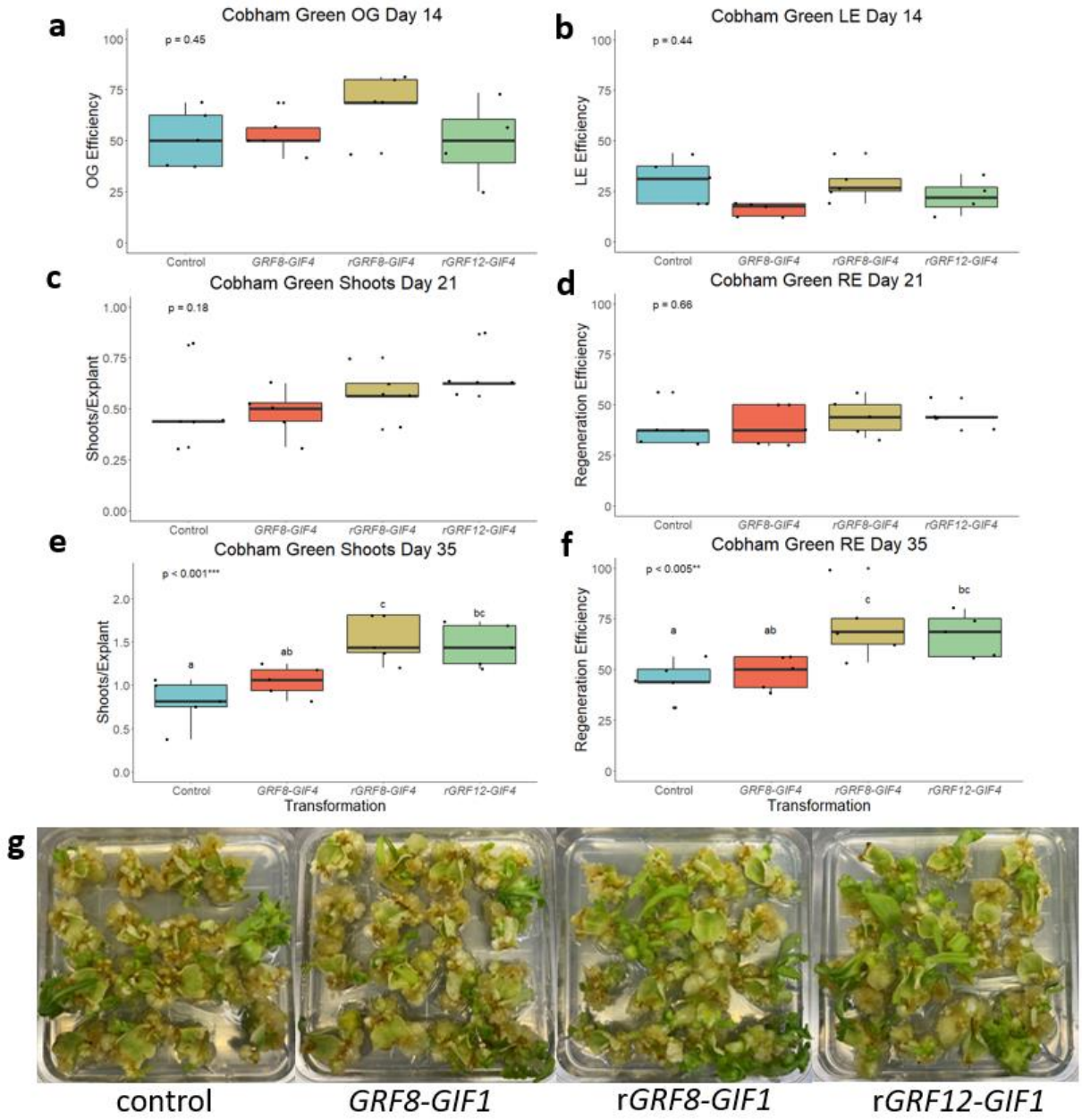


Figure 4.4. Regeneration rates of Cobham Green after introduction of the wildtype and miR396 resistant tomato *GRF-GIF* fusions. **a-f)** Boxplots representing frequencies of organized growth (a) and leaf emergence (b) at day 14, shooting frequency (c) and regeneration efficiency (d) at day 21, and shooting frequency (e) and regeneration efficiency (f) at day 35. Letters above each boxplot represent pairwise significance differences (TukeyHSD, $\alpha = 0.05$) and p-values were calculated using a one-way ANOVA ($\alpha = 0.05$). **g)** Tissue cultures of Cobham Green after 21 days on induction medium.

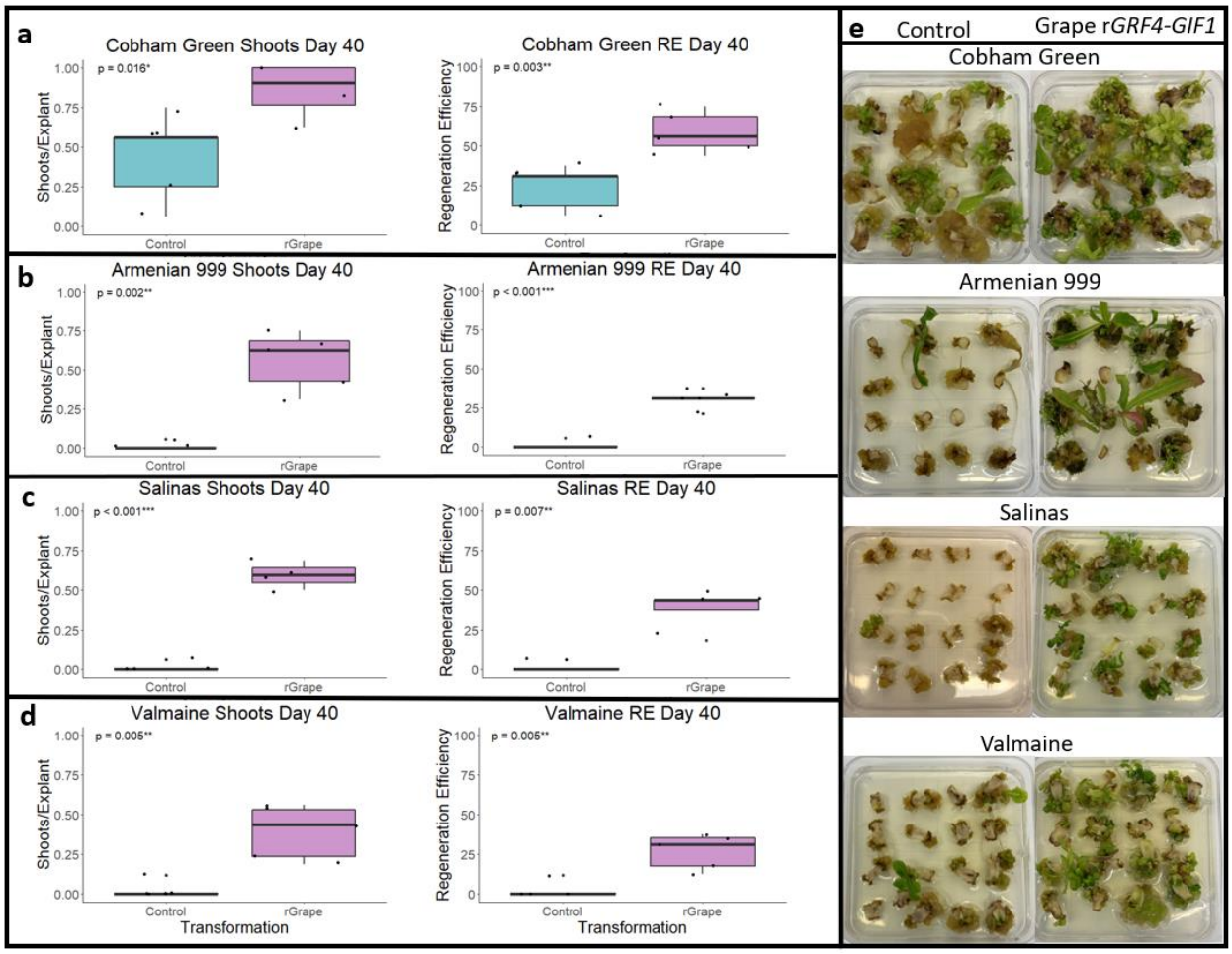


Figure 4.5. Regeneration rates of different lettuce genotypes after transformation with the grape r*GRF4-GIF1* or an empty vector control. **a-d**) Boxplots represent shooting frequencies and regeneration efficiency for each transformation of Cobham Green (a), Armenian 999 (b), Salinas (c), Valmaine (d) after 40 days on induction medium. P-values were calculated using a Welch's t-test. **e**) Forty-day old tissue cultures of each genotype after introduction of an empty vector control (left) and the grape r*GRF4-GIF1* fusion (right)

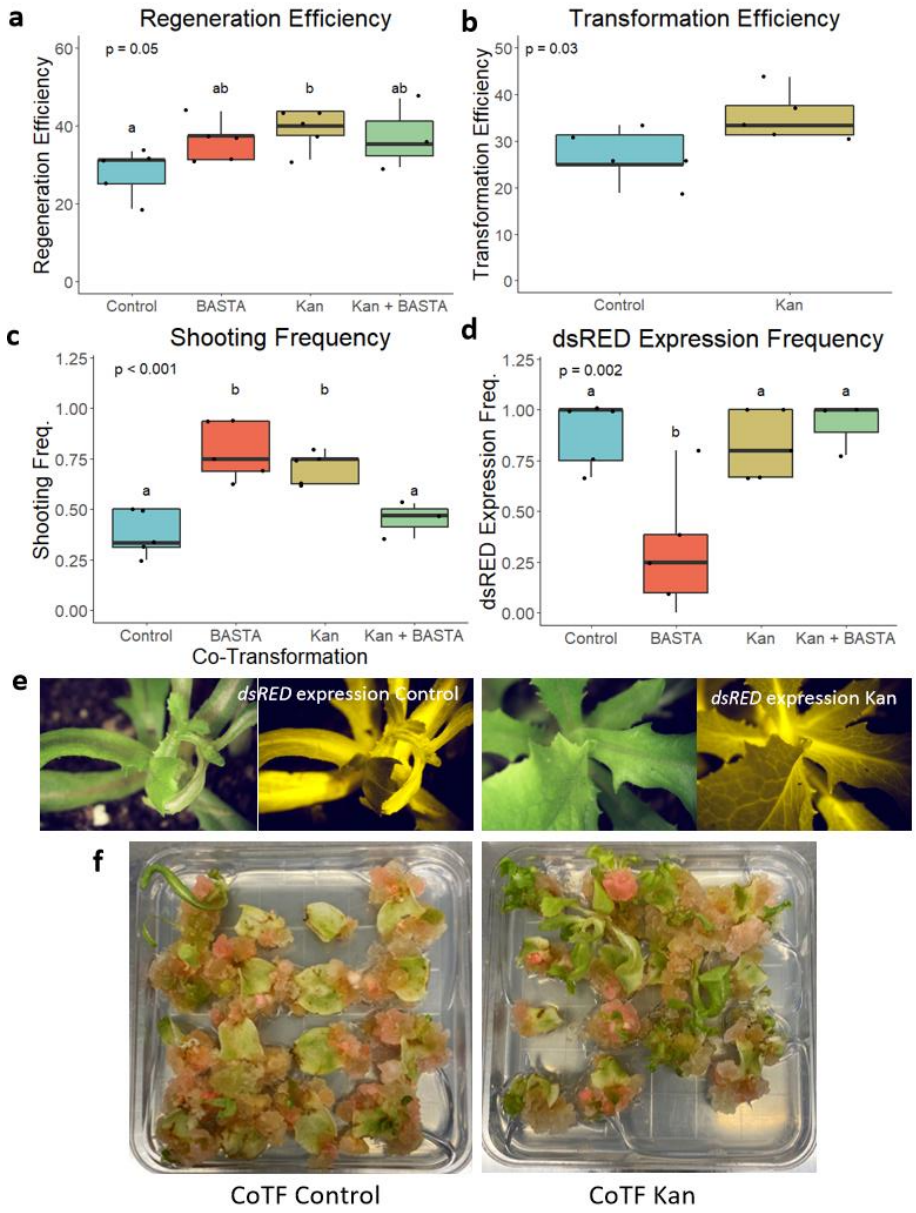


Figure 4.6. Summary of the co-transformation experiment in Cobham Green using *pLsUBI:dsRED:tLsUBI* + Grape r*GRF4-GIF1* selected on either kanamycin (Kan), BASTA^{TM} (BASTA) or kanamycin + BASTA^{TM} (Kan + BASTA) and *pLsUBI:dsRED:tLSUBI* + empty vector control (pTB005) selected on kanamycin (control). **a-d**) Boxplots representing regeneration efficiency (a), transformation efficiency (b), shooting frequency (c), and *dsRED* expression frequency (d) of each co-transformation. Letters above each boxplot represent pairwise significance differences (TukeyHSD, $\alpha = 0.05$) and p-values were calculated using a one-way ANOVA ($\alpha = 0.05$). **e**) *dsRED* expressing transgenics from the control (left) and Kan (right) co-transformations. **f**) Differences of regeneration in tissue culture between control (left) and Kan (right) co-transformations.

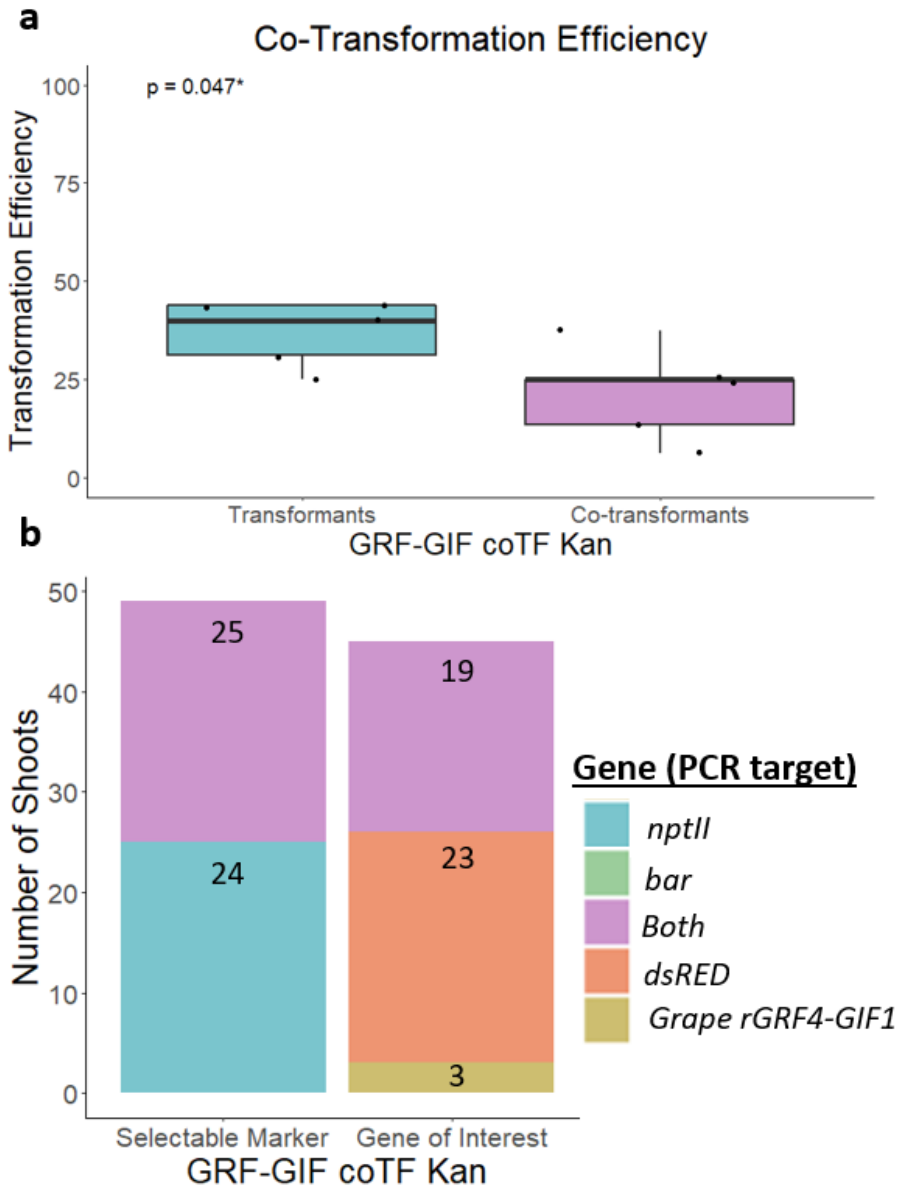


Figure 4.7. Transformation and co-transformation efficiencies of the GRF-GIF coTF Kan treatment. **a**) Transformation and co-transformation efficiencies. The p-value was calculated using a Welch's t-test ($\alpha = 0.05$). **b**) The number of shoots PCR positive for selectable marker (kanamycin and/or BASTA) and transgene (*dsRED* and/or GRF4-GIF1). Each color represents the proportion of shoots that showed amplification of each specific gene target and "both" is representative of the co-transformants. The black numbers refer to the number of shoots PCR positive for each condition.

Table 4.1. Means of final shooting frequencies (shoots per explant) and regeneration efficiencies (%) of each treatment from the GRF-GIF experiments in Armenian 999 and Cobham Green. Significant differences (p-values) of treatments for each experiment (one-way ANOVA, $\alpha = 0.05$) and pairwise comparisons between each treatment (TukeyHSD) for each experiment.

Variety	Construct	Mean Shooting Frequency	ANOVA	TukeyHSD	Mean Regeneration Efficiency	ANOVA	TukeyHSD
Armenian 999	pTB005	0.37		a	11.42		a
	Tomato	0.42		a	15.73		ab
	Pepper	0.57	0.0002***	a	14.71	0.00008***	ab
	Citrus	0.91		a	33.73		bc
	rGrape	1.68		b	53.01		c
	JD641	0.82		a	46.92		a
	Tomato WT *GRF8-GIF4*	1.33		ab	48.75		a
	Tomato r*GRF8-GIF4*	1.94	0.0008***	b	69.19	0.0004***	b
	Tomato r*GRF12-GIF4*	1.96		b	68.33		b
Cobham Green	pTB005	0.58		a	36.25		a
	Tomato	0.64	6.19×10^{-9}***	a	41.25	0.0002***	a
	Pepper	1.40		b	56.25		ab

	Citrus	1.51		b	55.0		ab
	rGrape	2.32		c	75.15		b
	JD641	0.80		a	45.00		a
	Tomato WT *GRF8-GIF4*	1.05		ab	48.24		ab
	Tomato r*GRF8-GIF4*	1.53	0.0006***	c	71.92	0.005**	c
	Tomato r*GRF12-GIF4*	1.46		bc	67.25		bc

Table 4.2. Calculated p-values (one-way ANOVA, $\alpha = 0.05$) after comparing final shooting frequencies and regeneration efficiencies of constructs, genotypes, and a construct by genotype interaction. P-values are shown for the first and second experiments testing the stimulation of regeneration using GRF-GIF fusions from different species and a comparison of the stimulation of wildtype and resistant GRF-GIFs with a mutated miRNA binding site.

Experiment	Antibiotic Selection	Factor	Shooting Frequency P-Value	Regeneration Efficiency P-Value
GRF-GIF stimulation from different species	BASTA	Construct	3.23×10^{-12}***	1.42×10^{-8}***
		Genotype	4.08×10^{-6}***	4.44×10^{-10}***
		Interaction	0.178	0.287
Wildtype vs. rGRF-GIF	Kanamycin	Construct	5.04×10^{-7}***	2.27×10^{-6}***
		Genotype	0.007**	0.953
		Interaction	0.409	0.962

Table 4.3. Co-transformation treatments and selection strategies.

Co-Transformation	Constructs	Selection
CoTF BASTA	*pLsUBI-dsRED-tLSUBI* + JD638	BASTA
CoTF Kan	*pLsUBI-dsRED-tLSUBI* + JD638	Kanamycin
CoTF Kan + BASTA	*pLsUBI-dsRED-tLSUBI* + JD638	BASTA + Kanamycin
CoTF control	*pLsUBI-dsRED-tLSUBI* + pTB005	Kanamycin

REFERENCES

Altpeter, F., Springer, N. M., Bartley, L. E., Blechl, A. E., Brutnell, T. P., Citovsky, V., Conrad, L. J., Gelvin, S. B., Jackson, D. P., Kausch, A. P., Lemaux, P. G., Medford, J. I., Orozco-Cárdenas, M. L., Tricoli, D. M., van Eck, J., Voytas, D. F., Walbot, V., Wang, K., Zhang, Z. J., and Neal Stewart Jr, C. (2016). Advancing Crop Transformation in the Era of Genome Editing OPEN. *The Plant Cell*, *28*, 1510–1520.

Boutilier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C. M., van Lammeren, A. A. M., Miki, B. L. A., Custers, J. B. M., and van Lookeren Campagne, M. M. (2002). Ectopic Expression of *BABY BOOM* Triggers a Conversion from Vegetative to Embryonic Growth. *The Plant Cell*, *14*(8), 1737–1749.

Debernardi, J. M., Mecchia, M. A., Vercruyssen, L., Smaczniak, C., Kaufmann, K., Inze, D., Rodriguez, R. E., and Palatnik, J. F. (2014). Post-transcriptional control of GRF transcription factors by microRNA miR396 and GIF co-activator affects leaf size and longevity. *The Plant Journal*, *79*(3), 413–426.

Debernardi, J. M., Tricoli, D. M., Ercoli, M. F., Hayta, S., Ronald, P., Palatnik, J. F., and Dubcovsky, J. (2020). A GRF–GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nature Biotechnology 2020 38:11*, *38*(11), 1274–1279.

Elhiti, M., Mira, M. M., So, K. K. Y., Stasolla, C., Hebelstrup, K. H., Loyola-Vargas, M., and Ochoa-Alejo, N. (2021). Synthetic Strigolactone GR24 Improves *Arabidopsis* Somatic Embryogenesis through Changes in Auxin Responses. *Plants*, 1*0*(12), 2720.

Gammoudi, N., Pedro, T. S., Ferchichi, A., and Gisbert, C. (2018). Improvement of regeneration in pepper: a recalcitrant species. *In Vitro Cellular and Developmental Biology - Plant*, *54*(2), 145–153.

Huang, W., He, Y., Yang, L., Lu, C., Zhu, Y., Sun, C., Ma, D., and Yin, J. (2021). Genome-wide analysis of growth-regulating factors (GRFs) in Triticum aestivum. *PeerJ*, *9*, e10701.

Jones, T., Lowe, K., Hoerster, G., Anand, A., Wu, E., Wang, N., Arling, M., Lenderts, B., and Gordon-Kamm, W. (2019). Maize Transformation Using the Morphogenic Genes . *Methods in Molecular Biology*, *1864*, 81–93.

Khatun, K., Hasan, A., Robin, K., Park, J.-I., Kumar Nath, U., Kim, C. K., Lim, K.-B., Nou, S., and Chung, M.-Y. (2017). Molecular Characterization and Expression Profiling of Tomato GRF Transcription Factor Family Genes in Response to Abiotic Stresses and Phytohormones. *International Journal of Molecular Sciences*, *18*(5), 1056.

Kong, J., Martin-Ortigosa, S., Finer, J., Orchard, N., Gunadi, A., Batts, L. A., Thakare, D., Rush, B., Schmitz, O., Stuiver, M., Olhoft, P., and Pacheco-Villalobos, D. (2020). Overexpression of the Transcription Factor GROWTH-REGULATING FACTOR5 Improves Transformation of Dicot and Monocot Species. *Frontiers in Plant Science*, *11*.

Li, S., Gao, F., Xie, K., Zeng, X., Cao, Y., Zeng, J., He, Z., Ren, Y., Li, W., Deng, Q., Wang, S., Zheng, A., Zhu, J., Liu, H., Wang, L., and Li, P. (2016). The OsmiR396c-OsGRF4-OsGIF1 regulatory module determines grain size and yield in rice. *Plant Biotechnology Journal*, *14*, 2134–2146.

Ma, J. Q., Jian, H. J., Yang, B., Lu, K., Zhang, A. X., Liu, P., and Li, J. N. (2017). Genome-wide analysis and expression profiling of the GRF gene family in oilseed rape *(Brassica napus* L.). *Gene*, *620*, 36–45.

Omidbakhshfard, M. A., Proost, S., Fujikura, U., and Mueller-Roeber, B. (2015). Growth-Regulating Factors (GRFs): A Small Transcription Factor Family with Important Functions in Plant Biology. *Molecular Plant*, *8*(7), 998–1010.

Rodriguez, R. E., Ercoli, M. F., Debernardi, J. M., and Palatnik, J. F. (2016). Growth-Regulating Factors, A Transcription Factor Family Regulating More than Just Plant Growth. *Plant Transcription Factors: Evolutionary, Structural and Functional Aspects*, 269–280.

Rodriguez, R. E., Mecchia, M. A., Debernardi, J. M., Schommer, C., Weigel, D., and Palatnik, J. F. (2010). Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396. *Development*, *137*(1), 103–112.

Rosenquist, M., Alsterfjord, M., Larsson, C., and Sommarin, M. (2001). Data Mining the *Arabidopsis* Genome Reveals Fifteen 14-3-3 Genes. Expression Is Demonstrated for Two out of Five Novel Genes 1. *Plant Physiology*, *127*(1), 142-149.

Stasolla, C., and Yeung, E. C. (2003). Recent advances in conifer somatic embryogenesis: improving somatic embryo quality. In *Plant Cell, Tissue and Organ Culture*, *74*, 15-25.

van der Knaap, E., Kim, J. H., and Kende, H. (2000). A Novel Gibberellin-Induced Gene from Rice and Its Potential Regulatory Role in Stem Growth. *Plant Physiology*, *122*(3), 695.

Wang, F., Qiu, N., Ding, Q., Li, J., Zhang, Y., Li, H., and Gao, J. (2014). Genome-wide identification and analysis of the growth-regulating factor family in Chinese cabbage (*Brassica rapa* L. ssp. pekinensis). *BMC Genomics*, *15*(1), 1–12.

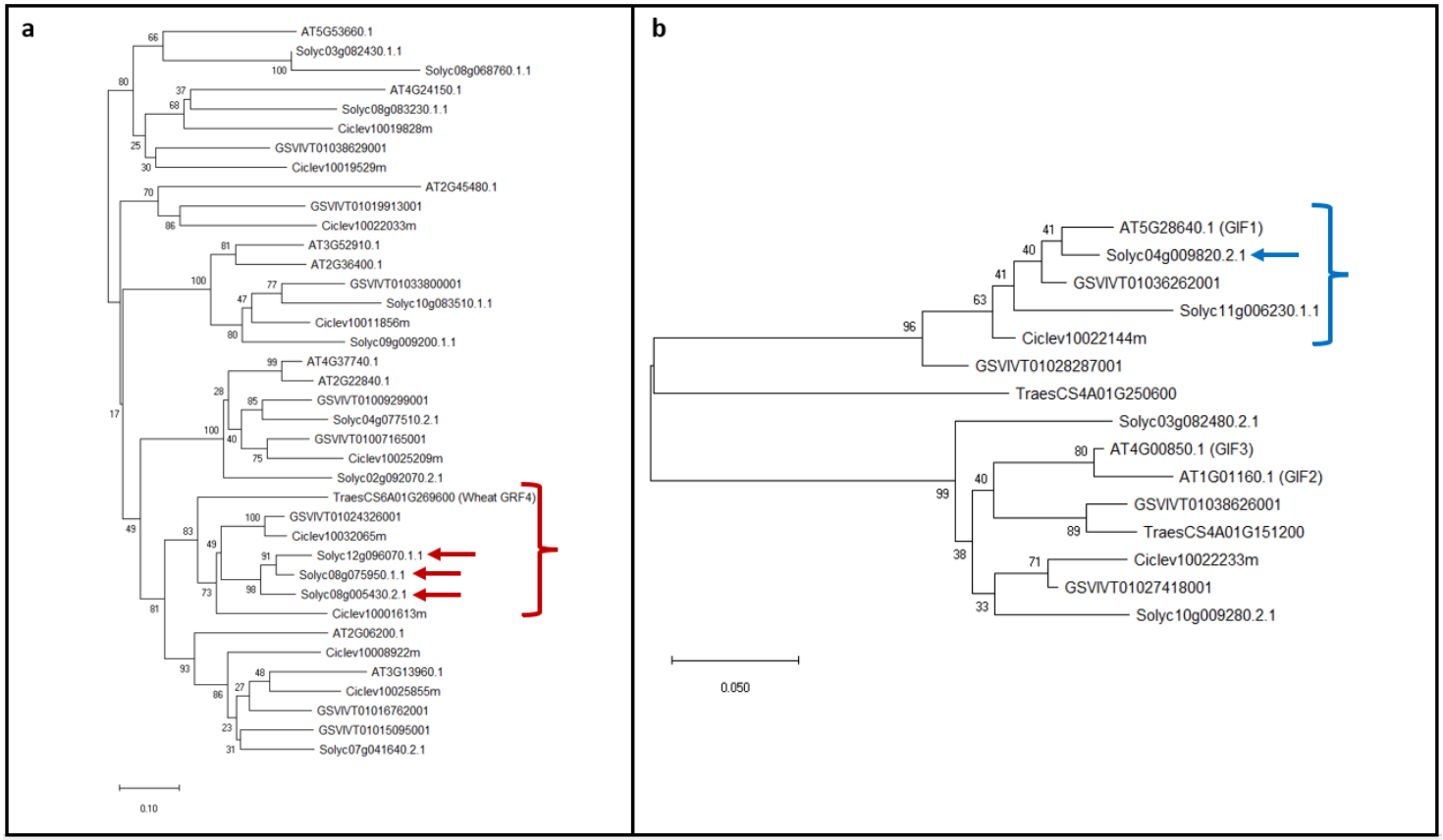
Wu, J., Zhang, X., Nie, Y., Jin, S., and Liang, S. (2004). Factors Affecting Somatic Embryogenesis and Plant Regeneration From a Range of Recalcitrant Genotypes of Chinese Cottons (*Gossypium Hirtsutum* L.). *In Vitro Cellular and Developmental Biology - Plants*, *40*, 371–375.

Zhang, B., Tong, Y., Luo, K., Zhai, Z., Liu, X., Shi, Z., Zhang, D., and Li, D. (2021). Identification of GROWTH-REGULATING FACTOR transcription factors in lettuce (*Lactuca sativa*) genome and functional analysis of *LsaGRF5* in leaf size regulation. *BMC Plant Biology*, *21*, 485.

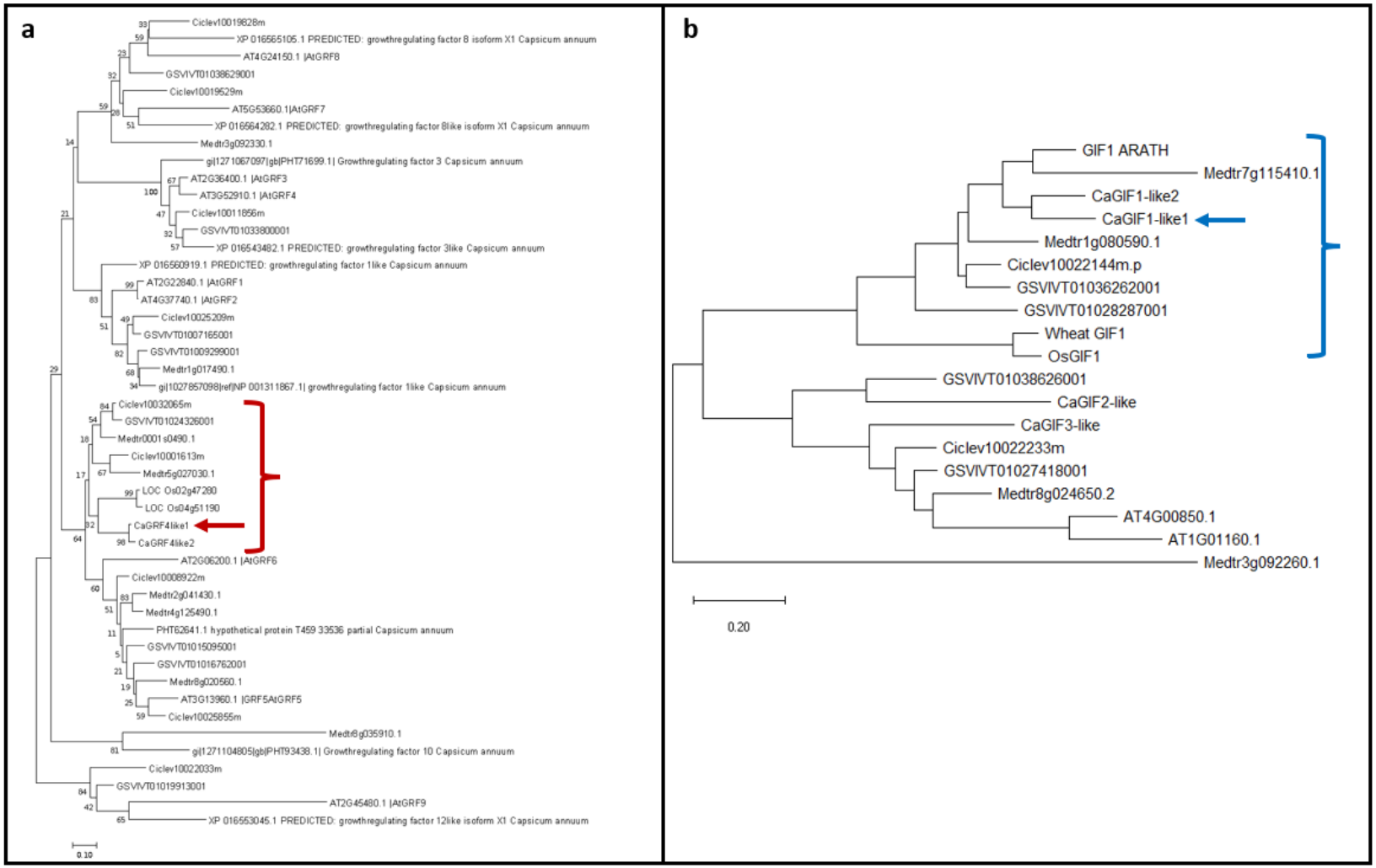
Zhang, D. F., Li, B., Jia, G. Q., Zhang, T. F., Dai, J. R., Li, J. S., and Wang, S. C. (2008). Isolation and characterization of genes encoding GRF transcription factors and GIF transcriptional coactivators in Maize (*Zea mays* L.). *Plant Science*, *175*(6), 809–817.

Zheng, Q., Zheng, Y., and Perry, S. E. (2013). AGAMOUS-Like15 Promotes Somatic Embryogenesis in *Arabidopsis* and Soybean in Part by the Control of Ethylene Biosynthesis and Response. *Plant Physiology*, *161*(4), 2113–2127.

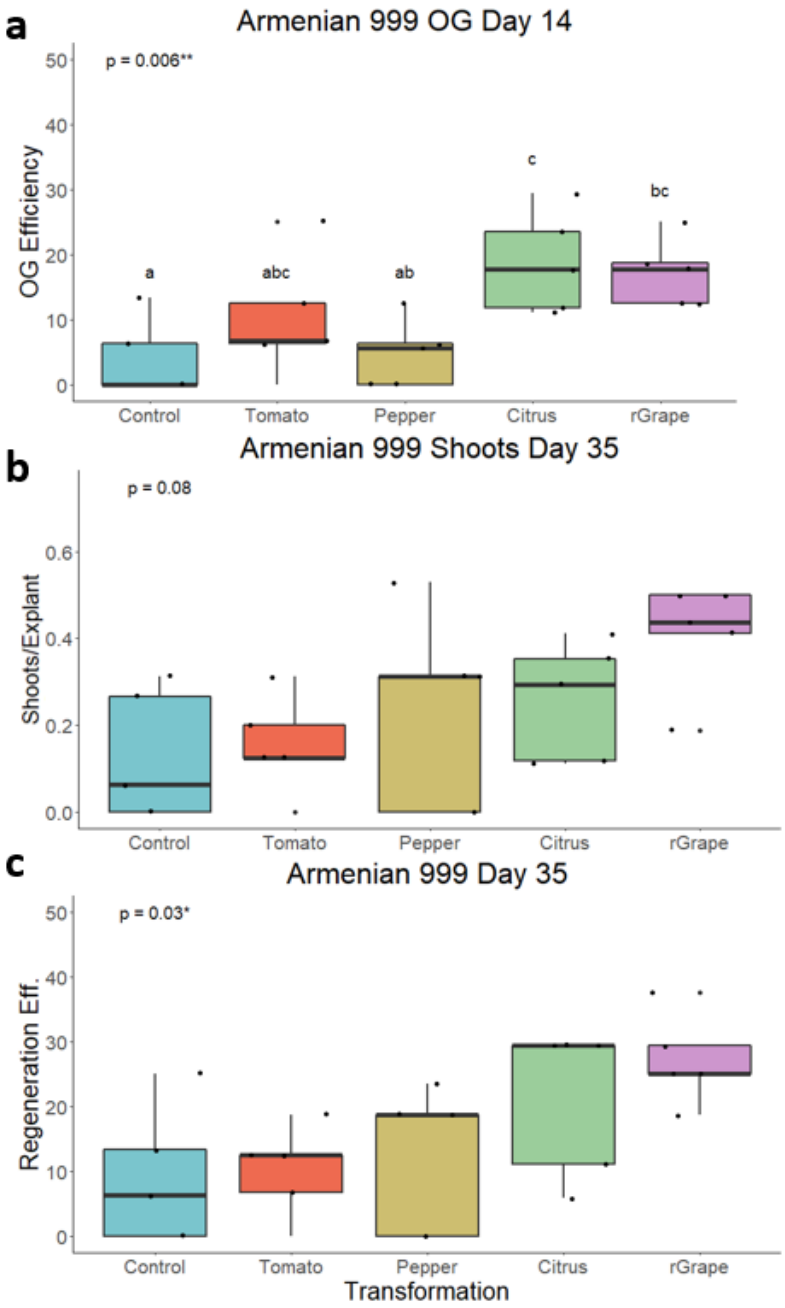
SUPPLEMENTARY FIGURES AND TABLES



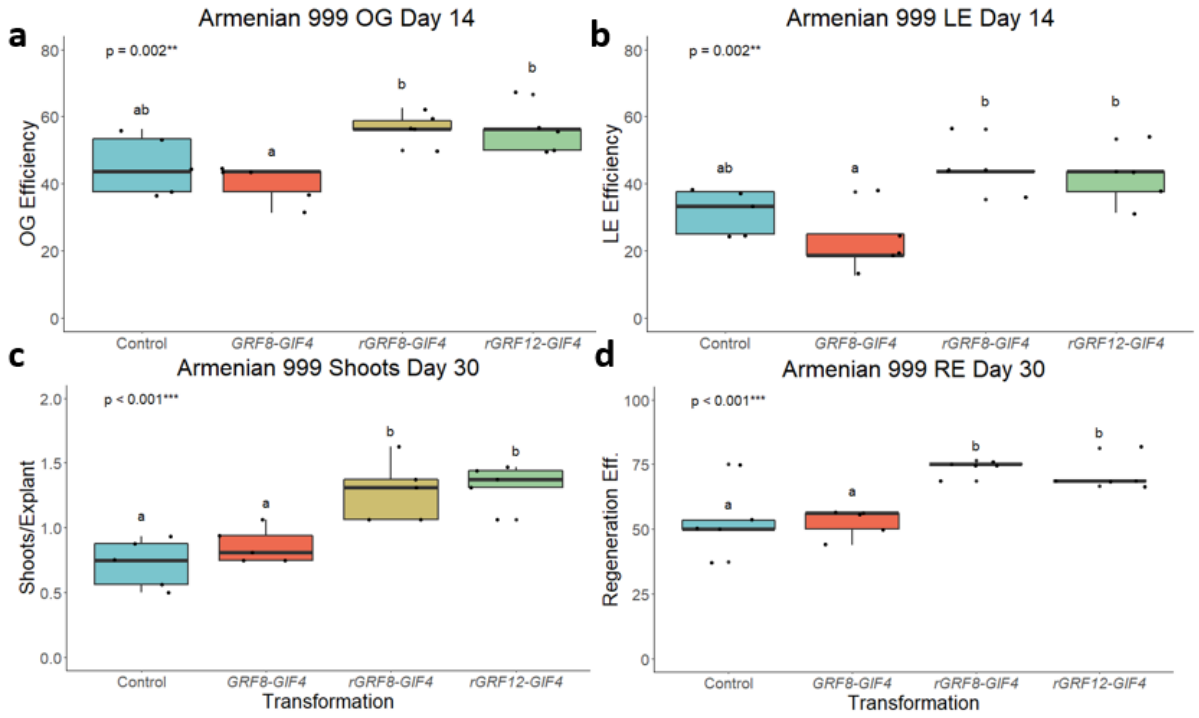
Supplementary Figure S4.1. Evolutionary relationships of *GRF* and *GIF* genes in analyzed taxa for identification of tomato *GRF* and *GIF* genes. **a**) The phylogenetic tree of GRF genes. The red bracket and arrows represents the tomato *GRF* genes, *GRF8* and *GRF12*, most closely related to the wheat and grape *GRF4* genes previously used. **b**) The phylogenetic tree of *GIF* genes used to identify the the closest tomato relative of the *Arabidopsis GIF1*. The blue bracket and arrow represents the tomato *GIF* gene, *GIF4*, selected for fusion constructs, and its closest relatives. These phylogenetic trees were produced by J. Derbernardi.



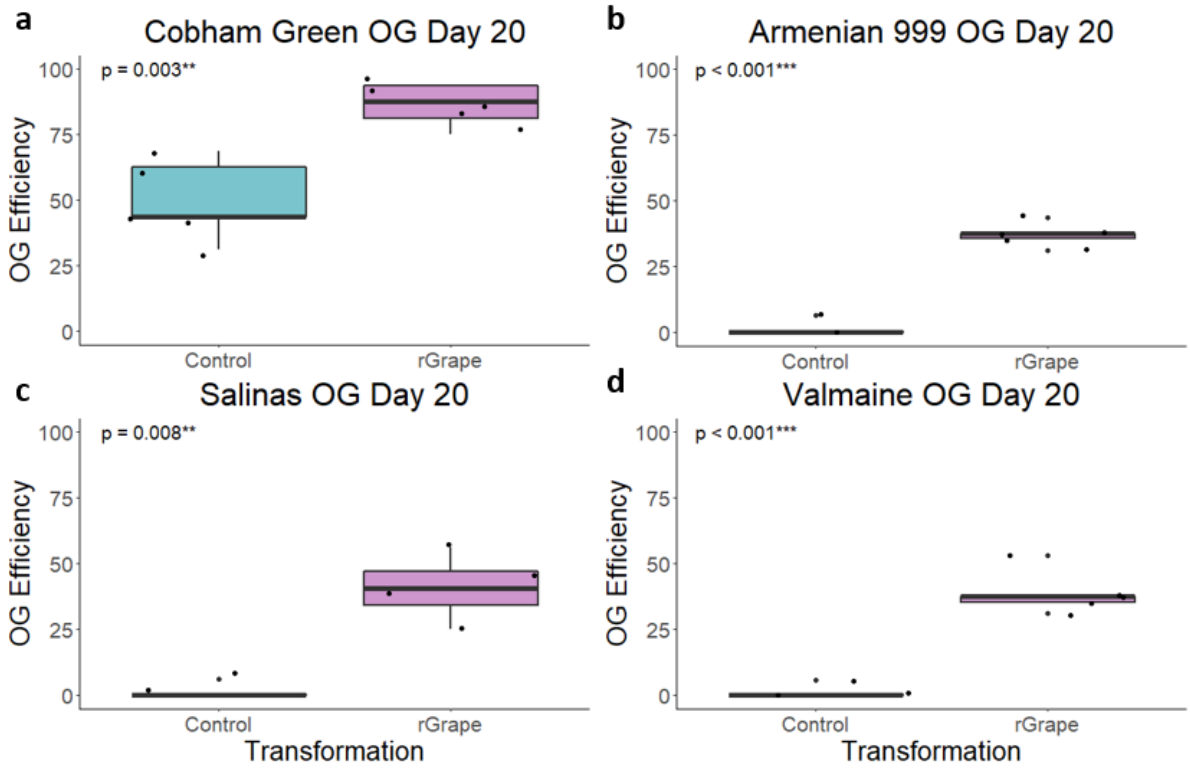
Supplementary Figure S4.2. Evolutionary relationships of *GRF* and *GIF* genes in analyzed taxa for identification of pepper *GRF* and *GIF* genes. **a**) The phylogenetic tree of *GRF* genes. The red bracket and arrow represents the pepper GRF gene most closely related to the grape *GRF4* previously used. **b**) The phylogenetic tree of *GIF* genes used to identify the the closest pepper relative of the *Arabidopsis GIF1*. The blue bracket and arrow represents the pepper *GIF* gene, *GIF1*-like1, selected for fusion constructs, and its closest relatives. These phylogenetic trees were produced by T. Hill.



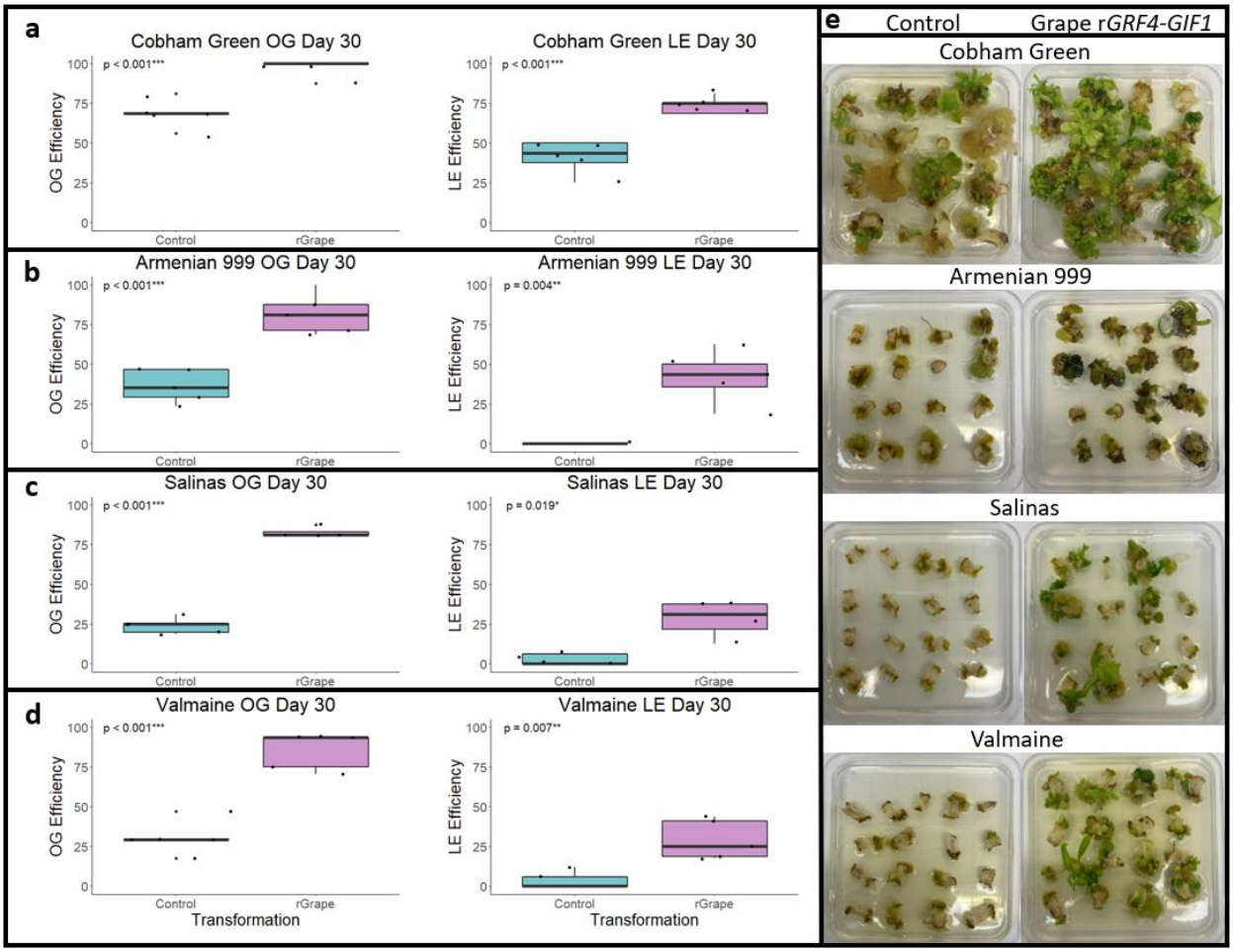
Supplementary Figure S4.3. Regeneration rates of Armenian 999 after transformation with *GRF-GIF* gene fusions from tomato, pepper, citrus, and grape. **a**) A boxplot representing frequencies of organized growth after 14 days on induction on medium; **b-c**) boxplots representing shooting freqeuncy (b) and regeneration efficiency (c) after 35 days on induction medium. Letters above each boxplot represent pairwise significance differences (TukeyHSD, $\alpha = 0.05$) and p-values were calculated using a one-way ANOVA ($\alpha = 0.05$).



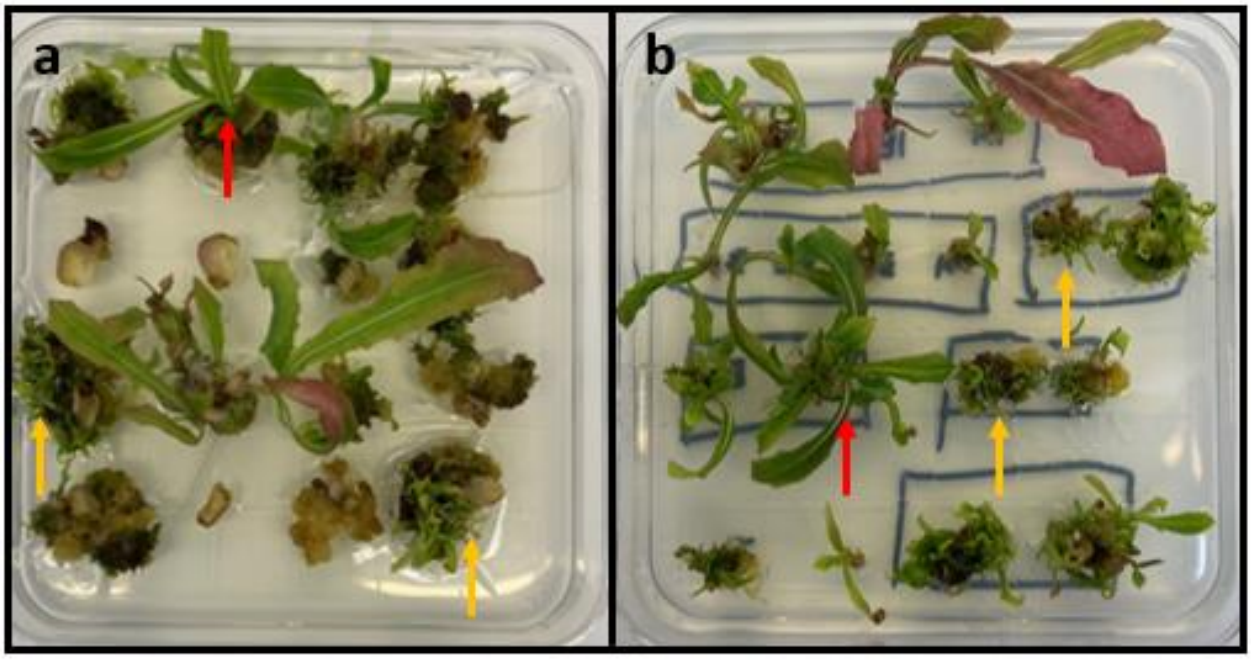
Supplementary Figure S4.4. Regeneration rates of Armenian 999 after introduction of the wildtype and miR396 resistant tomato *GRF-GIF* fusions. **a-d**) Boxplots representing efficienice of organized growth (a) and leaf emergernce (b) at day 14, and shooting frequency (c) and regeneration efficiency (d) at day 30. Letters above each boxplot represent pairwise significance differences (TukeyHSD, $\alpha = 0.05$) and p-values were calculated using a one-way ANOVA ($\alpha = 0.05$).



Supplementary Figure S4.5. Regeneration rates of Cobham Green (a), Armenian 999 (b), Salinas (c), and Valmaine (d) after introduction of the grape r*GRF4-GIF1* or an empty vector control. Boxplots represent organized growth frequencies after 20 days on induction medium. P-values were calculated using a Welch's t-test ($\alpha = 0.05$).



Supplementary Figure S4.6. Regeneration rates of different lettuce genotypes after introduction of the grape r*GRF4-GIF1* or an empty vector control. **a-d**) Boxplots represent shooting frequencies and regeneration efficiency for each transformation of Cobham Green (a), Armenian 999 (b), Salinas (c), Valmaine (d) after 30 days on induction medium. P-values were calculated using a Welch's t-test. **e**) Thirty-day old tissue cultures of each genotype after introduction of an empty vector control (left) and the grape r*GRF4-GIF1* fusion (right).



Supplementary Figure S4.7. Abnormal regeneration phenotype observed in Armenian 999 cultures after transformation with the grape r*GRF4-GIF1*. Red arrows represent normal shoot regeneration and yellow arrows represent abnormal shoot regeneration. a) Armenian 999 culture after 40 days on induction medium. b) Armenian 999 culture after 6 days on elongation medium (46 total days in culture).

Supplementary Table S4.1. Groups of transformations with *GRF-GIF* performed with construct name, construct components, and lettuce genotype. Blue highlighted transformations were performed to identify which GRF-GIF construct resulted in the highest increase in regeneration in lettuce. Orange highlighted boxes were performed to observe if miRNA396 resistant GRF-GIFs increased regeneration when compared to wildtype GRF-GIFs. The green highlighted transformations were performed to identify if the introduction of a GRF-GIF result in genotype independent regeneration in lettuce. Transformations were performed in Cobham Green (CG), Armenian 999, Valmaine, and Salinas.

Construct	Vector Backbone	Components	Genotype(s)	Selection
pTB005	pEG100	Empty Vector	CG, Armenian 999	BASTA
JD761	pEG100	Tomato *GRF8-GIF4*	CG, Armenian 999	BASTA
JD638	pEG100	Grape r*GRF4-GIF1*	CG, Armenian 999	BASTA
JD689	pEG100	Citrus *GRF-GIF*	CG, Armenian 999	BASTA
pTH1903	pEG100	Pepper *GRF4-GIF1*	CG, Armenian 999	BASTA
JD641	pGWB14	Empty Vector	CG, Armenian 999	Kan
JD746	pGWB14	Tomato *GRF8-GIF4*	CG, Armenian 999	Kan
JD747	pGWB14	Tomato r*GRF8-GIF4*	CG, Armenian 999	Kan
JD749	pGWB14	Tomato r*GRF12-GIF4*	CG, Armenian 999	Kan
pTB005	pEG100	Empty Vector	CG, Armenian 999, Valmaine, Salinas	BASTA
JD638	pEG100	Grape r*GRF4-GIF1*	CG, Armenian 999, Valmaine, Salinas	BASTA

Supplementary Table S4.2. Primer names, sequences, and PCR conditions used for amplification of the transgene for each transformation.

Primer Name	Primer Target	Primer sequence	Amplicon size (bp)	Annealing Temp. (°C)	Extension Time (sec)
TB166	Grape r*GRF4-GIF1* Fw	5’ ATGGACTCGCCTTTGGACAG3’	319	59.4	45
TB167	Grape r*GRF4-GIF1* Rv	5’GACCGAGAGCTGGTTGTTGA 3’			
TB68	*NPTII* (JD641) Fw	5’ GGTGCCCTGAATGAACTCCA 3’	448	59.0	45
TB69	*NPTII* (JD641) Rv	5’ AAAAGCGGCCATTTTCCACC 3’			
TB145	*Bar* (pTB005) Fw	5’ CAGTTCCCGTGCTTGAAGC 3’	307	58.9	45
TB146	*Bar* (pTB005) Rv	5’ CGCTATCCCTGGCTCGTC 3’			
TB109	*dsRED* Fw	5’ CCGACATCCCCGACTACAAG 3’	151	58.7	30
TB85	*dsRED* Rv	5’ ACGCCGATGAACTTCACCTT 3’			

Chapter 5: Fast fluorescent titration quantification of plasmid DNA with DNA attractive magnetic nanoparticles

ACKNOWLEDGEMENTS

This chapter presents joint work with Jennifer Lien (Ting Guo Lab)

I constructed the pEG100:dsRED plasmid, performed the plasmid preparations, qPCR titration experiments, and analyzed the qPCR titration data. Jennifer Lien performed all other experiments and drafted the manuscript.

This work has been published:

Lien, J., Bull, T., Michelmore, R.W., Guo, T., (2021). Fast Fluorescence Titration Quantification of Plasmid DNA with DNA Attractive Magnetic Nanoparticles. *Analytical Chemistry*, *93*(38), 12854-12861.

ABSTRACT

Fluorescence titration using magnetic nanoparticles (FTMN) was performed as a rapid, inexpensive, and simple method to quantify the amount of fluorophore-intercalated plasmid DNA on these DNA attractive nanoparticles. Binding of the propidium iodide (PI) intercalated DNA (PI/DNA) to polyethylenimine (PEI) coated monodisperse iron oxide magnetic nanoparticles (PEI-MNs) was confirmed with transmission electron microscopy (TEM) after the two species were mixed in water for less than a minute. The amount of DNA on PEI-MNs in aqueous solution, however, could not be easily determined using direct fluorescence measurements due to strong scattering by aggregated MNs, especially at high nanoparticle concentrations. Instead, fluorescence measurements were taken immediately after the solution of PI/DNA and PEI-MN mixtures were treated with a magnet to pull the PEI-MNs out of the solution. The detected fluorescence signal of the remaining free PI/DNA in the solution decreased as the concentration of PEI-MNs in the pre-treated solutions increased, resulting in a titration curve, which was used to

determine the amount of DNA on MNs, the dissociation constant, and binding energy after the concentration of PEI-MNs was calibrated with microwave-plasma atomic emission spectroscopy (MP-AES). Quantitative PCR (qPCR) was used to understand the binding of DNA to MNs and to measure the amount of free PI/DNA in solution, and the results were similar to that obtained with the FTMN method.

INTRODCUTION

Delivery of genetic materials into living cells for genome editing often relies on coating the surface of nanomaterials with DNA for improved uptake into cells (Mehier-Humbert et al., 2005; Mykhaylyk et al., 2009; Zhao et al., 2017; Cunningham et al., 2018; Joldersma and Liu, 2018; Svoboda et al., 2018; Rohiwal et al., 2020). For efficient and quantitative delivery, it is important to quantify the amount of genetic materials on the surface of these nanomaterials. Fast and inexpensive quantification is therefore critical to many applications such as delivering CRISPR/Cas9 reagents for genome editing. Magnetic nanoparticles (MNs) as preferred delivery vehicles have many advantages, one of which is to use magnetic fields to enhance the delivery, a process also known as magnetofection or magnetoporation (Du et al., 2018; Nguyen, et al., 2019; Bi et al., 2020).

MNs have been studied for targeted drug delivery, energy storage, magnetic resonance imaging (MRI), cancer hyperthermia therapy, spintronics, fluorescent-magnetic nanocomposites for biological imaging, synthetic pigments in ceramics, and capture and degradation of pollutants (Kudr et al., 2017). Their applications in analytical chemistry include sensing, magnetic separation, capture, and pre-concentration (Lakshmanan et al., 2014; Neto et al., 2020; Yang et al., 2020). Magnetic decantation, where an analyte can be separated from the solution by applying an external magnetic field, is another important application (Berensmeier, 2006). Many of these applications depend on the modification of the surface of nanomaterials, and a variety of ligands can be used to produce a surface with reactive functional groups of $-NH_2$, -COOH, and -SH (Zhu et al., 2018; Song et al., 2019). Another type of functionalization is encapsulation, usually by polymeric or inorganic materials (ex. SiO_2, Al_2O_3, Au) (Kang et al., 2009; Shan et al., 2015; Sheng et al., 2016). As a result, the use of external magnetic fields and surface modification helps enhance the

capability of MNs to deliver genetic materials. Two general approaches have been reported in the literature to facilitate such delivery, one using plasmid DNA (pDNA) conjugated to MNs for transfection and the other using a mixture of pDNA and MNs (Wang et al., 2009; Zhang et al., 2014; Rohiwas et al., 2020). Both approaches showed successful transfection.

Despite these advancements, analytical methods are still needed to quantitatively determine the amount of genetic material on the surface of these nanoparticles, for such quantification is not readily available. Fluorescein (FITC) and other fluorophores or even nanomaterials such as gold nanoparticles on the surface of MNs were measured, though without calibration and hence only offering relative quantity information (Bouzas-Ramos et al., 2018; Hu et al., 2018). One difficulty of directly detecting fluorescent molecules on MNs in solution originates from the strong scattering of the detection light by these nanoparticles or their aggregates. Although other methods, such as gel electrophoresis, surface enhanced Raman spectroscopy (SERS), electrochemistry, mass spectrometry, and even direct colorimetry, have been studied, some of the results lack quantitative information while others usually require large quantities of samples and specially synthesized samples (Kouassi, et al. 2006; Lin et al., 2007; Zhang et al., 2009; Liu et al., 2014; Wang et al., 2014; Rong et al., 2016).

In this chapter, we demonstrated a method that takes advantage of the intrinsic properties and capabilities of MNs and their movement in magnetic fields to achieve fluorescence titration using magnetic nanoparticles (FTMN) for rapid quantification of the content on MNs (Shan et al., 2015). This titration study also revealed information about the binding between pDNA and MNs

MEHODS

Chemicals

Polyethylenimine (PEI), branched average Mw 25,000 (PEI) (≤1% water), ferric chloride hexahydrate (puriss. p.a., ACS reagent, crystallized, 98.0-102% (Redox Titration assay), nitric acid (ACS reagent, 68-70%), hydrochloric acid (ACS reagent, 37%), sodium hydroxide pellets (certified ACS), ethylene glycol (EG) (ReagentPlus®, ≥99%), 1,2-ethylenediamine (ReagentPlus®, ≥99%), trisodium citrate dihydrate (ACS reagent, ≥99.0%), sodium acetate (anhydrous, ReagentPlus®, ≥99.0%), and ammonia (ACS Reagent, 28.0–30.0% NH_3 basis) were purchased from Sigma-Aldrich and used as received. Ethanol (200-proof, Koptec) and Milli-Q (MQ) water with a resistivity of 18.2 MΩ·cm passed through a 0.22 μm filter were used.

MN synthesis

Twenty-six nanometer diameter magnetic Fe_3O_4 nanoparticles (MNs) were prepared as follows: In brief, 0.5 g NaOH was dissolved in 10 mL of ethylene glycol by vortexing. $FeCl_3 \cdot 6H_2O$ (1 g) was transferred into a 100 mL flask to which 3 g of sodium acetate was added, followed by the addition of 10 mL of 1,2-ethylenediamine. While stirring the solution, the NaOH/EG solution was added, and the mixture was continually stirred for 1 hour, after which the contents were transferred to a Teflon-lined stainless-steel autoclave reactor and sealed. The reactor was heated to and maintained at 200°C for 8 hours and was removed from the oven and cooled to room temperature. The MNs were magnetically purified 3×with redispersion in 20 mL of MQ water. The particles were examined with transmission electron microscopy (TEM). An aliquot of the purified particles was diluted 50-fold and a magnet was used to remove the fast responding particles. This magnetic removal and redispersion was repeated several times per aliquot to

completely remove the larger particles. The final size distribution was confirmed by TEM and dynamic light scattering (DLS) and the particles were dried under vacuum at room temperature. The yield of nanoparticles as the final product was determined to be 7.3% by mass.

Surface amination of MNs

For a typical functionalization, 1.3 mg of PEI was dissolved in 250 μL of MQ water in an Eppendorf tube to which 0.7 mg of MNs was added. The solution was vortexed and then sonicated for 1 hour. The PEI-MNs were purified magnetically 3×, redispersing the pellet in 250 μL of water each time. The final sample was redispersed in 1 mL MQ water, and 20 μL of this stock was diluted to 1 mL for DLS measurements. The resulting PEI-MNs were used in this work.

Nanoparticle size measurements and magnetic separation experiments

DLS was used to determine the nanoparticle size by its hydrodynamic radius after surface functionalization. A Malvern Zetasizer (ZEN1690) was used for DLS measurements. After the first DLS measurement, a square magnet (4500 Gauss) was placed next to the cuvette to pull the particles in a MQ water solution for 2 minutes, after which the supernatant was transferred to a new cuvette for a second DLS measurement and the pellet was redispersed in 100 μL of MQ water to form the 2-minute magnetic separation nanoparticle stock. This magnetic separation/DLS measurement process was repeated four more times, with 4, 6, 8, and 30-minute magnetic separation times. This purification process generated five nanoparticle stock samples; the nanoparticles that were magnetically separated out at 2 minutes (fast) and 30 minutes (slow) were used for titration experiments. Both slow and fast PEI-MNs produced the fluorescence titration curve, but this paper focuses on the fast particles. The nanoparticle solutions were analyzed with Microwave Plasma-Atomic Emission Spectroscopy (MP-AES).

MP-AES

MP-AES was used to determine the iron concentration so that the nanoparticle concentration could be calculated prior to incubation with plasmid DNA. Standard solutions containing 0.05, 0.1, 0.5, 1, 5, and 10 ppm Fe were prepared. Each analyte sample was prepared by dissolving 5 μL of the PEI_MN stock solution with 20 μL of freshly prepared aqua regia. After 30 min dissolution, the solution was diluted to 1 mL with MQ water to produce the final MP-AES analyte solution. Measurements were acquired using a 4200 MP-AES equipped with a standard torch, nebulizer, and spray chamber. Analysis was performed at 371.993 nm (Fe), and the measurement was taken in three replicates at a pump speed of 15 rpm. The measurement results of the standard showed linearity between 0.05 to 10 ppm. The nanoparticle concentration was calculated to be 1.6 nM.

Sample preparation for fluorescence measurements

Propidium iodide (PI) was first intercalated into plasmid DNA and then PI/DNA was incubated with PEI-MNs. A typical PI/DNA-PEI-MN sample was prepared as follows. The DNA stock solution contained 790 ng/μL LMV-0-HC-GFP (14949 bps) plasmid DNA. In an Eppendorf tube, 1 μL of 0.045 mM PI was added to 1 μL of the DNA stock solution and the solution was diluted 50-fold. The resulting 100 μL solution was placed on a shaker overnight to allow PI to intercalate in DNA. Next, free PI was purified away from PI intercalated in plasmid DNA (PI/DNA) using a 30K MWCO 0.5 mL Amicon Ultra centrifugal filter unit, which was wetted prior to use. Purification of PI/DNA was performed once using centrifugation at 5,500 rpm for 10 minutes. Supernatants were decanted and the final concentrations of PI/DNA were adjusted to 112.5 nM/0.2 nM by diluting the purified PI/DNA in MQ water. The purified PI/DNA solution was

used for incubation with purified PEI-MNs, and the concentrations of PI/DNA during incubation were 28 nM/0.05 nM, respectively. PEI-MNs were diluted to prepare the DNA-MNs solutions. A typical set of samples prepared included 5, 8, 10, 12.5, 15, 16, 17, 20, 25, 30, 40, and 50× dilution factors of nanoparticles (abbreviated "×d"). For example, the 5×d PEI-MNs sample was prepared by adding 70 μL of 1 nM PEI-MNs to 192 μL MQ water, vortexing, and then adding 88 μL of 0.2 nM PI/DNA to produce a final PEI-MN concentration of 0.2 nM. All the samples were vortexed for 30 seconds, and then used in the fluorescence titration and magnetic separation experiments.

Fluorescence measurements

Fluorescence measurements were used to confirm DNA binding to MNs. PI was selected for its binding to DNA through intercalation, which does not interfere with DNA binding to MNs. The prepared solutions of PI/DNA on PEI-MNs were transferred to a micro fluorescence quartz cuvette (Hellma). Fluorescence spectra were acquired before and after pulling the PEI-MNs with a square magnet (4500 Gauss). Emission spectra were taken with an excitation wavelength of 535 nm and fluorescence was detected from 575 nm to 700 nm at a 1 nm increment and 0.5 s integration time. Three scans were taken to obtain the average. A blank was measured in triplicate, and its signal was subtracted from each spectrum. The sample uniformity and DNA loading were confirmed via TEM, which was taken for as-purified PEI-MNs and PI/DNA-PEI-MNs stained with 4% uranyl acetate.

Deconvolution of fluorescence and scattering

The DNA-MNs spectra had a strong scattering peak at shorter wavelengths and began to gently slope down at longer wavelengths. This scattering peak (λ_{max} 578 nm) was deconvoluted from the fluorescence peak (λ_{max} 615 nm) as follows. Three wavelengths, 580, 615, and 675 nm,

were chosen and the fluorescence signal intensity ratios for pure PI-DNA at these three wavelengths were determined to be 0.7:1:0.3 (580 nm:615 nm:675 nm). Similarly, the scattering intensity ratios for pure PEI-MNs were determined to be 1.6:1:0.6. Using these two sets of signal ratios as the basis sets, the unknown decimal percentage of fluorescence x or the scattering decimal percentage 1-x for any measured fluorescence spectrum can be solved analytically and the spectrum can be deconvoluted. A more detailed description is given in SI.

Sample preparation for qPCR experiments

DNA plasmids were prepared using the manufacturer's protocol for the Qiagen QiaPrep Spin Miniprep Kit. A 0.2 nM stock of pBS70-LMV-0-GFPHC (14,949 bps, noted as LMV here) was used to prepare DNA standard solutions containing a series of known amounts of DNA (0, 0.15, 0.35, 1.03, 3.1, and 9.2 ng). The DNA-MN samples for qPCR analysis were prepared as described in the fluorescence section. For the competitive titration and binding/exchange rate qPCR experiments, pEG100 dsRED (10,678 bps, noted as dsRED here) was prepared and used as the inactive DNA component. In brief, a series of samples were prepared containing a constant total amount of the two plasmid DNA, but varying the ratio of the two. The nanoparticle dilution factor was also held constant at 40×d in all samples. Typical solutions were prepared to give 40×d PEI-MNs with 100/0, 75/25, 50/50, 25/75, and 0%/100%, % dsRED/%LMV. Positive and negative DNA only control samples were also prepared of dsRED and LMV. For the binding/exchange rate experiments, the active LMV and inactive dsRED were first mixed for 5 minutes. Then the PEI-MNs were added along with an amount of MQ water to bring the final nanoparticle dilution factor to 40×. The solution was mixed for another 5 minutes, after which a magnet was used to pull the NPs aside. The supernatant was analyzed by qPCR.

Real-time quantitative PCR (qPCR)

Quantitative PCR (qPCR) was conducted to confirm the binding of DNA to the magnetic nanoparticles and to quantify the DNA. Primers (forward primer 5'-TAAAAGGACAGGGCCATCGC-3' and reverse primer 5'-AAACATCCTCGGCCACAAGT-3') for detection of the plasmid pBS70-LMV-0-GFPHC were designed using Geneious Prime software (German-Retana 2003). A conventional PCR reaction (GoTaq Green Master Mix) was performed to test the efficiency of the primers. The qPCR was performed using the PowerUp™ SYBR™ Green Master Mix protocol from Applied Biosystems (Thermo Fisher Scientific). Each reaction consisted of 12 µL containing 6 µL of the PowerUp™ SYBR™ Green Master Mix (2X), 0.6 µL of forward and reverse primers with 10 µM stock concentration, and 4.8 µL of the respective samples. The qPCR was run on the Applied Biosystems 7900HT Real-Time PCR thermocycler at the UC Davis Veterinary Medicine PCR Laboratory. Cycling conditions were run under the standard, three-step cycling mode with the following cycling conditions: initial steps of UDG activation (50°C for 2 minutes) and Dual-Lock™ DNA polymerase (95°C for 2 minutes) followed by 40 cycles of denaturation (95°C for 15 seconds), annealing (58°C for 15 seconds), and extension (72°C for 1 minute). Four replicates were run for each sample. Cycle threshold (Ct) values were used to calculate the quantity of DNA in each sample using a standard curve generated from a qPCR of six known DNA concentrations, of which four replicates were also measured.

RESULTS

MNs and their DNA conjugates obtained through synthesis and purification were examined with TEM and DLS. TEM results (**Figures 5.1**) showed a nearly monodispersed 26±5.8 nm MNs (**Figure 5.1**) after PEI functionalization and the 1st round of magnetic purification. The size was

similar after the 5th purification (**Supplementary Figure S5.1**). This size was chosen so that MNs can enter plant cells through their membranes (Wang et al., 2016). However, DLS data of PEI modified MNs (PEI-MNs) in aqueous solutions before the first purification indicated large sizes (350 nm diameter) (**Figure 5.1**) and the size decreased to a diameter of 160 nm (**Figure 5.1**) after the 5th round of purification. The coverage of PEI on these MNs was considered to be maximum because of the 10× excess of PEI used during incubation. PEI-MNs after the 1st and 5th round of magnetic purification were both used, with the main difference between the two being the magnetic pulling speed of these particles in cuvettes by magnets. The large DLS sizes of PEI-MNs in solution were considered to be caused by aggregation of PEI-MNs in solution, and this disparity between TEM and DLS measurements of MNs has been reported in the literature (Arsianti et al., 2010; Ota et al., 2013; Xie et al., 2013; Wang et al., 2014; Zhang et al., 2014; Xie et al., 2015). PEI-MNs free of aggregation without DNA were possible, although it was not the case here (Cruz-Acuna et al., 2016; Narayanasamy et al., 2018). Given that 19% of individual MNs were larger than 30 nm (**Figure 5.1**) and MNs as small as 30 nm diameter are ferrimagnetic (Zhang et al., 2020), it is understandable that aggregation would occur without external magnetic fields (Wang et al., 2009; Shen et al., 2018). From these studies, it was also clear that PI/DNA were attracted to PEI-MNs, which themselves aggregated in solution to form larger than 150 nm diameter particles.

Fluorescence from PI, PI/DNA, and PEI-MNs coated with PI/DNA were measured. Fluorescence yield increased by over 50 fold from free PI to PI/DNA (**Supplementary Figure S5.2**), which is similar to values published in the literature (Ren et al., 2000; Zhao et al., 2010). Under the conditions of 560:1 of PI:DNA, each 15-kb DNA was found to have over 500 PI per DNA. After mixing PEI-MNs and PI/DNA for one minute, the sample was examined with TEM (top panel of **Figure 5.2**), which showed DNA (stained with uranyl acetate) on the surface of PEI-

MNs. The lower panel of **Figure 5.2** shows free PI/DNA stained with uranyl acetate. The footprint of the DNA is on the order of tens of nm or greater, consistent with the literature, although fully relaxed plasmid DNA can be quite larger (Song et al., 2017).

The fluorescence of mixtures of PI/DNA on PEI-MNs was measured and found to be attenuated by PEI-MNs, which was more severe in solutions of high concentrations of PEI-MNs due to strong scattering from aggregated MNs with diameters ranging from 150–350 nm. In order to detect PI/DNA on PEI-MNs in these samples, PEI-MNs were pulled from the solution by a magnet (**Figure 5.2**), as reported in a previous study, so that unbound PI/DNA was detected without interference from NMs (Schaller et al., 2008). PEI-MNs were pulled for up to five minutes, after which the solution was largely free of 26 nm PEI-MNs or their aggregates, although it is possible that smaller PEI-MNs remained in the solution. Little fluorescence was detected before (black line in the top panel of **Figure 5.2**) and after (light purple dashed line in the same panel) pulling PEI-MNs from solutions of high concentrations of PEI-MNs (5× dilution or 5×d); this lack of fluorescence was caused by strong scattering by MN aggregates before pulling, and nearly all the PI/DNA being removed from the solution after pulling, respectively. The latter also demonstrated strong DNA-MN attraction. Without MNs, PI in 0.05 nM PI/DNA would generate a fluorescence signal of approximately 30,000 cps fluorescence signal. At 16×d, fluorescence was detectable before pulling (black line). The measured signal originated from the fluorescence of PI/DNA, peaked at 615 nm with a 40-nm FWHM, and the scattering from PEI-MNs, a relatively flat monotonic decay profile. The fluorescence profile resembles that of pure PI/DNA after pulling (dashed light purple line). The signal of the difference (solid blue line in the middle panel (16×d) of **Figure 5.2**, middle panel) between fluorescence signals from samples before and after pulling contained weak fluorescence from PI/DNA. This means that fluorescence of PI/DNA in the

aggregated MNs can be detected under this condition. As the PEI-MN concentration decreased further, fluorescence of PI in DNA became more visible. At 50×d, scattering was less than 10% of the measured signal and fluorescence was readily detected (**Figure 5.2**, bottom panel). As shown below, most PI/DNA was free in the 50×d solution, whereas all of the PI/DNA was bound to PEI-MNs in the 5×d solution. These results suggest that it is difficult to directly measure the amount of DNA on PEI-MNs due to strong scattering from these nanoparticles and their aggregates.

When PI/DNA was incubated with PEG-MNs, most PI/DNA was considered to be dissolved in solution and not on the surface of MNs because PEG does not attract DNA as strongly as PEI ligands. Fluorescence measurements revealed detectable fluorescence even without magnet pulling. After pulling, over 85% of PI/DNA was in the solution even at high MN concentrations (**Supplementary Figure S5.4**), which was different from the PEI-treated MNs in which PI/DNA was not detected in the solution.

The stock concentration of PEI-MNs was determined using MP-AES. The concentrations of PEI-MNs used in the titration experiment were a series of dilutions of the stock solution. The stock solution of PEI-MNs was determined to contain 27 ppm Fe. The concentration of 26 nm MNs was 1.60 nM in the stock solution (equivalent to 1×d). The concentration of PEI-MNs was 0.032 nM at 50× dilution of the stock solution.

To quantitatively determine the amount of PI/DNA on PEI-MNs, fluorescence as a function of concentration of monodisperse PEI-MNs was detected after the nanoparticles were pulled from the solution by the magnet, hence constituting a titration process (**Figure 5.3**). In this case, PI/DNA was the analyte and PEI-MNs were the titrant. Measurements were acquired with PEI-MNs pulled to the side of the cuvette by the magnet. PEI-MNs were not physically removed from the cuvette because they stayed on the wall during each measurement, which lasted for 60 seconds.

Fluorescence of free PI/DNA in solution, i.e., unbound PI/DNA, was measured (**Figure 5.3**) after PEI-MNs were pulled aside. This approach eliminated the interference from MNs on PI fluorescence measurements, which could be significant. The results presented here show a titration curve, with strong signals at low PEI-MN concentrations and nearly no signal at high PEI-MN concentrations.

The fluorescence titration curve shown in **Figure 5.3** quantitatively explains the interactions of PI/DNA with PEI-MNs at different nanoparticle concentrations. The curve can be used to determine the amount of DNA on MNs after the MN stock solution was calibrated by MP-AES. For high dilution factors (upper X-axis), i.e., for concentrations of less than 0.05 nM (lower X-axis) of PEI-MNs (**Figure 5.3**), there were not enough PEI-MNs in the solution (<0.05 nM of 26 nm MNs) to take up all the PI/DNA (0.05 nM); therefore, after the PEI-MNs were pulled aside, nearly all the PI/DNA was still in solution, resulting in the strong fluorescence signals (>2.7×10^4 cps) corresponding to >0.045 nM PI/DNA. As the dilution factor increased to *ca.* 15× or as the concentration of PEI-MNs reached 0.07 nM, the fluorescence signal was ~1.2×10^4 cps after PEI-MNs were pulled aside. Fluorescence from PI on the pulled particles was deduced to be approximately 5×10^3 cps (**Figure 5.3**) after deconvolution. The sum of these two signals was ~1.7×10^4 cps, which is approximately half of the signal of the free PI/DNA in solution. This outcome reveals fluorescence was detectable when PI/DNA was on the surface of PEI-MNs, although the signal was less than that from the actual amount of PI/DNA on the surface, again indicating that aggregated PEI-MNs partially mask the fluorescence of PI intercalated in DNA. This is possible because the aggregation may block the light from reaching PI in the aggregates. At even lower dilution factors (<10×d) or higher PEI-MN concentrations (>0.1 nM), the PI signal largely disappeared, indicating that little PI/DNA remained in the solution and all the PI/DNA

moved away to the wall with PEI-MNs. At these MN concentrations, the deconvoluted signal difference between samples before and after magnet pulling contained no PI fluorescence. Based on the detectable fluorescence signals from PI using the differences in deconvoluted signal between samples before and after magnet pulling, the maximum amount of detectable PI/DNA on PEI-MNs occurred at approximately 0.07 nM PEI-MNs (**Figure 5.3**). This is reasonable because at higher concentrations of MNs (in the volume of solution defined by the cuvette), scattering is too strong to detect 0.05 nM PI/DNA. At lower MN concentrations, there is little PI/DNA on the MNs. The two processes together result in the measured pattern shown in **Figure 5.3**. This trend demonstrates the necessity to measure DNA on MNs using the titration method described here.

Three more insights can be inferred from the titration curve. First, PI/DNA was attracted to PEI-MNs, as pulling MNs decreased the amount of PI/DNA in solution (**Figure 5.3**). Second, the titration curve suggests DNA coverage on MNs experienced a nonlinear response. For PEI-MN concentrations of less than 0.1 nM, fluorescence signals decrease approximately linearly at a rate of 1.2 nM^{-1} (PEI-MNs). This means each PEI-MN pulls away 1.2 DNA plasmids. As the PEI-NM concentration reaches ~0.07 nM, the fluorescence signal decreases more precipitately to near zero, with a slope of nearly 2.0 nM^{-1}. This indicates that each PEI-MN pulls away approximately 2.0 DNA molecules. This nonlinear response of PI/DNA conjugated to PEI-MNs may be explained by how PEI-MNs interact in solution. DLS measurements were performed on both the PEI-MNs without DNA (black diamonds, **Figure 5.3**) and the mixture of PI/DNA and PEI-MN (red triangles, **Figure 5.3**). It is interesting to note that adding PI/DNA to PEI-MNs changed the behavior of the particles in terms of the degree of their aggregation, as the DLS size for PEI-MNs with PI/DNA decreased as a function of PEI-MN concentration. DLS sizes of PEI-MNs without DNA increased as a function of PEI-MN concentration, which was more intuitively predictable. The decrease of

DLS size was possibly due to PI/DNA binding to PEI-MNs, resulting in the reduction of attraction among PEI-MNs. The size decrease of the aggregates explains the nonlinear response near 0.07 nM because smaller aggregates can take up additional PI/DNA on their surface. Third, it is possible to obtain the dissociation constant K_d from the titration curve (**Figure 5.3**), which would be 3.0 ± 1.8 pM at the equivalence point, assuming the binding between the DNA and MNs is first order with respect to all species. Using this K_d value, the binding energy is estimated to be -66 ± 0.6 kJ/mol. qPCR was performed to investigate PI/DNA-PEI-MN binding and the results confirmed the trend found in FTMN. Another set of separation experiments, including magnet pulling, were performed, and a similar fluorescence titration curve was obtained (**Figure 5.4**), which showed a similar trend as that shown in **Figure 5.3**. In the qPCR experiments, PI/DNA solutions were treated with a magnet and the supernatants were measured using qPCR, which detects the amount of free PI/DNA in solution after magnetic separation, making it possible to directly compare qPCR results with that of titration. Calibration of qPCR was performed (**Supplementary Figure S5.5**). The results (**Figure 5.4**) were identical to that of fluorescence measurements, suggesting that both detection methods are valid to determine the amount of free DNA after magnetically separating the nanoparticles. All the curves show the same onset point of 15×d or 0.07 nM.

An advantage of qPCR over fluorescence measurements is that qPCR can differentiate DNA payloads such as between two forms of DNA of pBS70-LMV-0-GFPHC, or LMV as noted here, from pEG100 dsRED, or dsRED as noted here. This feature allows the study of competitive binding. In one experiment, LMV and dsRED were premixed and the mixtures were incubated with PEI-MNs. The experiment was similar to measuring one of the titration points, which in this case had a dilution factor of 5×. At low LMV/dsRED ratios, low levels of LMV were in the solution and there was little LMV detected, whereas at high ratios, more LMV was detected because there

was more LMV in the mixture. The active DNA LMV was detected by qPCR, which showed increased amounts of LMV as the LMV in the mixture increased (**Figure 5.4**). This suggests that there was no preferential binding between the two types of DNA molecules. Competitive binding was also tested by incubating premixed LMV and dsRED with PEI-MNs at room temperature (RT) and 40°C (**Figure 5.4**), and it is clear that binding at these two temperatures was similar, meaning that once LMV was bound to the surface, it cannot be replaced by dsRED in solution. Competitive binding was further studied by first mixing active LMV with PEI-MNs. Then inactive dsRED was mixed with the LMV-incubated PEI-MNs (**Figure 5.4**). The results show no detected LMV as a function of dsRED concentration, once again supporting that bound DNA does not leave PEI-MNs. It should be noted that these measurements cannot be performed using the fluorescence method because fluorescence cannot differentiate dsRED from LMV, whereas qPCR can detect LMV but not dsRED, and can therefore differentiate the two.

Although FTMN and qPCR can both be used to obtain the titration curve, the two have different merits. FTMN is faster and less expensive, especially when MNs are available and multiple fluorescence measurements can be taken simultaneously. qPCR, on the other hand, can differentiate active from inactive DNA, a feature that has yet to be achieved with FTMN.

DISCUSSION

The results shown here demonstrate that it is possible to quickly determine the quantity of DNA on magnetic nanoparticles even when fluorescence measurements do not directly yield the needed information due to strong scattering by the nanoparticle aggregates. Instead, fluorescence measurements can determine the amount of unbound DNA in solution after magnetic nanoparticles are pulled from the solution by a magnet. The results give rise to a titration curve as a function of

magnetic nanoparticles for a fixed amount of DNA. The titration curve can be used to determine the amount DNA on nanoparticles if the concentration of magnetic nanoparticles is known, which can be obtained through atomic emission spectroscopy (AES) or other methods. The titration curve can also be used to estimate the dissociation constant and binding energy, which was approximately -66 kJ/mol. This process of fluorescence titration of PI/DNA using PEI-MNs through the action of magnetic pulling of MNs presents a rapid, simple, and inexpensive method to determine the binding of DNA to magnetic nanoparticles. qPCR results corroborate the fluorescence titration results and qPCR can help elucidate competitive binding between different DNA molecules. The work presented here only utilized one method of attraction of strong electrostatic force between DNA and MNs. Many other methods can be used, including weak electrostatic forces, van der Waals forces, direct chemical bonding, and cleavage.

FIGURES

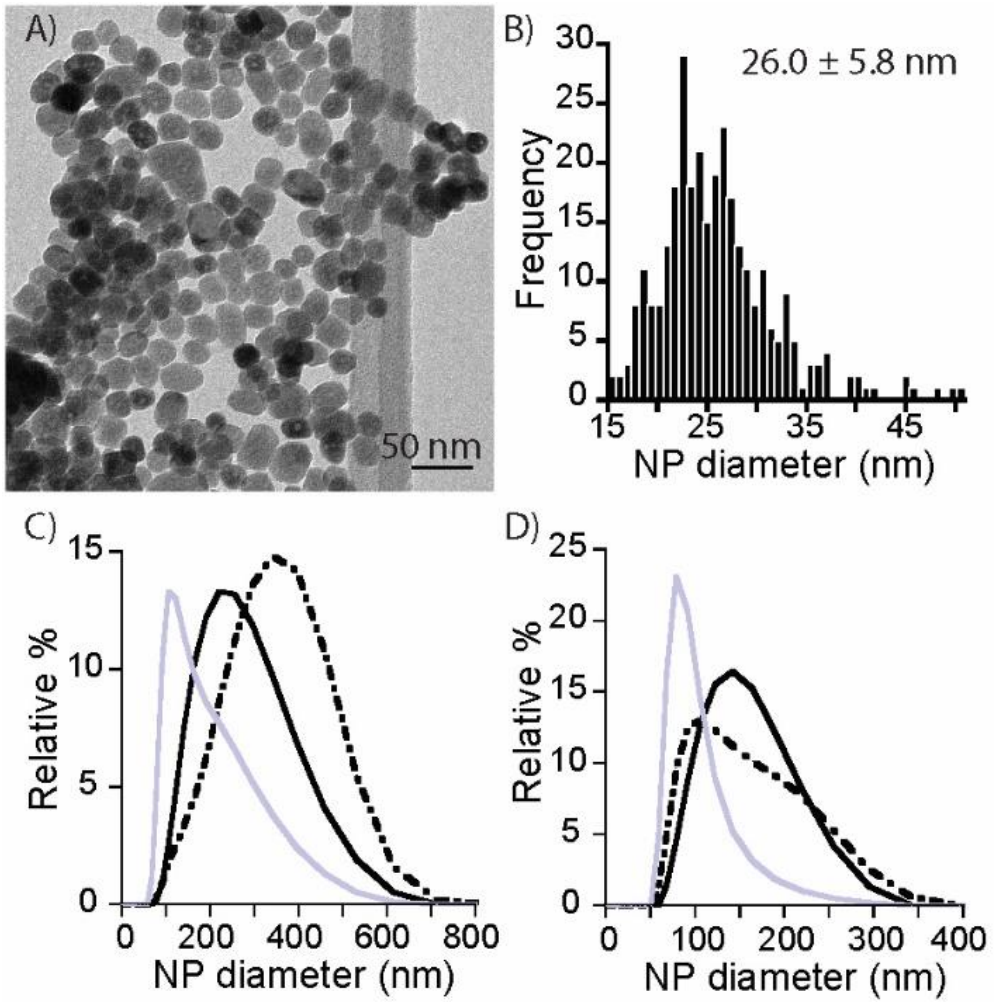


Figure 5.1. Transmission electron microscopy (TEM) characterization of functionalized and purified MNs and dynamic light scattering (DLS) spectra with number % (blue), volume % (dot dash black), and intensity % (black) traces. **A)** TEM image of PEI-MNs. **B)** Size distribution of the nanoparticles in A). **C)** DLS spectra of PEI-MNs pre-purified with Z-average of 192.4 nm diameter. **D)** DLS spectra of magnetically purified PEI-MNs after four magnetic purification rounds with Z-average 135.5 nm diameter.

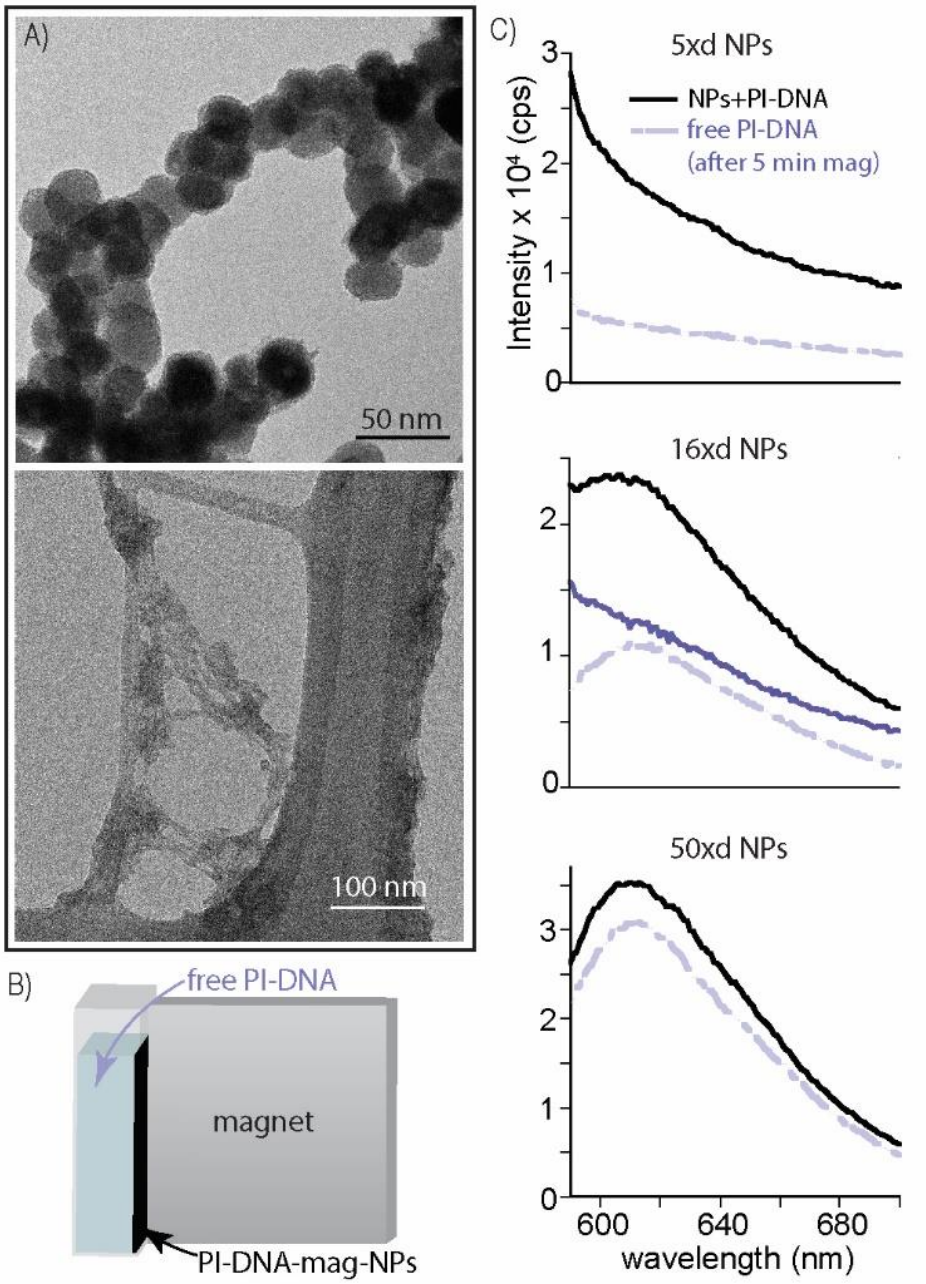


Figure 5.2. Characterization of interactions of PEI-MNs with PI/DNA**. A)** TEM images of PI/DNA on PEI-MNs (top panel) and PI/DNA (lower panel). PI/DNA (top) and DNA (bottom) were stained with 4% uranyl acetate. **B)** Magnet-pulling setup. **C)** Fluorescence spectra of PI/DNA and PEI-MNs in solutions at different PEI-MN concentrations of 5×, 16×, and 50× dilution factors measured without (black lines) and with (light purple dashed lines) magnetic separation to highlight the need for pulling the MNs out of solution prior to fluorescence measurement. The difference between the fluorescence spectra with and without magnetic pulling in the 16×d case is shown (middle panel, solid blue line). Deconvolution was performed to isolate the contribution of scattering from PEI-MNs and fluorescence from PI/DNA.

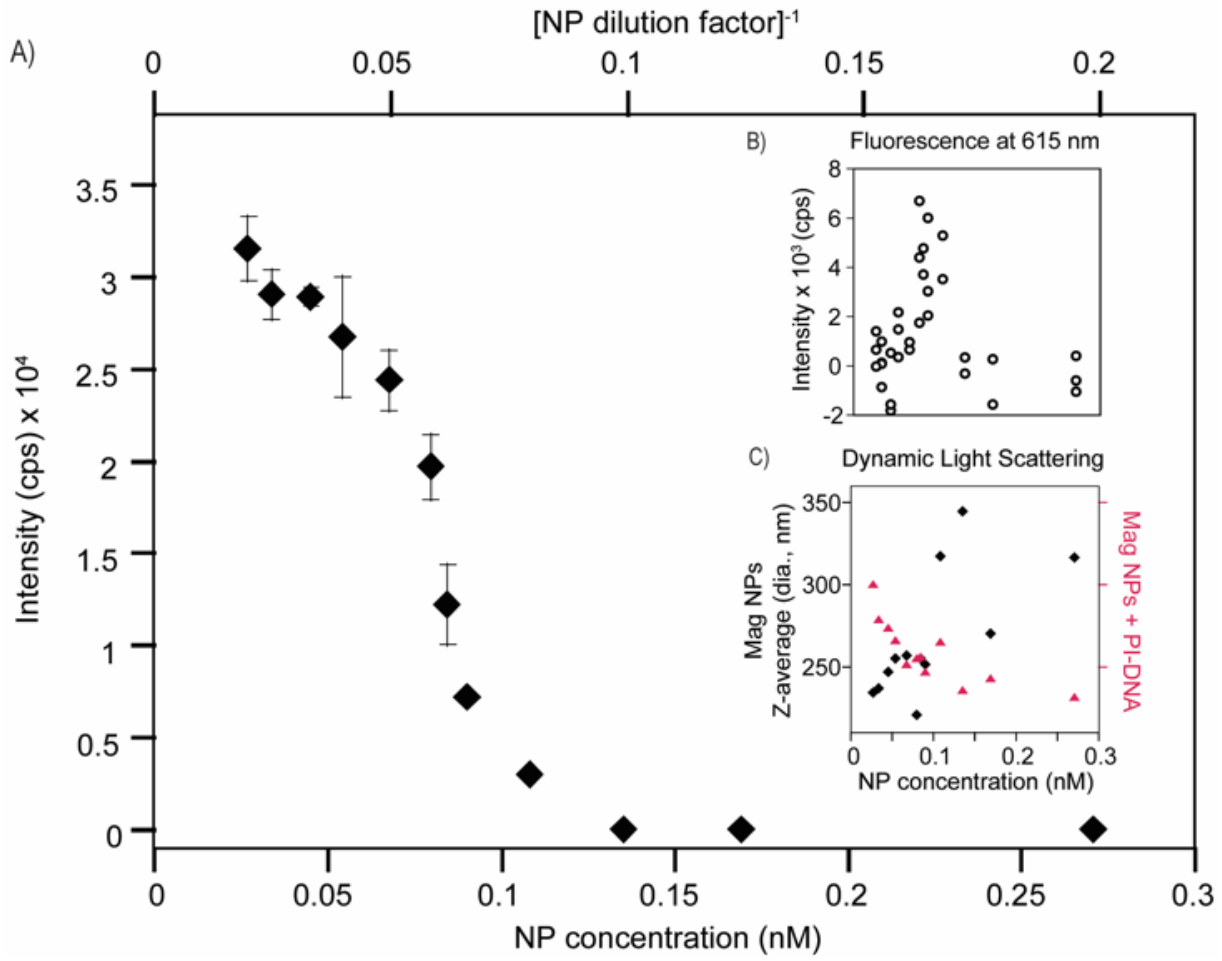


Figure 5.3. Fluorescence titration using magnetic nanoparticles (FTMN). **A)** The results of the fluorescence titration. The titration of PI-DNA using PEI-MNs at various concentrations (displayed as the dilution factor from a 1.6 nM stock solution of nanoparticles) is shown. **B)** The detectable fluorescence from PI after deconvolution of the signal difference between signals of fluorescence spectra before and after pulling PEI-MNs from the solution. **C)** The results of DLS measurements using the Z-average of PEI-MNs (left Y axis, black diamonds) and PEI-MNs mixed with PI/DNA (right Y axis, red triangles).

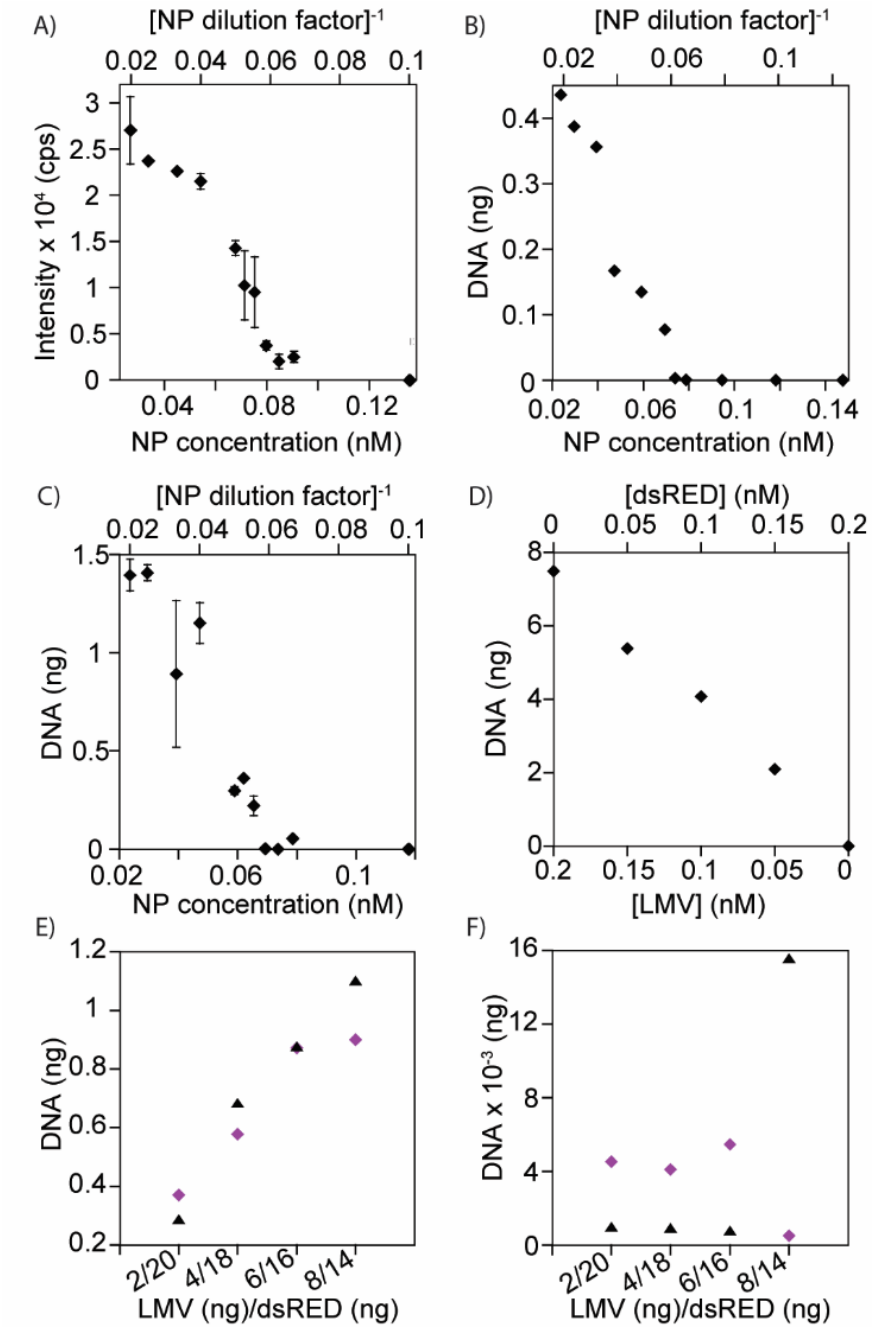


Figure 5.4. Fluorescence and qPCR titration. **A)** Fluorescence titration curve corresponding to the qPCR titration in B). This is similar to Figure 3A, indicating high repeatability. **B)** Titration using qPCR without removal of MNs. The trend is similar to Figures 3A and 4A. The onset points of titration occurred at approximately 16xd dilution factor or 0.07 nM. **C)** Titration using qPCR after removal of MNs, expressed in terms of nanoparticle concentrations, which are obtained using the MP-AES calibration factor and the dilution factor. **D)** Competitive titration (LMV-0-GFPHC is the active DNA, dsRED is the inactive DNA). **E)** Premixed binding experiment at room temperature (RT, purple diamonds) and 40°C (black triangles). **F)** Spiked binding experiment at RT (purple diamonds) and 40°C (black triangles).

REFERENCES

Arsianti, M., Lim, M., Marquis, C. P., Amal, R. (2010). Assembly of Polyethylenimine-Based Magnetic Iron Oxide Vectors: Insights into Gene Delivery. *Langmuir***,** *26* (10), 7314-7326.

Bi, Q.; Song, X.; Hu, A.; Luo, T.; Jin, R.; Ai, H.; Nie, Y., Magnetofection. (2020). Magic magnetic nanoparticles for efficient gene delivery *Chinese Chemical Letters*

Berensmeier, S. (2006). Magnetic particles for the separation and purification of nucleic acids *Applied Microbiology and Biotechnology*, *73* (3), 495-504.

Bouzas-Ramos, D., Trapiella-Alfonso, L., Pons, K., Encinar, J. R., Costa-Fernandez, J. M., Tsatsaris, V., Gagey-Eilstein, N. (2018). Controlling Ligand Surface Density on Streptavidin-Magnetic Particles by a Simple, Rapid, and Reliable Chemiluminescent Test. *Bioconjugate Chemistry*, *29* (8), 2646-2653.

Cunningham, F. J., Goh, N. S., Demirer, G. S., Matos, J. L., Landry, M. P. (2018). Nanoparticle-Mediated Delivery towards Advancing Plant Genetic Engineering *Trends Biotechnology***,** *36* (9), 882-897.

Cruz-Acuna, M. Maldonado-Camargo, L. Dobson, J. Rinaldi, C. (2016). From oleic acid-capped iron oxide nanoparticles to polyethyleneimine-coated single-particle magnetofectins *Journal of Nanoparticle Research, 18* (9).

Du, X. F. Wang, J. Zhou, Q. Zhang, L. W. Wang, S. J. Zhang, Z. X. Yao, C. P. (2018). Advanced physical techniques for gene delivery based on membrane perforation. *Drug Delivery***,** *25* (1), 1516-1525.

Joldersma, D., Liu, Z. C. (2018). Plant genetics enters the nano age? *Journal of Integrative Plant Biology*, *60* (6), 446-447.

Hu, P. Zhang, S. J. Wu, T. Ni, D. L. Fan, W. P. Zhu, Y., Qian, R. Shi, J. L. (2018). Fe-Au Nanoparticle-Coupling for Ultrasensitive Detections of Circulating Tumor DNA. *Advanced Materials*, *30* (31).

Kang, K. Choi, J., Nam, J. H., Lee, S. C., Kim, K. J., Lee, S. W., Chang, J. H. (2009). Preparation and Characterization of Chemically Functionalized Silica-Coated Magnetic Nanoparticles as a DNA Separator. *Journal of Physical Chemistry B*, *113* (2), 536-543.

Kouassi, G. K., Irudayaraj, J. (2006). Magnetic and gold-coated magnetic nanoparticles as a DNA sensor. *Analytical Chemistry, 78* (10), 3234-3241.

Kudr, J., Haddad, Y., Richtera, L., Heger, Z., Cernak, M., Adam, V., Zitka, O. (2017). Magnetic Nanoparticles: From Design and Synthesis to Real World Applications. *Nanomaterials-Basel, 7* (9).

Lakshmanan, R., Sanchez-Dominguez, M., Matutes-Aquino, J. A., Wennmalm, S., Rajarao, G. K. (2014). Removal of Total Organic Carbon from Sewage Wastewater Using Poly(ethylenimine)-Functionalized Magnetic Nanoparticles. *Langmuir, 30* (4), 1036-1044.

Lin, P. C., Tseng, M. C., Su, A. K., Chen, Y. J., Lin, C. C. (2007). Functionalized magnetic nanoparticles for small-molecule isolation, identification, and quantification. *Analytical Chemistry, 79* (9), 3401-3408.

Liu, Q., Li, J. Y., Liu, H. X., Tora, I., Ide, M. S., Lu, J. W., Davis, R. J., Green, D. L., Landers, J. P. (2014). Rapid, cost-effective DNA quantification via a visually-detectable aggregation of superparamagnetic silica-magnetite nanoparticles. *Nano Research, 7* (5), 755-764.

Mehier-Humbert, S., Guy, R. H. (2005). Physical methods for gene transfer: Improving the kinetics of gene delivery into cells. *Advanced Drug Delivery Reviews, 57* (5), 733-753.

Mykhaylyk, O. Z., Hammerschmid, E., Anton, M., and C. Plank. (2009). Recent Advances in Magnetofection and Its Potential to Deliver siRNA in Vitro. In *siRNA and miRNA Gene Silencing: From Bench to Bedside*, Sioud, M., Ed. Humana Press, pp 111-146.

Narayanasamy, K. K., Cruz-Acuna, M., Rinaldi, C., Everett, J., Dobson, J., Telling, N. D. (2018). Alternating current (AC) susceptibility as a particle-focused probe of coating and clustering behaviour in magnetic nanoparticle suspensions. *Journal of Colloid and Interface Science, 532*, 536-545.

Neto, D. M. A., J.S. Rocha, P.B.A. Fechine, Leonardo Vivas, Dinesh Pratap Singh, R.M. Freire. (2020). Magnetic Nanoparticles in Analytical Chemistry. In *Magnetochemistry – Materials and Applications*, Inamuddin, R. B., and A.M. Asiri, Ed. Materials Research Forum LLC: Millersville, PA, Vol. 66, pp 173-216.

Nguyen, A. H., Abdelrasoul, G. N., Lin, D. H., Maadi, H., Tong, J. F., Chen, G., Wang, R., Anwar, A., Shoute, L., Fang, Q., Wang, Z. X., Chen, J. (2018). Polyethylenimine-coated iron oxide magnetic nanoparticles for high efficient gene delivery. *Applied Nanoscience, 8* (4), 811-821.

Ota, S., Takahashi, Y., Tomitaka, A., Yamada, T., Kami, D., Watanabe, M., Takemura, Y. (2013). Transfection efficiency influenced by aggregation of DNA/polyethylenimine max/magnetic nanoparticle complexes. *Journal of Nanoparticle Research, 15* (5).

Ren, J. S., Jenkins, T. C., Chaires, J. B. (2000). Energetics bf DNA intercalation reactions. *Biochemistry-Us*, *39* (29), 8439-8447.

Rohiwal, S. S., Dvorakova, N., Klima, J., Vaskovicova, M., Senigl, F., Slouf, M., Pavlova, E., Stepanek, P., Babuka, D., Benes, H., Ellederova, Z., Stieger, K. (2020). Polyethylenimine based magnetic nanoparticles mediated non-viral CRISPR/Cas9 system for genome editing. *Scientific Reports-Uk*, *10* (1).

Rong, Z., Wang, C. W., Wang, J. F., Wang, D. G., Xiao, R., Wang, S. Q. (2016). Magnetic immunoassay for cancer biomarker detection based on surface-enhanced resonance Raman scattering from coupled plasmonic nanostructures. *Biosensors and Bioelectronics*, *84*, 15-21.

Schaller, V., Kraling, U., Rusu, C., Petersson, K., Wipenmyr, J., Krozer, A., Wahnstrom, G., Sanz-Velasco, A., Enoksson, P., Johansson, C. (2018). Motion of nanometer sized magnetic particles in a magnetic field gradient. *Journal of Applies Physics*, *104* (9).

Shan, Z., Jiang, Y. J., Guo, M. Y., Bennett, J. C., Li, X. H., Tian, H. F., Oakes, K., Zhang, X., Zhou, Y., Huang, Q. M., Chen, H. P. (2015). Promoting DNA loading on magnetic nanoparticles using a DNA condensation strategy. *Colloid and Surfaces B: Biointerfaces*, *125*, 247-254.

Shen, L. Z., Li, B., Qiao, Y. S. (2018). Fe3O4 Nanoparticles in Targeted Drug/Gene Delivery Systems. *Materials*, *11* (2).

Sheng, W., Wei, W., Li, J. J., Qi, X. L., Zuo, G. C., Chen, Q., Pan, X. H., Dong, W. (2016). Amine-functionalized magnetic mesoporous silica nanoparticles for DNA separation. *Applied Surface Science*, *387*, 1116-1124.

Song, H., Yu, M. H., Lu, Y., Gu, Z. Y., Yang, Y. N., Zhang, M., Fu, J. Y., Yu, C. Z. (2017). Plasmid DNA Delivery: Nanotopography Matters. *Journal of the American Chemical Society*, *139* (50), 18247-18254.

Song, W. X., Gregory, D. A., Al-janabi, H., Muthana, M., Cai, Z. Q., Zhao, X. B. (2019). Magnetic-silk/polyethyleneimine core-shell nanoparticles for targeted gene delivery into human breast cancer cells. *International Journal pf Pharmaceutics*, *555,* 322-336.

Svoboda, O., Fohlerova, Z., Baiazitova, L., Mlynek, P., Samouylov, K., Provaznik, I., Hubalek, J. (2018). Transfection by Polyethyleneimine-Coated Magnetic Nanoparticles: Fine-Tuning the Condition for Electrophysiological Experiments. *Journal of Biomedical Nanotechnology*, *14* (8d), 1505-1514.

Wang, P., Lombi, E., Zhao, F.-J., Kopittke, P. M. (2016). Nanotechnology: A New Opportunity in Plant Sciences. *Trends in Plant Science*, *21* (8), 699-712.

Wang, X. L., Zhou, L. Z., Ma, Y. J., Li, X., Gu, H. C. (2009). Control of Aggregate Size of Polyethyleneimine-Coated Magnetic Nanoparticles for Magnetofection. *Nano Research*, *2* (5), 365-372.

Wang, Y., Cui, H., Li, K., Sun, C., Du, W. (2014). A Magnetic Nanoparticle-Based Multiple-Gene Delivery System for Transfection of Porcine Kidney Cells. *Plos One*, *9* (8).

Xie, L., Jiang, W., Nie, Y., He, Y. Y., Jiang, Q., Lan, F., Wu, Y., Gu, Z. W. (2013). Low aggregation magnetic polyethyleneimine complexes with different saturation magnetization for efficient gene transfection *in vitro* and *in vivo*. *RSC Advances*, *3* (45), 23571-23581.

Xie, L., Jiang, Q., He, Y. Y., Nie, Y., Yue, D., Gu, Z. W. (2015). Insight into the efficient transfection activity of a designed low aggregated magnetic polyethyleneimine/DNA complex in serum-containing medium and the application *in vivo*. *Biomaterials Science-Uk*, *3* (3), 446-456.

Yang, Q., Dong, Y., Qiu, Y., Yang, X. Z., Cao, H., Wu, Y. (2020). Design of Functional Magnetic Nanocomposites for Bioseparation. *Colloid and Surfaces B: Biointerfaces*, *191*.

Zhang, X. L., Li, L. L., Li, L., Chen, J., Zou, G. Z., Si, Z. K., Jin, W. R. (2009). Ultrasensitive Electrochemical DNA Assay Based on Counting of Single Magnetic Nanobeads by a Combination of DNA Amplification and Enzyme Amplification. *Analytical Chemistry*, *81* (5), 1826-1832.

Zhang, L., Li, Y. C., Yu, J. C., Chen, Y. Y., Chan, K. M. (2014). Assembly of polyethylenimine-functionalized iron oxide nanoparticles as agents for DNA transfection with magnetofection technique. *Journal of Materials Chemistry B*, *2* (45), 7936-7944.

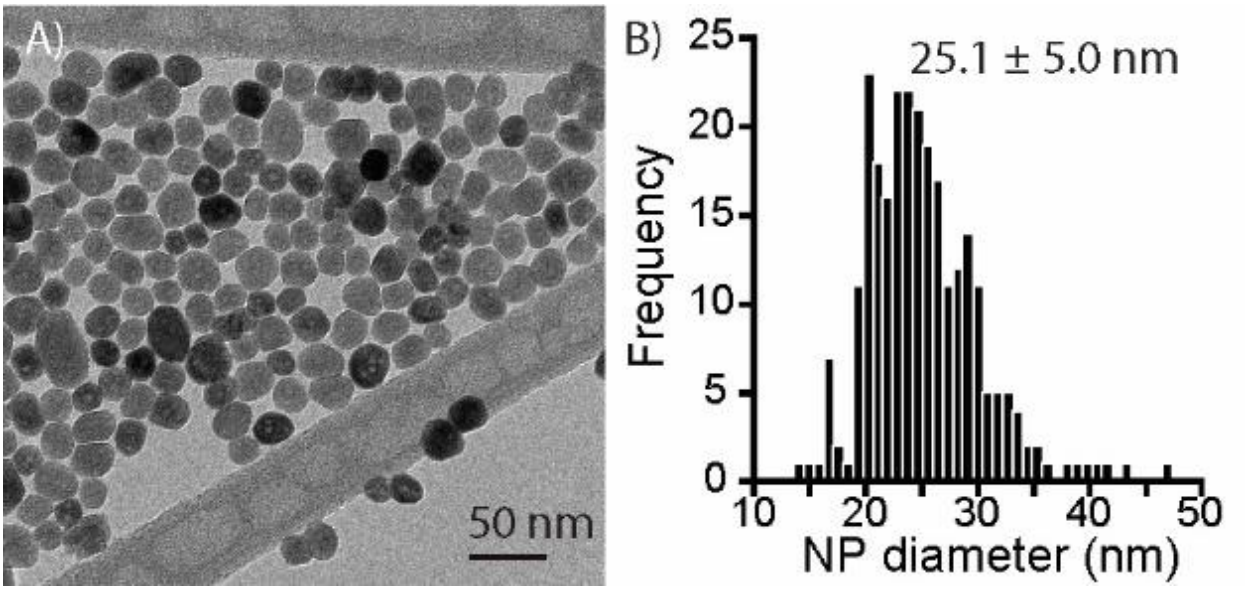
Zhang, T. Y., Xu, Q. H., Huang, T., Ling, D. S., Gao, J. Q. (2020). New Insights into Biocompatible Iron Oxide Nanoparticles: A Potential Booster of Gene Delivery to Stem Cells. *Small*.

Zhao, H., Oczos, J., Janowski, P., Trembecka, D., Dobrucki, J., Darzynkiewicz, Z., Wlodkowic, D. (2010). Rationale for the Real-Time and Dynamic Cell Death Assays Using Propidium Iodide *Cytometry Part A*, *77a* (4), 399-405.

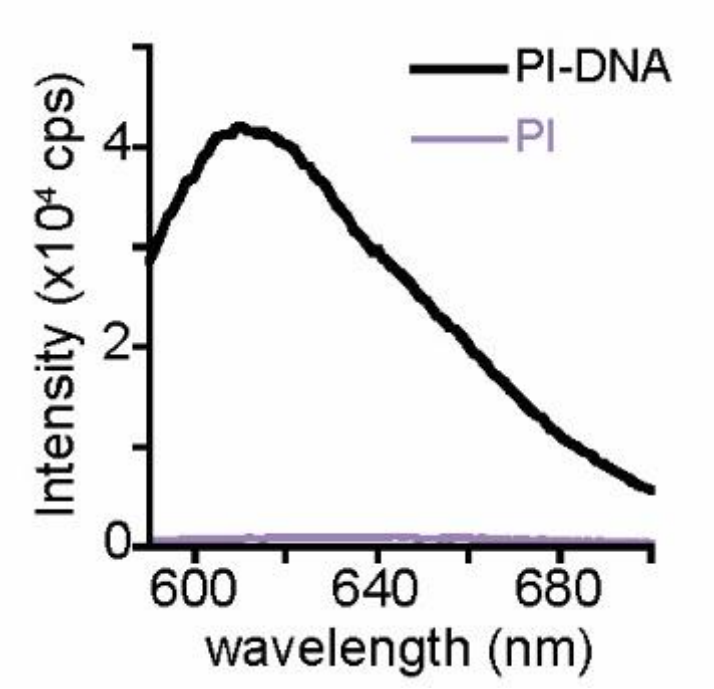
Zhao, X., Meng, Z. G., Wang, Y., Chen, W. J., Sun, C. J., Cui, B., Cui, J. H., Yu, M. L., Zeng, Z. H., Guo, S. D., Luo, D., Cheng, J. Q., Zhang, R., Cui, H. X. (2017). Pollen magnetofection for genetic modification with magnetic nanoparticles as gene carriers. *Nature Plants*, *3* (12), 956-964.

Zhu, N., Ji, H. N., Yu, P., Niu, J. Q., Farooq, M. U., Akram, M. W., Udego, I. O., Li, H. D., Niu, X. B. (2018). Surface Modification of Magnetic Iron Oxide Nanoparticles *Nanomaterials-Basel*, *8* (10).

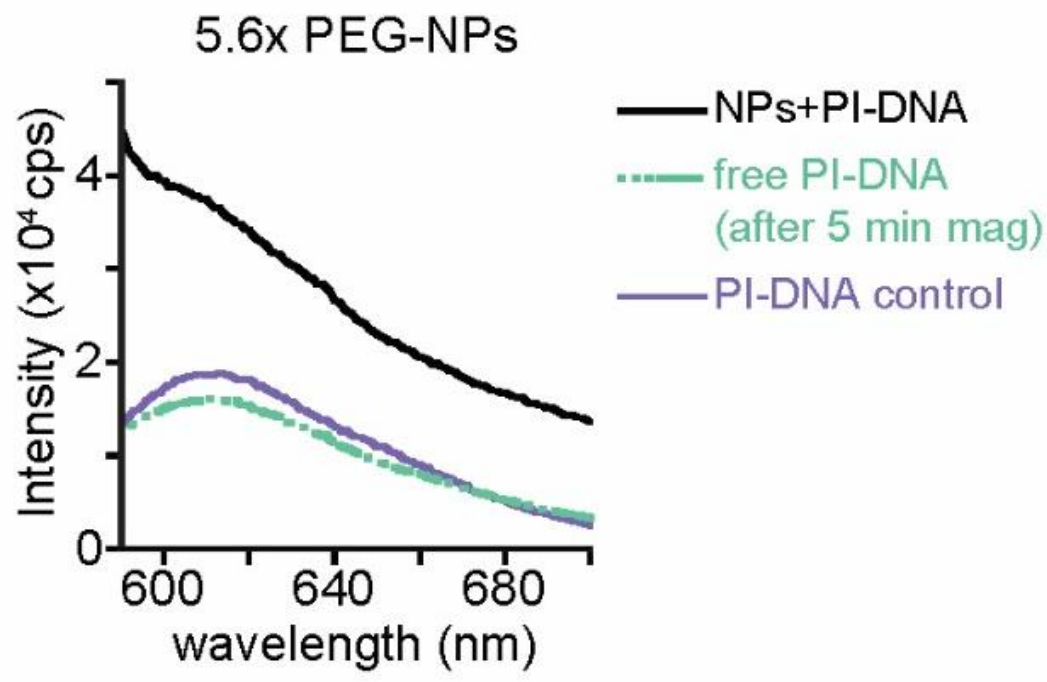
SUPPLEMENTARY FIGURES



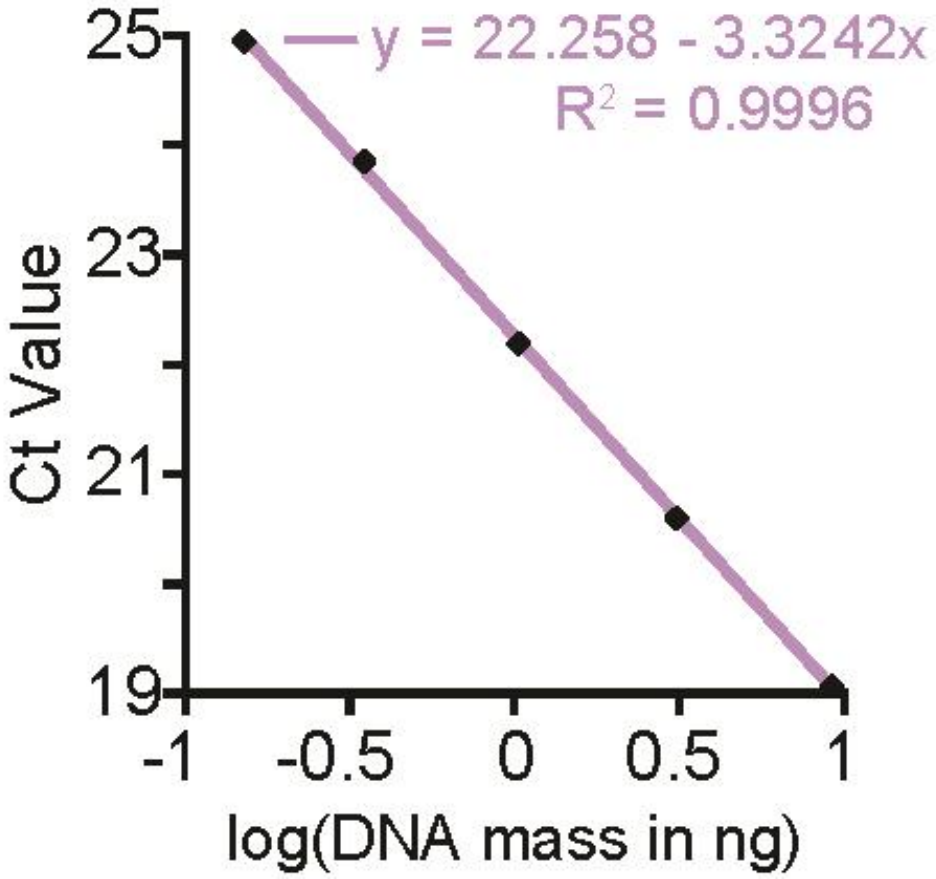
Supplementary Figure S5.1. The size of PEI-MNs was similar after the 1st and 5th round of magnetic purification. **A)** TEM of PEI-MNs after the 5th round of dilute magnetic purification. **B)** Histogram of size distribution of MNs after the 5th round of dilute magnetic purification obtained from analyzing over 300 particles in ImageJ.



Supplementary Figure S5.2. Fluorescence from freshly prepared PI and PI intercalated in plasmid DNA (PI/DNA). The fluorescence yield increased by approximately 50-fold from PI/DNA compared to free PI.



Supplementary Figure S5.3. Control experiments using PEG-MNs. When PI/DNA was incubated with PEG-MNs, PI/DNA was considered to be dissolved in solution but not on the surface of MNs because PEG does not attract DNA as strongly as PEI ligands. Fluorescence measurements (black line) revealed detectable fluorescence even without magnetic pulling. After pulling, over 85% of PI/DNA was in the solution, even at a high MN concentration of 5.6x dilution. This is different from the PEI-MN treatment in which PI/DNA was not detected in the solution (Figure 5.3).



Supplementary Figure S5.4. Calibration of qPCR. The Ct values show a strong linear correlation relationship with DNA concentrations covering those used in this work.

Chapter 6: Conclusions and future perspectives

CONCLUSIONS AND FUTURE PERSPECTIVES

In vitro plant regeneration has been used for successful genetic engineering and genome editing of high regenerating genotypes of lettuce (*Lactuca sativa* L.) (Bertier et al., 2018; Govindarajulu et al., 2015; Michelmore et al., 1987); however, like many other species, *in vitro* regeneration is genotype-dependent in lettuce (Chapter 3); therefore, studying molecular determinants and genetic loci important to regeneration in lettuce will lead to the improvement of regeneration and transformation efficiencies. Accomplishing this will increase the number of genotypes amenable to *Agrobacterium*-mediated transformation as well as genome editing. The work presented in this dissertation is the first step into identifying genetic loci and molecular players that affect regeneration and transformation rates in lettuce. I discussed potential targets for the manipulation of *de novo* organogenesis (Chapter 2) and identified multiple genetic loci significantly associated with different regeneration traits in lettuce including callus formation, organized growth, leaf development, and shoot regeneration (Chapter 3). I also implemented the ectopic expression of *GROWTH REGULATING FACTOR* (*GRF*) and *GRF-INTERACTING FACTOR* (*GIF*) gene fusions from other species to increase regeneration efficiency, transformation efficiency, and shooting frequency of multiple lettuce genotypes (Chapter 4).

De novo organogenesis is under the control of multiple genetic loci in lettuce. This is consistent with evidence of *in vitro* regeneration being a polygenic trait (Lall et al., 2004; Schiantarelli et al., 2001; Takeuchi et al., 2000; Trujillo-Moya et al., 2011). In total, eight QTLs on Chromosomes 1, 2, 3, 4, 8, and 9 were associated with *de novo* shoot organogenesis in lettuce. A major QTL that was consistently present across multiple regeneration traits was fine mapped to narrow down the QTL interval and the candidate gene search. Identifying and phenotyping of recombinants within the intervals of other QTLs would decrease the size of each interval. In

addition, higher resolution fine mapping could be performed on the major QTL on Chromosome 3 as well as the remaining QTLs to narrow down candidate gene searches.

A preliminary search of candidate genes under each QTL revealed multiple developmental genes known to be involved throughout *de novo* organogenesis and somatic embryogenesis in other species (Chapter 3). Many of these genes were within two mega bases of each QTL peak. Candidate genes of particular interest include the KNOTTED-like gene *KNAT6, AGAMOUS LIKE 104* (*AGL104*), *SQUAMOSA PROMOTER BINDING LIKE 6* (*SPL6*), *PIN-FORMED 4* (*PIN4*), and *MEAF6*. These genes will be cloned and tested *in planta* either through ectopic expression or genome editing in both parents of this mapping population. In addition to the candidate genes above, multiple predicted uncharacterized proteins were at or near peaks of every identified QTL. These genes could be cloned and characterized to identify whether they play roles in *de novo* organogenesis in lettuce. Gene validation through over expression or genome editing would be ideal because altered regeneration rates are easily detected.

In addition to validating candidate genes, using RNA sequencing (TAGseq) to identify differential expression of genes between genotypes would be informative. Gene expression profiles from bulked samples of high and low regenerating RILs during the shoot initiating step of *de novo* organogenesis could be used to identify candidate gene expression levels at this critical step in regeneration. Furthermore, using results from the TAGseq experiment, promoter-reporter gene fusions can be used to observe expression patterns of selected candidate genes. For example, *AGL104* has been shown to be involved in pollen maturation and tube growth in *Arabidopsis* but has not been characterized in lettuce (Adamczyk and Fernandez, 2009). Because *AGL104* is located near the peak of the major QTL identified on Chromosome 3, it is possible that it has functions during *de novo* organogenesis in lettuce. By cloning the native promoter of *AGL104* and

fusing it to reporter gene such as *dsRED* (Baird et al., 2000) or *RUBY* (Baird et al., 2000; He et al., 2020), I could visualize its native expression patterns to identify if it plays a role in *in vitro* regeneration and/or has functions in reproductive development as seen in *Arabidopsis*. Repeating this with other candidates and predicted, uncharacterized genes will allow us to gain a better understanding of the genes important to *in vitro* regeneration in lettuce. This would also be the first step towards building gene and molecular networks for *in vitro* regeneration in lettuce, allowing for the alteration of regeneration rates of recalcitrant lettuce genotypes.

Introduction of *GRF-GIF* chimeric gene fusions increases shooting frequency, regeneration efficiency, and transformation efficiency in lettuce. In Chapter 4, I used GRF-GIFs from grape, citrus, pepper, and tomato to test for their ability to increase regeneration rates in lettuce. I also showed that resistant GRF-GIFs with mutated miRNA 396 binding sites stimulated higher amounts of regeneration. However, I only tested resistant tomato GRF-GIFs against the wildtype GRF-GIF. The wildtype tomato GRF-GIF consistently showed regeneration rates similar to that of the control transformations, resulting in no significant increase in shooting frequency or regeneration efficiency. It would be interesting to test miRNA-resistant GRF-GIFs against wildtype GRF-GIFs that are now known to increase regeneration rates in lettuce, such as the pepper and citrus GRF-GIFs. In addition, it would be useful to clone and overexpress lettuce *GRF* and *GIF* genes to observe if endogenous gene fusions stimulate even higher shooting frequency and regeneration efficiency than GRF-GIFs from other species. These GRF-GIFs should be tested in other Compositae species such as recalcitrant sunflower (*Helianthus annuus* L.).

Furthermore, I developed a strategy for increased recovery of transgenic lettuce plants transformed with a gene of interest. I used a co-transformation method by transforming a gene of interest and the best perforation GRF-GIF (grape *rGRF4-GIF1*) into lettuce, which increased the

transformation and regeneration efficiencies. Previously, it has been reported that approximately 25% of shoots recovered in the absence of a selectable marker were transgenic (Debernardi et al., 2020). It would be interesting to test this strategy in lettuce as well. If proven successful, cloning the GRF-GIF chimeric coding sequence in place of the selectable marker could lead to increased transformation efficiency and higher recovery of transgenic or genome edited plants without the use of antibiotic selection.

Another strategy to combat genotypes recalcitrant to *in vitro* regeneration and transformation is the direct delivery of genes into germline cells, thereby eliminating tissue culture steps. One approach is with the introduction of nanoparticles carrying plasmid DNA into germline cells. In this dissertation (Chapter 5), I discussed a fast method for loading and quantifying large plasmid DNA onto magnetic nanoparticles. Next, we could optimize the introduction of the magnetic nanoparticles into plant cells followed by the release of the DNA for transient expression. Following optimization, we could introduce the magnetic nanoparticles into accessible tissues targeting germline cells (Appendix A). If proven successful, we could load Cas9 ribonucleoproteins (RNPs) and gRNAs for genome editing of germline cells.

The research presented in this dissertation is the first study to identify genetic loci associated with *de novo* organogenesis and to molecularly manipulate shooting frequencies and regeneration efficiencies in lettuce. The application of these results could help untangle genetic and molecular pathways of indirect *de novo* organogenesis in lettuce and increase regeneration of recalcitrant lettuce genotypes. Furthermore, ectopic expression of validated regeneration candidate genes could stimulate regeneration of other recalcitrant species of the Compositae family.

REFERENCES

Adamczyk, B. J., and Fernandez, D. E. (2009). MIKC* MADS Domain Heterodimers Are Required for Pollen Maturation and Tube Growth in *Arabidopsis*. *Plant Physiology*, *149*, 1713–1723.

Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2000). Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *PNAS*, 11984–11989.

Bertier, L. D., Ron, M., Huo, H., Bradford, K. J., Britt, A. B., and Michelmore, R. W. (2018). High-resolution analysis of the efficiency, heritability, and editing outcomes of CRISPR/Cas9-induced modifications of *NCED4* in lettuce (*Lactuca sativa*). *G3: Genes, Genomes, Genetics*, *8*(5), 1513–1521.

Debernardi, J. M., Tricoli, D. M., Ercoli, M. F., Hayta, S., Ronald, P., Palatnik, J. F., and Dubcovsky, J. (2020). A GRF–GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nature Biotechnology 2020 38:11*, *38*(11), 1274–1279.

Govindarajulu, M., Epstein, L., Wroblewski, T., and Michelmore, R. W. (2015). Host-induced gene silencing inhibits the biotrophic pathogen causing downy mildew of lettuce. *Plant Biotechnology Journal*, *13*, 875–883.

He, Y., Zhang, T., Sun, H., Zhan, H., and Zhao, Y. (2020). A reporter for noninvasively monitoring gene expression and plant transformation. *Horticulture Research*, *7*(152).

Lall, S., Nettleton, D., DeCook, R., Che, P., and Howell, S. H. (2004). Quantitative Trait Loci Associated With Adventitious Shoot Formation in Tissue Culture and the Program of Shoot Development in *Arabidopsis*. *Genetics*, *167*(4), 1883–1892.

Michelmore, R., Marsh, E., Seely, S., and Landry, B. (1987). Transformation of lettuce (*Lactuca sativa*) mediated by *Agrobacterium tumefaciens*. *Plant Cell Reports*, *6*, 439–442.

Schiantarelli, E., de la Peña, A., and Candela, M. (2001). Use of recombinant inbred lines (RILs) to identify, locate and map major genes and quantitative trait loci involved with *in vitro* regeneration ability in *Arabidopsis thaliana*. *Theoretical and Applied Genetics 2001 102:2*, *102*(2), 335–341.

Takeuchi, Y., Abe, T., and Sasahara, T. (2000). RFLP mapping of QTLs influencing shoot regeneration from mature seed-derived calli in rice. *Crop Science*, *40*(1), 245–247.

Trujillo-Moya, C., Gisbert, C., Vilanova, S., and Nuez, F. (2011). Localization of QTLs for *in vitro* plant regeneration in tomato. *BMC Plant Biology*, *11*.

Appendix A: Identification of accessible target tissues for gene delivery into lettuce (Lactuca sativa L.)

ABSTRACT

The development of plant transformation and biotechnologies has made it possible to introduce genes and editing reagents into plant cells for genetic engineering and genome editing. The most common method used is *Agrobacterium*-mediated transformation; however, this requires *in vitro* regeneration with few exceptions, which is genotype dependent. Development of methods for *in planta* gene delivery are desirable to eliminate tissue culture steps. The first step of *in planta* gene delivery is identifying accessible target tissues such as germline cells. In this appendix, I describe a time course of shoot apical meristem and reproductive apex development to identify time points that are most accessible for gene delivery in lettuce (*Lactuca sativa* L.).

INTRODUCTION

The development of plant transformation techniques, such as *Agrobacterium*-mediated transformation, has made it possible to introduce cloned genes into multiple plant species. However, this process involves time consuming and labor-intensive tissue culture methods with few exceptions. In addition, *in vitro* regeneration and *Agrobacterium*-mediated transformation is genotype dependent. In contrast, gene delivery into germline cells for transformation *in planta* would eliminate the need for tissue culture methods; therefore, identifying target tissues accessible for direct gene delivery into germline cells is desirable.

Methods of *in planta* transformations have been studied for multiple decades. One of the first methods of *in planta* transformation in *Arabidopsis* involved the inoculation of severed primary and secondary inflorescences with *A. tumefaciens* (Chang et al., 1994). This resulted in the development of shoots from severed sites. Other methods developed in *Arabidopsis* include the germination of seeds in *Agrobacterium* suspension cultures and the commonly used Floral-dip method (Clough and Bent, 1998; Katavic et al., 1994). More recently, *Agro*-injection into lateral meristems and developing embryos (e.g., seeds or pods) for direct gene delivery *in plana* have been demonstrated. In soybean, injection of *A. tumefaciens* into immature soybean pods resulted in transformed seeds, although at low frequencies (Zia et al., 2011). In *Nicotiana benthamiana*, genome edited plants were produced after the *Agrobacterium*-mediated delivery of editing reagents and developmental regulators (e.g., *BABY BOOM, WUSCHEL*) into shoots after removal of apical and axillary meristems (Maher et al., 2020). Although these methods have proven successful in other species, methods for direct gene delivery *in planta* have not yet been developed in lettuce (*Lactuca sativa* L.).

In this study, I describe target tissues potentially accessible for direct gene delivery *in planta*. In addition, I produce a developmental timeline of shoot apical meristem (SAM) maturation in young seedlings and pollination in flowers. The results of this study give us a better understanding of the accessibility of the SAM and reproductive apices for direct gene delivery in lettuce.

METHODS

Plant material

Seeds were surface sterilized with 20% Clorox for 20 minutes with constant agitation at 250 rpm. Sterile seeds were rinsed three times with 100 mL of sterile distilled water. For observance of the SAMs, *L sativa* cv. Green Towers seeds were sown on sterile filter paper soaked with approximately 3 mL of milliQ water in Magenta boxes. Seeds were incubated for different timepoints between 0 and 48 hours (0, 3, 6, 8, 12, 24, 48 hrs) and for 3 days and 6 days in a 15°C growth chamber under a 16/8 hour light/dark cycle. For observance of flowers and pollen, *L. sativa* cv. Salad Bowl flowers were artificially pollinated and flowers were collected 0, 15, 30, 45, and 60 minutes after pollination.

Preparation of SAM samples

Seeds were germinated for various time points described above. Seed samples for the 0-hour time point were soaked in sterile water at 4°C in the dark for 12 hours to soften the seed coat prior to removal. For 0, 3, 6, 8, 12, and 24-hour samples the seed coat was removed prior to fixing tissues. After each timepoint, seedlings were fixed in 5 mL of 100% methanol for 10 minutes followed by two 20-minute washes with 100% ethanol. After fixation, seedlings were hand

sectioned directly through the SAM or by carefully removing one cotyledon to leave an intact SAM on the remaining cotyledon. After sectioning, samples were returned to clean 100% ethanol and transferred to the UC Davis Biological Electron Microscopy Facility for further analysis.

Collection of reproductive apices

Flowers were collected for the observance of reproductive apices. Immediately after floral opening, five flowers were artificially pollinated, followed by collection of one flower at 0, 15, 30, 45, and 60 minutes after pollination (MAP). Immediately after collection, samples were fixed in 100% methanol for 20 minutes followed by two 20-minute washes in 100% ethanol. After the second wash, all samples were returned to clean 100% ethanol and transferred to the UC Davis Biological Electron Microscopy Facility for further analysis.

Scanning electron microscopy

After fixation, tissues were critically point dried using the Tousimis 931 critical point dryer (UC Davis Biological Electron Microscopy Facility). Samples were then mounted onto scanning electron microscope (SEM) specimen stubs and gold sputter coated for three cycles using the Pelco SC-7 sputtering system (UC Davis Biological Electron Microscopy Facility). Images of samples were taken using the Philips XL30 SEM (UC Davis Biological Electron Microscopy Facility).

RESULTS

Accessibility of shoot apical meristems

Shoot apical meristems were examined in young seedlings at multiple timepoints after imbibition. The SAM continued to stay accessible for potential gene delivery until 24 hours after

imbibition. Leaf primordia had not yet fully protected the SAM at 0, 3, 6, 8, and 12 hours after imbibition, with hours 0 to 6 appearing to be the most accessible (**Figures A1.1 to A1.3**). However, for all time points after 24 hours, the leaf primordia began to mature and close over the SAM (**Figures A1.3 and A1.4**), making it inaccessible for potential injection with *Agrobacterium* or other forms of gene delivery.

Accessibility of reproductive apices

Inflorescences and flowers (**Figure A1.5**) were collected for a period of 60 MAP to observe the progression of stigma bifurcation. After floral opening and pollination, stigmas began to split and progressed to fully bifurcated over a period of 1 hour (**Figure A1.6**). The collected flowers were also used to observe progression of pollen tube germination at 0, 15, 30, 45, and 60 MAP (**Figure A1.7**). Multiple pollen grains were in contact with the stigma and had germinated 60 MAP. In addition, 60 MAP, one pollen tube had penetrated the style (**Figure A1.7**).

DISCUSSION

Shoot apical meristems and reproductive apices both have the potential to be targeted for direct gene delivery *in planta*. Successful gene delivery into apices would introduce genes of interest into the germline, eliminating the need for tissue culture methods. By studying the development of shoot apical meristems and reproductive apices, the most accessible timepoints can be identified and used for future gene delivery studies.

In this experiment, the SAM appears to be the most accessible between 0 and 6 hours after imbibition. Interestingly, the SAM of the 0 hour imbibed seeds appeared to be more developed than the SAM of the 3 hour imbibed seeds. This is likely due to the 0 hour imbibition treatment

soaking in sterile water for 12 hours prior to fixation and sectioning. This was performed to soften and ease the removal of the seed coat. Although the seeds were in the dark and at 4°C, this likely began the process of germination leading to maturation of the SAM and surrounding tissues. Leaf primordia blocked accessibility of the SAM for all time points after 24 hours.

Stigma bifurcation progression was assessed, and fully bifurcated stigmas were observed 60 MAP. In addition, pollen germination was assessed and observed to start approximately 60 MAP. Both of these observations are consistent with previous reports of pollen germination in lettuce (Lukasz and Piosik, 2013). One pollen grain was observed to have penetrated the stigma 60 MAP. However, previous reports observed pollen tube penetration 180 MAP (Lukasz and Piosik, 2013). Combining data presented in this chapter and data from previous reports, pollen 60 to 180 MAP may be accessible for direct gene delivery. Further investigation on pollen pore accessibility needs to be conducted.

These results identified accessible timepoints of potential gene delivery into the SAM or pollen. Successful gene delivery into either the SAM or pollen would allow for direct transfer of genes of interest *in planta* and eliminate the need for tissue culture. Experiments should be performed to test gene delivery methods, such as delivery of plasmid DNA by magnetic nanoparticles, to identify the most reliable target tissue. In addition, more work should be performed to identify additional target tissues such as lateral apices, seeds, and embryos.

FIGURES

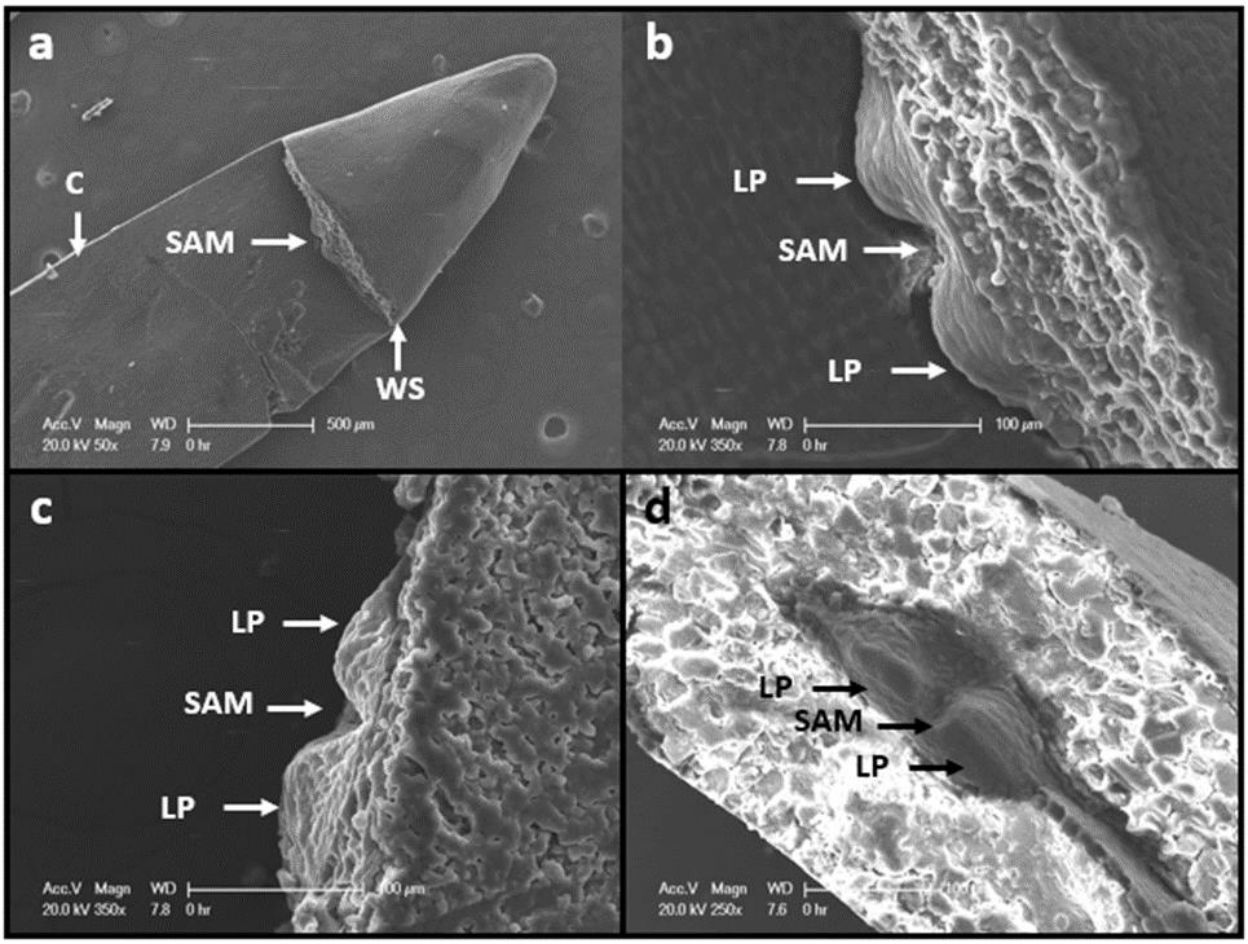


Figure A1.1. SEMs of shoot apical meristems 0 hours after imbibition. **a**) Macroscopic view of the cotyledon (c), shoot apical meristem (SAM), and the wound site (WS) after removal of one cotyledon. **b**) Magnification of the SAM and leaf primordia (LP) from image a. **c**) The SAM and LP from a sample with both cotyledons removed. **d**) Top view of the SAM and LP from a sample with both cotyledons removed.

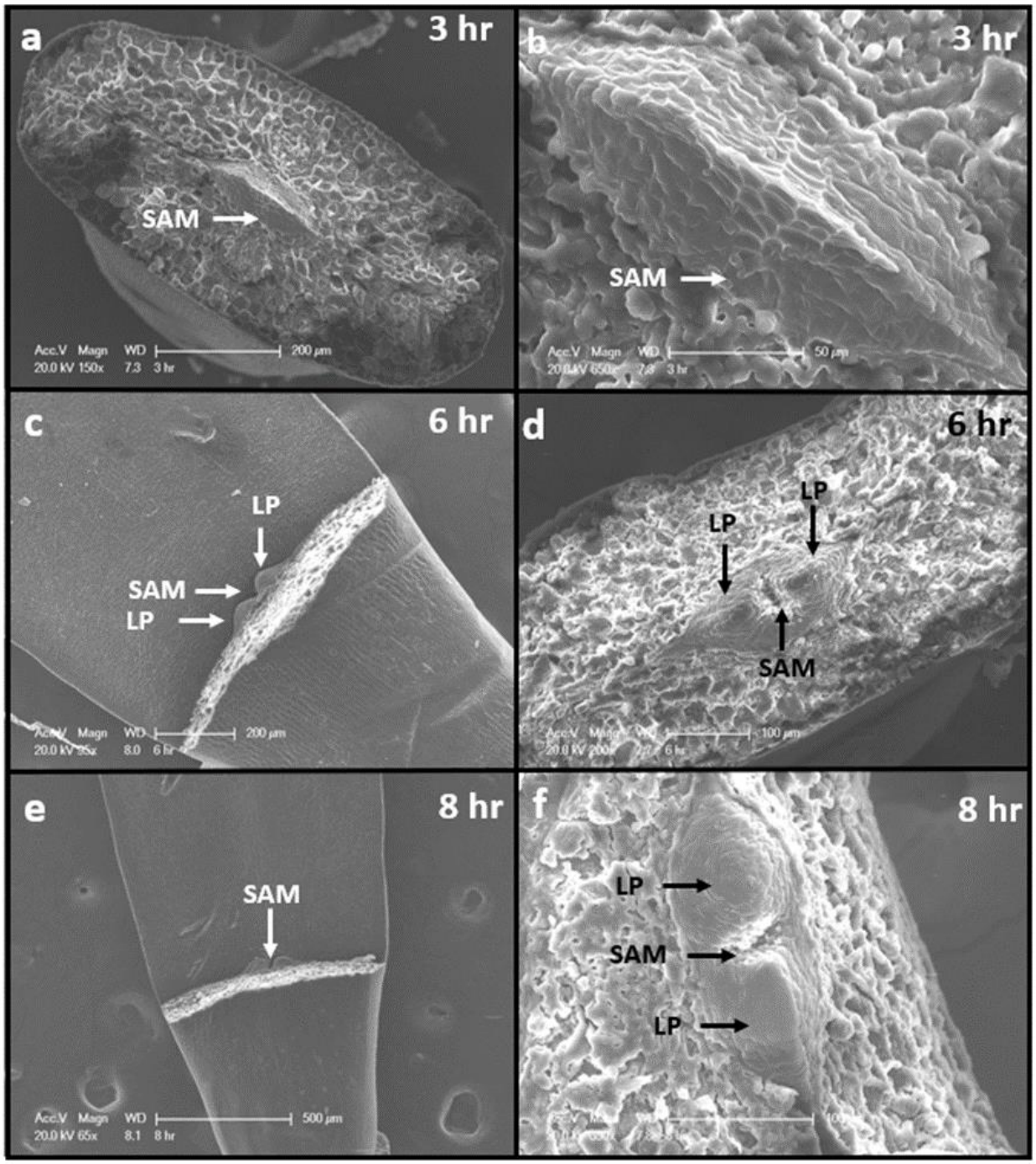


Figure A1.2. SEMs of shoot apical meristems 3, 6, and 8 hours after imbibition. **a**) Top view of the shoot apical meristem (SAM) after removal of both cotyledons 3 hours after imbibition **b**) Magnification of the SAM and from image a. **c**) The SAM and leaf primordia (LP) after removal of one cotyledon 6 hours after imbibition. **d**) Top view of the after removal of both cotyledons 6 hours after imbibition. **e**) The SAM after removal of one cotyledon 8 hours after imbibition. **f**) Top view of the SAM and LP after removal of both cotyledons 8 hours after imbibition.

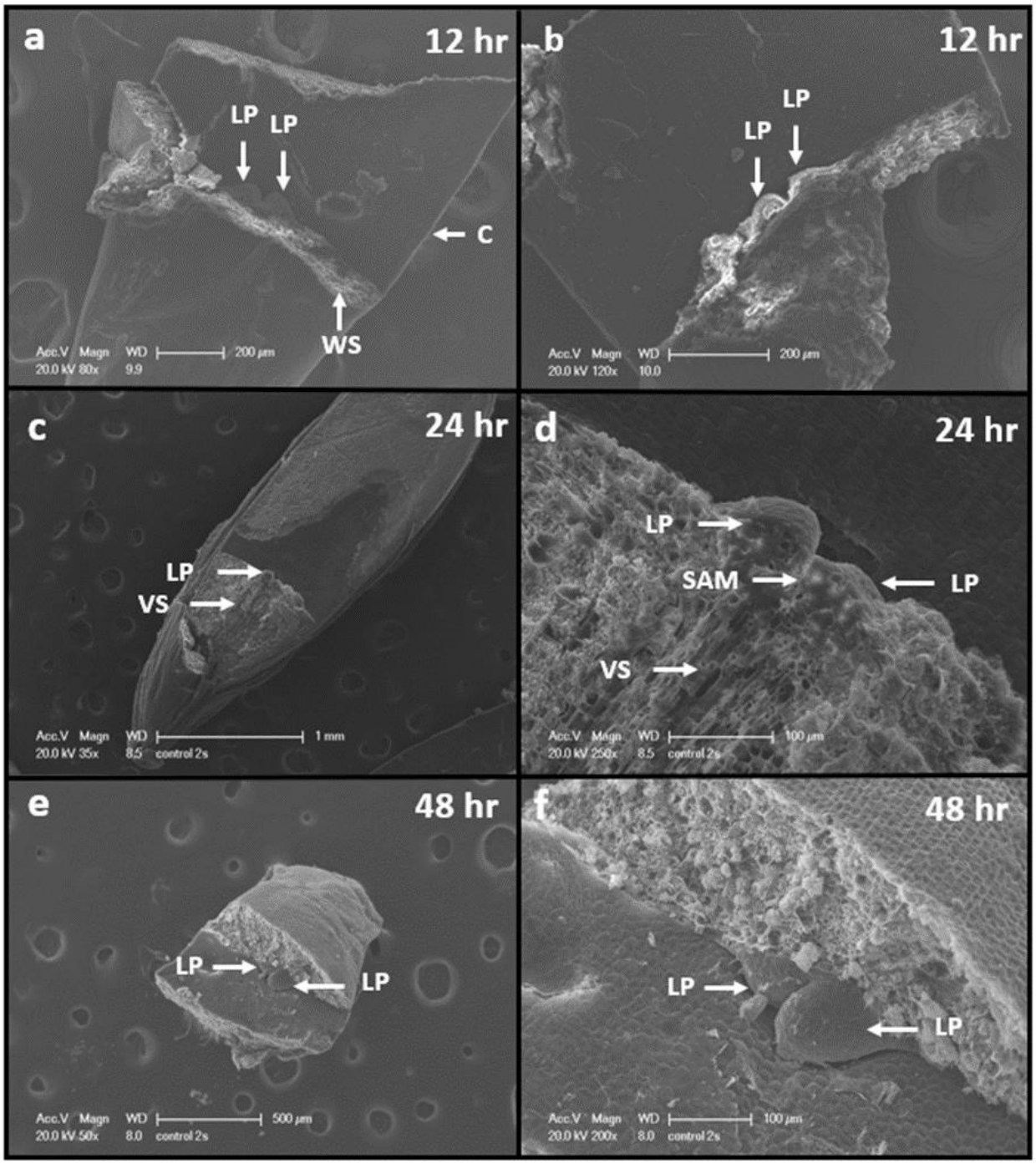


Figure A1.3. SEMs of shoot apical meristems 12, 24, and 48 hours after imbibition. **a**–**b**) The leaf primordia (LP) after removal of both cotyledons 12 hours after imbibition **c**) Longitudinal section of a seed 24 hours after imbibition. **d**) Magnified view of image c showing leaf primordia and vascular strands (VS) originating from the shoot apical meristem (SAM). **e**) LP after removal of both cotyledons 48 hours after imbibition. **f**) Magnified view of image f displaying LP covering the SAM, reducing access.

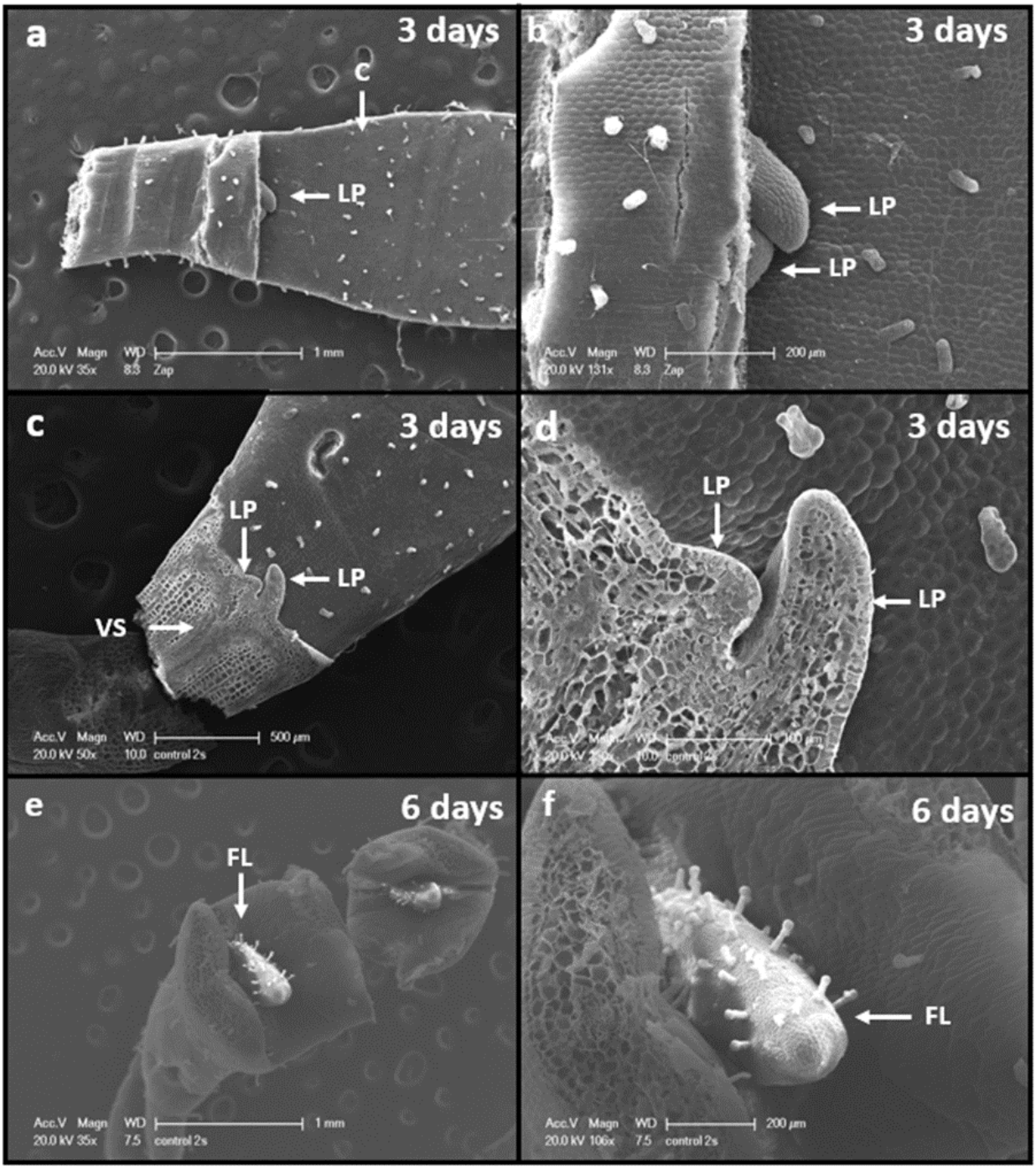


Figure A1.4. SEMs of shoot apical meristems 3 and 6 days after imbibition. **a**) The leaf primordia (LP) after removal of one cotyledon 3 days after imbibition. **b**) LP covering the SAM 3 days after imbibition. **c**) Longitudinal section of a seedling 3 days after imbibition showing LP and vascular strands (VS). **d**) Magnified view of image c showing leaf primordia 3 days after imbibition. **e**) First true leaf (FL) emerging 6 days after imbibition. **f**) Magnified view of image f displaying FL emergence 6 days after imbibition.

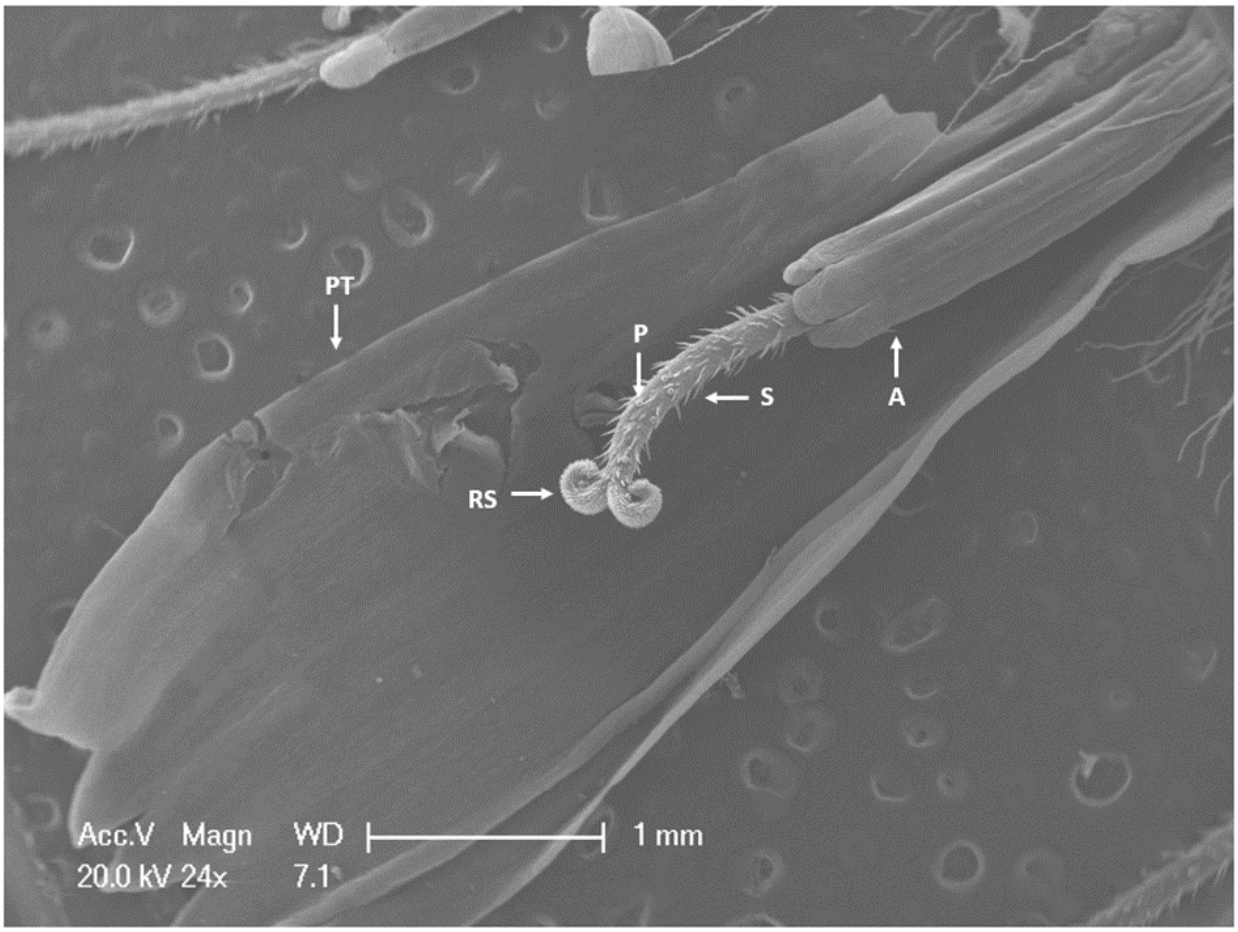


Figure A1.5. A single flower from an inflorescence of lettuce. Shown are the petal (PT), receptive stigma (RS), pollen grains (P), style (S), and anthers (A).

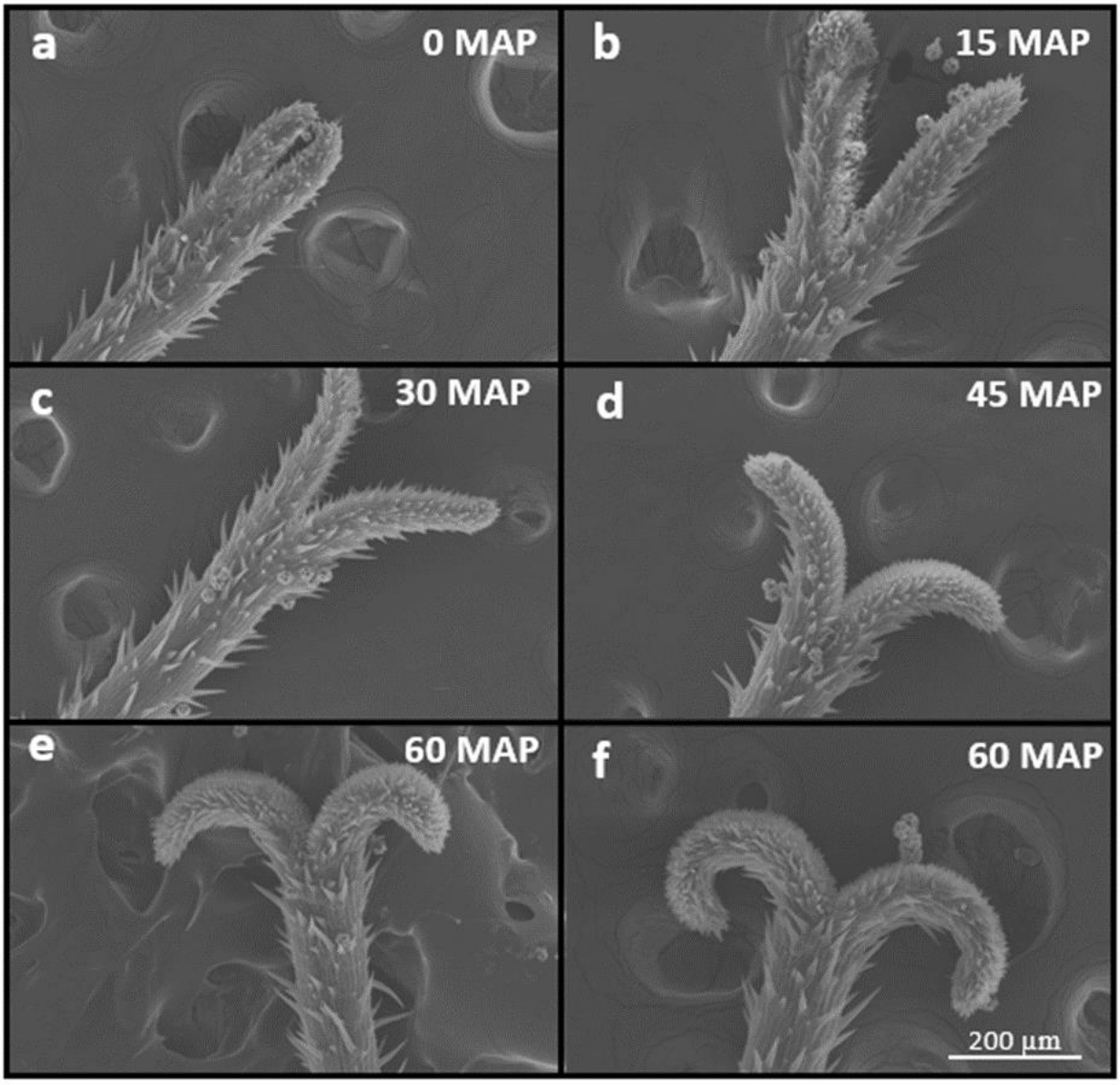


Figure A1.6. Time course of the stigma bifurcation after floral opening. Images represent the stigma immediately after artificial pollination/floral opening (a), 15 minutes after pollination (MAP) (b), 30 MAP (c), 45 MAP (d), and 60 MAP (e,f).

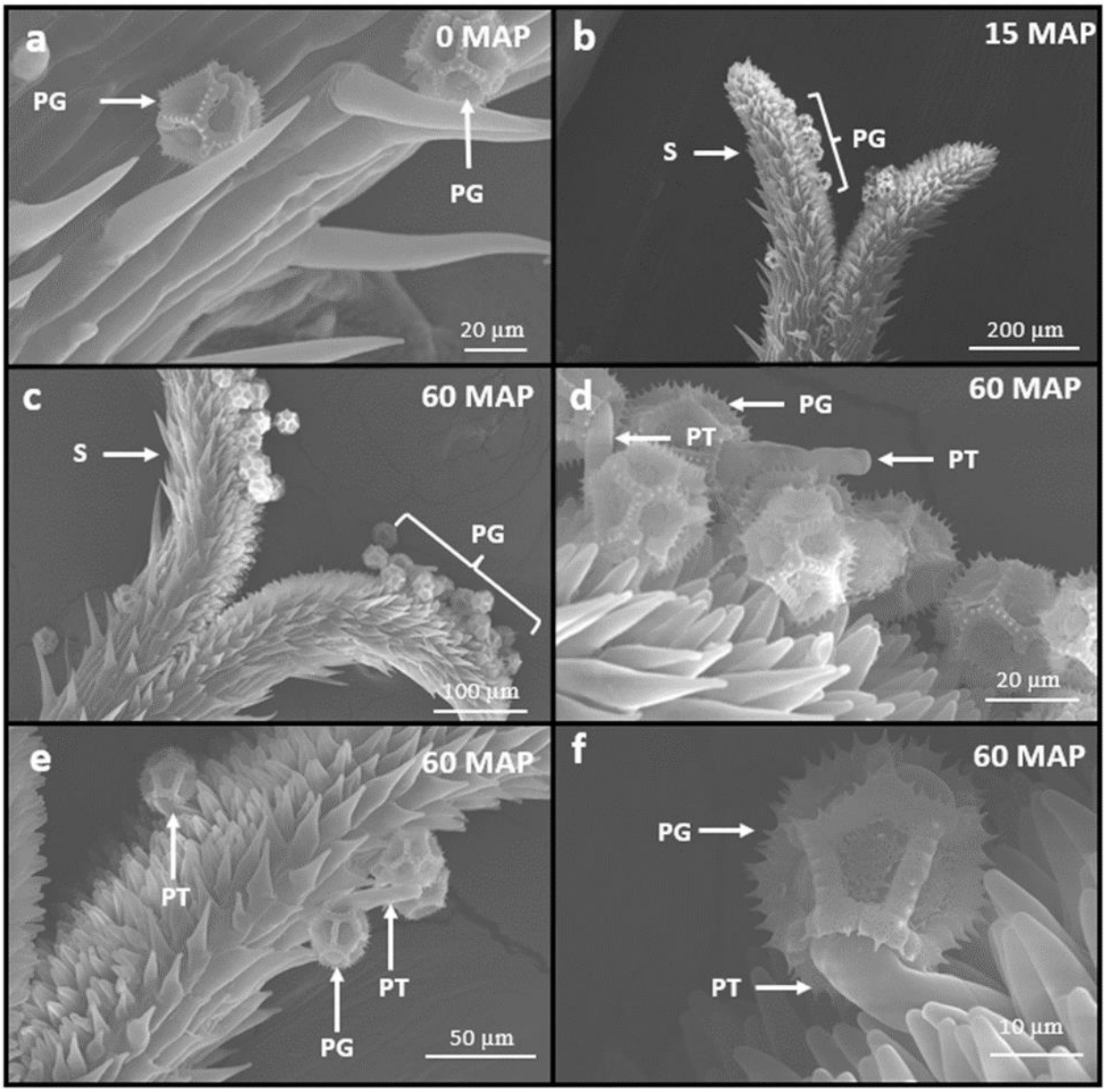


Figure A1.7. Progression of pollen germination over a period of 60 minutes. **a**) Two pollen grains (PGs) attached to the stigma immediately after pollination. **b**) Multiple PGs attached to the stigma (S) 15 minutes after pollination. **c–f**) Images representing PG germination and pollen tube (PT) growth 60 minutes after germination.

REFERENCES

Chang, S. S., Park, S. K., Kim, B. C., Kang, B. J., Kim, D. U., and Nsm, H. G. (1994). Stable genetic transformation of *Arabidopsis thaliana* by *Agrobacterium* inoculation *in planta*. In *The Plant Journal*, *5*(4).

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, *16*(6), 735-743.

Katavic, V., Haughn, G. W., Reed, D., Martin, M., and Kunst, L. (1994). *In planta* transformation of *Arabidopsis thaliana*. *Molecular and General Genetics MGG*, *245*, 363–370.

Lukasz, P., and Piosik, L. (2013). Haploid embryos of lettuce (*Lactuca sativa*) induced by alien pollen or chemical factors. *African Journal of Biotechnology* , *12*(4), 345–352.

Maher, M. F., Nasti, R. A., Vollbrecht, M., Starker, C. G., Clark, M. D., and Voytas, D. F. (2020). Plant gene editing through *de novo* induction of meristems. *Nature Biotechnology*, *38*, 84–89.

Zia, M., Bibi, Y., and Nisa, S. (2014). Does Agro-injection to soybean pods transform embryos? *Plant Omics Journal*, *4*(7), 384-390.

Appendix B: Quantitative trait loci mapping for seed traits and seed viability using the Armenian 999 (*Lactuca serriola* L.) x PI251246 (*L. sativa* L.) recombinant inbred line (RIL) population

ABSTRACT

Seed traits, such as seed weight, seed color, and seed viability, have an impact on crop productivity and marketability. In the last two decades, quantitative trait locus (QTL) mapping has been used to study genetic loci associated with seed traits in lettuce. In addition, domestication of crops typically leads to changes in seed weight, seed texture, and seed color. In this study, a mapping population derived from a cross between a wild lettuce (*Lactuca serriola*) and a cultivated lettuce (*L. sativa*) was used to identify QTLs associated with seed weight, seed color, and seed viability. Across all traits, a total of seven QTLs were identified, which explained 9.5 to 20.9% of the phenotypic variance. The QTLs identified were similar to previous QTLs reported for the same traits, and one QTL co-localized with a QTL associated with callus formation using the same mapping population and genetic map. The results from this study provide evidence supporting existing QTLs and reveal new QTLs associated with seed traits in lettuce.

INTRODUCTION

Cultivated lettuce (*Lactuca sativa*) is consumed mainly as a fresh leafy vegetable; however, some varieties, such as *L. sativa* oil seed PI2512146, are still produced in Egypt for seed oil content (Harlan, 1986; Ryder, 1999). Prior to domestication, lettuce grew wild with one of the direct ancestors being *L. serriola.* Wild lettuce varieties tend to have smaller seeds and exhibit seed shattering (Harlan, 1986; Ryder, 1999). As seen with other species, the domestication of lettuce has led to an increase in seed size and consequently seed weight (Hartman et al., 2013). Domestication of lettuce has also led to changes in seed color and seed texture.

Other seed traits, such as seed quality and seed viability play major roles in crop productivity and uniformity in lettuce. Multiple environmental factors (e.g., temperature, humidity, light) influence seed longevity and seed aging in lettuce. Previous studies have identified mechanisms that affect seed aging in other species such as lipid peroxidation, ROS-induced DNA damage, and DNA methylation (Michalak et al., 2015; Ratajczak et al., 2015; Waterworth et al., 2015). The most widely used method to test for seed viability is through germination assays to assess the quality and normality of the cotyledons. Specifically in lettuce, seed testers assay for "red cotyledon," which is considered as an abnormal phenotype that results in the removal of seeds prior to planting.

Seed aging and other seed related traits are under genetic control in lettuce. Previously, the *L. sativa* cv Salinas x *L. serriola* accession UC96US23 mapping population was used to identify multiple QTLs associated with seed related traits including seed longevity, seed weight, seed length, and seed width (Argyris et al., 2005; Hartman et al., 2013; Johnson et al., 2000; Schwember and Bradford, 2010a; Wei et al., 2014). QTLs associated with domestication traits in lettuce tend

to cluster rather than be uniformly distributed over the genome; however, specific seed-related traits (seed weight, seed length, and seed weight) have not been shown to co-locate together (Hartman et al., 2013). QTLs were also revealed for seed longevity under conventional and controlled deterioration storage conditions (Schwember and Bradford, 2010b). In addition, mapping populations have been used to study seed traits in many other species including tomato, soybean, rice, and sunflower (Brummer et al., 1997; Doganlar et al., 2000b; Miura et al., 2002).

In this study, I used a population developed from a cross between an oil seed of *L. sativa* and an accession of *L. serriola* to identify QTLs associated with seed weight, seed color, and seed viability. Three QTLs were revealed for seed weight, one QTL for seed color, and two QTLs for seed longevity. The identification and analysis of seed weight, seed color, and seed longevity in multiple mapping populations allows for a better understanding of genetic loci regulating seed traits in lettuce. In addition, the knowledge gained from this study can be applied to other important seed and oil crops of the Compositae family such as sunflower.

METHODS

Analysis of seed viability

An F_7 population of 236 recombinant inbred lines (RILs) was produced from single seed descent from a cross between *L. serriola* accession Armenian 999 and *L. sativa* oil seed PI251246. For seed viability assays, a total of 84 RILs were randomly selected from the population. The 84 selected RILs were randomly assigned to six batches. Each batch consisted of 14 RILs and both parents, for a total of 16 genotypes per batch. For each batch, seeds were randomly assigned a position (positions 1 through 16) on the Petri plate. After phenotyping the first six batches, the

original 84 RILS were randomly assigned to six more batches, two more times for a total of three replications per RIL and 18 replications per parental genotype.

Seeds were surface sterilized using 20% bleach and 5 µL of Tween20. The seeds were soaked for 20 minutes with constant agitation (250 rpm) and rinsed with 100 mL of sterile distilled water. Twenty seeds were sown on 1/2x Hoagland's medium (0.815 g/L Hoagland's salt mixture + 8 g/L Phytoagar™) in Corning® square bioassay Petri dishes (Product No. 431301) and incubated in a grow room at 21°C for 4 days under a 9/15 hr light/dark cycle and LED (Fluence Bioengineering, INC) lights providing approximately 13,000 lux.

After four days, seed viability was scored based on the health of the emerged cotyledons (germination score). Each seedling was scored on a scale from one to three, with one being little to no germination present and necrotic cotyledons, and three being healthy, green cotyledons (**Figure A2.1**). Phenotypes for each replication were calculated by averaging the score of 20 of the seedlings for each RIL and parental genotype. The mean of the three replications was used for the QTL analysis.

Seed weight and seed color phenotyping

Seed weight was calculated for each RIL of the population. The weight of twenty-five seeds was calculated and divided by 25 to get the mean group weight of the seeds. This was repeated five more times for a total of six group means per RIL. The group means were then averaged to get a final mean used for QTL mapping. Seed color was scored on a binary light/dark scale with 0 referring to dark seeds and 1 referring to light seeds (**Figure A2.1**).

Data analysis

Descriptive statistics were calculated for each trait assessed in the population. An analysis of variance (one-way ANOVA, $\alpha = 0.05$) was performed to evaluate the effect of genotype on seed weight, seed color, and seed viability. A one-way ANOVA was also used to identify environmental effects (batch effects) on seed viability. Means were used directly for the QTL analysis. Correlations between seed weight and seed germination were calculated using Pearson's correlation coefficient. Mean weights and germination scores were calculated for both dark and light seeds.

Genetic map and QTL analysis

A genetic map developed using unique k-mer markers from each parent was used. This genetic map is described in detail in Chapter 3. All QTL analyses were performed using the package rqtl in R version 3.6.2 (2019-12-12). Initially, simple interval mapping using the scanone function was used to identify putative QTLs. Traits showing putative QTLs with LOD scores exceeding the genome-wide log of odds (LOD) ratio statistic threshold ($\alpha = 0.05$) determined using a 1,000-iteration permutation test were further analyzed using multiple QTL (MQTL) mapping. MQTL mapping using imputed data (1,000-iteration imputation) was performed using the scantwo and stepwiseqtl functions in R. The stepwise QTL uses a stepwise regression model up to 10 steps of forward selection followed by backward elimination to the null model. The model with the highest LOD score was selected. Details of individual terms of each QTL model were calculated using the makeqtl and fitqtl functions in R. Final QTL model terms were selected if the LOD scores exceeded the genome-wide pLOD threshold and represented at least 5% of the phenotypic variance explained (PVE). QTLs with a PVE greater than 10% were considered major

QTLs. To identify the locations of each QTL, a 1.5-LOD support interval was calculated using the refineqtl function in R. QTLs were named using "q" followed by three letters representing the phenotype (SWT for seed weight, SCL for seed color, and SVB for seed viability) and the chromosome. For example, a QTL identified on Chromosome 4 for seed weight would be named *qSWT4*.

RESULTS

Analysis of seed traits in lettuce

The seed weight, seed color, and seed viability were studied using a RIL population derived from a cross between *L. serriola* accession Armenian 999 and *L. sativa* oil seed PI251246. Continuous distributions were observed for both seed weight and seed viability, with a seed viability showing a bimodal distribution (**Figure A2.2**). Approximately 48% and 52% of the population was observed to have light and dark colored seeds, respectively. Light colored seeds had a significantly larger seed weight ($p < 0.001$, $\alpha = 0.05$) and a higher germination score ($p < 0.001$, $\alpha = 0.05$) than dark seeds (**Figure A2.2**). On average, PI251246 had a germination score 1.25 times higher than Armenian 999, which differed significantly ($p < 0.001$, $\alpha = 0.05$) (**Figure A2.3**). The seed germination assay of the RIL population resulted in seed viability values ranging from 0.133 to 2.133 with means differing significantly based on genotype ($p < 0.001$, $\alpha = 0.05$) (**Figure A2.3**). An ANOVA showed no presence of a significant batch effect within the germination assays ($p = 0.223$, $\alpha = 0.05$). A correlation test resulted in a Pearson's correlation coefficient of 0.456 between seed weight and seed viability (**Figure A2.2**).

Genetic loci associated with seed traits

Simple interval mapping was used to detect putative QTLs associated with seed weight, seed color, and seed viability. Putative QTLs were revealed on Chromosomes 2, 4, 7, and 9 for seed weight, Chromosome 2 for seed color, and Chromosome 6 for seed viability (**Figure A2.4**). Further analysis of each trait was conducted using MQTL mapping

Seed weight

Multiple QTL mapping initially revealed four QTLs associated with seed weight on Chromosomes 3, 4 (two QTLs), and 7 (**Figure A2.4**). However, after further analysis, the QTL on Chromosome 3 explained only 3.7% of the phenotypic variance and was dropped from the model. The final model, consisting of three terms, had a LOD score of 25.0 and a PVE of 38.8% (**Table A2.1**). One major QTL was identified on Chromosome 4, *qSWT4.1*, which had a LOD score of 12.5 and explained 17.1% of the phenotypic variance. A second QTL on Chromosome 4, *qSWT4.2*, and a QTL on Chromosome 7, *qSWT7*, had LOD scores of 7.38 and 7.42, respectively, and each explained 9.5% of the phenotypic variation. Support intervals were calculated using a 1.5 LOD-drop interval for each QTL. Intervals for *qSWT4.1*, *qSWT4.2*, and *qSWT7* covered 5.995, 15.471, and 8.146 cM, respectively. For all QTLs detected, the PI25146 allele contributed to higher seed weight (**Table A2.2**).

Seed color

The model identified after MQTL mapping contained 10 individual terms; however, after selecting only QTLs with a LOD score that exceeded the pLOD genome-wide significance threshold and explained at least 5% of the phenotypic variation, only one term remained (**Figure**

A2.4). This QTL was located on Chromosome 2 and had a LOD score of 4.5 and a PVE of 16.5%. The 1.5 LOD support interval covered 25.776 cM (**Table A2.1**). The PI251246 allele was responsible for contributing to lighter colored seeds (**Table A2.2**).

Seed viability

Multiple QTL mapping initially revealed 10 QTLs associated with seed viability, but after further analysis only three QTLs were selected for the final model (**Figure A2.4**). These QTLs were located on Chromosomes 2, 4, and 6. The QTL on Chromosome 2, *qSVB*, had a LOD score of 2.9 and a PVE of 10.2%. The QTL on Chromosome 4, *qSVB4*, had a LOD score of 3.6 and a PVE of 12.9%. The QTL revealed on Chromosome 6 had the largest LOD (5.5) and PVE (20.9%) of all QTLs identified. Support intervals for each QTL ranged from 14.176 to 90.659 cM. For all QTLs detected, the PI251246 allele contributed to better seed viability (**Table A2.2**).

DISCUSSION

Seed traits, such as seed viability and seed longevity, play roles in crop productivity in lettuce. In addition, seed weight and seed color have impacts in crop marketability of seed crops such as rice, maize, soybean, and peanut; therefore, studying different seed traits in lettuce could help us identify genetic markers to improve these useful traits. Using the *L. serriola* accession Armenian 999 x *L. sativa* oil seed PI251246 mapping population, significant QTLs were identified for seed weight, seed color, and seed viability. Previous studies have identified significant QTLs associated with both seed weight and seed viability (reported as seed longevity) (Argyris et al., 2005; Hartman et al., 2013; Schwember and Bradford, 2010a, 2010b), but no studies have mapped genetic loci associated with seed color in lettuce. In addition, previous reports of these traits were studied using the *L. sativa* cv. Salinas x *L. serriola* accession UC96US23 population.

Significant QTLs revealed for seed weight in this study differ from QTLs identified in other mapping populations. Three significant QTLs were identified on Chromosomes 4 and 7 using the Armenian 999 x PI251246 mapping population. A previous study investigating domestication traits in lettuce using the Salinas x UC96US23 mapping population revealed one QTL on Chromosome 1 associated with seed weight (Hartman et al., 2013). This QTL explained approximately 17% of the phenotypic variation, which is similar to *qSWT4.1*. In addition to *qSWT4.1*, two more QTLs, *qSWT4.2* and *qSWT7*, were identified that both explained 9.5% of the phenotypic variation. Multiple significant QTLs identified for seed weight in lettuce is also consistent with seed weight being under the control of multiple genetic loci in other crops such as soybean, mung bean, chickpea, rapeseed, and tomato (Doganlar et al., 2000a; Hossain et al., 2010; Humphry et al., 2005; Teng et al., 2009; Yang et al., 2012).

QTLs associated with seed viability are similar to those previously reported for seed longevity and plant quality (Argyris et al., 2005; Schwember and Bradford, 2010b). In this study, three significant QTLs were identified on Chromosomes 2, 4, and 6. Previously, QTLs significantly associated with seed longevity after high resolution mapping were identified on Chromosomes 1 and 4. One QTL identified on Chromosome 4, *Slg4.1*, was located at a similar genetic position as *qSVB4*. In addition, other QTLs identified for plant quality (*Pq2.2*) and high temperature germination (*Htg6.1*) had similar genetic positions to *qSVB2* and *qSVB6* (Argyris et al., 2005). However, little can be inferred between the QTLs in this study and the QTLs previously identified because they were revealed using different genetic maps and mapping populations.

A QTL identified for seed color in lettuce co-localizes with a QTL identified for callus production using the same mapping population (Chapter 3). One QTL was identified for seed color, *qSCL2*, which explained approximately 17% of the phenotypic variation. The peak of the

QTL revealed for callus production, *qCFB2* (Chapter 3), was located 1.713 cM away from the peak of *qSCL2*. In addition, the 1.5-LOD intervals for each QTL overlapped. This is interesting because it connects an *in vitro* regeneration trait to seed color in the same mapping population; therefore, we could potentially predict callus formation based on the seed color of each RIL.

Genetic loci for germination, seed longevity, and seed weight have been studied previously using the Salinas x UC96US23 mapping population Argyris et al., 2005; Hartman et al., 2013; Schwember and Bradford, 2010a, 2010b. Using the Armenian 999 x PI251246 population, similar and new QTLs were identified for these traits. In addition, a QTL was revealed for seed color, which co-localized with a QTL identified previously for callus formation. The information presented here could be investigated further to identify candidate genes for each trait discussed. In addition, other seed traits, such as seed length and seed width, could be measured and compared to previous reports in the Salinas x UC96US23 mapping population. Because PI251246 is an oil seed cultivar, oil content could also be measured and studied in this population. Furthermore, this information could be used to identify the potential co-localization with other regeneration QTLs, which would allow us to predict *in vitro* regeneration phenotypes based on seed traits.

FIGURES AND TABLES

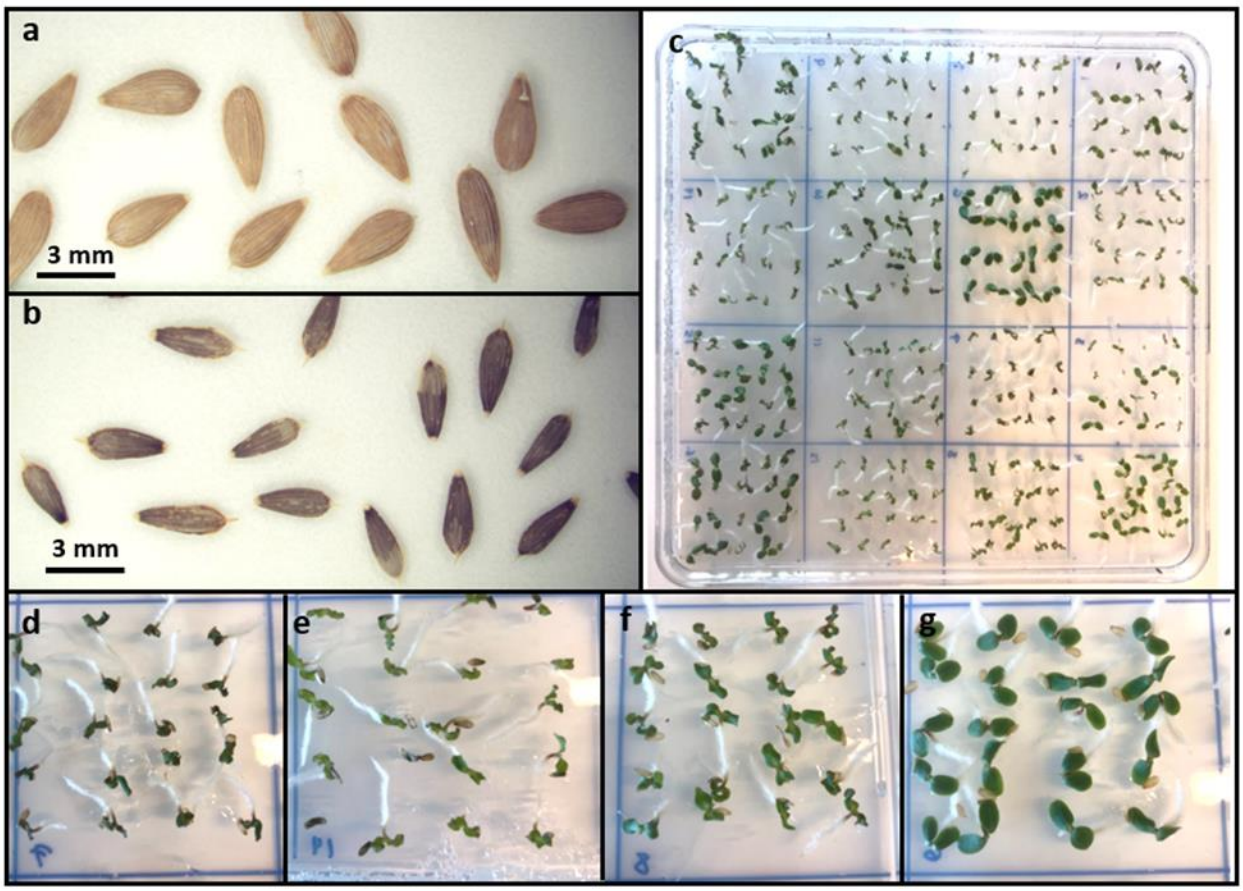


Figure A2.1. Phenotypes of seed color and seed viability analyzed in this study. **a, b)** Light (a) and dark (b) colored seeds from two different RILs of the Armenian 999 x PI25146 population. **c)** The Petri plate layout of one batch from the seed germination assay. **d–g)** Magnified sections from image c displaying different seed germination phenotypes scored as 0 to 3 from left to right.

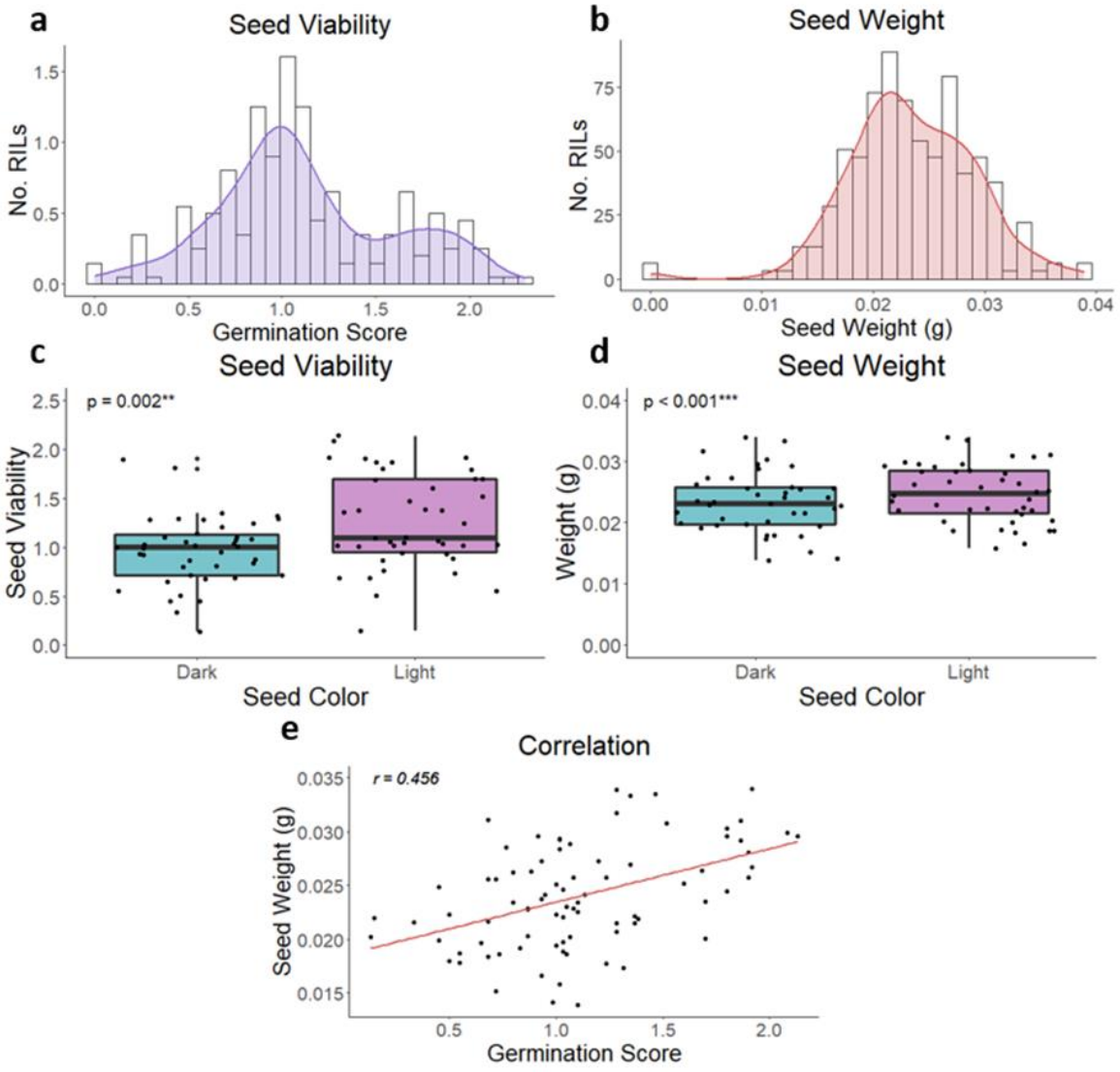


Figure A2.2. The distributions and relationships of seed weight, seed color, and seed viability. **a)** The distribution and density plots of seed viability. **b**) The distribution and density plots of seed weight. **c, d)** Boxplots representing mean seed viability scores (c) and mean seed weight (d) based on seed color; p-value is the result of a Welch's t-test ($\alpha = 0.05$). **e)** The correlation between seed weight and seed viability (germination score) using Pearson's correlation coefficient (r).

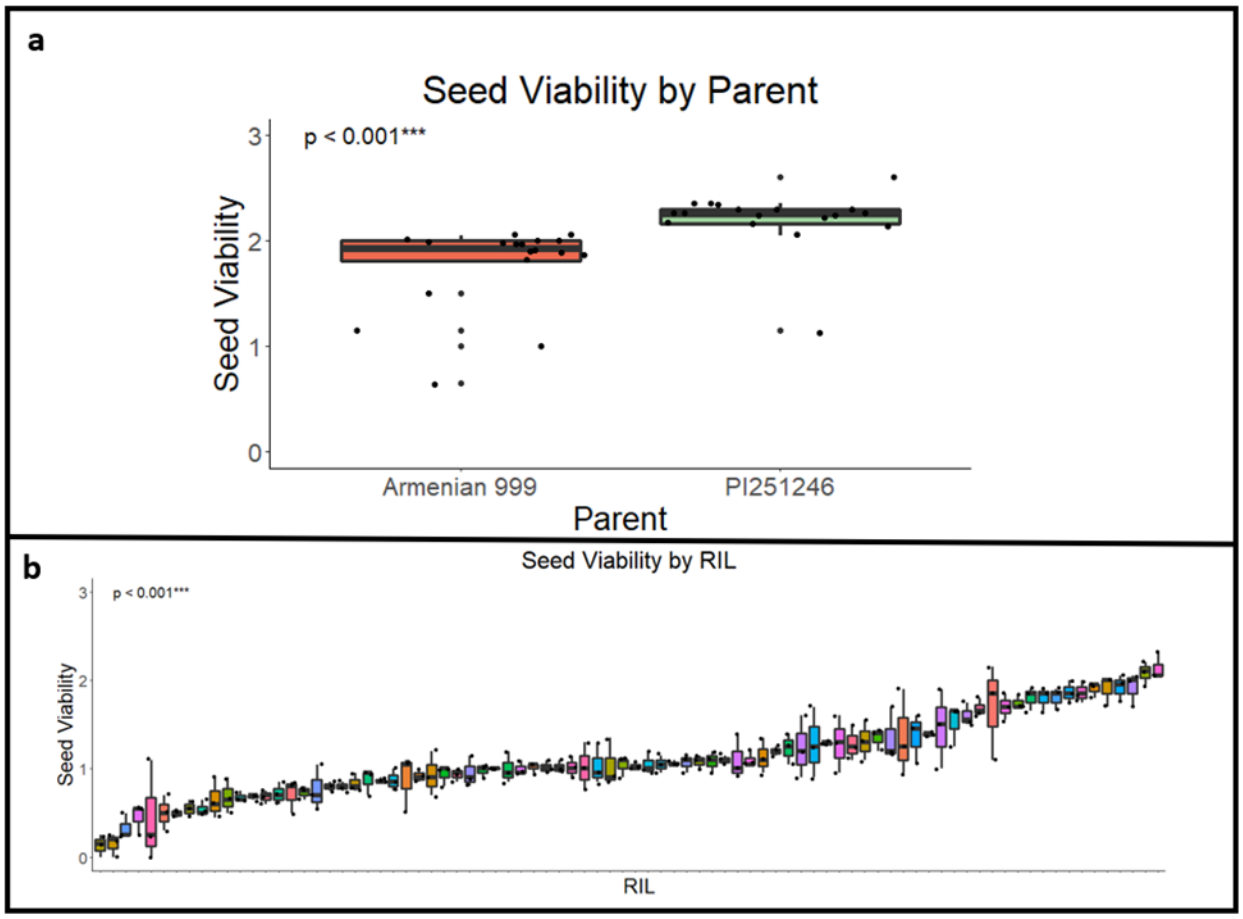


Figure A2.3. Seed viability genotypes of Armenian 999, PI251246, and the RIL population. **a)** Boxplot showing means and variance of seed viability of both parents. The p-value is the result of a Welch's t-test ($\alpha = 0.05$). **b)** Boxplot displaying seed viability means and variances of 84 RILs. The p-value is the result of a one-way ANOVA ($\alpha = 0.05$).

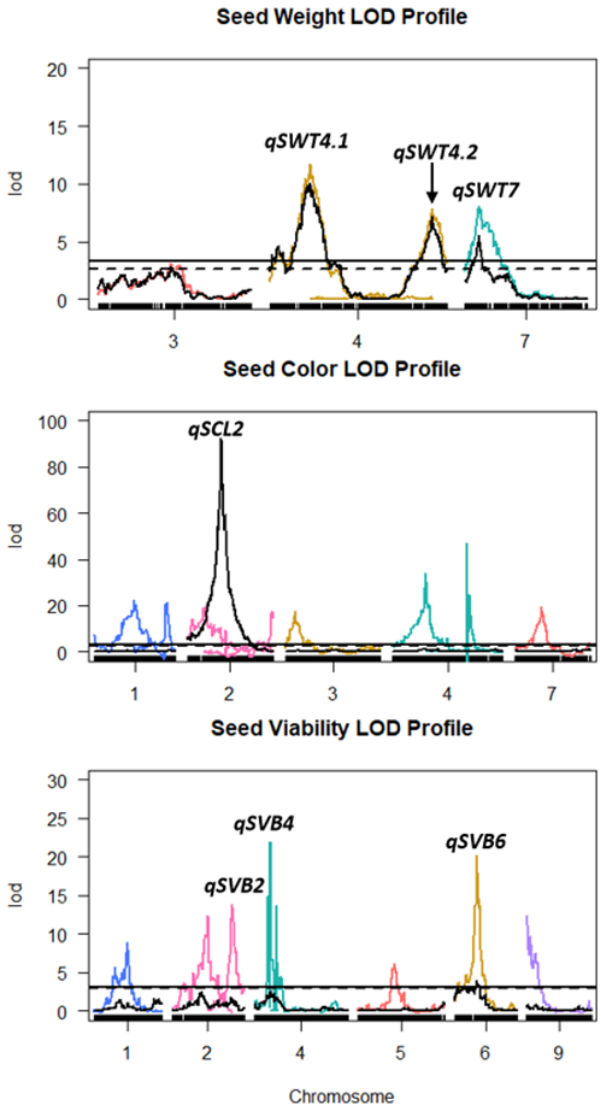


Figure A2.4. Significant associations between seed trait phenotypes and positions in the lettuce genome. The LOD profiles of seed weight, seed color, and seed viability after simple interval mapping (black profiles) and multiple QTL mapping (colored profiles). Only chromosomes with significant associations are shown for each phenotype. Labeled QTLs were selected from the initial MQTL mapping models based on two criteria: 1) LOD scores exceeded the genome wide pLOD threshold and 2) QTLs represented greater than 5% phenotypic variance explained (PVE). Black horizontal lines represent the MQTL mapping genome-wide pLOD significance threshold (solid) and simple interval mapping genome-wide significance threshold (dashed) calculated using a 1,000-iteration permutation test.

Table A2.1. Putative QTLs (peak exceeds or is near the genome wide LOD significance threshold) identified for seed weight (SWT), seed color (SCL), and seed viability (SVB) using simple interval mapping.

Trait	LOD threshold $\alpha = 0.05$	Chr	LOD	Genetic Position (cM)	Marker Nearest Peak
SWT	3.27	2	4.503	79.540	7250_61♂
		4	10.075	54.482	16108_62 ♀
		7	5.516	18.907	7535_88 ♀
		9	3.261	56.468	1063_70♀
SLC	3.17	2	92.272	73.943	5127_68♀
SVB	3.15	6	3.954	56.956	13724_71 ♀

♀ Female marker, ♂ Male marker

Table A2.2. LOD scores and PVE for the full QTL models and individual QTL terms for seed weight (SWT), seed color (SCL), and seed viability (SVB). Markers, genetic (G. peak) positions, physical (P. peak) positions, and a 1.5 LOD interval is represented for each QTL.

Full Model				**QTL**							**1.5 LOD Interval**	
Trait	Model	LOD	PVE	QTL	Chr	LOD	PVE	Marker Close to Peak	G. peak	P. peak	Start (cM)	End (cM)
SWT	y ~ Q2 + Q3 + Q4	25.0	38.8	*qSWT4.1*	4	12.5	17.1	16108_62♀	54.482	307,416,549	52.345	58.340
				qSWT4.2	4	7.4	9.5	2487_62 ♂	218.615	41,144,185	213.472	228.889
				qSWT7	7	7.4	9.5	7535_88 ♀	18.907	28,731,359	15.039	23.185
SCL	y ~ Q3	4.5	16.5	*qSCL2*	2	4.5	16.5	12737_66♀	43.029	59,228,938	28.000	53.776
SVB	y ~ Q5 + Q9	9.8	42.1	*qSVB2*	2	2.9	10.2	15785_73♀	153.000	197,878,602	72.661	163.320
				qSVB4	4	3.6	12.9	10348_61♀†	41.223-43.360	332,002,495	30.453	70.000
				qSVB6	6	5.5	20.9	13724_71♀	56.956	138,940,201	51.371	65.547

†Multiple markers at peak, ♀ Female marker, ♂ Male marker

Table A2.3. Estimated effects of significant QTLs identified for seed weight (SWT), seed color (SCL), and seed viability (SVB).

QTL	Chr	Estimated Effect	Allele †
qSWT4.1	4	0.0023351	PI251246
qSWT4.2	4	0.0017495	PI251246
qSWT7	7	0.0017512	PI251246
qSCL2	2	0.47132	PI251246
qSVB2	2	0.14199	PI251246
qSVB4	4	0.15927	PI251246
qSVB6	6	0.20268	PI251246

† Parental allele contributing to higher regeneration rates

REFERENCES

Argyris, J., Truco, M. J., Ochoa, O., Knapp, S. J., Still, D. W., Lenssen, G. M., Schut, J. W., Michelmore, R. W., and Bradford, K. J. (2005). Quantitative trait loci associated with seed and seedling traits in *Lactuca*. *Theoretical and Applied Genetics*, *111*(7), 1365–1376.

Brummer, E. C., Graef, G. L., Orf, J., Wilcox, J. R., and Shoemaker, R. C. (1997). Mapping QTL for Seed Protein and Oil Content in Eight Soybean Populations. *Crop Science*, *37*(2), 370–378.

Doganlar, S., Frary, A., and Tanksley, S. D. (2000). The genetic basis of seed-weight variation: tomato as a model system. *Theoretical and Applied Genetics 2000 100:8*, *100*(8), 1267–1273.

Han, R., Lavelle, D., Truco, M. J., and Michelmore, R. (2021a). Quantitative Trait Loci and Candidate Genes Associated with Photoperiod Sensitivity in Lettuce (*Lactuca* spp.). *Theoretical and Applied Genetics*, *134*(10), 3473–3487.

Han, R., Wong, A. J. Y., Tang, Z., Truco, M. J., Lavelle, D. O., Kozik, A., Jin, Y., and Michelmore, R. W. (2021b). Drone phenotyping and machine learning enable discovery of loci regulating daily floral opening in lettuce. *Journal of Experimental Botany*, *72*(8), 2979–2994.

Hartman, Y., Hooftman, D. A. P., Eric Schranz, M., and van Tienderen, P. H. (2013). QTL analysis reveals the genetic architecture of domestication traits in Crisphead lettuce. *Genetic Resources and Crop Evolution*, *60*(4), 1487–1500.

Johnson, W. C., Jackson, L. E., Ochoa, O., van Wijk, R., Peleman, J., st. Clair, D. A., and Michelmore, R. W. (2000). Lettuce, a shallow-rooted crop, and *Lactuca serriola*, its wild progenitor, differ at QTL determining root architecture and deep soil water exploitation. *Theoretical and Applied Genetics 2000 101:7*, *101*(7), 1066–1073.

Michalak, M., Plitta-Michalak, B. P., Naskręt-Barciszewska, M., Barciszewski, J., Bujarska-Borkowska, B., and Chmielarz, P. (2015). Global 5-methylcytosine alterations in DNA during ageing of Quercus robur seeds. *Annals of Botany*, *116*(3), 369–376.

Miura, K., Lin, S. Y., Yano, M., and Nagamine, T. (2002). Mapping quantitative trait loci controlling seed longevity in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics 2002 104:6*, *104*(6), 981–986.

Ratajczak, E., Małecka, A., Bagniewska-Zadworna, A., and Kalemba, E. M. (2015). The production, localization and spreading of reactive oxygen species contributes to the low vitality of long-term stored common beech (*Fagus sylvatica* L.) seeds. *Journal of Plant Physiology*, *174*, 147–156.

Schwember, A. R., and Bradford, K. J. (2010a). A genetic locus and gene expression patterns associated with the priming effect on lettuce seed germination at elevated temperatures. *Plant Molecular Biology*, *73*(1–2), 105–118.

Schwember, A. R., and Bradford, K. J. (2010b). Quantitative trait loci associated with longevity of lettuce seeds under conventional and controlled deterioration storage conditions. *Journal of Experimental Botany*, *61*(15), 4423–4436.

Waterworth, W. M., Bray, C. M., and West, C. E. (2015). The importance of safeguarding genome integrity in germination and seed longevity. *Journal of Experimental Botany*, *66*(12), 3549–3558.

Wei, Z., Julkowska, M. M., Laloë, J. O., Hartman, Y., de Boer, G. J., Michelmore, R. W., van Tienderen, P. H., Testerink, C., and Schranz, M. E. (2014). A mixed-model QTL analysis for salt tolerance in seedlings of crop-wild hybrids of lettuce. *Molecular Breeding*, *34*(3), 1389–1400.

Appendix C: Development of a dsRED-tagged TRV-RNA2 construct

ABSTRACT

Direct gene delivery into plants without the use of tissue culture methods is desirable to combat genotype dependent regeneration of many species. One approach is the use of viruses to deliver genes and editing reagents into plant cells without the requirements of tissue culture. Tobacco Rattle Virus (*Tobravirus*; TRV) is an RNA virus that has been used for virus-induced gene silencing (VIGS) in multiple crops, and more recently has been used to deliver guide RNAs into plant cells for genome editing. However, the carrying capacity of TRV and other RNA viruses is limited, making it difficult to produce edits without the presence of stable expressing Cas9 plant lines. Combining compatible RNA viruses for the simultaneous delivery of gRNAs and Cas9 (or other nucleases) is an intriguing approach. Another RNA virus, Lettuce Mosaic Virus (LMV), has been studied extensively in lettuce (*Lactuca sativa* L.), and GFP-tagged constructs of both LMV and TRV are available. In this study, I replaced the GFP tag of TRV-RNA2 with dsRED for distinguishability between the TRV and LMV reporter constructs. By using both viruses tagged with different fluorescent reporters, we can investigate the synergistic effects and movement patterns of co-agroinfiltration of both TRV and LMV.

INTRODUCTION

Efficient technologies for gene delivery in plants is desirable in order to combat genotype-dependent plant regeneration and *Agrobacterium*-mediated transformation. One approach for the elimination of *in vitro* regeneration is through direct gene delivery into plant cells using nanoparticles (Demirer et al., 2019; Lien et al., 2021). A second approach is through viral-mediated gene delivery into plant cells (Ali et al., 2015a; Ali et al., 2015b; Nagalakshmi et al., 2022.; Watson et al., 2012). Virus induced gene silencing (VIGS) exploits the posttranscriptional gene silencing (PTGS) mechanism of plants for viral-mediated silencing of endogenous genes. VIGS has been successfully implemented in multiple species such as *Nicotiana benthamiana*, tomato, and *Arabidopsis* (Watson et al., 2012). In addition, tagging a virus, specifically Tobacco Rattle Virus (TRV), with a fluorescent reporter such as GFP, has allowed for easily traceable viral movement in *N. benthamiana*, *Arabidopsis*, and strawberry (Burch-Smith et al., 2006; Tian et al., 2013).

TRV is an RNA *Tobravirus* composed of two parts (bipartite): 1) TRV-RNA1, which encodes the movement and 2) replicase protein and TRV-RNA2, which encode the coat protein and two non-structural proteins (Tian et al., 2013). TRV-based VIGS vectors have been constructed by cloning the RNA1 and RNA2 into two separate T-DNA vectors and expressed from duplicated cauliflower mosaic virus (35S) promoters and the nopaline synthase (NOS) terminator (Liu et al., 2002). In addition, a TRV-GFP reporter vector was developed by fusing the GFP coding sequence to the 3' end of the coat protein before the terminator. More recently, using *A. tumefaciens* co-infiltration of TRV-RNA1 and TRV-RNA2 mediated delivery of editing reagents into plants stably expressing Cas9 has demonstrated successful genome editing in *Arabidopsis*, *Solanum* spp., *N. benthamiana* and *N. attenuata* and epigenome editing in *Arabidopsis* (Ali et al., 2015a; Ali et al., 2015b; Ghoshal Id et al., 2020; Oh and Kim, 2021).

Although the TRV system is capable of delivering gRNAs, it has a limited cargo carrying capacity. It was recently demonstrated that simultaneous use of compatible RNA viruses delivering either the DNA encoding the Cas12a nuclease or the guide RNAs produced transgene-free genome editing in *N. benthamiana* (Uranga et al., 2021). Another RNA virus, lettuce mosaic virus (LMV; genus *Potyvirus*), has been tagged with both GFP (further referred to as LMV-GFP) and ß-glucuronidase (GUS) to investigate viral accumulation and movement in susceptible and resistant lettuce cultivars (German-Retana et al., 2000). In this appendix, I modify the TRV-RNA2-GFP construct by replacing the GFP coding sequence with dsRED. This will allow for co-infiltrations of both LMV-GFP and TRV-RNA1/TRV-RNA2-dsRED to track viral movement. In addition, we can investigate synergistic effects between TRV and LMV when both are present and replicating in the same cell.

METHODS

The constructs pYL192-TRV-RNA1 and SPDK2686-TRV-RNA2-GFP were received from the Dinesh-Kumar Lab at UC Davis. SPDK2682-TRV-RNA2-GFP and pEG100-dsRED (Chapter 5) were used as templates for cloning. SPDK2686-TRV-RNA2-dsRED was constructed using the In-Fusion Cloning protocol from Takara Bio.

Vector backbone digestion

SPDK2686-TRV-RNA2-GFP was digested with XBaI (New England BioLabs Inc.) in CutSmart Buffer at 37°C for one hour (**Figure A3.1**). Gel electrophoresis with a 1% agarose gel was used to separate the digestion products and the correct size band was gel purified using the Macherey-Nagel NucleoSpin Gel and PCR Clean-up Kit (Product ID# 740609). The purified fragment was used as the backbone vector for In-Fusion cloning.

In-Fusion cloning and bacterial transformation

Individual DNA fragments were amplified using the Takara Cloneamp HiFi PCR Premix Protocol with SPDK2682-TRV-RNA2-GFP and pEG100-dsRED (Chapter 5) used as DNA templates (**Figure A3.1**). Primer sets and PCR conditions are described in **Table A3.1**. An In-Fusion cloning reaction was performed following the manufacturer's protocol and incubated at 50°C for 15 minutes. To linearize any plasmid still retaining the GFP sequence, the In-Fusion reaction mixture was digested with SalI (New England BioLabs Inc.) in Cutsmart buffer at 37°C for one hour. Immediately after digestion, 3 µL of the In-Fusion mixture was transformed into Stellar competent *Escherichia coli* cells using the Clontech Stellar™ Competent Cells protocol PT5055-2. Fifty microliters of the transformation mixture were plated onto Luria-Bertani (LB) agar (20 g/L LB Agar powder, 18 g/L Bacto™ Agar) supplemented with 50 mg/L of spectinomycin and incubated overnight at 37°C.

Colony PCR

Twenty colonies were screened for the presence of SPDK2686-TRV-RNA2-dsRED using primers TB24 and TB25 (**Table A3.1**). Cells of each colony were suspended in 5 µL of sterile distilled water, and 1 µL of the suspension was used for colony PCR using Promega GoTaq® Green Master Mix. One colony showing amplification of the correct size (colony 9) was cultured in liquid LB supplemented with 50 mg/L spectinomycin in an orbital shaker (200 rpm) for 16-18 hours at 37°C. After 16 to 18 hours, plasmid DNA was extracted from the suspension cultures using the QIAGEN QIAprep® Spin Miniprep protocol (Cat. #27104). Plasmid DNA was Sanger sequenced at the College of Biological Sciences UC DNA Sequencing Facility using primers TB24 and TB25. After Sanger sequencing confirmation, plasmid DNA was sent to Massachusetts General Hospital

Center for Computational and Integrative Biology (MGH CCIB) DNA Core for whole plasmid sequencing.

Agrobacterium-mediated transient assay

Sequence confirmed SPDK2686-TRV-RNA2-dsRED plasmid was electroporated into *A. tumefaciens* strain C58. Transformed *A. tumefaciens* (50 µL) was plated on LB agar supplemented with 50 mg/L rifampicin and 50 mg/L spectinomycin and incubated for two days at 28°C. Twenty colonies were screened for the plasmid using colony PCR as described above and primers TB24 and TB25.

Three separate cultures for agroinfiltration were started by inoculating 20 mL of LB supplemented with appropriate antibiotics with pYL192-TRV-RNA1 (50 mg/L Rifampicin and 50 mb/L kanamycin), SPDK2686-TRV-RNA2-GFP (50 mg/L Rifampicin and 50 mg/L spectinomycin), or one positive colony (colony 9) of SPDK2686-TRV-RNA2-dsRED (50 mg/L Rifampicin and 50 mg/L spectinomycin). Cultures were incubated in an orbital shaker (200 rpm) at 28°C for 16–18 hours. All cultures were pelleted, washed, and resuspended in 10 mM $MgCl_2$ to an OD_{600} between 0.4 and 0.5. Leaves of four-week-old *Lactuca sativa* cv. Cobham Green and *N. benthamiana* were infiltrated using a previously optimized protocol for lettuce (Wood et al., 2020; Wroblewski et al., 2005). Immediately prior to agroinfiltration, pYL192-TRV-RNA1/SPDK2686-TRV-RNA2-GFP and pYL192-TRV-RNA1/SPDK2686-TRV-RNA2-dsRED cultures were made by mixing individual components in a 1:1 ratio. Plants were incubated for three days in a 24°C growth room to allow for transient expression of the transgenes. After three days, expression of *dsRED* in infiltrated leaves was observed using confocal microscopy.

RESULTS

All 20 Stellar *E. coli* colonies screened after transformation with the digested In-Fusion mixture showed amplification of the plasmid target (**Figure A3.2**). Sanger sequencing indicated a G-to-A point mutation at position 419 in the dsRED coding sequence (**Figure A3.2**); however, whole plasmid sequence verification of the SPDK2686-TRV-RNA2-dsRED resulted in no mutations, and assembly of fragments occurred as expected. Transient expression of both GFP and dsRED was observed in leaves three days after co-agroinfiltration of TRV-RNA1 and TRV-RNA2 (**Figure A3.3**).

DISCUSSION

TRV has been used to deliver reagents for VIGS and guide RNAs for genome and epigenome editing in *Arabidopsis* (Ali et al., 2015a; Ali et al., 2015b; Ghoshal Id et al., 2020; Oh and Kim, 2021). Although successful genome editing through TRV-mediated delivery of gRNAs has been demonstrated, the carrying capacity of RNA viruses is limited; therefore, the use of a second RNA virus for delivery of gRNAs and Cas9 simultaneously would result in transgene free, genome-edited plants.

In this appendix, I modified TRV-RNA2-GFP by replacing GFP with dsRED. Successful expression of dsRED after co-agroinfiltration of TRV-RNA1 and TRV-RNA-dsRED was observed. Having LMV and TRV tagged with different reporters will allow us to simultaneously study infection and movement patterns of both TRV-dsRED and LMV-GFP. In addition, after further optimization of conditions, we could explore the simultaneous delivery of gRNAs using TRV and Cas9 using LMV.

FIGURES AND TABLES

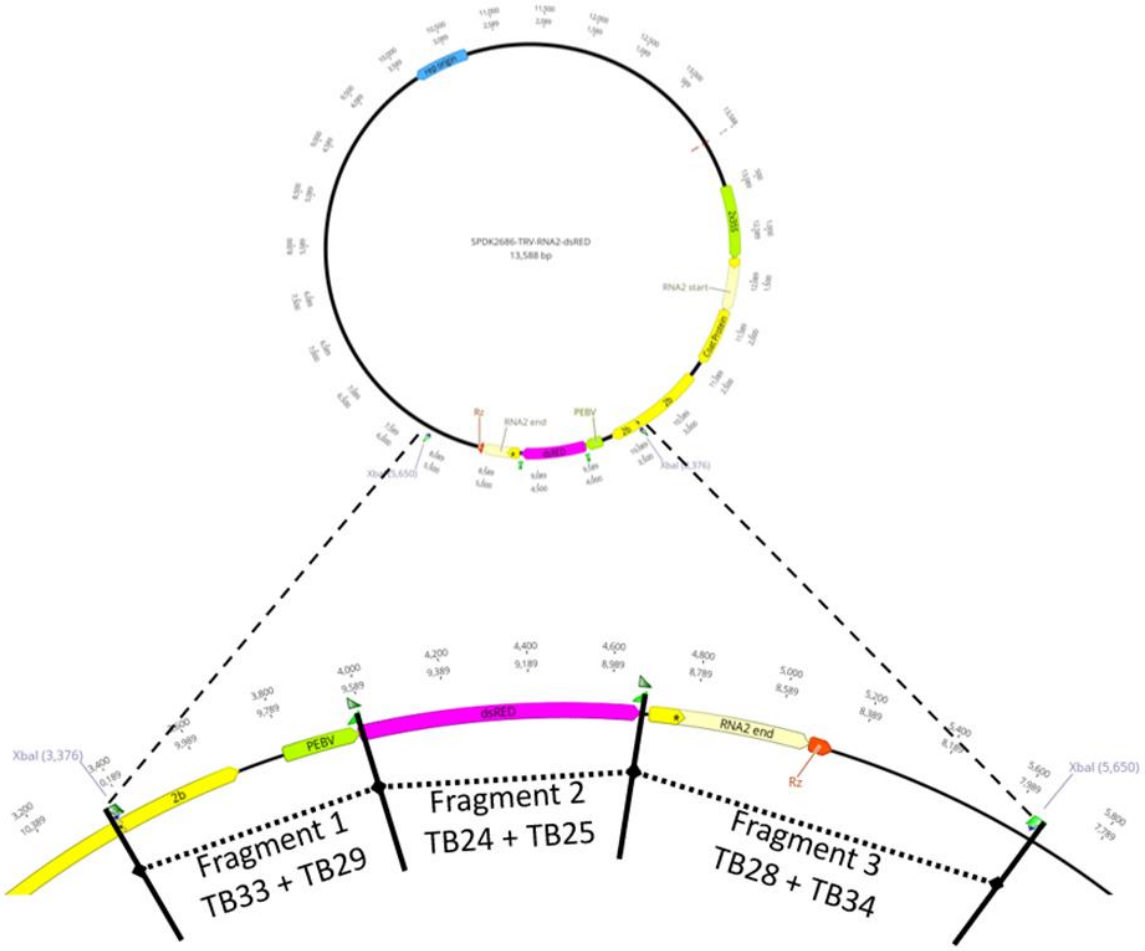


Figure A3.1. Restriction sites (XbaI) for vector digestion and fragments amplified for In-Fusion cloning. Numbers under each fragment correspond to the primer sets used to amplify each fragment.

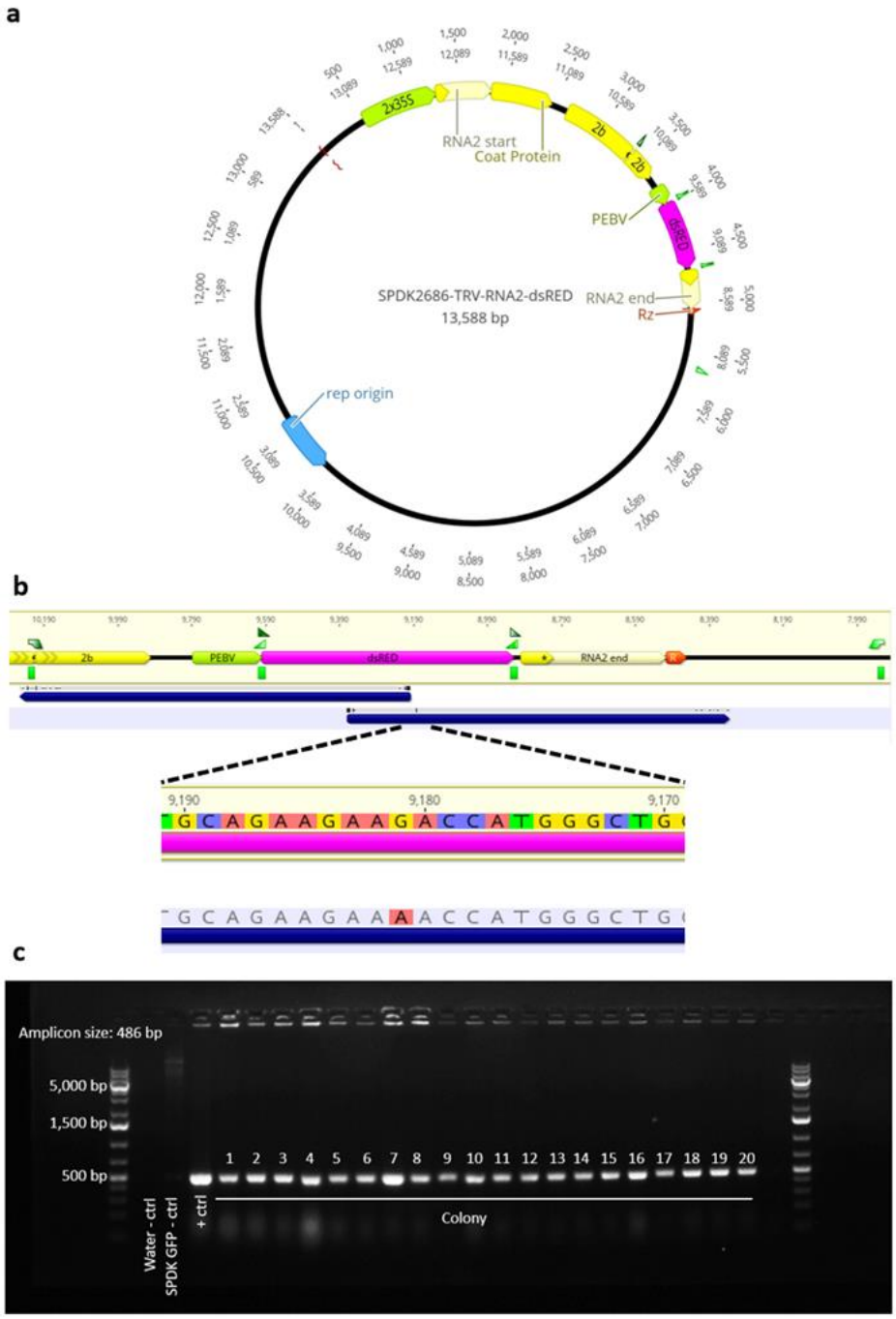


Figure A3.2. SPDK-TRV-RNA2-dsRED construct cloning and sequencing results. **a)** Plasmid map and annotations of SPDK-TRV-RNA2-dsRED. **b)** Observation of a G-to-A point mutation in the dsRED coding sequence after Sanger sequencing of clones. **c)** An agarose gel showing amplification of dsRED after transformation of Stellar competent cells with the digested In-Fusion cloning mixture and colony PCR.

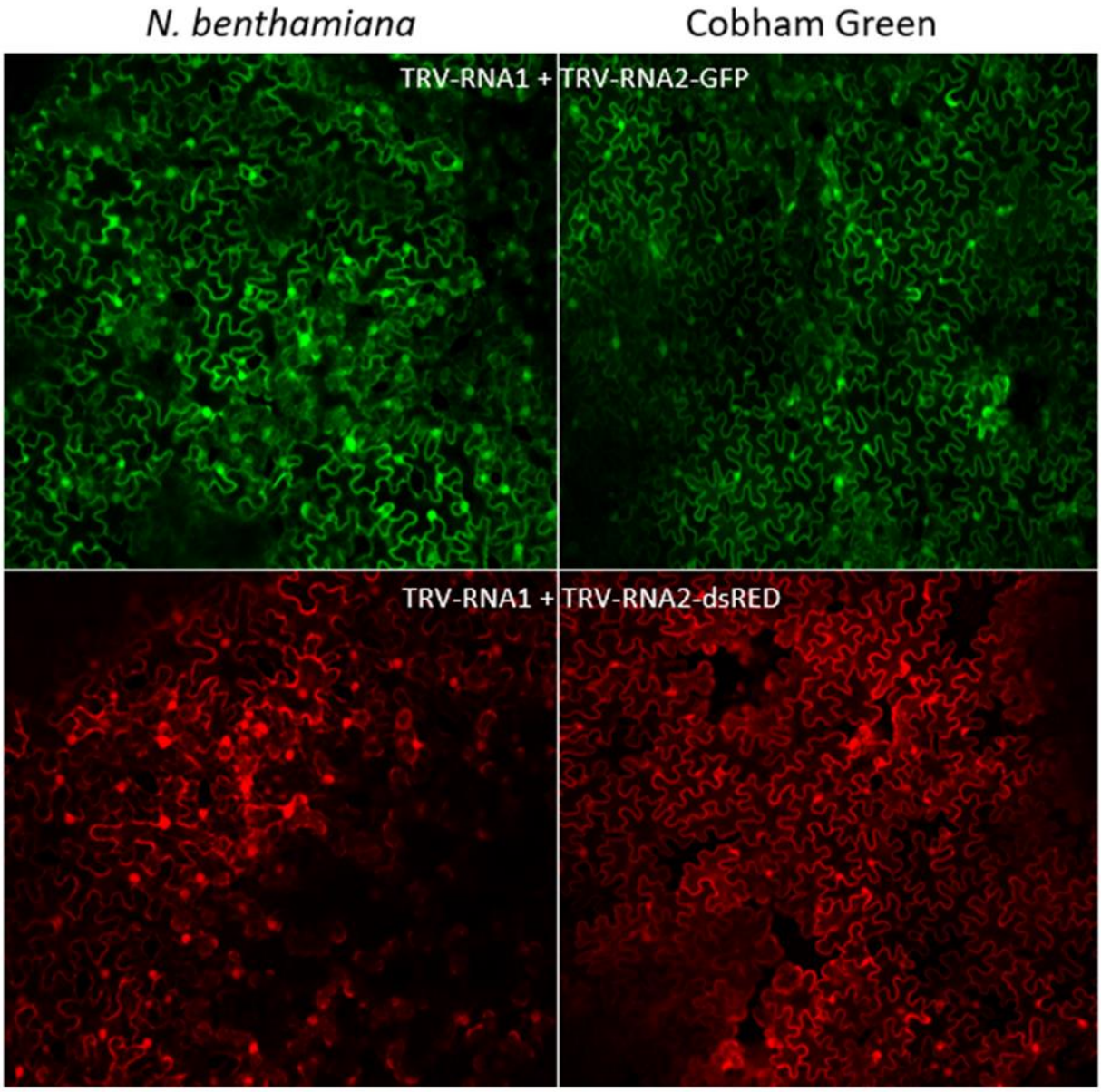


Figure A3.3. Transient expression of GFP and dsRED after co-agroinfiltration of TRV constructs into *N. benthamiana* (left) and *L. sativa* cv. Cobham Green (right). Co-infiltration was performed by mixing pYL192-TRV-RNA1 and SPDK2686-TRV-RNA2-GFP (top) or pYL192-TRV-RNA1 and SPDK2686-TRV-RNA2-dsRED (bottom).

Table A3.1. Primer names, sequences, and PCR conditions used to amplify each fragment used for In-Fusion cloning.

Primer	Primer sequence	Target	Amplicon Size (bp)	Annealing Temp (°C)	Extension Time (sec)
TB33	5’ CCGAAAGGAACAAAATTTCTAGACAAACACGTTCCTGA 3’	Fragment 1	640	55	60
TB29	5’ GCGCGCCATCTCGTTAACTCGGGTAAGTGA 3’				
TB24	5’ GTTAACGAGATGGCGCGCTCCTCCAAGAA 3’	Fragment 2	699	55	60
TB25	5’ TCGAATTCACTACAGGAACAGGTGGTGGC 3’				
TB28	5’ TTCCTGTAGTGAATTCGAGCTCGGTACCC 3’	Fragment 3	1,011	55	60
TB34	5’ ATCCAAGCTCAAGCTGCTCTAGAGCGCTCTAGCCAATACG 3’				

REFERENCES

Ali, Z., Abul-Faraj, A., Li, L., Ghosh, N., Piatek, M., Mahjoub, A., Aouida, M., Piatek, A., Baltes, N. J., Voytas, D. F., Dinesh-Kumar, S., and Mahfouz, M. M. (2015a). Efficient Virus-Mediated Genome Editing in Plants Using the CRISPR/Cas9 System. *Molecular Plant*, *8*, 1288–1291.

Ali, Z., Abul-Faraj, A., Piatek, M., Magdy, and Mahfouz, M. (2015b). Plant Signaling and Behavior Activity and specificity of TRV-mediated gene editing in plants. *Plant Signaling and Behavior*, *10*(10).

Burch-Smith, T. M., Schiff, M., Liu, Y., and Dinesh-Kumar, S. P. (2006). Efficient Virus-Induced Gene Silencing in *Arabidopsis*. *Plant Physiology*, *142*, 21–27.

Demirer, G. S., Zhang, H., Goh, N. S., González-Grandío, E., and Landry, M. P. (2019). Carbon nanotube-mediated DNA delivery without transgene integration in intact plants. *Nature Protocols*, *14*, 2954–2971.

German-Retana, S., Candresse, T., Alias, E., Delbos, R.-P., and le Gall, O. (2000). Effects of Green Fluorescent Protein or β-Glucuronidase Tagging on the Accumulation and Pathogenicity of a Resistance-Breaking Lettuce mosaic virus Isolate in Susceptible and Resistant Lettuce Cultivars. *Molecular Plant-Microbe Interactions MPMI*, *13*(3).

Ghoshal Id, B., Vong Id, B., Picard, C. L., Feng, S., Tam Id, J. M., and Jacobsen Id, S. E. (2020). A viral guide RNA delivery system for CRISPR-based transcriptional activation and heritable targeted DNA demethylation in *Arabidopsis thaliana*. *PLoS Genetics.*

Lien, J., Bull, T., Michelmore, R. W., and Guo, T. (2021). Fast Fluorescence Titration Quantification of Plasmid DNA with DNA Attractive Magnetic Nanoparticles. *Analytical Chemistry*, *93*(38), 12854–12861.

Liu, Y., Schiff, M., Marathe, R., and Dinesh-Kumar, S. P. (2002). Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus. *The Plant Journal*, *30*(4), 415–429.

Nagalakshmi, U., Meier, N., Liu, J.-Y., Voytas, D. F., and Dinesh-Kumar, S. P. (2022). High efficiency multiplex biallelic heritable editing in Arabidopsis using an RNA virus. *Plant Physiology*.

Oh, Y., and Kim, S. G. (2021). RPS5A Promoter-Driven Cas9 Produces Heritable Virus-Induced Genome Editing in *Nicotiana attenuata*. *Molecules and Cells*, *44*(12), 911.

Tian, J., Pei, H., Zhang, S., Chen, J., Chen, W., Yang, R., Meng, Y., You, J., Gao, J., and Ma, N. (2013). TRV-GFP: a modified Tobacco rattle virus vector for efficient and visualizable analysis of gene function. *Journal of Experimental Botany*, *65*(1), 311-322.

Uranga, M., Vazquez-Vilar, M., Orzáez, Di., and Daròs, J. A. (2021). CRISPR-Cas12a Genome Editing at the Whole-Plant Level Using Two Compatible RNA Virus Vectors. *CRISPR Journal*, *4*(5), 761–769.

Watson, J. M., Wang, M.-B., Bachan, S., and Dinesh-Kumar, S. P. (2012). Tobacco Rattle Virus (TRV)-Based Virus-Induced Gene Silencing. *Methods in Molecular Biology*, *894*, 83–92.

Wood, K. J., Nur, M. I., Gil, J. I., Fletcher, K. I., Lakeman, K., Gann, D., Gothberg, A. I., Khuu, T., Kopetzky, J. I., Naqvi, S. I., Pandya, A., Zhang, C., Maisonneuve, B. I., Pel, M., and Michelmore, R. I. (2020). Effector prediction and characterization in the oomycete pathogen *Bremia lactucae* reveal host-recognized WY domain proteins that lack the canonical RXLR motif. *PLoS Pathogens*.

Wroblewski, T., Tomczak, A., and Michelmore, R. (2005). Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis*. *Plant Biotechnology Journal*, *3*, 259–273.