Doctoral thesis of

Carmine De Paola

Enhancing *Nicotiana benthamiana* as chassis for Molecular Farming: targeting flowering time for increased biomass and recombinant protein production

Directors:

Diego Orzáez

Marta Vázquez-Vilar

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Valencia, March 2024

Index

Agradecimientos 5

Abbreviations 7

Index of figures 9

Index of tables 9

Summary 10

Resumen 11

Resum 12

General introduction 13

Plant Molecular Farming 13

Nicotiana benthamiana as host 14

Approaches to improve the chassis 17

Suppression of gene silencing 17

Storage compartments 17

Chaperones activity 18

Protease activity 18

Remodulation of host cell proteins abundance 19

Strategies to avoid toxic metabolites 19

Self-catalysing residual biomass 20

Glycoengineering 20

Genome editing 21

Multiplexing 23

Genome editing in the context of biofactory breeding 24

Growth habits and developmental phase transitions in the context of biofactories 25

Mechanisms controlling flowering time: the predominant role of *FT* genes 26

The *SPL* genes and their role in controlling plant architecture and developmental phase transition 27

FTs and *SPLs* as possible targets for biofactory breeding 29

Objectives 30

Chapter 1. Comparative analysis of the *SQUAMOSA PROMOTER BINDING-LIKE (SPL)* gene family in *Nicotiana benthamiana* and *Nicotiana tabacum* 31

Abstract 31

Introduction 32

Materials and Methods 34

Search of SPL genes in *N. benthamiana* and *N. tabacum* cv. K326 34

Reannotation of *N. tabacum* cv. K326 SPLs members 34

Identification of SBP conserved domains and motif analysis 34

Gene structure and miRNA156 complementary regions ... 35
Phylogenetic tree and family classification ... 35
Plant material, growing conditions and samples collection for RNA extraction ... 35
N. benthamiana phenotyping ... 36
RNA extraction and sequencing ... 36
Expression analysis ... 36
Plasmid assembly ... 36
Plant material and genetic transformation ... 37
Genomic DNA extraction and editing efficiency evaluation ... 38
Results ... 39
Phylogenetic analysis of the SPL family in *N. tabacum* and *N. benthamiana* ... 39
Protein structure of SPL genes: the SBP domain ... 40
Gene structure of *SPL* genes: exon-intron structure, putative miRNA156 target site ... 43
Expression analysis of *SPL* genes in leaves ... 44
Gene editing of NbSPL13 gene subfamily resulted in delayed flowering ... 46
Discussion ... 50
Conclusions ... 53
Supplementary material ... 54
Chapter 2. Extension of vegetative phase in *Nicotiana benthamiana* using genome editing of *FLOWERING LOCUS T (FT)* and *SQUAMOSA PROMOTER BINDING-LIKE (SPL)* genes. ... 68
Abstract ... 68
Introduction ... 69
Materials and methods ... 72
Protein identification and phylogenetic tree construction ... 72
Gene structure ... 72
Plasmid assembly ... 72
Plant material and stable transformation ... 72
Transient expression and fluorescence detection ... 73
Vacuum infiltration and fluorescence detection ... 74
Genomic DNA extraction and editing efficiency evaluation ... 74
Phenotyping ... 74
Results ... 75
Generation of FT5 and SPL13 knock-out lines ... 75
Stacking FT4 mutations on FT5 and SPL13 knock-out lines ... 79
Comparative analysis of flowering time and biomass-related parameters in FT4, FT5 and SPL13 mutant combinations ... 81
Analysis of recombinant protein expression in mutant *N. benthamiana* lines ... 86

Discussion 91
Supplementary material 95
General discussion 102
Conclusions 107
Bibliography 108

Agradecimientos

Habiendo llegado al final de mi doctorado, puedo decir que estos años pasados en Valencia han sido una de las mejores experiencias de mi vida hasta el momento. A pesar de los momentos difíciles que este camino ha conllevado, los buenos momentos han sido muchos más, gracias a todas las personas que he conocido. Reconozco que durante estos últimos años he crecido mucho como profesional y sobre todo como persona.

Siempre he querido pasar parte de mi vida viviendo en España, y agradezco el día que tuve la oportunidad de hacerlo, para poder sumergirme en el idioma, la cultura y la mentalidad de este país en el que siempre me he sentido bienvenido.

Estoy muy contento de haber conocido a todos mis compañeros del Lab 2.09-10, personas extraordinarias, muchos de los cuales se han convertido en amigos. Gracias a vosotros, esta experiencia ha sido mucho menos dura de lo esperado.

Son muchas las personas a las que tengo que dar las gracias, espero no olvidarme de ninguna. En primer lugar, gracias a Diego, por acogerme en tu laboratorio y darme la oportunidad de llevar a cabo mi tesis aquí. Gracias a Marta, por supervisarme y transmitirme tus conocimientos. Gracias Asun, por tu paciencia y por responder siempre a las dudas que me surgían. A Silvia G., por ayudarme en las etapas iniciales con todos aquellos documentos que nos pedían, por aclarar siempre cualquier duda que tenía sobre Crispr y nuestro trabajo en general, y por estar siempre disponible para una cerveza o una cena en compañía. Gracias a Víctor, por haberme ayudado muchísimo con cosas de bioinformática, las cervezas, las comidas y los almuerzos pasados juntos. A Asier, gracias por todo: las innumerables risas, la compañía, por introducirme a los memes de Belén Esteban y a todos los clásicos del trash español, por ayudarme mucho a mejorar mi castellano (aunque "por/para" no lo pillo del todo todavía), las cenas, los botellones en tu casa, las fiestas en The Muse y Deseo, los viajes juntos, la Isla. Gracias por presentarme a Carlos y a todos tus amigos, con vosotros lo he pasado genial muchísimas veces, y gracias por hacer que el tiempo que pasamos en el labo fuera mucho más llevadero. A Maria (sin tilde), gracias por estar siempre disponible para un café y un almuerzo a media mañana, por todas tus anécdotas sobre tu loca juventud y tu abuela, las cervezas juntos, las cenas, las risas, las fiestas, por apoyarme cuando me encontraba mal de la tripa durante nuestros viajes juntos, y por salvarme de aquella gaviota que quería atacarme en aquella playa de Cádiz, gracias a tu poderoso grito. Gracias al equipo de la isla de las tentaciones (Maria, Martita, Asier, Jose), por las noche en las que nos juntábamos para ver este maravilloso programa lleno de trash, chonis y canis, y por las mañanas siguientes en las que teníamos nuestro propio debate sobre el capítulo anterior. Gracias por darme a conocer este importante pilar del trash español. Gracias a Camilo por estar siempre disponible para ayudarme con GraphPad y las estadísticas de los gráficos, a Martita por ayudarme varias veces con mis experimentos cuando tenía

mucho trabajo, y por supuesto por salir juntos, las cervezas, las cenas, por ver la Isla y comentarla juntos. Gracias a Elena G. por las tardes pasadas en Olegari, las tardes de fiesta en Play, por proponerme siempre ir a comer a la hora guiri cuando yo también tenía hambre. Gracias a Sara, por las cervezas que tomamos juntos y por los chistes y las risas en los almuerzos y las comidas. Gracias a Rubén y Bea, vecinos de escritorio, siempre dispuestos a compartir bromas y charlitas entre experimento y experimento, por estar siempre disponibles para aconsejarnos mutuamente sobre las actividades del labo y apoyarnos en los momentos difíciles. Entre los que sólo pasaron unos meses en el laboratorio, gracias a Martina, Elisa y Ciro por salir juntos, por las risas y por hacerme "relajar" al poder hablar en italiano con vosotros y hacer bromas que sólo entre italianos podemos entender. En cuanto al personal del IBMCP, gracias a Marisol por su disponibilidad e inmensa amabilidad al ayudarme durante mis sesiones de microscopía, a Victoria por estar siempre disponible para facilitar nuestro trabajo en el invernadero, a Eugenio y Ana por su ayuda con la secuenciación. Fuera del laboratorio, me gustaría agradecer en general a toda la gente que he conocido aquí en Valencia, por permitirme descansar de mi rutina diaria durante los fines de semana y hacer mi experiencia con la tesis mucho más llevadera: Eli, Giuseppe, Jorge, Mikel, Fabio, Filippo, Sergio, Tommaso, Andrea, gracias por las fiestas en casa de Eli, las cervezas y spritzes tomados en el cómic y la boba, las fiestas en Deseo, las cenas, los tapeos y las tardes en la playa.

Infine, alla mia bellissima famiglia: mamma, papà, Antonio, Clelia. Grazie per l'amore incondizionato che mi date, e per appoggiarmi sempre, nei momenti belli e in quelli più difficili. Vi voglio un bene immenso.

Carmine (Nuccio)

Abbreviations

6-BAP: cytokinin 6-benzylaminopurine

AGO2: Argonaute2

AP1: APETALA 1

ATC: ARABIDOPSIS THALIANA CENTRORADIALIS

BFT: BROTHER OF FT AND TFL1

BiP1: Binding protein 1

Cas9: CRISPR associated protein 9

CDS: coding sequence

CO: CONSTANS

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

crRNA: CRISPR RNA

CycB: LYCOPENE BETA CYCLASE

DSB: double-strand break

eGFP: enhanced green fluorescent protein

ER: endoplasmic reticulum

ER: ERECTA

FAS: FASCIATED,

FD: FLOWERING LOCUS D

FT: FLOWERING LOCUS T

FUL: FRUITFULL

FW2.2: FRUIT WEIGHT 2.2

GA: gibberellic acid

GFP: green fluorescent protein

GMO: genetically modified organism

gRNA: guide RNA

HCP: host cell proteins

HDR: Homology Directed Repair

KO: knock-out

LFY: LEAFY

MAX1: MORE AXILLARY GROWTH1

MFT: MOTHER OF FT AND TFL1

miR156: microRNA 156

miRNA: micro RNA

MULT: MULTIFLORA

NGT: new genomic techniques

NHEJ: Non-Homologous End Joining

O: OVATE,

P4H: prolyl-4-hydroxylase

PDI: protein disulfide isomerase

PEBP: phosphatidylethanolamine-binding protein

PI: protease inhibitors

PI: protease inhibitors

PMF: Plant Molecular Farming

RDR1: RNA-dependent RNA polymerase 1

RISC: RNA-induced silencing complex

RNAi: RNA interference

RuBisCO: ribulose-1,5-bisphosphate carboxylase/oxygenase

SAM: shoot apical meristem

SDN: site-directed nuclease

sgRNA: single guide RNA

siRNA: small interfering RNAs

SOC1: SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1

SP: SELF-PRUNING

SP5G: SELF-PRUNING 5G

SPL: SQUAMOSA PROMOTER BINDING-LIKE

TFL1: TERMINAL FLOWER 1

TSF: TWIN SISTER OF FT

TU: transcriptional unit

UPR: unfolded protein response

UTR: untranslated region

VIGE: virus-induced genome editing

WT: wild type

Index of figures

Figure 1. Transient expression in *N. benthamiana* by agroinfiltration 15
Figure 2. Scheme representing stable transformation in *N. benthamiana*. 16
Figure 3. Scheme representing genome editing. 23
Figure 4. Example of a vector for CRISPR/Cas multiplexing. 24
Figure 5. Decrease of miRNA156 level. 28
Figure 6. Cascade of activation involving *FT*, *SPL* genes and SOC1. 29
Figure 7. Phylogenetic tree with SPL genes from *N. tabacum* cv. K326, *N. benthamiana* and *A. thaliana*. 40
Figure 8. Motifs composition and SBP domain of NbSPLs and NtSPLs. 42
Figure 9. Gene structure of *NbSPLs* (left) and *NtSPLs* (right). 44
Figure 10. Expression profiles of *NbSPL* and *NtSPL* genes in leaf at different developmental stages. 45
Figure 11. Generation of *N. benthamiana* lines edited in *NbSPL13_1a* gene. 47
Figure 12. Phenotype of *NbSPL13* edited lines. 49
Figure 13. Edition of the genes NbSPL13_1a, NbFT5_1a/1b and NbFT4 in *N. benthamiana*. 78
Figure 14. Phenotyping of SPL13 3-4, FT5 E2-3 and FT5 SPL13 22E-5 *N. benthamiana* lines. 79
Figure 15. Efficiency of the sgRNAs. 81
Figure 16. Phenotyping of *N. benthamiana* mutant lines. 83
Figure 17. Scheme of inflorescence architecture in *N. tabacum, A. thaliana* and *N. benthamiana*. 84
Figure 18. Phenotyping of *N. benthamiana* plants grown in high density. 86
Figure 19. GFP expression in different *N. benthamiana* genotypes. 89
Figure 20. Recombinant GFP production in *N. benthamiana* plants grown in high density. 90

Index of tables

Table 1. List of target genes, sgRNAs and mutations. 75

Summary

The term Plant Molecular farming (PMF) refers to the production of industrially relevant and commercially valuable recombinant products in plants. Its purpose is to provide a safe and cost-effective approach for the manufacturing of recombinant bioproducts at a large scale. Plants of the *Nicotiana* genus, especially *Nicotiana tabacum* and *Nicotiana benthamiana*, have become increasingly important as production platforms for PMF due to their advantages such as high biomass yield, ease of transformation, and robust protein expression. However, at present, there is room for improvement for *N. tabacum* and *N. benthamiana* as ideal hosts for molecular farming. Breeding goals such as delaying or abolishing flowering to enhance plant biomass could convert *N. benthamiana* into a prime chassis for molecular farming purposes. This objective was the focus of this research. In the first chapter, a genome-wide analysis of SQUAMOSA PROMOTER BINDING-LIKE (SPL) genes was performed. These genes are involved in vegetative phase transition and flowering time, on this species and its close relative *N. tabacum*, identifying 49 SPL genes in *N. tabacum* and 43 SPL genes in *N. benthamiana*. The SPL genes of the two species were classified into eight phylogenetic groups according to the SPL classification in *Arabidopsis thaliana*. The exon-intron gene structure and the DNA-binding domains were highly conserved between homeologues and orthologues, and the potential targets of microRNA156, involved in vegetative phase transition, were also identified. The expression of SPL genes in leaves was analysed by RNA-seq at three different growth stages, revealing that genes not under miR156 control were in general constitutively expressed at high levels, whereas miR156-regulated genes showed lower expression levels, often developmentally regulated. The *N. benthamiana SPL13_1a* gene was selected as target for a CRISPR/Cas9 knockout experiment. The full knock out of this single gene lead to a significant delay in flowering time of 2-5 days and increased branching. In the second chapter, more CRISPR/Cas9 gene editions are performed in *N. benthamiana* with the objective of flowering abolition. Floral inducers FLOWERING LOCUS T 4 and 5 (*NbFT4* and *NbFT5_1a/1b*) were knocked out alone and in combination with *NbSPL13_1a*. In the most edited line FT4-FT5-SPL13 40-1 flowering time was doubled compared to wild type plants. However, total abolition of flowering was not achieved. The delayed flowering had consequences on various aspects of plant growth, that were quantified through various parameters: highly edited lines had increased biomass, height, number of leaves and total leaves area compared to the less edited ones and wild type. Moreover, the generated lines were evaluated for their potential to express heterologous proteins. Unexpectedly, they were not able to maintain high expression levels after week five. In the future, knockouts in other important players in flowering initiation, such as *NbSPL9/15* and *NbSPL3/4/5*, will be stacked in our lines.

Resumen

Plant Molecular Farming (PMF) es la producción de proteínas de interés industrial y valor comercial en plantas. Su objetivo es proporcionar un enfoque seguro y rentable para la producción de proteínas recombinantes a gran escala. Las plantas del género *Nicotiana*, especialmente *Nicotiana tabacum* y *Nicotiana benthamiana*, han adquirido una importancia creciente como plataformas de producción de PMF debido a sus ventajas, como el alto rendimiento de biomasa, la facilidad de transformación y la expresión robusta de proteínas. Sin embargo, en la actualidad *N. tabacum* y *N. benthamiana* no son hospedadores ideales para el cultivo molecular. Los objetivos de mejora genética, como retrasar o suprimir la floración para aumentar la biomasa de la planta, podrían convertir a *N. benthamiana* en un chasis de primera para fines de cultivo molecular. En este trabajo de investigación nos centramos en este objetivo. En el primer capítulo, realizamos un análisis de todo el genoma de los genes SQUAMOSA PROMOTER BINDING-LIKE (SPL), implicados en la transición de fase vegetativa y el tiempo de floración, en esta especie y en su pariente cercana *N. tabacum*, identificando 49 genes SPL en *N. tabacum* y 43 genes SPL en *N. benthamiana*. Los genes SPL de las dos especies se clasificaron en ocho grupos filogenéticos de acuerdo con la clasificación de SPL en *Arabidopsis thaliana*. La estructura génica exón-intrón y los dominios de unión al ADN se conservaron en gran medida entre homeólogos y ortólogos, y también se identificaron las dianas potenciales del microARN156, implicado en la transición de fase vegetativa. La expresión de genes SPL en hojas se analizó mediante RNA-seq en tres fases de crecimiento diferentes, revelando que los genes que no estaban bajo el control de miR156 se expresaban en general de forma constitutiva a niveles altos, mientras que los genes regulados por miR156 mostraban niveles de expresión más bajos, a menudo regulados por el desarrollo. Seleccionamos el gen *SPL13_1a* de *N. benthamiana* como diana para un experimento de knockout CRISPR/Cas9. El knock out completo de este único gen condujo a un retraso significativo en el tiempo de floración de 2-5 días y a un aumento de la ramificación. En el segundo capítulo, mostramos más ediciones de genes CRISPR/Cas9 realizadas en *N. benthamiana* con el objetivo de la abolición de la floración. Se eliminaron los inductores florales FLOWERING LOCUS T 4 y 5 (NbFT4 y NbFT5_1a/1b) solos y en combinación con NbSPL13_1a. En la línea más editada FT4-FT5-SPL13 40-1 el tiempo de floración se duplicó en comparación con las plantas de tipo silvestre. Sin embargo, no se logró la abolición total de la floración. El retraso de la floración tuvo consecuencias en varios aspectos del crecimiento de la planta, que cuantificamos a través de diversos parámetros: las líneas altamente editadas presentaron un aumento de la biomasa, la altura, el número de hojas y el área foliar total en comparación con las menos editadas y el tipo silvestre. Además, se evaluó el potencial de las líneas generadas para expresar proteínas heterólogas. Inesperadamente, no fueron capaces de mantener altos niveles de expresión después de la quinta semana. En el futuro, se apilarán en nuestras líneas knockouts en otros actores importantes en el inicio de la floración, como NbSPL9/15 y NbSPL3/4/5.

Resum

Plant Molecular Farming (PMF) és la producció de proteïnes d'interès industrial i valor comercial a plantes. El seu objectiu és proporcionar un enfocament segur i rendible per a la producció de proteïnes recombinants a gran escala. Les plantes del gènere *Nicotiana*, especialment *Nicotiana tabacum* i *Nicotiana benthamiana,* han adquirit una importància creixent com a plataformes de producció de PMF a causa dels seus avantatges, com ara l'alt rendiment de biomassa, la facilitat de transformació i l'expressió robusta de proteïnes. No obstant això, actualment la *N. tabacum* i la *N. benthamiana* no són hostes ideals per al cultiu molecular. Els objectius de millora genètica, com ara endarrerir o suprimir la floració per augmentar la biomassa de la planta, podrien convertir *N. benthamiana* en un xassís de primera per a fins de cultiu molecular. En aquest treball de recerca ens centrem en aquest objectiu. Al primer capítol, realitzem una anàlisi de tot el genoma dels gens SQUAMOSA PROMOTER BINDING-LIKE (SPL), implicats en la transició de fase vegetativa i el temps de floració, en aquesta espècie i en el seu parent proper *N. tabacum*, identificant 49 gens SPL a *N. tabacum* i 43 gens SPL a *N. benthamiana*. Els gens SPL de les dues espècies es van classificar en vuit grups filogenètics d'acord amb la classificació de SPL a *Arabidopsis thaliana*. L'estructura gènica exón-intron i els dominis d'unió a l'ADN es van conservar en gran mesura entre homeòlegs i ortòlegs, i també es van identificar les potencials dianes del microARN156, implicat en la transició de fase vegetativa. L'expressió de gens SPL en fulles es va analitzar mitjançant RNA-seq en tres fases de creixement diferents, revelant que els gens que no estaven sota el control de miR156 s'expressaven en general de forma constitutiva a nivells alts, mentre que els gens regulats per miR156 mostraven nivells més baixos d'expressió, sovint regulats pel desenvolupament. Seleccionem el gen *SPL13_1a* de *N. benthamiana* com a diana per a un experiment de knockout CRISPR/Cas9. El knock out complet d'aquest gen va conduir a un retard significatiu en el temps de floració de 2-5 dies ia un augment de la ramificació. Al segon capítol, mostrem més edicions de gens CRISPR/Cas9 realitzades a *N. benthamiana* amb l'objectiu de l'abolició de la floració. Es van eliminar els inductors florals FLOWERING LOCUS T 4 i 5 (NbFT4 i NbFT5_1a/1b) sols i en combinació amb NbSPL13_1a. A la línia més editada FT4-FT5-SPL13 40-1 el temps de floració es va duplicar en comparació amb les plantes de tipus silvestre. Tot i això, no es va aconseguir l'abolició total de la floració. El retard de la floració va tenir conseqüències en diversos aspectes del creixement de la planta, que vam quantificar a través de diversos paràmetres: les línies altament editades van presentar un augment de la biomassa, l'alçada, el nombre de fulles i l'àrea foliar total en comparació amb les menys editades i el tipus silvestre. A més, es va avaluar el potencial de les línies generades per expressar proteïnes heteròlogues. Inesperadament, no van ser capaços de mantenir alts nivells dexpressió després de la cinquena setmana. En el futur, s'apilaran a les nostres línies knockouts en altres actors importants a l'inici de la floració, com NbSPL9/15 i NbSPL3/4/5.

General introduction

Plant Molecular Farming

Plant Molecular Farming (PMF) refers to the production of industrially relevant and commercially valuable recombinant products in plants (Eidenberger *et al*., 2023). Its purpose is to provide a safe and cost-effective approach for the manufacturing of bioproducts at large scale. This research field holds great promise for the efficient and cost-effective production of a wide range of therapeutic proteins, vaccines, antibodies, and other bioactive molecules. The use of whole plants or plant cell cultures as biofactories offers numerous potential advantages over traditional production methods in other organisms, such as microbial fermentation or animal cell culture, particularly in terms of safety and scalability. Many plant species are considered generally safe for recombinant production due to their long history of consumption by humans (Stoger *et al*., 2014), and plant-based production systems are safer since they do not support the replication of mammalian viruses that could arise in mammalian cell cultures (Hundleby *et al*., 2018). Another significant advantage of plant molecular farming is the potential for scalability and the cost-effectiveness of protein production. Several authors indicate that fermentation of cell cultures is more expensive than the growth of plants, which can be up-scaled to the level of open-field cultivation (Fischer & Buyel, 2020). Another benefit of plant-based protein production is the simplicity, since sterility is not required during production (Buyel, 2019; Fischer & Buyel, 2020). Despite these advantages, PMF has drawbacks, compared to other production platforms, that need to be considered. Historically, targeted genetic manipulation has been more common in other organisms, such as yeast (Green & Tibbetts, 1980), while in plants only the advent of CRISPR/Cas and similar technologies offered the possibility of efficient site-directed mutagenesis (Doudna & Charpentier, 2014; J.-F. Li *et al*., 2014). Long regeneration times of stable transgenic plant lines (6-18 months) often dissuade industry from adopting production in plants (Sack *et al*., 2015) and obstacles in the present regulatory approval process contribute to this trend (Ma *et al*., 2015; Tusé *et al*., 2020). Moreover, the highest registered levels of recombinant protein production in plants were around 4 g/kg for GFP, influenza antigens and monoclonal antibodies (Shoji *et al*., 2012; Yamamoto *et al*., 2018; Zischewski *et al*., 2016) while production yields of antibodies in mammalian cells often exceeds 25 g/L (W. C. Yang *et al*., 2016). Other problems may include the activation of endogenous proteases in plants during transient expression by infiltration (Grosse-Holz *et al*., 2017), the difference in glycosylation patterns (Fischer *et al*., 2021; Strasser, 2016) and other undesired modifications due to oxidation and proteolysis during downstream processing (Buyel *et al*., 2021).

A diverse array of recombinant proteins has been successfully synthesized across a broad spectrum of plant hosts, and these products are readily accessible in the market today. An early example of this breakthrough was witnessed when Merck pioneered the production of avidin and β-glucuronidase in

transgenic maize (Hood *et al*., 1999). More recently, ORF Genetics achieved the commercial synthesis of growth factors in transgenic barley (Magnusdottir *et al*., 2013), while Protalix Biotherapeutics accomplished the commercialization of glucocerebrosidase, produced in carrot cells (Zimran *et al*., 2018), and is developing the production of pegunigalsidase alfa in the same host (van der Veen *et al*., 2020). In addition to the production of recombinant proteins, PMF also includes the production of metabolites. An early example is the increased level of carotenoids in genetically engineered tomato (Enfissi *et al*., 2005).

Nicotiana benthamiana as host

Plant molecular farming has gained significant momentum in recent years with certain plant species, particularly those belonging to the *Nicotiana* genus, such as tobacco (*Nicotiana tabacum*) and *Nicotiana benthamiana*, emerging as prominent production platforms. These plants offer several advantages, including their non-food crop status, robust protein expression capability, ease of transformation and high biomass yield, making them ideal candidates for molecular farming applications (Tremblay *et al*., 2010).

The utilization of tobacco as a biofactory for molecular farming began in the early 1990s when its potential for expressing foreign genes and producing specific proteins of interest was recognized. Hiatt *et al*. (1989) successfully assembled and produced functional antibodies in tobacco (Hiatt *et al*., 1989) and Mason *et al*. (1996) produced in this chassis the Norwalk virus capsid protein, which elicited an immunogenic response in mice (Mason *et al*., 1996).

The successful expression of antibodies and viral proteins in tobacco plants highlighted the versatility of this plant species as a bioreactor for molecular farming. In addition to tobacco, *N. benthamiana*, a tobacco relative originally from Australia, has rapidly emerged as an excellent host for molecular farming purposes. *N. benthamiana* possesses several advantageous characteristics that make it an ideal choice for protein expression. Firstly, it is well-suited for transient expression through a technique called agroinfiltration (Fig. 1), which allows for rapid and efficient protein production (Sheludko *et al*., 2007). This method involves introducing *Agrobacterium tumefaciens* carrying the gene of interest into the plant, resulting in high levels of protein expression. Moreover, *N. benthamiana* has several key features that contribute to its suitability as a molecular farming host. It has an appropriate size for indoor production and a shorter life cycle compared to other related plant species, enabling faster protein production (Bally *et al*., 2015). The leaf anatomy of *N. benthamiana* facilitates the infiltration of genetic material, allowing efficient expression of the introduced genes. Moreover, the endogenous gene silencing mechanisms of the *N. benthamiana* LAB strain, which is commonly used in research, are partially defective due to a natural frame-shift insertion in the *RNA-dependent RNA polymerase 1 (RDR1)* gene. This defect favours the use of DNA delivery vectors and enhances the expression of proteins from viral vectors (Bally *et al*.,

2015; S.-J. Yang *et al.*, 2004). These favourable characteristics have made *N. benthamiana* a popular choice as a production chassis for pharmaceuticals. As a consequence of this increased interest in the *N. benthamiana* host, numerous resources have been generated in recent years, including a multi-omic resource containing genomic, transcriptomic and epigenomic data (Ranawaka *et al.*, 2023). In addition, comprehensive metabolomic profiles have been produced, including those describing plant developmental stages (Drapal *et al.*, 2021a), specific responses to the agroinfiltration process (Drapal *et al.*, 2021b). All this is included in a chemotype core collection of the *Nicotiana* genus including six accessions of *Nicotiana benthamiana* and comprising over 360 identified metabolites of a wide range of chemical classes as well as thousands of unknown compounds with dedicated spectral and chromatographic properties (Drapal *et al.*, 2022). Together, this collaborative characterization effort serves as basis for future improvements of the chassis.

Numerous examples demonstrate its effectiveness and rapidness in producing important bioactive molecules by transient expression. For instance, *N. benthamiana* has been utilized to produce a HIV neutralizing antibody (Hamorsky *et al.*, 2013) and influenza virus-like particles capable of eliciting an immune response in mice (D'Aoust *et al.*, 2008). More recently, *N. benthamiana* has played a crucial role as a production platform for antigens and vaccines targeting the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of COVID-19. SARS-CoV-2 antigens were successfully expressed in *N. benthamiana* to support the development of diagnostic tests and potential vaccine candidates (Maharjan & Choe, 2021; Ruocco & Strasser, 2022).

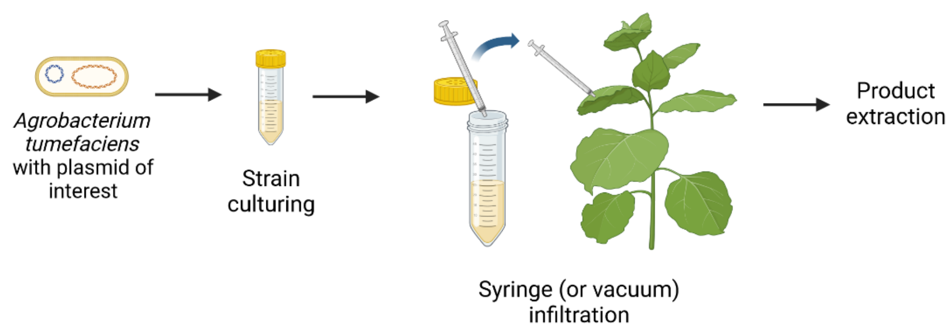


Figure 1. Transient expression in *N. benthamiana* by agroinfiltration.

While transient expression in *N. benthamiana* offers the advantage of rapid protein production, it may not guarantee long-term and stable production of pharmaceuticals. To address this concern, stable transformation of *N. benthamiana* can be a viable alternative for molecular farming, as it allows for the generation of transgenic lines whose seeds can be stored and sown as needed (Fig. 2). This approach ensures consistent and sustainable production of desired molecules and proteins. Several examples of successful stable transformation for molecular farming purposes in *N. benthamiana* have been reported. For instance, Forestier *et al.* (2021) developed transgenic *N. benthamiana* lines producing high-value

diterpenes (Forestier *et al*., 2021), the production of volatile moth sex pheromones was developed by Mateos-Fernández *et al*. (2021) who generated stable *N. benthamiana* lines harbouring the pheromones biosynthetic pathway with the aim of sustainable pest control (Mateos-Fernández *et al*., 2021). More recently, Forestier *et al*. (2023) achieved the production of casbene in *N. benthamiana* through heat-inducible promoters (Forestier *et al*., 2023), and Limkul *et al*. (2015) successfully achieved the production of human glucocerebrosidase, a therapeutic enzyme used in the treatment of Gaucher's disease, in stable transgenic *N. benthamiana* lines (Limkul *et al*., 2015). Besides the whole plant approach, transgenic *N. benthamiana* and *N. tabacum* cell suspensions have been also widely used as efficient recombinant protein production platforms. Indeed, most of the early studies in bioreactors and cell cultures have been done with *Nicotiana tabacum* BY2 cells (Verdú-Navarro *et al*., 2023).

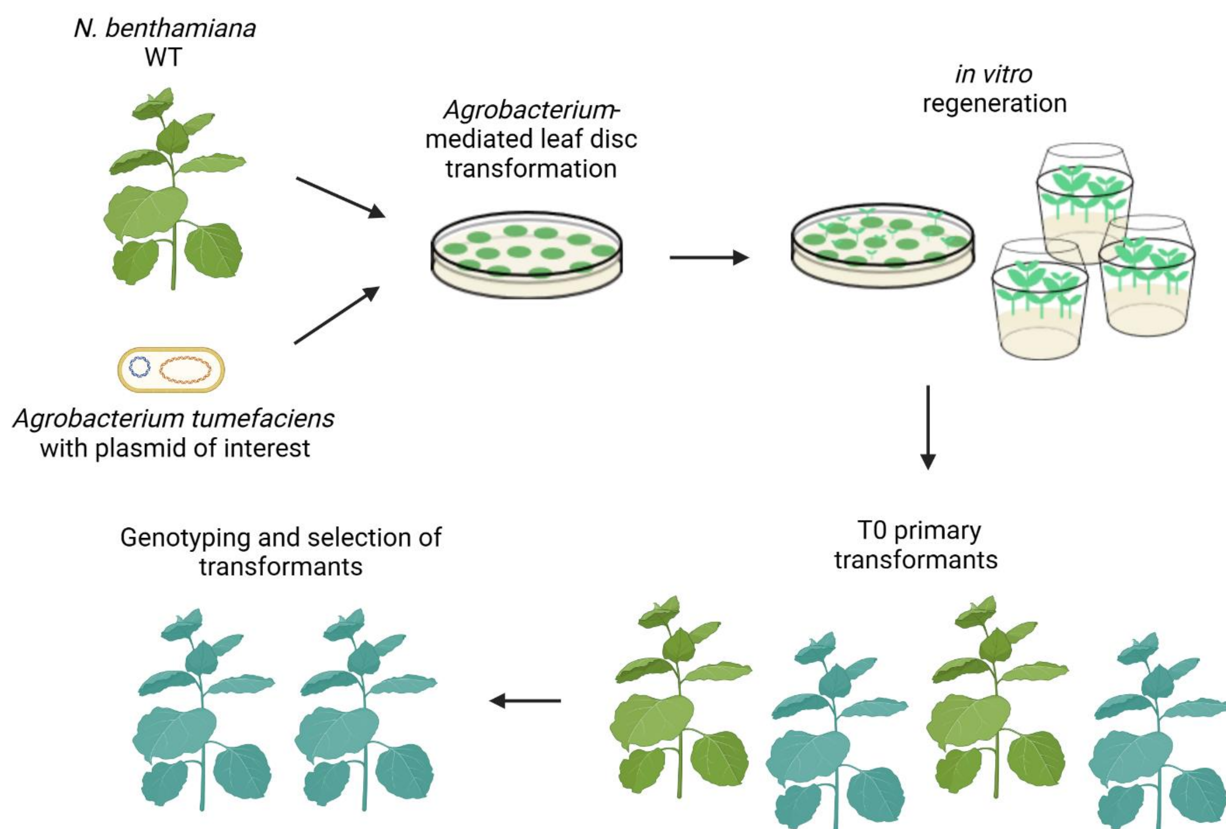


Figure 2. Scheme representing stable transformation in *N. benthamiana*.

Discs are cut from leaves and kept in contact with A. tumefaciens carrying the plasmid of interest. Then, the T0 shoots generated in vitro are grown, genotyped and selected. Blue plants represent the transformed ones.

Overall, the unique attributes of *N. benthamiana*, including its efficient agroinfiltration-based transient expression system and defective gene silencing mechanisms, have positioned it as a leading choice for

molecular farming applications, enabling the production of diverse pharmaceuticals and contributing to advancements in biotechnology.

Approaches to improve the chassis

As already mentioned, PMF has advantages, but also drawbacks compared to protein and compound production in other organisms. Several approaches have been introduced in order to tackle these problems and also to further improve the capacity of the production hosts as biofactories. Some of these strategies are presented in this section.

Suppression of gene silencing

The yield of recombinant proteins can be reduced due to gene silencing induced by transgenes. Viral silencing suppressors can prevent this mechanism by being co-expressed with the transgene in *N. benthamiana* (Arzola *et al*., 2011). For instance, the P19 suppressor derived from the tomato bushy stunt virus (TBSV) has been extensively used to increase the expression of recombinant proteins (Garabagi *et al*., 2012). This is achieved by its ability to bind to small interfering RNAs (siRNA) and prevent the assembly of the RNA-induced silencing complex (RISC) (Garabagi *et al*., 2012). Repressing endogenous genes involved in RNA silencing can also improve recombinant protein production. The use of CRISPR/Cas9 in *N. benthamiana* to inactivate Argonaute2 (AGO2) increased the expression of a viral vector encoding GFP (Ludman *et al*., 2017). The knockout in *N. benthamiana* of RNA-dependent RNA polymerase 6 (RDR6) using CRISPR/Cas9 disrupted post-transcriptional gene silencing and increased recombinant GFP expression (Matsuo & Atsumi, 2019). As previously mentioned, the *N. benthamiana* LAB strain possesses a natural frameshift insertion in its RDR1 gene, which causes an altered response to viral infections and makes it an ideal chassis for the expression of viral vectors (S.-J. Yang *et al*., 2004).

Storage compartments

Another approach to enhance protein production is to boost the storage capacity of organelles, such as the endoplasmic reticulum (ER) or vacuoles. The ER is crucial in protein synthesis, folding, modification, and storage. A well-developed ER correlates with the capacity to secrete recombinant proteins (Margolin *et al*., 2020; Zhu *et al*., 2019). Promoting membrane synthesis is a strategy to increase protein production (de Ruijter *et al*., 2016; Schuck *et al*., 2009). The production of larger quantities of phospholipids and the expansion of the ER can be induced by eliminating the phosphatidic acid phosphatase (PAP or PAH). In *Arabidopsis*, the knock-out of PAH1/2 promoted ER proliferation (Craddock *et al*., 2015).

Plant vacuoles, which typically function as lytic compartments, account for a significant portion of the cell volume. In seeds and other storage organs, proteins can be stored in vacuoles (Arcalis *et al*., 2014;

Herman & Larkins, 1999; Marty, 1999; Takaiwa *et al.*, 2017). Normally the central vacuole of leaf cells is considered an unstable environment for recombinant proteins due to their nature as lytic compartments, but there are cases in which recombinant proteins have accumulated in this organelle (Marin Viegas *et al.*, 2017). In *N. benthamiana*, Ocampo *et al.* (2016) achieved the targeting of monoclonal antibodies to vacuoles (Ocampo *et al.*, 2016), and human glucocerebrosidase was successfully targeted to vacuoles in carrot cells to gain a specific N-linked glycan structure (Shaaltiel *et al.*, 2007). Although the lytic nature of the vacuole can be a drawback, its characteristics can be modified by genetic engineering: the overexpressed transcriptional regulator LEAFY COTYLEDON2 (LEC2) altered leaf morphology and reduced the size of the lytic vacuole in *Arabidopsis*, replacing it with protein storage vacuoles (Feeney *et al.*, 2013). Nevertheless, substantial interventions in cellular metabolism, such as transforming vacuoles into protein storage vacuoles, require further understanding and consideration of potential side effects.

Chaperones activity

The production of recombinant proteins in the endoplasmic reticulum often leads to ER stress and an increased expression of chaperones, which are essential for protein folding (Strasser, 2018). Predictably, this gave rise to the idea of upregulating specific chaperones to boost the generation of both endogenous and recombinant proteins (Buyel *et al.*, 2021). In rice, excessive overexpression of the chaperone Binding protein 1 (BiP1) resulted in altered seeds and reduced protein accumulation (Wakasa *et al.*, 2011). Nevertheless, a minimal increase in the expression of this chaperone led to optimal recombinant protein yields, suggesting the need for balanced modification of BiP1 levels for improved production (Wakasa *et al.*, 2011). Similar variable outcomes were observed in yeast and mammalian cells when overexpressing chaperones like BiP and protein disulfide isomerase (PDI) (Damasceno *et al.*, 2007; Klabunde *et al.*, 2007; Kunert & Reinhart, 2016). In *N. benthamiana*, the expression of human chaperones, such as calreticulin, enhanced the accumulation of human viral glycoproteins (Margolin *et al.*, 2020).

Protease activity

Endogenous proteases can degrade the recombinant proteins produced in the plants. For example, IgG frequently suffer proteolytic degradation (Donini *et al.*, 2015; Puchol Tarazona *et al.*, 2020). Strategies to mitigate proteolysis include RNAi-mediated gene silencing to downregulate proteases (N.-S. Kim *et al.*, 2008; Mandal *et al.*, 2014) and the utilization of genome editing. For example, in moss, knocking out specific proteases through CRISPR/Cas9 technology has been successful in increasing recombinant protein yields (Hoernstein *et al.*, 2018). Additionally, inducible promoters can be harnessed to spatiotemporally control the expression of genome-edited proteases, avoiding interference with plant growth (Lowder *et al.*, 2017; X. Wang *et al.*, 2020). Broad-spectrum protease inhibitors (PIs) can also be co-expressed with the target protein to counteract proteolytic degradation. For example, Jutras *et al.*

(2016) achieved antibody accumulation in *N. benthamiana* through the simultaneous expression of the tomato protease inhibitor SlCYS8 (Jutras *et al*., 2016), while Grosse-Holz *et al*. (2018) increased the accumulation of α-galactosidase and erythropoietin by co-expressing protease inhibitors from *N. benthamiana* and humans (Grosse-Holz *et al*., 2018). Moreover, genome editing can be employed to modify the promoters of endogenous PI genes, allowing spatiotemporal regulation of PI expression (Mandal *et al*., 2016; Pillay *et al*., 2012).

Remodulation of host cell proteins abundance

It is beneficial in PMF to divert cellular resources from the synthesis of host cell proteins (HCPs) to the production of the recombinant protein of interest. Reducing the abundance of HCPs simplifies downstream processing and increases synthesis capacity (Buyel *et al*., 2016; Opdensteinen *et al*., 2019). Strong constitutive promoters are frequently used to maximize transcription rates. Among the most common ones are the Cauliflower mosaic virus 35S and ubiquitin promoters (W. Liu & Stewart, 2016). Nevertheless, a significant portion of cellular resources is still dedicated to HCP synthesis, with proteins like ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) representing a substantial proportion of total proteins in leaves (Buyel *et al*., 2015). An optimal PMF platform should possess the ability to control the shutdown of HCP synthesis over time, utilizing inducible RNAi constructs or transcriptional repressors based on inactivated Cas9. RuBisCO is a key target for shutdown strategies, as it is greatly abundant in the plant biomass (Robert *et al*., 2015). Other proteins can be considered as targets for this strategy, as the ones related to stress response, cell division, photosynthesis or cell growth (Mahalik *et al*., 2014; Sharma *et al*., 2020). In seed crops, reducing endogenous storage protein accumulation can increase product yields by creating space in storage organelles and reducing competition for translation and assembly in the endoplasmic reticulum (Takaiwa, 2013). CRISPR/Cas has been used to target storage protein genes in various crops (A. Li *et al*., 2018; Lyzenga *et al*., 2019; Sánchez-León *et al*., 2018).

Strategies to avoid toxic metabolites

Employing non-edible plants, such as tobacco, for PMF purposes offers the benefit of reducing the risk of contaminating the food chain with the transgenes (Breyer *et al*., 2012; Commandeur *et al*., 2003). Even so, these plants may still produce toxic metabolites, such as nicotine. Purification steps in biopharmaceutical production remove small molecules and impurities below the limit of detection, but organic solvent-based techniques may be required for technical protein formulations (Fu *et al*., 2010; Ma *et al*., 2015; McNulty *et al*., 2020). Developing a plant molecular farming chassis that does not produce toxic compounds is desirable. In tobacco, the knockout of all berberine bridge enzyme-like (BBL) genes using CRISPR/Cas9 resulted in an almost complete elimination of nicotine production (Schachtsiek & Stehle, 2019). Nonetheless, targeting key enzymes involved in secondary metabolism may have unintended side effects: Kaiser *et al*. (2002) overexpressed homospermidine synthase for the reduction of spermidine levels in tobacco, but this manipulation resulted in a stunted phenotype (Kaiser

et al., 2002). Rather than overexpressing or knocking out the gene of interest, a possible alternative can be the spatiotemporal regulation of its corresponding transcription factor in order to govern the concentration of the target metabolite (Hayashi *et al.*, 2020).

Self-catalysing residual biomass

The concept of self-catalysed processing of residual biomass in PMF processes can enhance the economic viability by utilizing the remaining plant biomass as a cascade biorefinery (Buyel, 2019). The predominant component in residual biomass is cellulose (Sheen, 1983), and its degradation can be performed by exoglucanases and endoglucanases, which convert cellulose into oligo- and monosaccharides (Bornscheuer *et al.*, 2014). Restricting enzyme expression to specific organs or inducing it at a precise moment, such as pre-harvest, can reduce the rigidity of the fibres, facilitating the extraction of the product (Vicuna Requesens *et al.*, 2019). Genome editing techniques can be employed to alter the promoters of endogenous cellulase, enabling precise control over spatiotemporal expression of these genes (Buyel *et al.*, 2021).

Glycoengineering

Plants and animals produce complex N-linked glycans, but there are differences between plant and human glycans, such as the presence of core α-1,3-fucose and β-1,2-xylose in plants (Montero-Morales & Steinkellner, 2018). One of the objectives that were pursued in the context of PMF was the glycoengineering, i.e. the "humanization" of the glycoproteins produced in plants, to avoid possible immunogenic effects in humans. Nagels *et al.* (2011) focused on the generation of transgenic *N. benthamiana* lines capable of producing complex multiantennary N-glycans. These lines produced humanized N-glycan structures, which are crucial for therapeutic glycoproteins with enhanced functionality and reduced immunogenicity (Nagels *et al.*, 2011). Jansing *et al.* (2019) utilized stable transformation of *N. benthamiana* to produce humanized non-immunogenic glycoproteins lacking β-1,2-xylose and core α-1,3-fucose. (Jansing *et al.*, 2019). The same result was also achieved in BY-2 cells of tobacco (Hanania *et al.*, 2017; Mercx *et al.*, 2017) and in SR-1 tobacco plants (Göritzer *et al.*, 2022). O-linked glycans, although less studied than N-linked glycans, are important for the biological activity of produced recombinant proteins (Buyel *et al.*, 2021). In plants, O-linked glycosylation starts with prolyl-4-hydroxylation, and subsequently involves the addition of arabinogalactan or arabinose (Strasser, 2016). Knockout of *prolyl-4-hydroxylase* (*P4H*) genes has been achieved in *Physcomitrella patens* and research is ongoing to target P4H isoforms in *N. benthamiana* (Parsons *et al.*, 2013; Schoberer & Strasser, 2018).

Genome editing

Many of the strategies aiming at the improvement of heterologous protein production in plants are now possible thanks to genome editing through CRISPR/Cas9 technology. The discovery of CRISPR/Cas9 has had a profound impact on genetic engineering, particularly in relation to the field of molecular farming (Belhaj *et al.*, 2013). CRISPR/Cas stands for "Clustered Regularly Interspaced Short Palindromic Repeats" and "CRISPR-associated protein". This system is a natural defence mechanism found in bacteria against bacteriophages. It involves the use of a nuclease called Cas9, guided to cut the DNA of invading viruses by interacting with two RNA molecules: a non-specific molecule known as tracrRNA and a specific molecule called CRISPR RNA (crRNA). The crRNA is derived from a sequence within the bacteriophage itself. Once the presence of a viral sequence is detected, it is inserted into the bacterial genome to create a permanent genetic record of past infections. The recognition of the viral DNA occurs through an interaction between the viral DNA and the crRNA, which is flanked by a sequence called PAM (protospacer adjacent motif) (Bortesi & Fischer, 2015). When these elements are recognized, the target DNA molecule is cleaved. The CRISPR/Cas systems are classified into two main classes: Class 1, which includes enzyme types I, III, and IV, and Class 2, which includes enzyme types II, V, and VI (Makarova *et al.*, 2015). The widely used CRISPR/Cas9 system belongs to Class 2 and utilizes a type II enzyme. Other Cas proteins were later discovered to cleave RNA instead of DNA, such as Cas13 (Abudayyeh *et al.*, 2016).

The remarkable simplicity of CRISPR/Cas9 technology lies in its requirement of only two components: the Cas9 protein, which introduces a double-strand break (DSB) in the DNA, and a guide RNA (gRNA) that directs Cas9 to a specific location in the genome. The gRNA is an engineered sequence that combines the tracrRNA and the crRNA, incorporating a unique 20-nucleotide sequence. Targeting Cas9 to different genomic loci involves designing these 20 specific nucleotides, always ensuring their presence before a PAM sequence at the target site.

Following the DNA break, the cell activates a series of mechanisms to repair the damaged DNA molecule. The repair outcome depends on the DNA repair pathway triggered by the cell (Fig. 3). Depending on the repair system employed by the cell, and the type of modification that results from such repair, the genetic modifications obtained by CRISPR (or by any other site-directed nucleases, SDNs) can be categorized as follows:

SDN-1: In this approach, the nuclease cleaves the DNA molecule, and the repair mechanism rejoins the broken ends using the Non-Homologous End Joining (NHEJ) repair system. This repair process is error-prone and can lead to various mutations at the target site, such as nucleotide substitutions or the addition/loss of a few nucleotides (indels). SDN-1 is commonly employed in plant species to induce loss-of-function mutations (Belhaj *et al.*, 2013; Bortesi & Fischer, 2015; Podevin *et al.*, 2013; Scheben *et al.*, 2017);

SDN-2: This technique involves the nuclease-mediated cleavage of the DNA molecule, accompanied by the utilization of an exogenous DNA molecule as a template for repair. The repair process occurs through a system called Homology Directed Repair (HDR). SDN-2 allows for precise and controlled mutations, including nucleotide substitutions or the addition/loss of one or a few nucleotides. Compared to SDN-1, this method results in the generation of either a variant of an existing gene found in nature or the creation of a new variant (Bortesi *et al*., 2016; Bortesi & Fischer, 2015; Podevin *et al*., 2013);

SDN-3: In SDN-3, the nuclease cleaves the DNA molecule at a predetermined site, followed by the integration of a new sequence into that site (Bortesi *et al*., 2016; Bortesi & Fischer, 2015; Podevin *et al*., 2013). The nature and origin of the inserted sequence determine whether a cisgenic, intragenic or transgenic plant is obtained: in cisgenesis, the gene of interest flanked by its own promoter and terminator is derived from another variety of the same species or from a sexually compatible one, whereas in intragenesis the gene of interest can be combined with regulatory elements of the same species (or a sexually compatible one) belonging to a different locus (Marone *et al*., 2023). In transgenesis, the gene of interest comes from a non-sexually compatible species (Marone *et al*., 2023). Nowadays, cisgenic and intragenic plants have a greater potential to be used in PMF, since they are comparable to varieties obtained through conventional breeding methods, therefore they can be more easily accepted in the market. On the other hand, transgenic varieties have to face stricter regulatory approvals in many countries and are nowadays totally forbidden in the European Union (Marone *et al*., 2023).

Overall, CRISPR/Cas9 technology has opened up a wide range of possibilities for precise genetic modifications in both plant and biomedical research, allowing to manipulate genetic material with unprecedented efficiency and accuracy. This breakthrough has transformed the field of genetic engineering and holds great potential for advancements in various areas of science and medicine.

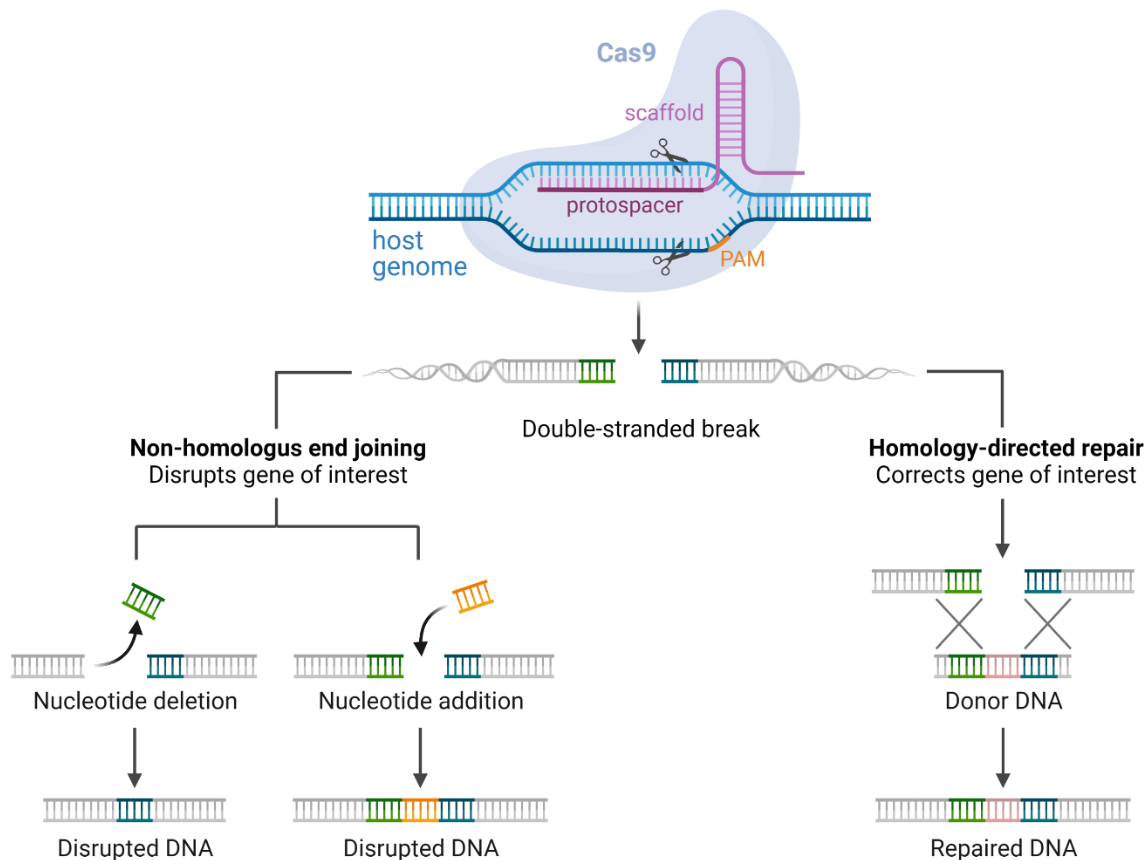


Figure 3. Scheme representing genome editing.

The Cas protein is directed by pairing to a specific site on the host genome, where it performs a double stranded break. The repair outcome depends on the DNA repair pathway triggered by the cell. Adapted from the templates of Biorender (https://www.biorender.com/).

Multiplexing

One of the outstanding advantages of CRISPR/Cas technology is the ability to deliver multiple gRNAs simultaneously, a strategy known as multiplexing, that enables targeting many loci at the same time. CRISPR/Cas multiplex editing in plant breeding paves the way for pyramiding favorable independent traits at unprecedented speed (Y. Zhang *et al.*, 2020). This capacity has been exemplified with the domestication of wild tomato by editing six genes involved in yield and productivity resulting in increased fruit size and number (Zsögön *et al.*, 2018), or with its adaptation to urban agriculture by editing genes that resulted in compact tomato plants with precocious fruits (Kwon *et al.*, 2020). In addition, multiplexing has the ability to uncover valuable traits which have remained elusive to breeding due to redundancy in large gene families. This is more evident in polyploid plants, which account for some of the most important crop species. Remarkable examples are low gluten wheat obtained upon mutation of 35 genes of the highly redundant α-gliadin family (Sánchez-León *et al.*, 2018), glyco-

engineered *N. benthamiana* plants with knockouts (KOs) in two xylosyl and four fucosyltransferase genes (Jansing *et al*., 2019) , or semi-dwarf rapeseed with increased yield with biallelic mutations in the two *MORE AXILLARY GROWTH1 (MAX1)* homeologue genes (Zheng *et al*., 2020).

Multiplex CRISPR/Cas constructs minimally involve three transcriptional units (TUs): a plant selection marker, the Cas nuclease, and at least two gRNA. Additional gRNAs can be expressed either from several promoters as separate TUs, or from a single promoter as a polycistronic transcript that is further processed resulting in the active gRNAs (Fig. 4). Cas9 has no ability to process gRNA tandem arrays, although recent studies in viral vectors seem to provide exceptions to this general rule (Uranga *et al*., 2021). Therefore, processable spacers need to be included in the array, so they can be rightly processed and trimmed into single functional units. Among the different spacer strategies described, the tRNA spacer method described by Xie *et al*. (2015) is being widely used in plant editing (K. Xie *et al*., 2015). This method relies on endogenous plant RNase P and RNase Z required to process the tRNAs flanking each spacer-scaffold unit.

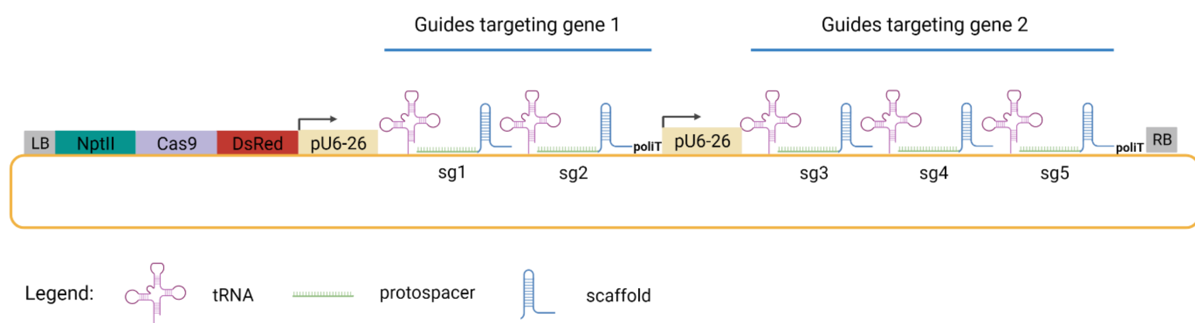


Figure 4. Example of a vector for CRISPR/Cas multiplexing.

In this case, the vector contains the plant selection marker nptII, which confers resistance to kanamycin, the Cas9 enzyme, the DsRed fluorescent protein for visual selection, and two transcriptional units comprising two guides targeting gene 1, and three guides targeting gene 2, each one under the control of U6-26 promoter. The sgRNAs transcriptional units are assembled following the tRNA strategy (K. Xie *et al*., 2015).

Genome editing in the context of biofactory breeding

Genome editing has been a powerful tool for biofactory breeding purposes during the last years. As already discussed earlier, CRISPR/Cas was used for different objectives in the context of chassis improvement. Briefly, the generation of humanised and non-immunogenic glycan profiles on secreted proteins by knocking out glycosyltransferases genes (Hanania *et al*., 2017; Jansing *et al*., 2019; Mercx *et al*., 2017) and *prolyl-4-hydroxylase* (*P4H*) genes (Parsons *et al*., 2013; Schoberer & Strasser, 2018), the elimination of nicotine in tobacco through the edition of the *BBL* gene family (Schachtsiek & Stehle, 2019), the increased recombinant protein expression by knocking out *RDR6* in *N. benthamiana* (Matsuo

& Atsumi, 2019) and specific proteases in moss (Hoernstein *et al*., 2018), the disruption of storage proteins in seed crops (A. Li *et al*., 2018; Lyzenga *et al*., 2019; Sánchez-León *et al*., 2018) to create space in storage organelles (Takaiwa, 2013). In addition to these examples, another approach involving CRISPR/Cas in biofactory breeding was the modulation of oxidases activity. During the extraction of recombinant proteins from plants, phenolic compounds are produced. Polyphenol oxidases (PPOs) in plants can form covalent complexes between these compounds and the produced proteins, causing protein aggregation and reduction in yield and product quality during extraction (Twyman *et al*., 2003). Knocking out *PPO* genes using genome editing offers an efficient solution to prevent phenolic oxidation. The knockout of a potato *PPO* gene mediated by CRISPR/Cas9 significantly reduced PPO enzyme activity in this plant without collateral effects (González *et al*., 2020). This approach could be used also in other chassis as tobacco and *N. benthamiana*. Nevertheless, spatiotemporal regulation may be necessary to avoid interference with normal plant growth, since PPOs are necessary for plant defence (Buyel *et al*., 2021).

Growth habits and developmental phase transitions in the context of biofactories

Increasing the yield of harvested biomass is crucial for PMF. This can be achieved by enhancing the assimilation of carbon or by implementing improved agronomical practices to optimize the yield potential of plants. Efforts to enhance efficiency of photosynthesis and carbon gain have involved conventional plant transformation methods and synthetic pathway introduction with remarkable gains in biomass (Głowacka *et al*., 2018; Kromdijk *et al*., 2016; South *et al*., 2019). Nonetheless, genome editing is a more effective technology for the mutation and manipulation of regulatory sequences to enhance or modulate the expression of genes, thereby improving photosynthetic efficiency and biomass production (Long *et al*., 2015). Genome-edited plants with enhanced photosynthetic efficiency and increased biomass production can serve as desirable PMF hosts. While plants can handle elevated levels of recombinant protein, biomass accumulation and growth can be affected by the need for protein synthesis capacity (Oey *et al*., 2009; J. A. Schmidt *et al*., 2019). Improvements in carbon assimilation, growth and energy conversion should not compromise the yield of recombinant protein. An approach to address this issue is to concentrate the recombinant protein in a specific district of biomass, such as seeds. This strategy can achieve both higher product yields and increased biomass accumulation (Takaiwa *et al*., 2017).

Modifying the height and architecture of plants is another strategy to impact biomass accumulation and facilitate bioprocessing in PMF. Cultivation conditions, particularly lighting, can be controlled to some extent to optimize plant properties, but the ideal conditions vary among plant species, leading to greater costs in process development and the requirement for advanced equipment, such as inter-lighting or wavelength-adjustable LED modules (Y. Park & Runkle, 2018; Poorter *et al*., 2012; Tewolde *et al*., 2018).

However, manipulating plant height through cultivation conditions may have unintended effects on secondary metabolites production that could complicate downstream processing (Buyel *et al*., 2015; Darko *et al*., 2014).

A promising way to increase biomass in biofactory chassis is to enhance branching, in order to have more available leaf surface to be infiltrated, as in the case of *Nicotiana* genus plants. The knockout through gene editing of the two *BnaMAX1* genes in rapeseed, homologs of *Arabidopsis* gene *MAX1,* resulted in lines with increased branching, and this had a positive impact in seed productivity (Zheng *et al*., 2020). Beside gene editing, cultural practices can also impact branching. Goulet *et al*. (2019) observed that the primary contributors to the yield of recombinant proteins in agroinfiltrated *N. benthamiana* are the young leaves in axillary stems. Spraying plants with the cytokinin 6-benzylaminopurine (6-BAP) increased branching (Goulet *et al*., 2019), and moreover, prolonging the photoperiod from 16 to 24 hours resulted in higher biomass and a considerably higher titre of influenza virus hemagglutinin H1 (Goulet *et al*., 2019).

Another aspect of biofactory breeding related to development that would need improvement is the biosafety: pollen and seeds dispersal must be avoided to prevent outcrossing with agricultural crops (Buyel *et al*., 2021). For this purpose, abolished flowering would be a desirable trait for biofactory breeding. In addition, since the optimal time frame for infiltrating *N. benthamiana* is week 5-6, when wild type plants flower, the prolongation of juvenility would be beneficial for PMF purposes, so that the expression capability of heterologous proteins would be at its maximum potential for longer periods. These are the reasons why plants of *Nicotiana* genus with suppressed flowering and prolonged juvenility are desirable for molecular farming goals.

One strategy for the improvement of growth habits is the manipulation of genes affecting phase transition, especially the ones related to flowering time and juvenility. The *SQUAMOSA PROMOTER BINDING-LIKE (SPL)* and *FLOWERING LOCUS T (FT)* genes are among the master regulators controlling these traits.

Mechanisms controlling flowering time: the predominant role of *FT* genes

Flowering plants employ various mechanisms to regulate floral development and enhance their reproductive success. The key regulatory proteins that coordinate these pathways were initially identified in *Arabidopsis thaliana*, and belong to the phosphatidylethanolamine-binding protein (PEBP) family (Chardon & Damerval, 2005; Hedman *et al*., 2009; Karlgren *et al*., 2011). Within the plant kingdom, ancestral PEBP genes have undergone duplication and divergence, resulting in regulatory proteins with antagonistic functions classified into three clades: in *Arabidopsis*, the FT-like clade comprises FLOWERING LOCUS T (FT) and TWIN SISTER OF FT (TSF), the TFL1-like clade includes TERMINAL FLOWER 1 (TFL1), BROTHER OF FT AND TFL1 (BFT), and ARABIDOPSIS THALIANA

CENTRORADIALIS (ATC), and the MFT-like clade consists of MOTHER OF FT AND TFL1 (MFT). The proteins of the FT-like and MFT-like clades are floral promoters (Chardon & Damerval, 2005; Hedman *et al*., 2009; Karlgren *et al*., 2011). In long-day conditions, FT is expressed in Arabidopsis in the leaf phloem companion cells and causes the beginning of floral development in the shoot apical meristem (SAM). This activation occurs since the transcription factor CONSTANS (CO) is stabilized by the light, and then activates FT (Samach *et al*., 2000; Valverde *et al*., 2004). Subsequently, the FT protein is transported to the SAM via sieve elements and interacts with the transcription factor FLOWERING LOCUS D (FD). This interaction activates downstream targets, like the floral meristem identity gene APETALA 1 (AP1) and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) (Abe *et al*., 2005; Corbesier *et al*., 2007; Wigge *et al*., 2005). Also TSF, another FT-like protein, and MFT act as floral promoters in *Arabidopsis* (Yamaguchi *et al*., 2005; S. Y. Yoo *et al*., 2004). In contrast, TFL1-like proteins, including TFL1 itself, BFT, and ATC, generally hinder floral development and function as floral repressors, frequently exhibiting overlapping roles (Mimida *et al*., 2001; S. J. Yoo *et al*., 2010). In the shoot apex, FT, TFL1, and other PEBP proteins interact with FD, resulting in an antagonistic regulation of downstream target genes (Abe *et al*., 2005; Hanano & Goto, 2011; Wigge *et al*., 2005). Chardon and Damerval (2005) conducted a phylogenetic analysis of the PEBP gene family in cereals and identified 19 FT-like genes in rice, 20 in wheat, 9 in barley, 30 in maize and 5 in sorghum, highlighting the conservation of FT-like genes across cereal species (Chardon & Damerval, 2005). While FT-like proteins have been identified in various plant species, their functions can differ. Harig *et al*. (2012) investigated the FT-like proteins in tobacco and identified four paralogs: NtFT1, NtFT2, NtFT3, and NtFT4. Surprisingly, they revealed distinct roles for each paralog. NtFT1, NtFT2, and NtFT3 were found to act as floral repressors, impeding the onset of flowering. Conversely, NtFT4 emerged as a floral inducer, facilitating the transition to flowering in tobacco plants (Harig *et al*., 2012). Further investigations led to the discovery of an additional FT-like protein in tobacco, which was named NtFT5. Functional analysis of *NtFT5* demonstrated that overexpression of this gene in tobacco plants resulted in an early flowering phenotype, highlighting its role as a crucial floral activator (G. Wang *et al*., 2018). In a groundbreaking advancement, Schmidt *et al*. (2020) employed CRISPR/Cas9 technology to explore the role of *NtFT5* as the primary floral inducer in tobacco. Through gene editing, they successfully knocked out *NtFT5* in tobacco plants. The edited tobacco plants exhibited a complete inability to flower in long-day conditions (F. J. Schmidt *et al*., 2020), providing evidence for the pivotal role of NtFT5 in the flowering process of this species.

The *SPL* genes and their role in controlling plant architecture and developmental phase transition

SQUAMOSA PROMOTER BINDING-LIKE PROTEINs (SPL) is a family of transcription factors specific to plants. Various aspects of plant physiology and development are controlled by *SPL* genes, such as

vegetative phase transition (Xu *et al.*, 2016), leaf initiation rate and shoot and inflorescence branching (Schwarz *et al.*, 2008; Wu & Poethig, 2006), flowering time (Gandikota *et al.*, 2007; Xu *et al.*, 2016), floral organ development and fertility (Xing *et al.*, 2010), fruit development and ripening (Ferreira e Silva *et al.*, 2014), the development of pollen sac (Unte *et al.*, 2003), root (Yamasaki *et al.*, 2009; Yu *et al.*, 2015) and trichomes (Yu *et al.*, 2010). The SBP domain, consisting of approximately 78 amino acid residues, is a characteristic of the proteins of this family (Birkenbihl *et al.*, 2005; Cardon *et al.*, 1999) and some SPL genes possess a conserved microRNA 156 (miR156) binding site. Multiple studies affirmed that miR156 levels play a pivotal role in regulating phase transition in plants. MiR156 tends to accumulate significantly during the seeding and juvenile phases, then its abundance decreases notably in the adult phase (Xu *et al.*, 2016; T. Zhang *et al.*, 2015). Simultaneously, as miR156 levels decline, there is a corresponding increase in the expression of miR156-targeted SPLs with age, ultimately culminating in the transition to reproductive phase (H. Wang & Wang, 2015) (Fig. 5).

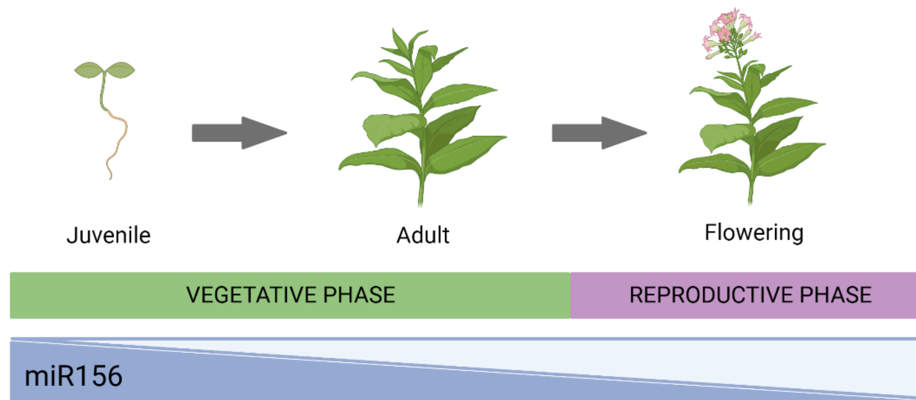


Figure 5. Decrease of miRNA156 level.

The level of miRNA156 decreases proceeding from vegetative to reproductive phase, allowing the expression of the SPL genes (H. Wang & Wang, 2015).

The two classes of genes described above, the *FT* and *SPL* genes, both contribute to promote flowering. The connection between them is the transcription factor SOC1. In Arabidopsis, under long day conditions, the FT-FD module activates SOC1, which in turn binds directly to the promoters of SPL3, SPL4 and SPL5. These genes activate downstream targets such as *APETALA1 (AP1)*, *FRUITFULL (FUL)* or *LEAFY (LFY),* floral meristem identity genes that start the floral transition (Balanzà *et al.*, 2014). The FT-FD module can also directly bind to SPL3/4/5 promoters. Under short day conditions, gibberellic acid (GA) directly activates SOC1, which in turn activates SPL3/4/5 and floral meristem identity genes (Jung *et al.*, 2012) (Fig. 6).

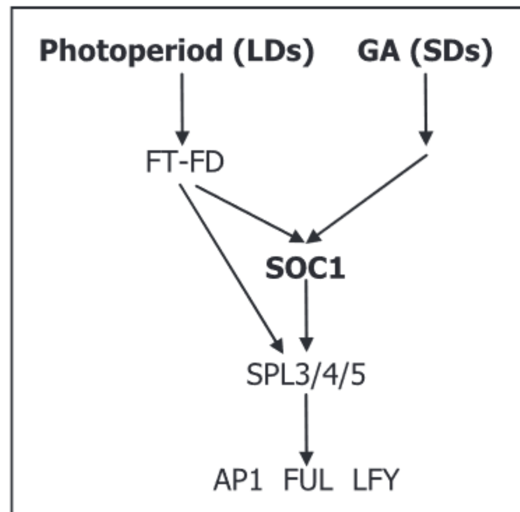


Figure 6. Cascade of activation involving *FT*, *SPL* genes and SOC1.

This cascade of activation ultimately leads to floral transition (Jung *et al*., 2012).

FTs and *SPLs* as possible targets for biofactory breeding

The reasons for considering *FT* and *SPL* genes as targets for biofactory breeding, particularly in the case of *N. benthamiana*, has been outlined above. One of the most advantageous traits that research on molecular farming may promote in *N. benthamiana* is the abolition of flowering. By creating a chassis with this specific feature, the suitability of this plant for molecular farming purposes could significantly increase. Such a chassis, devoid of the ability to flower, would exhibit several desirable characteristics for large-scale bioproduction. Firstly, a non-flowering *N. benthamiana* would direct its energy and resources away from reproductive processes, allowing to allocate more of its metabolic means towards vegetative growth. As a result, the plant would accumulate significantly higher levels of biomass compared to WT plants. The increased biomass could be employed for higher yields of biofactory products, such as pharmaceuticals, enzymes, or metabolites. Moreover, the inability to flower would prevent the release of pollen and seeds, reducing the risk of unwanted outcrossing with cultivated or wild plant species in the vicinity (Buyel *et al*., 2021). This containment measure is crucial for maintaining the genetic integrity of the engineered *N. benthamiana*, especially if grown in open fields or greenhouses alongside other crops.

To achieve these benefits, in this work *SPL13* was first identified as a candidate gene in *N. benthamiana* and then CRISPR/Cas technology was utilized to specifically edit *SPL13, FT4* and *FT5* genes in this species. By knocking out these genes, the flowering process would be compromised, leading to a possible non-flowering phenotype.

Objectives

The aim of this work was to improve *Nicotiana benthamiana* as a chassis for molecular farming by generating new mutant lines with delayed flowering time. To this end, we aimed to characterize the *Nicotiana benthamiana* gene families *SQUAMOSA PROMOTER BINDING-LIKE (NbSPL)* and *FLOWERING LOCUS T (NbFT)* and to generate loss-of-function mutants of selected members of both families using CRISPR/Cas9 gene editing tools. In particular, the specific objectives of this thesis were:

- To perform a genome-wide analysis of *SQUAMOSA PROMOTER BINDING-LIKE (SPL)* gene family in *N. benthamiana* and tobacco, identifying those gene family members that could have a greater influence in controlling flowering time (Chapter 1).
- To knockout selected *NbSPL* members, analysing the effect in delaying *N. benthamiana* flowering time (Chapter 1).
- To knockout selected *NbFT* family members analysing the effect in delaying *N. benthamiana* flowering time (Chapter 2).
- To combine mutations in selected *NbSPL* and *NbFT* family members and to analyse the effect in *N. benthamiana* flowering time and recombinant protein production (Chapter 2)

Chapter 1. Comparative analysis of the *SQUAMOSA PROMOTER BINDING-LIKE (SPL)* gene family in *Nicotiana benthamiana* and *Nicotiana tabacum*

Carmine De Paola, Victor Garcia-Carpintero, Marta Vázquez-Vilar, Kacper Kaminski, Asun Fernández-del-Carmen, Nicolas Sierro, Nikolai V. Ivanov, Giovanni Giuliano, Peter Waterhouse, Diego Orzáez.

Plant Science, 2023. DOI: 10.1016/j.plantsci.2023.111797

My contribution to this chapter was essential. I contributed to the search for the *SPL* genes in the *N. benthamiana* genome, I performed the RNA extraction in *N. benthamiana* leaf samples, I generated, genotyped and phenotyped the *NbSPL13* lines. I wrote and corrected the text of the chapter and generated its figures.

Abstract

SQUAMOSA PROMOTER BINDING-LIKE (SPL) proteins constitute a large family of transcription factors known to play key roles in growth and developmental processes, including juvenile-to-adult and vegetative-to-reproductive phase transitions. This makes SPLs interesting targets for precision breeding in plants of the *Nicotiana* genus used as e.g. recombinant biofactories. The identification of 49 *SPL* genes in *Nicotiana tabacum* cv. K326 and 43 *SPL* genes in *Nicotiana benthamiana* LAB strain is reported. These genes were classified into eight phylogenetic groups according to the SPL classification in *Arabidopsis*. Exon-intron gene structure and DNA-binding domains were highly conserved between homeologues and orthologues. Thirty of the *NbSPL* genes and 33 of the *NtSPL* genes were found to be possible targets of microRNA 156. The expression of *SPL* genes in leaves was analysed by RNA-seq at three different stages, revealing that genes not under miR156 control were in general constitutively expressed at high levels, whereas miR156-regulated genes showed lower expression, often developmentally regulated. The *N. benthamiana SPL13_1a* gene was selected as target for a CRISPR/Cas9 knock-out experiment. It is shown here that a full knock-out in this single gene leads to a significant delay in flowering time, a trait that could be exploited to increase biomass for recombinant protein production.

Keywords: Plant biofactories; *Nicotiana benthamiana;* tobacco; SQUAMOSA PROMOTER BINDING-LIKE; CRISPR/Cas9; flowering time

Abbreviations: SPL: SQUAMOSA promoter binding-like; miRNA: micro-RNA; CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats; Cas9: CRISPR associated protein 9.

Introduction

SQUAMOSA PROMOTER BINDING-LIKE PROTEINs (*SPL*) is a family of plant-specific transcription factors. *SPL* genes control many aspects of plant development and physiology, including vegetative phase transition (Xu *et al*., 2016), flowering time (Gandikota *et al*., 2007; Xu *et al*., 2016), leaf initiation rate and shoot and inflorescence branching (Schwarz *et al*., 2008; Wu & Poethig, 2006), fruit development and ripening (Ferreira e Silva *et al*., 2014), floral organ development and fertility (Xing *et al*., 2010), pollen sac development (Unte *et al*., 2003), trichome development (Yu *et al*., 2010), root development (Yamasaki *et al*., 2009; Yu *et al*., 2015), and stress responses (Mao *et al*., 2016; Zeng *et al*., 2019). The first two *SPL* genes were discovered in *Antirrhinum majus* and were named SQUAMOSA PROMOTER BINDING PROTEIN (AmSBP1 and AmSBP2) due to their *in vitro* binding activity with the promoter of the floral meristem identity gene SQUAMOSA (Klein *et al*., 1996). After this discovery, *SPL* genes were found in green algae, mosses, gymnosperms, and angiosperms (Preston & Hileman, 2013).

Proteins belonging to the SPL family are characterized by the presence of the SBP domain, a DNA binding domain of approximately 78 amino acid residues. This domain contains two zinc-finger motifs, Cys–Cys–Cys–His (Zn1) and Cys–Cys–His–Cys (Zn2), with the second motif partially overlapping a nuclear localization signal at the C-terminal of the SBP domain (Birkenbihl *et al*., 2005; Cardon *et al*., 1999). Additionally, some *SPL* genes contain conserved microRNA 156 (miR156) binding sites. MicroRNAs are non-coding RNAs that can complementarily bind to target sites and repress expression via mRNA cleavage or repression of translation (Rogers & Chen, 2013). The miR156 complementary sites are present in the coding region or in the 3' untranslated region (3'-UTR) of several *SPL* genes. Numerous studies confirmed that the level of miR156 is responsible for phase transitions in plants - MiR156 accumulates at high levels at seedling stage and during the juvenile phase, while its amount is significantly reduced in the adult phase (Xu *et al*., 2016; T. Zhang *et al*., 2015). In parallel to the decrease of miR156 levels, the expression level of miR156-targeted *SPLs* rises with age, ultimately leading to the reproductive phase transition (H. Wang & Wang, 2015).

The specific role of each *SPL* gene in the plant has been extensively studied in *Arabidopsis thaliana*. *Arabidopsis* possesses 16 *SPL* genes (Cardon *et al*., 1999). These can be divided in eight clades: *AtSPL1/12/14/16, AtSPL2/10/11, AtSPL3/4/5, AtSPL6, AtSPL7, AtSPL8, AtSPL9/15*, and *AtSPL13* (Preston & Hileman, 2013; Z. Yang *et al*., 2008). In *Arabidopsis*, 10 out of the 16 *SPL* genes are targets of miR156 (Gandikota *et al*., 2007; Wu & Poethig, 2006). Among the ones regulated by miR156, *AtSPL2/9/10/11/13/15* contribute to both the juvenile-to-adult phase transition (vegetative phase change) and the vegetative-to-reproductive phase transition (reproductive phase change or flowering), with *AtSPL9/13/15* being more important than *AtSPL2/10/11* (Xu *et al*., 2016). *AtSPL3/4/5* do not play a major role in vegetative phase change or flowering but promote the floral meristem identity transition (Xu *et al*., 2016). *AtSPL6* does not have a major function in vegetative morphogenesis (Xu *et al*., 2016),

but it can positively regulate a subset of defence genes and plays a role in effector-triggered immunity (Padmanabhan *et al*., 2013). Among the ones not regulated by miRNA156, *AtSPL7* is a central regulator of copper homeostasis and plays a major role in cadmium response (Gielen *et al*., 2016; Yamasaki *et al*., 2009), *AtSPL8* plays pivotal roles in regulating pollen sac development, male fertility, and gibberellin (GA) biosynthesis and signalling (Unte *et al*., 2003; Xing *et al*., 2010; Y. Zhang *et al*., 2007) and *AtSPL14* regulates plant development and sensitivity to fumonisin B1 (Stone *et al*., 2005).

Genome-wide analyses of the *SPL* gene family were performed in many plant species other than *Arabidopsis* (Cardon *et al*., 1999), including rice (Z. Yang *et al*., 2008), maize (Mao *et al*., 2016), cotton (Cai *et al*., 2018), barley (Tripathi *et al*., 2018) , tomato (Salinas *et al*., 2012), citrus (Zeng *et al*., 2019), poplar (C. Li & Lu, 2014), Chrysanthemum (Song *et al*., 2016) , Moso Bamboo (Pan *et al*., 2017), Petunia (Zhou *et al*., 2018) and Tartary Buckwheat (M. Liu *et al*., 2019). In tobacco, there is one report in which Han *et al*. identified and characterized 15 SPL genes in *N. tabacum* L. cv. Qinyan95 (Han *et al*., 2016). However, a genome-wide identification of *SPL* genes in the plants of the *Nicotiana* genus is not available.

Nicotiana benthamiana and *Nicotiana tabacum* are two plant species of the *Nicotiana* genus with suitable properties to be engineered into efficient biofactories for high value-added compounds production (Derevnina *et al*., 2019; Goodin *et al*., 2008; Sierro *et al*., 2014). Both species have a high metabolic versatility and a non-food status. Moreover, several biotechnological tools have been developed for their genetic manipulation. Among others, transient recombinant gene expression via Agroinfiltration is widely used for research and bioproduction purposes as it results in high yields of recombinant protein in both species. As an initial step towards exploiting the potential of *SPL* targeted mutagenesis for the breeding of *N. tabacum* and *N. benthamiana* biofactories, in this study the genome-wide identification and characterization of the *SPL* genes in *N. benthamiana* LAB strain and in *N. tabacum* cv. K326 is reported. Fourty-nine (48 with SBP) putative *SPL* genes in *N. tabacum* cv. K326 were identified. The high quality of the recently released new version of the *N. benthamiana* genome (Ranawaka *et al*., 2023) also allowed us to identify 43 (37 with SBP) candidate *SPL* genes in *N. benthamiana* LAB strain. Additionally, the gene structure, the conserved motifs, and the expression profile of the identified SPLs were analysed. It is also described how the CRISPR/Cas9-directed knock out of a single gene in the smallest *SPL* subfamily in *N. benthamiana*, the *SPL13* clade comprising three genes (only one with SBP), leads to a consistent delay of flowering initiation time of approximately five days, a change that could impact production yields.

Materials and Methods

Search of SPL genes in *N. benthamiana* and *N. tabacum* cv. K326

SPL families from *Arabidopsis thaliana* (TAIR10), *Solanum lycopersicum, Nicotiana tomentosiformis* and *Nicotiana sylvestris* were retrieved and protein sequences were used to search for homologs in the *N. benthamiana* (LAB330, version 3.02 https://www.nbenth.com/) (Ranawaka *et al*., 2023) and *N. tabacum* cv. K326 (Nitab v4.5 Genome Scaffolds Edwards 2017, https://solgenomics.net/tools/blast/) (Edwards *et al*., 2017) reference genomes using TBLASTN. Matches with a 50% identity and 50% of coverage were checked for annotated gene models. Gene models were aligned to the National Center for Biotechnology Information (NCBI) non-redundant protein sequences database (Sayers *et al*., 2022) using BLASTP (Altschul *et al*., 1990). Gene models with top matches against annotated SPL proteins were kept as *N. benthamiana* and *N. tabacum* cv. K326 SPL family members. In order to verify the completeness of the SPLs search for both *N. benthamiana* and *N. tabacum* cv. K326*,* the smallest SPL member of each species was selected and aligned against all protein sequences for each genome using BLASTP.

Reannotation of *N. tabacum* cv. K326 SPLs members

SPLs identified for *N. tabacum* cv. K326 were aligned to the NBCI non-redundant protein sequence dataset (Sayers *et al*., 2022) using BLASTP (Altschul *et al*., 1990). For each one, the top match originating from *N. tabacum*, *N. sylvestris* or *N. tomentosiformis* was selected, using NCBI's RefSeq curated gene models when possible and compared with K326 gene model. If either number of exons and/or sequence length was different, a new gene model for K326 was searched with Exonerate V2.2 (Slater & Birney, 2005), using the protein from NCBI as query and the K326 source scaffold (Edwards *et al*., 2017) or the sequences obtained from the updated version (publicly not available) of the *N. tabacum* reference genome as target for the new model. Updated *SPL* gene models are listed in Table S2. If the new gene model could not fit neither in K326 scaffold nor in the updated version of the *N. tabacum* reference genome, the Nitab v4.5 original model was kept. New gene models were verified with RNA-seq expression data from Solgenomics by visual inspection. Finally, subgenome donors for *N. tabacum* cv. K326 SPL sequences were assigned by BLASTP against NCBI non-redundant protein database.

Identification of SBP conserved domains and motif analysis

Identified SPL protein sequences from both *N. benthamiana* and *N. tabacum* cv. K326 were scanned for SBP domains using MOTIF Search (https://www.genome.jp/tools/motif/) with default parameters. The conserved motifs along SPL protein sequences were detected by MEME software (http://meme.nbcr.net/meme/cgibin/meme.cgi) and SBP sequence visualization was performed using multiple alignment program MAFFT version 7 (http://mafft.cbrc.jp/alignment/server/; (Katoh *et al*., 2019)).

Gene structure and miRNA156 complementary regions

For *N. benthamiana,* coding sequences (CDS) and 3' untranslated regions (3'UTRs) were retrieved from *N. benthamiana* LAB330 v3.02 gene models. For *N. tabacum* cv. K326*,* CDS sequences from the old and new gene models were aligned against Nitab v4.5 Genome Scaffolds Edwards 2017 (https://solgenomics.net/tools/blast/) (Edwards *et al*., 2017), retrieving 5,000 bp downstream of the CDS. Both CDS sequences and downstream sequences were scanned for miRNA156-complementary regions using a *N. tabacum* miRNA database with psRNATarget (Dai *et al*., 2018) and default parameters. MiR156 binding sites found in the CDS were kept for both species. All miR156 matches in the 3'UTRs from *N. benthamiana* were also kept, while for *N. tabacum* cv. K326 matches farther than 500 bp downstream the CDS were discarded. Browser Extensible Data (BED) files containing *SPL* gene models for each *Nicotiana* species were created with an in-house python script, manually adding 500bp UTRs windows to *N. tabacum* cv. K326 gene models. These files were graphically represented using GSDS (Gene Structure Display Server) (http://gsds.cbi.pku.edu.cn/).

Phylogenetic tree and family classification

A multiple sequence alignment (MSA) between *A. thaliana*, *N. benthamiana* and *N. tabacum* cv. K326 SPL protein sequences was performed with MUSCLE v3.8.31 (Edgar, 2004). A phylogenetic tree was constructed with iqtree v1.6.12 (Nguyen *et al*., 2015) using model VT+F+R4 and a bootstrap value of 1000. All SPLs were named according to the *A. thaliana* TAIR10 SPL gene belonging to the same clade. Then, in *N. tabacum* cv. K326 the average protein length for each pair in the clade was calculated, number 1 was assigned to the pair with the highest value, number 2 to the following, and so on. Each pair member was marked as "a" if the most likely subgenome donor was *N. sylvestris* and "b" if it was *N. tomentosiformis*. *N. benthamiana* members were named considering their phylogenetic relationship to *N. tabacum* cv. K326 SPLs: if named with "a" or "b" the most similar subgenome is the *N. sylvestris* or *N. tomentosiformis*, respectively, but if named with "x" or "y" a similar subgenome could not be established, with "x" indicating one donor and "y" another one. Additionally, genes not having a partner were named with an "U", standing for "unique".

Plant material, growing conditions and samples collection for RNA extraction

N. benthamiana LAB strain and *N. tabacum* cv. K326 plants were grown under a 16-h light (24ºC)/8-h dark (20ºC) regime in growth chamber. For *N. benthamiana*, the whole fifth true leaf from the main axis was collected. Three biological replicates were made, each of them coming from a pool of three leaves. Samples were collected at four weeks (Juvenile stage: J), five weeks (Pre-flowering stage: P), six weeks (Flowering stage: F). For *N. tabacum* cv. K326, the whole fifth true leaf was collected at 60 days (J), 75 days (P), 100 days (F). Three biological replicates were made, each of them coming from an individual plant.

N. benthamiana phenotyping

For each plant, the day (post sowing) in which the first flower bud was visible was registered, and considered as flowering time. Regarding branching, the lateral branches of each plant were counted, a first time when WT plants flowered, and a second time, when each SPL13-edited plant flowered. All secondary growth axes, including nascent axillary meristems, emerging from the main axis were recorded as lateral branches.

RNA extraction and sequencing

Leaf samples were ground in liquid nitrogen and stored at -80 °C. RNA was extracted with GeneJET RNA purification kit from Thermofisher (USA) following the manufacturer instructions. Extracted RNA samples were prepared with Universal Plus mRNA-Seq with NuQuant. Libraries were later sequenced with an Illumina NovaSeq® 6000 System. Paired end (PE) 2x150bp sequencing was performed with NovaSeq6000 - Dual Index - Paired End - S4 - XP protocol. Sequencing data generated was demultiplexed by Illumina BaseSpace® Clarity LIMS (© Illumina, Inc., USA)

Expression analysis

Sequence reads were quality checked using FastQC v. 0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Raw reads were quality trimmed and Illumina adaptors were removed with Trimmomatic version 0.39 (Bolger *et al*., 2014). Next, HISAT2 v2-2.2.1 (D. Kim *et al*., 2019) was used for mapping the reads. *N. benthamiana* and *N. tabacum* reads were aligned against the *N. benthamiana* genome, version 3.3, (https://www.nbenth.com/) (Ranawaka *et al*., 2023) and against the *N. tabacum* genome, version Nitab v4.5, available at Solgenomics (https://solgenomics.net/organism/Nicotiana_tabacum/genome) (Edwards *et al*., 2017), respectively. Transcript abundances were calculated for *N. benthamiana* considering annotated gene models v3.02 (Ranawaka *et al*., 2023) while for *N. tabacum* cv. K326 considering annotated gene models version Nitab v4.5 (Edwards *et al*., 2017) and using StringTie 2.1.6 (Pertea *et al*., 2015). From these counts a gene expression table of raw read counts was generated. Genes from this table were filtered out if expression was not found for each development stage. Then expression levels were normalized by trimmed mean of M-values with EdgeR (Robinson *et al*., 2010) and an expression table in CPM units was generated.

Plasmid assembly

Constructs used for transformation were assembled using GoldenBraid (Vazquez-Vilar *et al*., 2020, 2021). For the assembly of guide RNAs on level 0, two partially complementary primers were designed at https://gbcloning.upv.es/do/crispr/multi_cas9_gRNA_domesticator_1 using as input the sequences of Table S1. Primers were included in a BsmBI restriction–ligation reaction together with pUPD2 and the corresponding level − 1 tRNA-scaffold plasmid (GB1208 for sgSPL1.5 and GB1207 for sgSPL1.6). Later, multipartite *Bsa*I restriction–ligation reactions from level 0 parts and binary *Bsa*I or *BsmBI* restriction–ligation reactions were performed to obtain all the level ≥ 1 assemblies. All plasmids were

validated by restriction enzyme (RE) analysis. The sequences of all level ≥ 1 constructs can be found entering their IDs (displayed at Table S3) at https://gbcloning.upv.es/search/features/.

Plant material and genetic transformation

The *N. benthamiana* LAB strain was used for transformation with *Agrobacterium tumefaciens* following a standard protocol (Horsch *et al.*, 1985). Briefly, fully expanded leaves of WT plants were sterilized with 5% commercial bleach for 10 minutes followed by four consecutive washing steps with sterile deionised water. Leaf discs (d= 0.8 cm) were cut with a cork borer and incubated overnight in co-culture plates (4.9 g/L MS supplemented with vitamins (Duchefa, The Netherlands https://www.duchefa-biochemie.com/), 3% sucrose (Sigma-Aldrich, USA, https://www.sigmaaldrich.com/), 0.8% Phytoagar (Duchefa, The Netherlands), 1 mg/L BAP (Sigma-Aldrich, USA), 0.1 mg/L NAA (Sigma-Aldrich, USA), pH=5.7). Leaf discs were incubated for 15 minutes with a culture of *A. tumefaciens* LBA4404 harboring plasmid GB3298 (OD600=0.3). Discs were returned to the co-cultivation plates and incubated for two days in darkness. Next, discs were transferred to selection medium (4.9 g/L MS supplemented with vitamins (Duchefa, The Netherlands), 3% sucrose (Sigma-Aldrich, USA), 0.8% Phytoagar (Duchefa, The Netherlands), 1 mg/L BAP (Sigma-Aldrich, USA), 0.1 mg/L NAA (Sigma-Aldrich, USA), 500 mg/L carbenicillin, 100 mg/L kanamycin, pH=5.7). Discs were transferred to fresh medium every seven days until shoots appeared (four-six weeks). Shoots were cut and transferred to rooting medium (4.9 g/L MS supplemented with vitamins (Duchefa, The Netherlands), 3% sucrose (Sigma-Aldrich, USA), 0.8% Phytoagar (Duchefa, The Netherlands), 500 mg/L carbenicillin, 100 mg/L kanamycin, pH=5.7) until roots appeared.

Transient expression asssays were performed as described in Moreno-Giménez *et al.* (2022) with minor differences. Five-weeks old *N. benthamiana* WT and NbSPL13_1a/1a/1b/1b (biallelic mutation for both homeologous genes) were used for Agroinfiltration. Bacterial suspensions were adjusted to an optical density of 0.05 at 600 nm (OD600). For enhanced GFP (eGFP) expression with the TMV-based expression system ICON (Giritch *et al.*, 2006), the bacterial suspensions harboring the MagnICON® Integrase (pICH14011), the MagnICON® 5' module (pICH17388) and the 3' eGFP module (GB4294, eGFP cloned in a vector adapted for BsaI cloning from MagnICON® pICH7410 (Diego-Martin *et al.*, 2020)) were mixed in equal volumes. For eGFP expression with a geminiviral replicon system based on the Bean Yellow Dwarf Virus (BeYDV) (Dahan-Meir *et al.*, 2018), equal volumes of bacterial suspensions harbouring plasmids GB3598 and GB4312 were mixed. Leaf samples were collected at four and seven days post infiltration (dpi). For the determination of fluorescence, 0.5 cm diameter disc were excised from Agroinfiltrated leaves and transferred to a black 96-well microplate. Subsequently, enhanced GFP (eGFP) fluorescence was determined using microplate reader Victor™ X5 (Perkin Elmer, USA) following the manufacturer instructions.

Genomic DNA extraction and editing efficiency evaluation

150 mg of leaf material was used for genomic DNA extraction with the CTAB (cetyl trimethylammonium bromide) method (Murray & Thompson, 1980) . The genomic regions flanking the nuclease target sites were PCR amplified using MyTaq™ DNA Polymerase (Bioline, https://www.bioline.com/) and primers listed on Table S4. The PCR amplicons were confirmed on a 1% agarose gel electrophoresis and purified with ExoSAP-IT™ PCR Product Cleanup Reagent (ThermoFisher Scientific, https://www.thermofisher.com) following the manufacturer's indications prior to Sanger sequencing. Chromatograms of Cas9-edited genomic DNA were analyzed using Inference of CRISPR Edits (ICE) v2 tool from Synthego (https://ice.synthego.com/). All analyses were manually curated.

Results

Phylogenetic analysis of the SPL family in *N. tabacum* and *N. benthamiana*

Putative SPL protein sequences from *N. tabacum* cv. K326 and *N. benthamiana* were retrieved as described in Materials and Methods. These sequences were aligned together with SPLs from *A. thaliana* and a phylogenetic tree was constructed using maximum likelihood method, with a bootstrap value of n = 1000 iterations (Fig. 7). All *SPL* genes were named as explained in Materials and Methods. In total, 49 *N. tabacum* cv. K326 and 43 *N. benthamiana* genes were identified and classified together with their *A. thaliana* homologues in eight differentiated clusters, namely *SPL1/12, SPL8, SPL13, SPL6, SPL7, SPL2/10/11, SPL9/15*, and *SPL3/4/5*. The different groups were named according to the *Arabidopsis SPL* gene appearing in the same clade. As it can be observed in the tree, all newly catalogued *Nicotiana* genes were distributed in one of the *Arabidopsis*-defined groups, but no *Nicotiana SPL* genes were found belonging to the *SPL14 Arabidopsis*-defined group. There was a markedly unequal distribution of the number of genes per clade. For example, the *SPL3/4/5* clade comprised 13 genes in *N. tabacum* cv. K326 and 12 in *N. benthamiana*, while the *SPL7* clade comprised only two genes in each species. Given the allotetraploid nature of both *Nicotiana* species, in most cases *SPL* genes were grouped in pairs of homeologous genes. When possible, each *SPL* of the same homeologous pair was assigned to the putative parental genome, which for *N. tabacum* cv. K326 are *N. sylvestris* and *N. tomentosiformis* (Sierro *et al*., 2014). Regarding *N. benthamiana*, performing this task was not so straightforward, due to the uncertainty about its origins. It was already hypothesized that *Noctiflorae* and *Sylvestres* sections were implicated in the genesis of *N. benthamiana* (Chase *et al*., 2003; Clarkson *et al*., 2004; Knapp *et al*., 2004), while a more recent study suggested that it could have originated from an introgression of *Petunioides* section into a member of *Noctiflorae* section that later hybridized with a member of *Sylvestres* section (Schiavinato *et al*., 2020). Given its complex origin, we did not have sufficient information to separate the subgenomes of *N. benthamiana*, and therefore it was not possible to assign each *SPL* gene from the same homeologous pair to a specific parental genome. To facilitate their identification, the genes in the tree were given names that paired with their closest homeologs (e.g., 1a-1b, 2a-2b, 3x-3y, etc). Genes having the same number and letter for *N. benthamiana* and *N. tabacum* cv. K326 were homologous, allegedly coming from a common ancestor (for example *NtSPL7_1a* and *NbSPL7_1a*). In some cases, there were *SPL* genes that did not have an homeologous partner, probably due to loss during evolution. Therefore, they were named with a number followed by U, standing for “unique”.

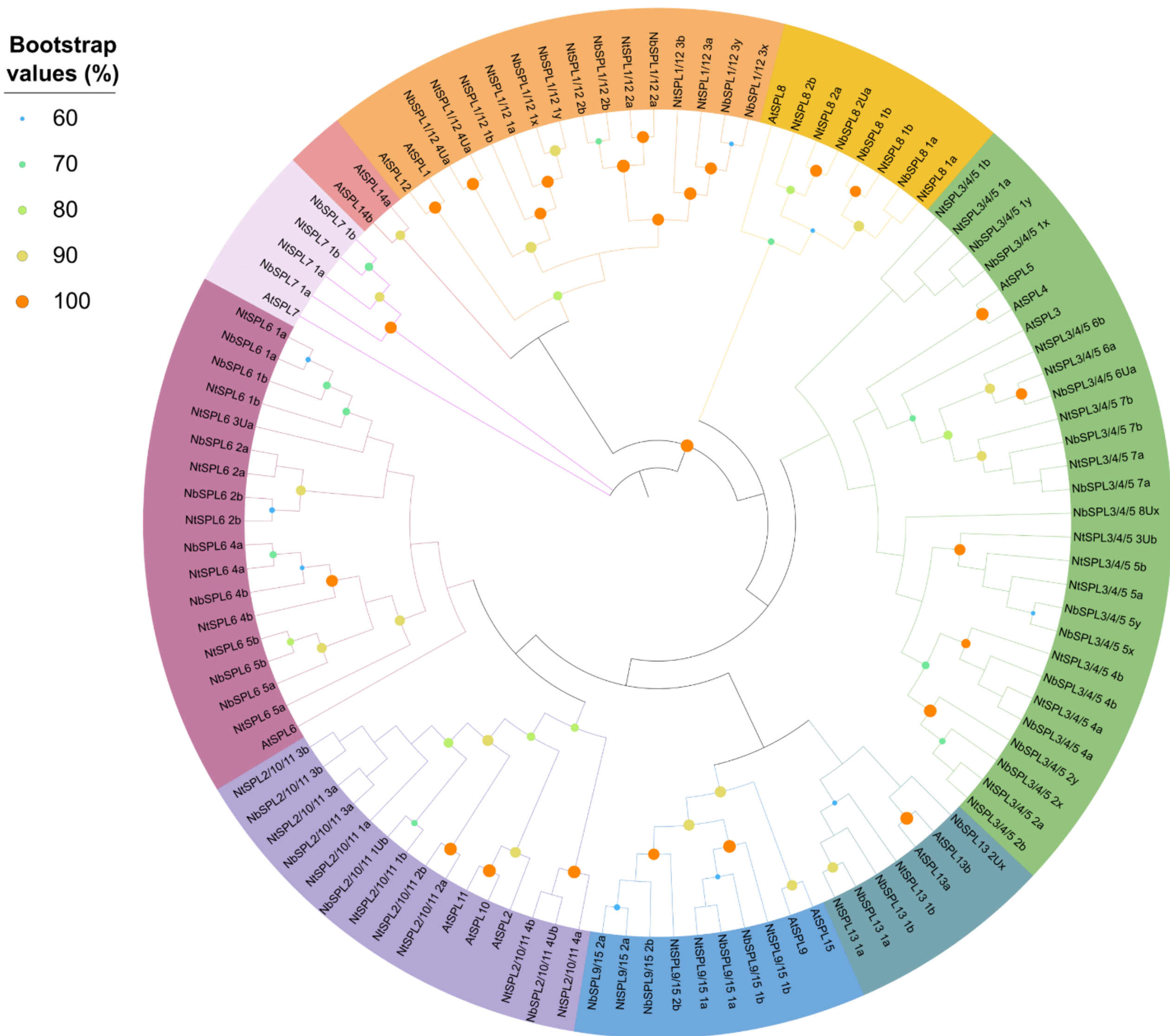


Figure 7. Phylogenetic tree with *SPL* genes from *N. tabacum* cv. K326, *N. benthamiana* and *A. thaliana*.

Phylogenetic tree was constructed from 16 *A. thaliana*, 49 *N. tabacum* cv. K326 and 43 *N. benthamiana* SPL proteins using the maximum-likelihood method with a bootstrap value of n = 1000 iterations. Bootstrap support values are represented in percentages.

Protein structure of SPL genes: the SBP domain

The length and structure of the *N. benthamiana* and *N. tabacum* cv. K326 SPL proteins was analysed and shown in Table S2. Among all the clades, groups SPL3/4/5 and SPL13 were the ones with the smallest proteins. All *N. benthamiana* and *N. tabacum* cv. K326 SPLs in these groups had a range of amino acids from 100 to 200 except for NbSPL3/4/5_1y (395 aa), NtSPL3/4/5_1a (402 aa) and NtSPL3/4/5_1b (380 aa). The largest proteins were those in clade SPL1/12 with 900–1000 amino acids, being the only exception NbSPL1/12_4Ua with 362 amino acids.

To gain a better understanding of the SPL protein characteristics in *N. benthamiana* and *N. tabacum* cv. K326, their amin8o acid sequences were analysed using the MEME software to find conserved motifs

along them (Fig. 8A, B). In total, ten conserved motifs were identified, but only members in the SPL1/12 group contained all ten motifs. Motifs 1, 2 and 3 represent together the SBP domain, which is close to the N-terminus and is key for the function of the SPL genes. Some proteins had a shorter SBP domain. This occurred in NtSPL13_1a, NtSPL3/4/5_5b, NtSPL6_5a, NtSPL8_2b, NbSPL2/10/11_1Ub and in the pair NbSPL3/4/5_5. Some proteins that were initially identified as SPLs using the whole sequence similarity criteria, turned out not to contain SBP domains, and were not included in fig. 8. For *N. benthamiana*, these were NbSPL1/12_4Ua, NbSPL13_1b, NbSPL13_2Ux, NbSPL3/4/5_1x, NbSPL3/4/5_8Ux and NbSPL6_1a. In *N. tabacum* cv. K326, only protein NtSPL2/10/11_3b was missing the SBP domain. Only proteins of clades SPL1/12 and SPL7 had additional motifs to 1, 2 and 3. Seven extra motifs were present in the case of SPL1/12 and two in SPL7, most of them with unknown function. Sequence alignment of the SBP domains of all NbSPLs and NtSPLs showed several highly conserved amino acids and a conserved structure (Fig. 8C). Fig. 8D shows the SBP domain of SPL9/15 from both species as an example. In this alignment the three motifs mentioned above can be clearly identified: the first Zinc-finger motif (Zn-1) Cys-Cys-Cys-His, the second one (Zn-2) Cys-Cys-His-Cys, and the nuclear localization signal (NLS).

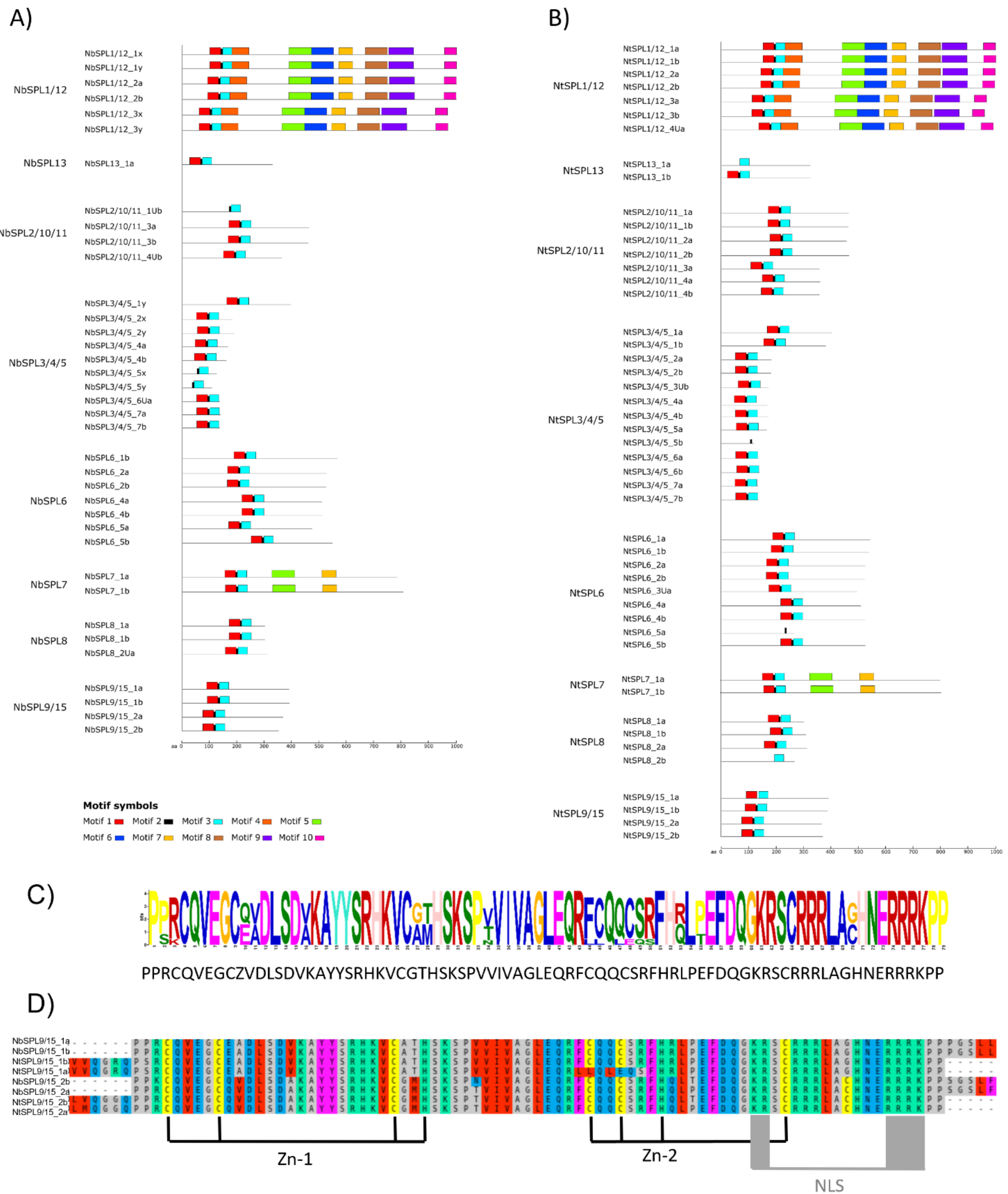


Figure 8. Motifs composition and SBP domain of NbSPLs and NtSPLs.

(A) Motifs along the NbSPL protein sequences, detected with MEME. Only *SPL* genes with SBP domain are shown. (B) Motifs along the NtSPL proteins sequences, detected with MEME. Only *SPL* genes with SBP domain are shown. (C) Motif logo and consensus sequence of the SBP domain of SPL9/15 proteins. Bits represent the conservation of sequence at a certain position. (D) Alignment of the SBP domains of *N. benthamiana* and *N. tabacum* cv. K326

SPL9/15 proteins. Multiple sequences alignment was performed using MAFFT version 7. The two Zn-finger like structures (Zn-1 and Zn-2) and the nuclear localization signal (NLS) are indicated.

Gene structure of *SPL* genes: exon-intron structure, putative miRNA156 target site

Nucleotide sequences of *SPL* genes were analysed for their exon-intron structure and for the presence of miR156 binding sites. As observed in Fig. 9, each clade had a well-defined exon-intron structure which was fairly conserved in all its members, but which differed strongly from other clades. Similarly, the presence/absence and the position of the putative target site for miR156 was a group-defining feature. Clades *SPL1/12*, *SPL7* and *SPL8* lacked miR156 target sites, while in the remaining clades most of their members contain a possible target position for miRNA156. The few exceptions to this general rule (e.g., *NbSPL2/10/11_1Ub, NtSPL6_5a* or *Nt_SPL3/4/5_5b*) were genes classified within a given clade but lacking the miRNA156 putative target site. These genes had also non-conserved exon-intron structures and were likely to be truncated genes or pseudogenes. Interestingly, both clade-specific gene structures and miRNA156 positions were strikingly well conserved between the two species, suggesting a strong selection and therefore a functional significance for both types of features. As expected, *N. tabacum* cv. K326 contained the same or more genes per clade than *N. benthamiana,* with the only exception of the *SPL13* clade, with three members in *N. benthamiana* and only two in *N. tabacum* cv. K326. However, as it can be observed also in Fig. 9 (left), only one of the representatives of the *NbSPL13* group had the expected exon-intron structure (and the SBP domain), whereas the remaining two had a smaller gene size and lacked a SBP domain, suggesting a possible ongoing process of pseudogenization for these two genes (*SPL13_1b* and *SPL13_Ux*).

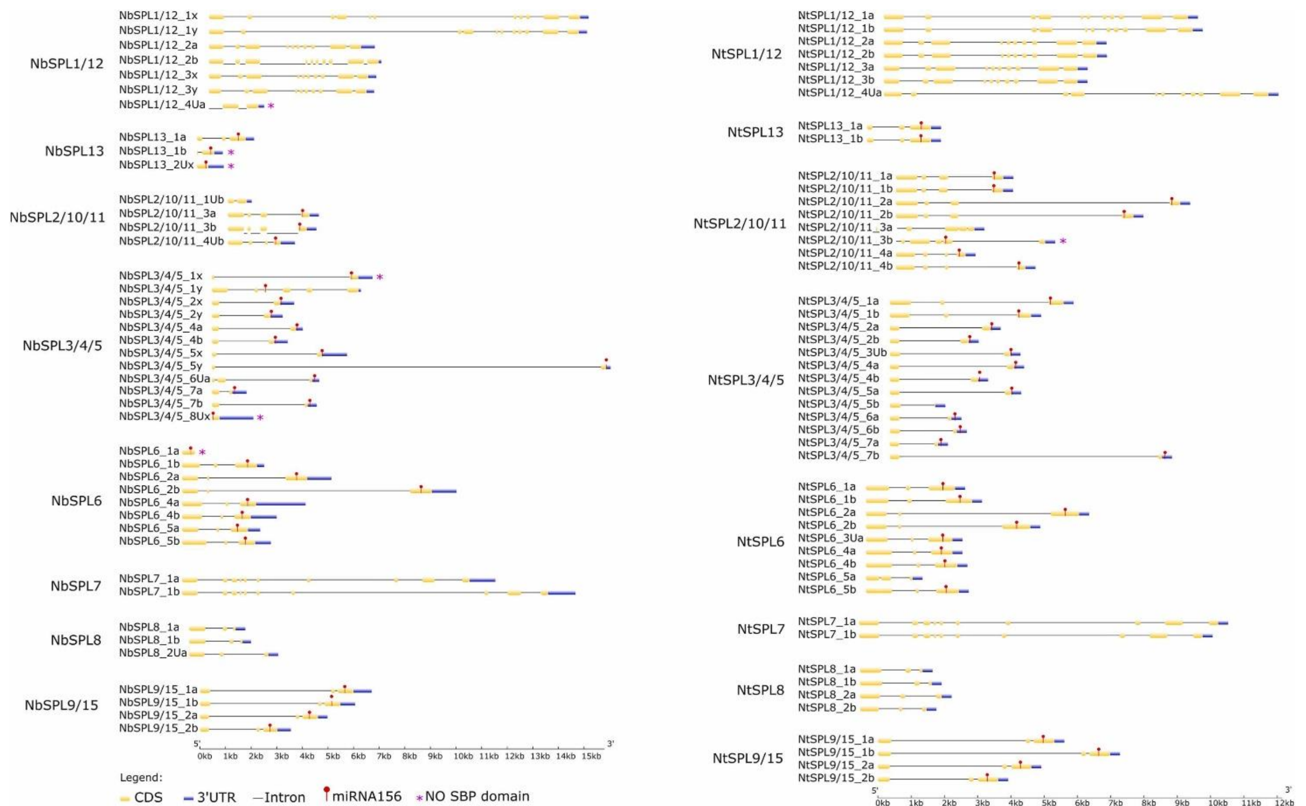


Figure 9. Gene structure of *NbSPLs* (left) and *NtSPLs* (right).

Exons, introns, CDS, 3' UTRs and miRNA156 annealing sites are shown. Genes with no SBP domain are indicated with a purple asterisk.

Expression analysis of *SPL* genes in leaves

To determine the temporal expression patterns of *NbSPL* and *NtSPL* genes, an RNA-seq analysis was performed on *N. tabacum* cv. K326 and *N. benthamiana* leaf samples collected at three different growth stages. All RNA samples were isolated from the same leaf in different plants (leaf number five counting from the first true leaf that appears during development), but at different developmental stages, namely juvenile (J), pre-flowering (P), and flowering (F), the latter with first flower primordia already present (Fig. 10A and B). Comparisons of the expression of the same gene in the different developmental stages were done using the CPM-normalized log2-transformed expression data as depicted in fig. 10C and D, while comparisons of the expression levels between genes were done with the FPKM-normalized log2-transformed expression data as depicted in Fig. S1.

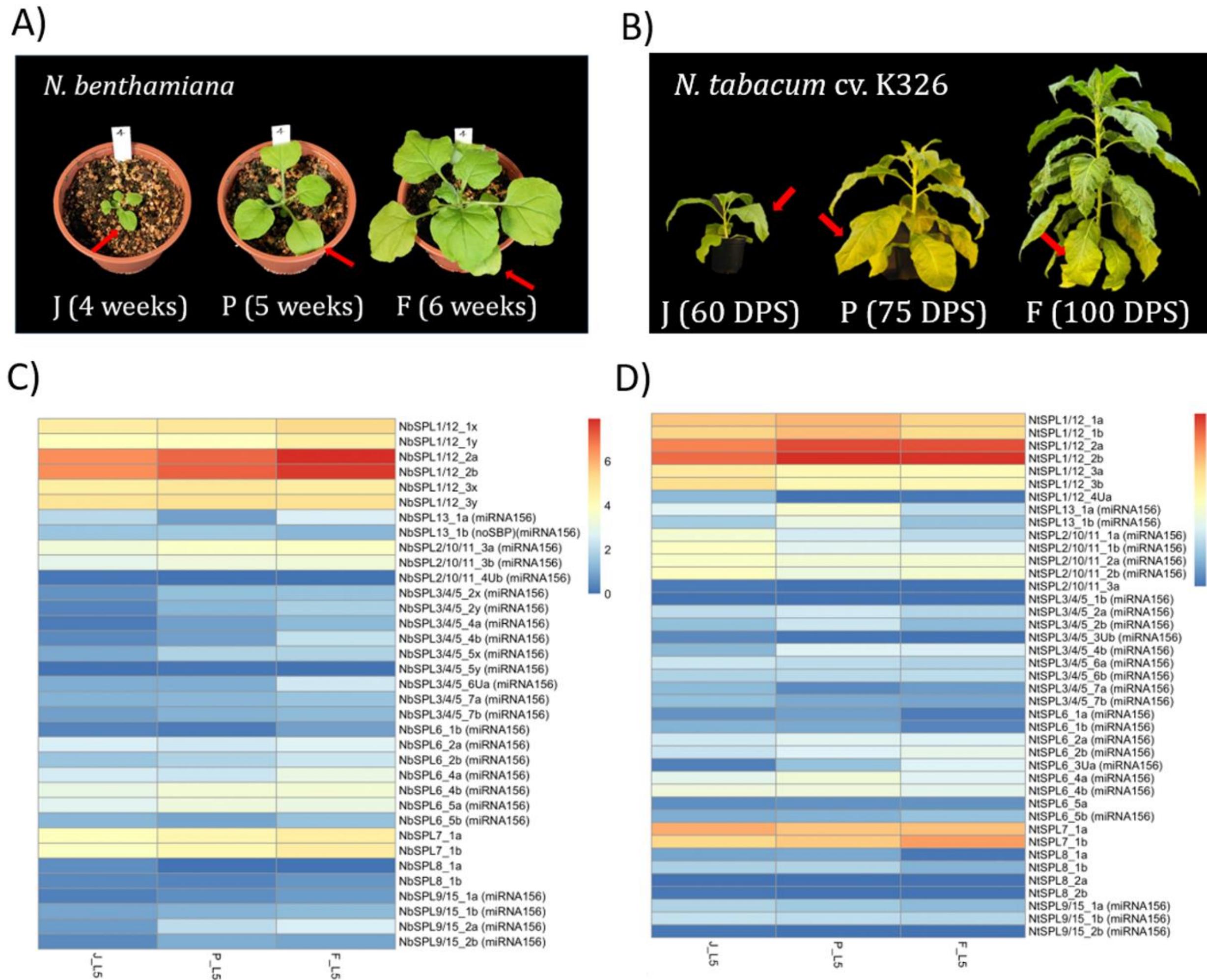


Figure 10. Expression profiles of *NbSPL* and *NtSPL* genes in leaf at different developmental stages.

Pictures of *N. benthamiana* (A) and *N. tabacum* cv. K326 (B) plants at juvenile (J), pre-flowering (P) and flowering (F) stages. Red arrows indicate the fifth true leaf. **C)** CPM-normalized log2-transformed expression values for *NbSPL* genes, based on transcriptome data. Each rectangle represents the mean of three replicates **D)** CPM-normalized log2-transformed expression values for *NtSPL* genes, based on transcriptome data. Each rectangle represents the mean of 3 replicates.

In general, several similarities in the *SPL* expression patterns could be found between the two species (Fig. 10C and D). For instance, *SPL1/12* and *SPL7* were expressed at high levels for both *N. benthamiana* and *N. tabacum* cv. K326 (Fig. 10C and D, Fig. S1 and Table S2). In particular, the *SPL1/12_2* pair showed the highest expression in both species and, surprisingly, their abundance increased with time despite not being targets of miRNA156. Transcripts of *SPL1/12_1* and *SPL1/12_3* pairs were relatively less abundant but still detectable at high levels. As an exception, the *N. tabacum* cv. K326 gene *NtSPL1/12_4Ua* was expressed at very low levels in all stages. The second group by relative transcript abundance was *SPL7*. Both in *N. tabacum* cv. K326 and *N. benthamiana*, the two *SPL7* homeologues were highly expressed in all stages, although in *N.*

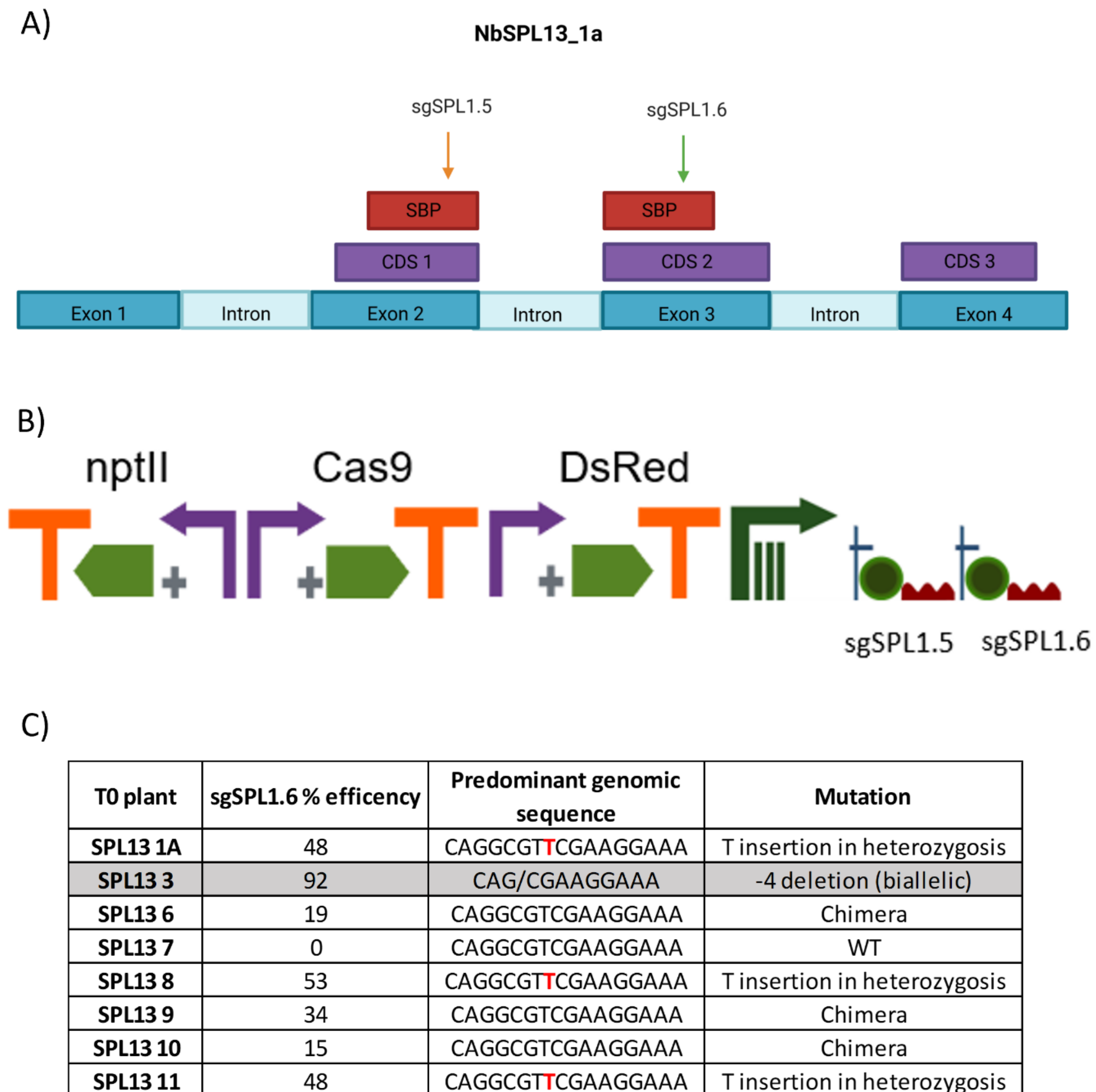
benthamiana their expression was relatively lower. The high expression levels in all stages of these two clades could be partially explained by the fact that they were not repressed by miRNA156. However, this is not a general rule for all non-miRNA156 targets; *SPL8* genes were not regulated by miRNA156, but were poorly expressed in all stages for both species.

SPLs in groups under miRNA156 control were expressed at much lower levels, and this complicated the interpretation of the developmental fluctuations. In general, miRNA156 putative targets showed strong developmental variations in expression levels, although they did not necessarily follow the same trend in *N. tabacum* cv. K326 and *N. benthamiana*, something that could reflect different timings in developmental transitions between the two species. It was expected that miRNA156-regulated *SPLs* would increase their expression along the three developmental timepoints. This seemed to be the case for most genes in *N. benthamiana* (Fig. 10C), where this general trend could be observed in most members of *NbSPL3/4/5* and *NbSPL9/15* clades, and to a lesser extend also in the *NbSPL9/15* and *NbSPL6* clades, although not for *NbSPL2/10/11*. Interestingly, the strongest variation in the only canonical representative of the *NbSPL13* clade, the *NbSPL13_1a* gene, was observed for the transition between pre-flowering and flowering stages. In *N. tabacum* cv. K326, age-associated upregulation trends in miRNA156-regulated genes were less obvious, as different behaviours were observed in different members within the same clade (Fig. 10D). For instance, *NtSPL2/10/11_1a* and *NtSPL2/10/11_1b* were clearly downregulated as the leaves aged, whereas the other two representatives of the same clade showed constant expression levels. Similarly, *NtSPL6_3Ua* levels increased dramatically with age, whereas clade partners *NtSPL6_1a* and *NtSPL6_1b* followed the opposite trend. Interestingly, and contrary to what was observed in *N. benthamiana*, expression of the *NtSPL13* clade peaked at pre-flowering stage and was reduced drastically at flowering.

Gene editing of NbSPL13 gene subfamily resulted in delayed flowering

SPL genes are involved in various physiological processes whose manipulation could be advantageous for plant biofactories, like the extension of the pre-flowering phase. To obtain a first indication of the phenotypic effects that mutagenesis in the SPL family could have, the *NbSPL13_1a* gene in *N. benthamiana* was selected as target for a knockout. *NbSPL13_1a* expression profile showed strong upregulation in the transition from pre-flowering to flowering stages, suggesting a role of this gene in flowering regulation. Furthermore, *NbSPL13* resulted as the smallest functional family in *N. benthamiana,* thus facilitating knock out strategy. As described earlier, among the three putative *NbSPL13* members, only *NbSPL13_1a* was likely to contribute to phase transition as it was the only one containing all canonical elements, namely an SPB domain and a miRNA156 putative target site. *NbSPL13_1b* had no SPB domain, whereas *NbSPL13_2Ux* was considered a pseudogene as there were no RNA reads associated to this gene. Two guide RNAs (gRNAs) were designed for

targeting *NbSPL13_1a*: the sgSPL1.5 gRNA targeting exon one and the sgSPL1.6 gRNA for exon two (Table S1, Fig. 11A). Although both gRNAs targeted also *NbSPL13_1b* at exons one and two, they were unlikely to produce an effect since according to the latest annotation of this gene they fall outside the predicted CDS.



T0 plant	sgSPL1.6 % efficency	Predominant genomic sequence	Mutation
SPL13 1A	48	CAGGCGTTCGAAGGAAA	T insertion in heterozygosis
SPL13 3	92	CAG/CGAAGGAAA	-4 deletion (biallelic)
SPL13 6	19	CAGGCGTCGAAGGAAA	Chimera
SPL13 7	0	CAGGCGTCGAAGGAAA	WT
SPL13 8	53	CAGGCGTTCGAAGGAAA	T insertion in heterozygosis
SPL13 9	34	CAGGCGTCGAAGGAAA	Chimera
SPL13 10	15	CAGGCGTCGAAGGAAA	Chimera
SPL13 11	48	CAGGCGTTCGAAGGAAA	T insertion in heterozygosis

Figure 11. Generation of *N. benthamiana* lines edited in *NbSPL13_1a* gene.

A) Structure of *NbSPL13_1a*, exons are indicated in blue, introns in light blue, CDSs in violet, and SBP domain in red. Arrows represent sgRNAs sgSPL1.5 and sgSPL1.6. B) Vector GB3298 used for transformation to edit *NbSPL13* genes. Violet arrows: promoters of nptII, Cas9 and DsRed. Orange Ts: terminators. The green arrow represents U6–26 promoter, followed by TU constituted by guides sgSPL1.5, sgSPL1.6. C) Table representing the T0 obtained from the transformation. For each plant are reported the editing efficiency values of sgSPL1.6 guide reported by Synthego, its predominant genomic sequence at the cut site, and the corresponding mutation. Letters highlighted in red represent base insertions. Plant 3 is highlighted, since it was the one selected for following with T1 generation.

A CRISPR/Cas9 construct was assembled that includes the nptII, Cas9 and DsRed TUs together with the gRNA TU (GB3298, Fig. 11B). The two gRNAs were expressed from a single U6–26 promoter using the tRNA strategy (K. Xie *et al*., 2015). This construct was used to transform a WT *N. benthamiana* plant, and eight primary transformants were obtained. All of them carried mutations at the sgSPL1.5-targeted site. Fig. 11C shows the percentage of mutations observed in the T0 generation for the sgSPL1.6-targeted site. One out of the eight regenerated plants was not edited, three of them were considered chimeras (percent editing reported by Synthego (https://ice.synthego.com/) below 40%), three of them carried heterozygous mutations (percent editing reported by Synthego around 50%) and only one plant showed biallelic mutations in *NbSPL13_1a* (>90%). The plant with biallelic mutations in *NbSPL13_1a* (nr. 3) carried a four-nucleotides deletion at the targeted site, but also contained a heterozygous mutation in *NbSPL13_1b*. The T1 offspring of this plant was grown in a growth chamber together with WT plants and all of them were phenotyped by scoring the flowering time and the number of lateral branches at two different time points. To account for any unexpected effect of *NbSPL13_1b* mutations, each T1 plant in the analysis was genotyped for *NbSPL13_1b*. In total, four groups of plants were analysed: WT, *NbSPL13_1a/1a* (biallelic mutation in *NbSPL13_1a* and not edited in *NbSPL13_1b*), *NbSPL13_1a/1a/1b* (biallelic mutation in *NbSPL13_1a* and heterozygous mutation for *NbSPL13_1b*), *NbSPL13_1a/1a/1b/1b* (biallelic mutation for both homeologous genes).

The analysis of *NbSPL13* mutated lines clearly indicated a role of *NbSPL13_1a* in the control of flowering time. Plants of all mutant lines flowered almost simultaneously, showing an average delay of 4.5 days compared to WT (Fig. 12A). As expected, the status of *NbSPL13_1b* gene, whereas mutated in biallelic or heterozygous form, had no effect in the flowering time. As mentioned before, the gRNAs targeting this gene were unlikely to produce loss-of-function mutants since they targeted the 5'UTR sequence but not the CDS. The number of lateral branches at WT flowering time (38.25 ± 0.87 dps) was also recorded, at that time no significant differences were observed between mutant and WT plants (Fig. 12B). However, when branching was recorded for each plant at their respective flowering times, significant differences were observed between WT and mutant plants, with six branches on average for WT and ten on average for *SPL13_1a* knock-out lines (Fig. 12C). To discard that *NbSPL13* mutations could affect negatively recombinant protein production capacity, syringe-Agroinfiltration assays were performed using two different expression systems, namely a geminivirus and a TMV-based vector (see Materials and Methods). No significant differences in eGFP fluorescence were observed between mutant and WT lines (see Fig. S3).

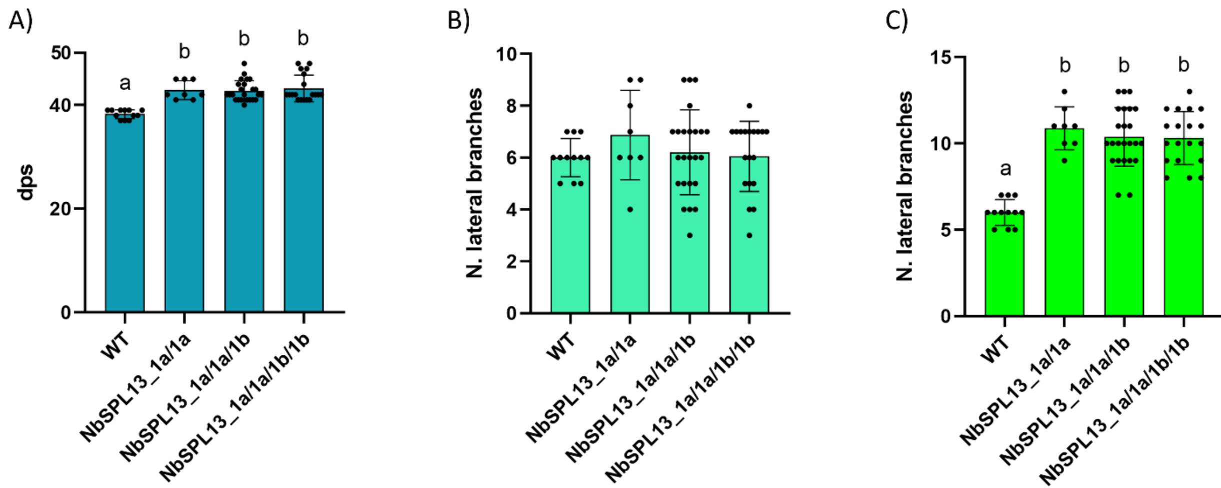


Figure 12. Phenotype of *NbSPL13* edited lines.

Phenotyping parameters of the plants grown in growth chamber. To account for any unexpected effect of *NbSPL13_1b* mutations, each T1 plant in the analysis was genotyped for *NbSPL13_1b*. In total four groups of plants were analysed: WT, *NbSPL13_1a/1a* (biallelic mutation in *NbSPL13_1a* and not edited in *NbSPL13_1b*), *NbSPL13_1a/1a/1b* (biallelic mutation in *NbSPL13_1a* and heterozygous mutation for *NbSPL13_1b*), *NbSPL13_1a/1a/1b/1b* (biallelic mutation for both homeologous genes). A) Flowering time expressed as days post sowing (dps). B) Number of lateral branches at flowering time of the WT plants. C) Number of lateral branches at flowering time of each plant. Each bar represents mean value ± SD (n = 12 for WT, n = 8 for *NbSPL13_1a/1a*, n = 22 for *NbSPL13_1a/1a/1b* and n = 19 for *NbSPL13_1a/1a/1b/1b*). A one-way ANOVA test was performed ($p < 0.05$). Groups marked with the same letter have no significant differences among them, if marked with different letters they differ significantly.

Discussion

SPLs are plant-specific transcription factors that play an important role in many aspects of plant development, including branching, leaf initiation rate or flowering time (Ferreira e Silva *et al*., 2014; Preston & Hileman, 2013; H. Wang & Wang, 2015; Xu *et al*., 2016; Yu *et al*., 2015). Many of them are subjected to miRNA156 control, whose expression level decreases throughout plant growth, allowing to express *SPL* genes in later stages of plant development (H. Wang & Wang, 2015; Xu *et al*., 2016; T. Zhang *et al*., 2015). Therefore, the miR156/SPL module has been proposed as a toolset for crop improvement as *SPL* genes are relevant targets for improving agronomic traits such as ideal plant architecture, better yield and optimal flowering time (H. Wang & Wang, 2015; T. Zhang *et al*., 2015). *Nicotiana* plants are routinely used at lab scale for added-value compounds manufacturing. Despite their several advantages, including their non-food status, fast growth, or amenability for genetic transformation and Agroinfiltration, they are not yet optimal biofactories (Alkanaimsh *et al*., 2019). Breeding objectives for *Nicotiana* towards this aim include biomass increase, delay in flowering time or more lateral branching, associated to higher yield upon Agroinfiltration (Goulet *et al*., 2019). Thus, manipulating *SPL* genes may represent a reasonable approach for breeding *Nicotiana* as improved biofactories.

In this work, a genome-wide analysis of SPL genes was performed in the two more relevant *Nicotiana* species for biofactory use, namely *N. benthamiana* and *N. tabacum* cv. K326 (Bally *et al*., 2018; Herpen *et al*., 2010; Ma *et al*., 2015; Molina-Hidalgo *et al*., 2021). Phylogenetic tree analysis showed that *NbSPLs* and *NtSPLs* clustered into eight groups, observing a similar number of *N. benthamiana* and *N. tabacum* cv. K326 genes within each group: seven *SPL1/12* genes, three *SPL13* genes for *N. benthamiana* and two for *N. tabacum* cv. K326, four *SPL2/10/11* genes for *N. benthamiana* and eight for *N. tabacum* cv. K326, 12 *SPL3/4/5* genes for *N. benthamiana* and 13 for *N. tabacum* cv. K326, eight *SPL6* genes for *N. benthamiana* and nine for *N. tabacum* cv. K326, two *SPL7* genes for both, three *SPL8* genes for *N. benthamiana* and four for *N. tabacum* cv. K326, and four *SPL9/15* genes for both. Only for group *SPL2/10/11* there was a difference greater than one between the number of proteins for both species, with eight *NtSPLs* while only four *NbSPLs*. The lower total number of SPL genes in *N. benthamiana* as compared to *N. tabacum* cv. K326 was to be expected as the result of the diploidization process, provided that the earlier is an ancient allotetraploid, whereas tetraploidization was a relatively recent phenomenon in the latter (Edwards *et al*., 2017; Schiavinato *et al*., 2020; Sierro *et al*., 2014).

The new version of the *N. benthamiana* genome (https://www.nbenth.com/) (Ranawaka *et al*., 2023) and the Edwards v4.5 version of the *N. tabacum* cv. K326 genome (Edwards *et al*., 2017) facilitated the genome-wide identification of the *NbSPLs* and *NtSPLs*. Gene models were confirmed with our own RNAseq data for most *SPLs*. However, the transcript levels of some genes were below detection levels, and further research is needed to validate the structural annotation of those genes. Several conserved motifs were identified in the SPL family members for both species. While the purpose of most motifs

remained unknown, the three motifs that consistently appeared in all potentially functional SPLs were the ones comprising the SBP domain. This domain is crucial for their binding to DNA and consequently for their role as a transcription factors (Xu *et al*., 2016). In the SBP domain of NbSPLs and NtSPLs two zinc finger motifs and one nuclear localisation signal were identified, as it occurs in SPL proteins of other species (Cai *et al*., 2018; Cardon *et al*., 1999; C. Li & Lu, 2014; M. Liu *et al*., 2019; Mao *et al*., 2016; Pan *et al*., 2017; Salinas *et al*., 2012; Song *et al*., 2016; Tripathi *et al*., 2018; Z. Yang *et al*., 2008; Zeng *et al*., 2019; Zhou *et al*., 2018). Six of the *N. benthamiana* and one of the *N. tabacum* cv. K326 proteins initially identified as SPLs in our search, did not contain a SBP domain and therefore it is unlikely that they can function as a transcription factors. On the other hand, it is very likely that those genes classified in the same clade play similar biological functions, given the high sequence homology among them, their similar gene structure, and the presence of the same conserved motifs.

As an additional step towards the functional characterization of the *SPL* family in *Nicotiana* genus, a transcriptomic analysis of equivalent leaves at three developmental time points was performed: juvenile, pre-flowering and early post-flowering. Some of the analyzed genes showed undetectable expression levels. For example, transcripts were not detected in those genes lacking the SBP domain with exception of *NbSPL13_1b*. The remaining genes were expressed in all developmental stages at different levels. Some genes not under miRNA156 control, such as those in groups *SPL1/12* and *SPL7*, were broadly expressed at high levels in both species. However, this is not true for *SPL8* - it was not regulated by miRNA 156, but had lower expression. A similar behaviour can be observed also in other species. *SPL1*, *SPL12* and *SPL7* exhibit a high expression in leaves of *A. thaliana* (Moreno *et al*., 2022) and cotton (Cai *et al*., 2018). In leaves of chrysanthemum *SPL1* and *SPL7* are highly expressed, while *SPL12* is expressed at low levels (Song *et al*., 2016). Regarding *SPL8*, it shows a very low level of expression in *A. thaliana* (Moreno *et al*., 2022) and cotton (Cai *et al*., 2018), while in chrysanthemum, on the contrary, is highly expressed (Song *et al*., 2016).

In *N. benthamiana* and *N. tabacum* cv. K326, unexpectedly, not all groups of miRNA156-regulated *SPLs* showed a progressive increase in expression with plant age. In general, they were all expressed at much lower levels compared to *SPL1/12* and *SPL7*. Some groups globally exhibited a progressively increasing expression as expected, whereas others showed different behaviours. This was true also for *SPL13* subfamily in *N. benthamiana*: *NbSPL13_1a* decreased from juvenile to pre-flowering stage, but then increased at flowering. Its partner *NbSPL13_1b* with no SBP domain stayed at similar expression level from juvenile to pre-flowering stage and decreased at flowering. The low level of expression for most of *SPL* genes could be explained by the fact that all SPL proteins belonging to the same group have redundant functions, so that the expression of a single *SPL* gene is not required to be high. The unexpected behaviour of various *SPL* genes regulated by miR156 could be due to a complex transcriptional behaviour occurring in this large family, which may include tissue-specific regulation or complex time fluctuations that could have escaped from the general experimental setup followed here.

A more detailed expression profile analysis for each gene family should follow to provide a more complete picture of the entire *SPL* regulatory complex in these two species.

Following the structural characterization and the gene expression analysis of the *NbSPL* genes, we wanted to explore the potential of SPLs as targets for breeding new biofactory-oriented traits. Particularly, we decided to study whether the knock-out of the smallest functional SPL group in *N. benthamiana*, *NbSPL13*, had an influence on the plant architecture and flowering time.

Previous studies have conducted loss-of-function experiments targeting *SPL* genes in different species, revealing various phenotypic effects. In rice, CRISPR/Cas-mediated knockout of individual *SPL* genes resulted in defects in plant height, reduced panicle size and altered grain length (Jiang *et al*., 2020). In barley, loss of *HvSPL8* function confers smaller leaves angle: this feature allows lower leaves to be shed by more light, increasing photosynthesis rate and therefore productivity (S. Yang *et al*., 2022). In *N. tabacum* cv. K326 the overexpression of miRNA156, with the subsequent down-regulation of the miR156-targeted *SPL* genes, resulted in delayed flowering and a higher leaves production rate (Feng *et al*., 2016) and the generation of biallelic mutations in nine *SPL* genes with CRISPR/Cas caused delayed flowering, leaf juvenility and more branching (Vazquez-Vilar *et al*., 2021). On the other hand, gain of function mutations in *OsSPL14* in rice confers an ideal plant architecture with thick culm, large panicle and reduced tillers (Jiao *et al*., 2010; J. Wang *et al*., 2017).

Regarding *SPL13*, previous studies in alfalfa showed that this gene has an important role in drought stress tolerance (Arshad *et al*., 2017; Feyissa *et al*., 2019) and vegetative-floral transition. Gao at al. demonstrated in 2018 that *SPL13* silencing caused a delay in flowering time and increased number of lateral branches in this species (R. Gao *et al*., 2018). Using CRISPR/Cas9, we obtained a plant carrying a biallelic deletion of four nucleotides in the first exon of *NbSPL13_1a* and a heterozygous mutation in the 5'UTR of *NbSPL13_1b*. The progeny of this plant was phenotyped for flowering time and all T1 plants exhibited a significant delay in flowering of 4–5 days with respect to WT.

It is a common practice both in academic and industrial experimental setups, to perform Agroinfiltration before the plants reach the flowering stage, as it is well known that expression levels drop dramatically afterwards (Sheludko *et al*., 2007). In this regard, plant varieties with delayed flowering time could give time to accumulate more productive biomass in the same growing area. Interestingly, in our study late flowering was accompanied by an increase in the total number of lateral branches. Interestingly, Goulet *et al*. (2019) previously reported that young leaves in axillary stems are the main contributors to recombinant protein yield upon Agroinfiltration (Goulet *et al*., 2019). Therefore, we anticipate that the combination late flowering and increased branching in *NbSPL13_1a* KO phenotype should lead to gains in recombinant protein yield per batch. However, as with field trials in traditional breeding of food crops, the full advantage conferred by this new trait to the *N. benthamiana* biofactory can only be confirmed in a real-life scenario, which in this case implies a pilot vacuum infiltration experiment in an industrial setup. Syringe-infiltration experiments showed no differences in expression in *NbSPL13* KO

compared to WT. However, these experiments do not appraise the potential advantages of the mutants, that would need to be assessed at a larger scale. Furthermore, pilot experiments should be followed by a techno economical evaluation that integrate not only calculations of yield per batch, but also the number of batches per time unit, since yield gains per batch need to compensate the longer growing times associated with late flowering.

In our view this work shows how genomic insights in large gene families can inform and orient breeding strategies, especially in polypoid crops. It also illustrates how the breeding of plant biofactories can benefit from new precision techniques and vice versa, since this type of industrial crops, usually grown under contained conditions, offer minimal ethical, legal and/or environmental restrictions for commercial implementation even in the most restrictive economic zones as the EU.

Conclusions

Our work aimed at the identification of Squamosa Promoter Binding-Like (SPL) genes in *N. tabacum* cv. K326 and *N. benthamiana.* Fourty-nine *SPL* genes were found in *N. tabacum* cv. K326 and 43 in *N. benthamiana* LAB strain, and classified into eight phylogenetic groups according to the SPL classification in *Arabidopsis*. Homeologues and orthologues showed a conserved exon-intron gene structure and a conserved DNA-binding domain. Thirty of the *NbSPL* genes and 33 of the *NtSPL* genes were found to be putative targets of microRNA 156. Their expression in leaves was analysed by RNA-seq at three different stages, revealing that genes not under miR156 control were in general constitutively expressed at high levels, while miR156-regulated genes showed lower expression levels, often developmentally regulated. *NbSPL13_1a* was chosen for editing and its knock-out led to a delay in flowering time, a trait that could be exploited to increase biomass for recombinant protein production. In our view this work showed how genomic insights in large gene families can inform and orient breeding strategies, especially in polypoid crops.

Supplementary material

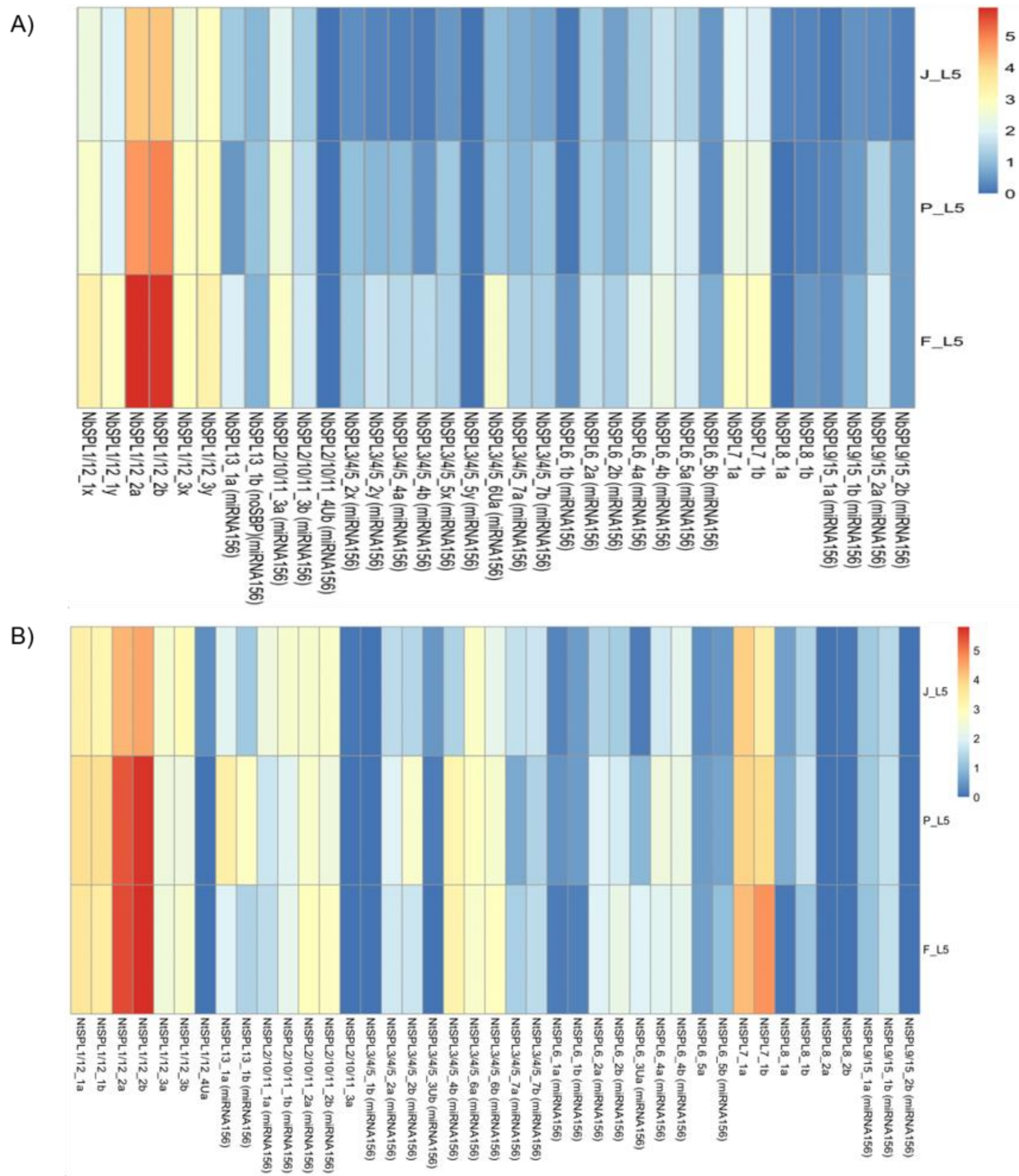


Figure S1. A) FPKM-normalized log2-transformed expression values for NbSPL genes, based on transcriptome data. Each rectangle represents the mean of three replicates B) FPKM-normalized log2-transformed expression values for NtSPL genes, based on transcriptome data. Each rectangle represents the mean of three replicates.

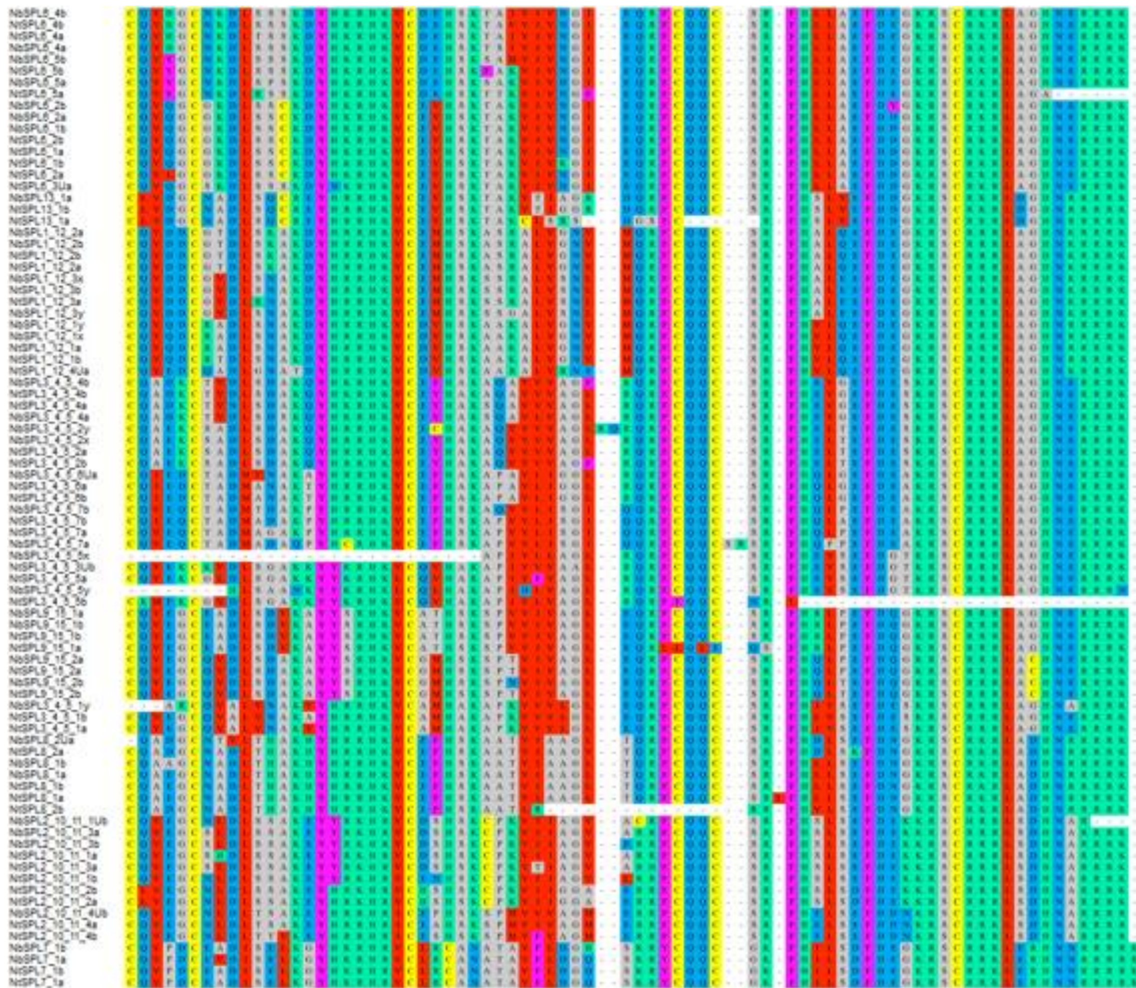


Figure S2. SBP domains of NtSPLs and NbSPLs

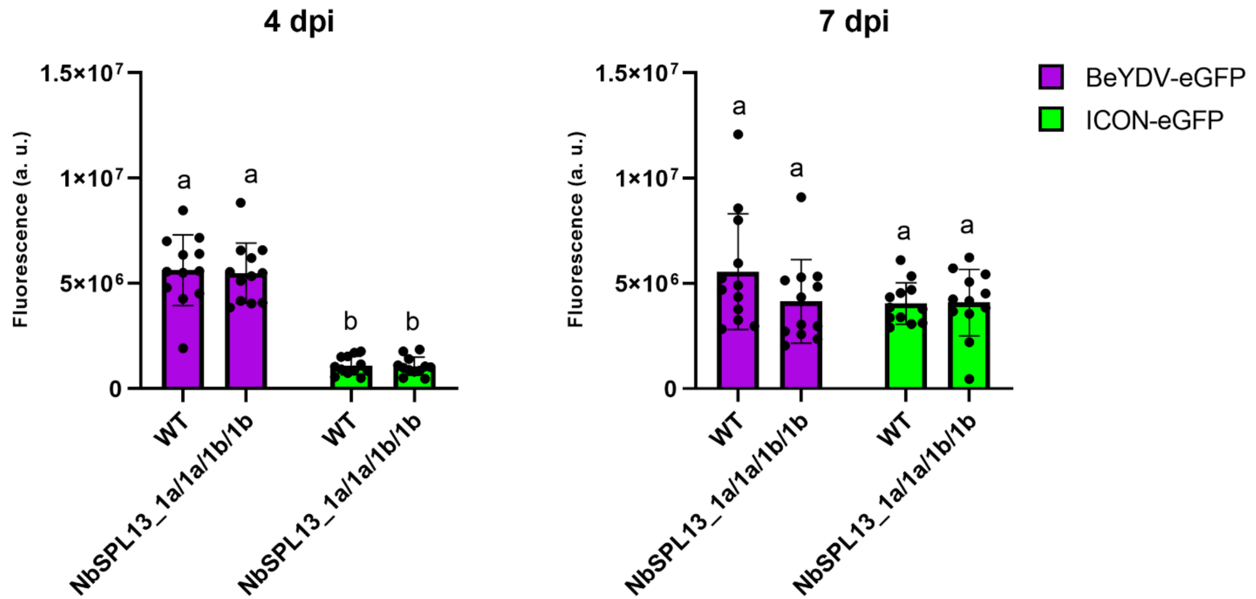


Figure S3. GFP expression in WT and NbSPL13 plants. WT and NbSPL13_1a/1a/1b/1b (biallelic mutation for both homeologous genes) *N. benthamiana* plants were Agroinfiltrated with two viral systems harbouring the CDS of the enhanced green fluorescent protein (eGFP), BeYDV-eGFP and ICON-eGFP. Fluorescence was detected after four (left) and seven days (right). A one-way ANOVA test was performed (n=12, $p<0.05$). Bars marked with the same letter have no significant differences among them, if marked with different letters they differ significantly.

gRNA name	Targeted genes and position	Protospacer sequence	PAM
sgSPL1.5	NbSPL13_1a-nc134/NbSPL13_1b-nc136 (theoretical)	GGACCTCACAAACTTTATGG	CGG
sgSPL1.6	NbSPL13_1a-c288/NbSPL13_1b-c290 (theoretical)	ATGGACATAACAGGCGTCGA	AGG

Table S1. Guide RNA sequences used for *NbSPL13* genes edition. gRNA targeted positions were determined as distance of the Cas9 cutting site to the ATG for coding sequences. “c” indicates that the gRNA is designed on the coding strand and “nc” that gRNA is designed on the non-coding strand. For *NbSPL13 _1b*, guides were designed onto its theoretical CDS, constructed onto the homeologous one.

NbSPL	LAB330 name	Chromosome	N. exons	SBP domain	Protein length (aa)	Top match species (excluding tobacco)	Stats	miR156 site
NbSPL1/12_1x	gene.7156.1.1.p1	NbLab330C15	11	Yes	1000	*Nicotiana attenuata*	Identities:965/1001(96%), Positives:978/1001(97%), Gaps:1/1001(0%)	no
NbSPL1/12_1y	gene.37344.0.2.p1	NbLab330C10	11	Yes	1000	*Nicotiana attenuata*	Identities:966/1001(97%), Positives:979/1001(97%), Gaps:1/1001(0%)	no
NbSPL1/12_2a	gene.23674.3.0.p1	NbLab330C08	10	Yes	999	*Nicotiana sylvestris*	Identities:970/1001(97%), Positives:983/1001(98%), Gaps:2/1001(0%)	no
NbSPL1/12_2b	gene.63653.1.1.p1	NbLab330C18	10	Yes	998	*Nicotiana attenuata*	Identities:968/999(97%), Positives:984/999(98%), Gaps:1/999(0%)	no
NbSPL1/12_3x	gene.59226.0.0.p1	NbLab330C05	10	Yes	967	*Nicotiana sylvestris*	Identities:917/968(95%), Positives:933/968(96%), Gaps:2/968(0%)	no
NbSPL1/12_3y	gene.9986.0.0.p1	NbLab330C12	10	Yes	968	*Nicotiana sylvestris*	Identities:919/968(95%), Positives:936/968(96%), Gaps:1/968(0%)	no
NbSPL1/12_4Ua	chr14.g9026.t1	NbLab330C04	3	No	362	*Nicotiana attenuata*	Identities:345/361(96%), Positives:352/361(97%), Gaps:0/361(0%)	No

NbSPL13_1a	gene.4997.0.0.p1	NbLab330C15	3	Yes	328	*Nicotiana sylvestris*	Identities:316/329(96%), Positives:320/329(97%), Gaps:1/329(0%)	yes
NbSPL13_1b	gene.74426.0.1.p1	NbLab330C14	2	No	172	*Nicotiana sylvestris*	Identities:154/168(92%), Positives:162/168(96%), Gaps:1/168(0%)	yes
NbSPL13_2Ux	gene.77593.0.1.p1	NbLab330C14	1	No	147	*Nicotiana sylvestris*	Identities:132/142(93%), Positives:133/142(93%), Gaps:0/142(0%)	yes
NbSPL2/10/11_1Ub	chr11.g10631.t1	NbLab330C16	2	Yes	213	*Nicotiana sylvestris*	Identities:194/228(85%), Positives:196/228(85%), Gaps:27/228(11%)	no
NbSPL2/10/11_3a	gene.4971.0.1.p1	NbLab330C15	4	Yes	461	*Nicotiana attenuata*	Identities:441/461(96%), Positives:444/461(96%), Gaps:1/461(0%)	yes
NbSPL2/10/11_3b	gene.74412.0.0.p1	NbLab330C14	4	Yes	458	*Nicotiana attenuata*	Identities:436/456(96%), Positives:441/456(96%), Gaps:1/456(0%)	yes
NbSPL2/10/11_4Ub	gene.36101.0.0.p1	NbLab330C10	4	Yes	361	*Nicotiana sylvestris*	Identities:332/362(92%), Positives:342/362(94%), Gaps:3/362(0%)	yes
NbSPL3/4/5_1x	gene.31143.0.0.p1	NbLab330C13	2	No	141	*Nicotiana sylvestris*	Identities:117/134(87%), Positives:121/134(90%), Gaps:4/134(2%)	yes

NbSPL3/4/5_1y	chr08.g12650.t1	NbLab330C10	6	Yes	395	*Nicotiana sylvestris*	Identities:360/403(89%), Positives:370/403(91%), Gaps:8/403(1%)	yes
NbSPL3/4/5_2x	gene.9663.0.0.p1	NbLab330C15	2	Yes	182	*Nicotiana tomentosiformis*	Identities:175/183(96%), Positives:178/183(97%), Gaps:3/183(1%)	yes
NbSPL3/4/5_2y	gene.25398.0.0.p1	NbLab330C03	2	Yes	187	*Nicotiana attenuata*	Identities:178/187(95%), Positives:181/187(96%), Gaps:3/187(1%)	yes
NbSPL3/4/5_4a	gene.63701.0.2.p1	NbLab330C18	2	Yes	166	*Nicotiana attenuata*	Identities:153/167(92%), Positives:157/167(94%), Gaps:1/167(0%)	yes
NbSPL3/4/5_4b	gene.16698.0.0.p1	NbLab330C01	2	Yes	160	*Nicotiana sylvestris*	Identities:143/165(87%), Positives:150/165(90%), Gaps:9/165(5%)	yes
NbSPL3/4/5_5x	gene.60667.0.1.p1	NbLab330C05	2	Yes	125	*Nicotiana attenuata*	Identities:75/83(90%), Positives:79/83(95%), Gaps:0/83(0%)	yes
NbSPL3/4/5_5y	gene.8997.0.0.p1	NbLab330C15	2	Yes	107	*Nicotiana sylvestris*	Identities:93/106(88%), Positives:96/106(90%), Gaps:0/106(0%)	yes
NbSPL3/4/5_6Ua	gene.38814.0.0.p1	NbLab330C10	2	Yes	137	*Nicotiana sylvestris*	Identities:132/136(97%), Positives:133/136(97%), Gaps:0/136(0%)	yes

NbSPL3/4/5_7a	gene.93876.0.1.p1	NbLab330C19	2	Yes	138	*Nicotiana tomentosiformis*	Identities:128/138(93%), Positives:130/138(94%), Gaps:0/138(0%)	yes
NbSPL3/4/5_7b	gene.83021.0.3.p1	NbLab330C02	2	Yes	136	*Nicotiana tomentosiformis*	Identities:126/136(93%), Positives:128/136(94%), Gaps:0/136(0%)	yes
NbSPL3/4/5_8Ux	chr06.g9674.t1	NbLab330C03	1	No	100	*Nicotiana attenuata*	Identities:89/100(89%), Positives:90/100(90%), Gaps:0/100(0%)	yes
NbSPL6_1a	gene.68624.0.0.p1	NbLab330C04	1	No	169	*Nicotiana sylvestris*	Identities:161/169(95%), Positives:164/169(97%), Gaps:0/169(0%)	yes
NbSPL6_1b	gene.90250.0.2.p1	NbLab330C19	3	Yes	564	*Nicotiana attenuata*	Identities:495/564(88%), Positives:510/564(90%), Gaps:24/564(4%)	yes
NbSPL6_2a	gene.53347.0.0.p1	NbLab330C16	3	Yes	524	*Nicotiana attenuata*	Identities:499/524(95%), Positives:511/524(97%), Gaps:1/524(0%)	yes
NbSPL6_2b	gene.50476.0.1.p1	NbLab330C06	3	Yes	523	*Nicotiana attenuata*	Identities:499/523(95%), Positives:508/523(97%), Gaps:0/523(0%)	yes
NbSPL6_4a	gene.17776.0.0.p1	NbLab330C01	2	Yes	508	*Nicotiana sylvestris*	Identities:479/508(94%), Positives:489/508(96%), Gaps:0/508(0%)	yes

NbSPL6_4b	gene.96230.0.0.p1	NbLab330C15	2	Yes	510	*Nicotiana sylvestris*	Identities:463/510(91%), Positives:481/510(94%), Gaps:2/510(0%)	yes
NbSPL6_5a	gene.2346.0.1.p1	NbLab330C07	3	Yes	472	*Nicotiana attenuata*	Identities:424/464(91%), Positives:432/464(93%), Gaps:5/464(1%)	yes
NbSPL6_5b	gene.43637.0.2.p1	NbLab330C17	3	Yes	572	*Nicotiana tomentosiformis*	Identities:423/555(76%), Positives:449/555(80%), Gaps:43/555(7%)	yes
NbSPL7_1a	gene.85643.0.1.p1	NbLab330C09	10	Yes	781	*Nicotiana attenuata*	Identities:751/804(93%), Positives:763/804(94%), Gaps:23/804(2%)	no
NbSPL7_1b	gene.70154.0.8.p1	NbLab330C04	10	Yes	804	*Nicotiana sylvestris*	Identities:771/804(96%), Positives:779/804(96%), Gaps:2/804(0%)	no
NbSPL8_1a	gene.76268.0.2.p1	NbLab330C14	3	Yes	300	*Nicotiana attenuata*	Identities:292/300(97%), Positives:298/300(99%), Gaps:0/300(0%)	no
NbSPL8_1b	gene.39637.0.0.p1	NbLab330C10	3	Yes	300	*Nicotiana attenuata*	Identities:292/300(97%), Positives:296/300(98%), Gaps:0/300(0%)	no
NbSPL8_2Ua	gene.22921.0.3.p1	NbLab330C08	3	Yes	309	*Nicotiana sylvestris*	Identities:262/296(89%), Positives:270/296(91%), Gaps:4/296(1%)	no

NbSPL9/15_1a	gene.47557.0.1.p1	NbLab330C06	3	Yes	388	*Nicotiana attenuata*	Identities:372/388(96%), Positives:376/388(96%), Gaps:1/388(0%)	yes
NbSPL9/15_1b	gene.53813.0.1.p1	NbLab330C16	3	Yes	390	*Nicotiana attenuata*	Identities:369/390(95%), Positives:373/390(95%), Gaps:3/390(0%)	yes
NbSPL9/15_2a	gene.74848.0.3.p1	NbLab330C14	3	Yes	366	*Nicotiana sylvestris*	Identities:359/366(98%), Positives:362/366(98%), Gaps:0/366(0%)	yes
NbSPL9/15_2b	gene.7983.0.0.p1	NbLab330C15	3	Yes	350	*Nicotiana attenuata*	Identities:322/351(92%), Positives:334/351(95%), Gaps:1/351(0%)	yes

Table S2-A. *SPL* genes in *N. benthamiana* with their characteristics: original name of LAB330 annotation, chromosome, number of exons, presence or absence of SBP domain, protein length as number of amino acids, the top probable subgenome donor and its statistics, presence or absence of miRNA156 target site.

NtSPL	Edwards 2017 name	Protein length Edwards 2017 (aa)	Protein length new model (aa)	N. exons Edwards 2017	N. exons new model	Updated gene model	Subgenome donor	SBP domain	miRNA156 site
NtSPL1/12_1a	Nitab4.5_0008312g0040.1	1000	1000	12	11	Yes	*Nicotiana sylvestris*	Yes	No
NtSPL1/12_1b	Nitab4.5_0000745g0150.1	1001	1001	12	11	Yes	*Nicotiana tomentosiformis*	Yes	No
NtSPL1/12_2a	Nitab4.5_0002994g0100.1	960	999	10	10	Yes	*Nicotiana sylvestris*	Yes	No
NtSPL1/12_2b	Nitab4.5_0000222g0290.1	997	997	10	10	No	*Nicotiana tomentosiformis*	Yes	No
NtSPL1/12_3a	Nitab4.5_0003324g0070.1	908	967	10	10	Yes	*Nicotiana sylvestris*	Yes	No
NtSPL1/12_3b	Nitab4.5_0000363g0110.1	936	960	11	10	Yes	*Nicotiana tomentosiformis*	Yes	No
NtSPL1/12_4Ua	Nitab4.5_0000225g0020.1	976	992	11	11	Yes	*Nicotiana sylvestris*	Yes	No
NtSPL13_1a	Nitab4.5_0002299g0030.1	324	324	3	3	No	*Nicotiana sylvestris*	Yes	Yes
NtSPL13_1b	Nitab4.5_0001010g0010.1	326	326	3	3	No	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL2/10/11_1a	Nitab4.5_0003900g0020.1	463	463	4	4	No	*Nicotiana sylvestris*	Yes	Yes
NtSPL2/10/11_1b	Nitab4.5_0000861g0050.1	462	462	4	4	No	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL2/10/11_2a	Nitab4.5_0002558g0020.1	456	456	4	4	No	*Nicotiana sylvestris*	Yes	Yes
NtSPL2/10/11_2b	Nitab4.5_0000067g0130.1	465	465	4	4	No	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL2/10/11_3a	Nitab4.5_0001315g0110.1	357	462	5	4	Yes	*Nicotiana sylvestris*	Yes	No
NtSPL2/10/11_3b	Nitab4.5_0007217g0040.1	466	466	6	6	No	*Nicotiana tomentosiformis*	No	Yes
NtSPL2/10/11_4a	Nitab4.5_0006792g0010.1	348	360	5	4	Yes	*Nicotiana sylvestris*	Yes	Yes
NtSPL2/10/11_4b	Nitab4.5_0000019g0050.1	207	357	2	4	Yes	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL3/4/5_1a	Nitab4.5_0002061g0020.1	270	402	2	3	Yes	*Nicotiana sylvestris*	Yes	Yes
NtSPL3/4/5_1b	Nitab4.5_0000210g0190.1	240	380	2	3	Yes	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL3/4/5_2a	Nitab4.5_0001752g0040.1	182	182	2	2	No	*Nicotiana sylvestris*	Yes	Yes
NtSPL3/4/5_2b	Nitab4.5_0000638g0040.1	251	181	3	2	Yes	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL3/4/5_3Ub	Nitab4.5_0002330g0030.1	172	172	2	2	No	*Nicotiana tomentosiformis*	Yes	Yes

NtSPL3/4/5_4a	Nitab4.5_0006721g0040.1	170	167	2	2	Yes	*Nicotiana sylvestris*	Yes	Yes
NtSPL3/4/5_4b	Nitab4.5_0004959g0040.1	198	171	2	2	Yes	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL3/4/5_5a	Nitab4.5_0008703g0010.1	165	165	2	2	No	*Nicotiana sylvestris*	Yes	Yes
NtSPL3/4/5_5b	Nitab4.5_0000327g0110.1	113	119	2	2	Yes	*Nicotiana tomentosiformis*	Yes	No
NtSPL3/4/5_6a	Nitab4.5_0003348g0050.1	139	136	2	2	Yes	*Nicotiana sylvestris*	Yes	Yes
NtSPL3/4/5_6b	Nitab4.5_0003942g0050.1	148	140	2	2	Yes	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL3/4/5_7a	Nitab4.5_0007487g0020.1	133	133	2	2	No	*Nicotiana sylvestris*	Yes	Yes
NtSPL3/4/5_7b	Nitab4.5_0002219g0060.1	136	136	2	2	No	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL6_1a	Nitab4.5_0002041g0010.1	469	542	3	3	Yes	*Nicotiana sylvestris*	Yes	Yes
NtSPL6_1b	Nitab4.5_0001797g0070.1	471	538	3	3	Yes	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL6_2a	Nitab4.5_0009912g0020.1	500	524	3	3	Yes	*Nicotiana sylvestris*	Yes	Yes
NtSPL6_2b	Nitab4.5_0002467g0020.1	526	522	3	3	Yes	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL6_3Ua	Nitab4.5_0010172g0010.1	481	494	3	3	Yes	*Nicotiana sylvestris*	Yes	Yes
NtSPL6_4a	Nitab4.5_0010273g0010.1	508	508	3	3	No	*Nicotiana sylvestris*	Yes	Yes
NtSPL6_4b	Nitab4.5_0000027g0060.1	523	523	3	3	No	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL6_5a	Nitab4.5_0000509g0040.1	265	265	3	3	No	*Nicotiana sylvestris*	Yes	No
NtSPL6_5b	Nitab4.5_0001118g0090.1	561	524	4	3	Yes	*Nicotiana tomentosiformis*	Yes	Yes
NtSPL7_1a	Nitab4.5_0000700g0020.1	560	802	5	10	Yes	*Nicotiana sylvestris*	Yes	No
NtSPL7_1b	Nitab4.5_0000059g0380.1	737	806	10	10	Yes	*Nicotiana tomentosiformis*	Yes	No
NtSPL8_1a	Nitab4.5_0002061g0030.1	301	301	3	3	No	*Nicotiana sylvestris*	Yes	No
NtSPL8_1b	Nitab4.5_0000210g0160.1	311	308	3	3	Yes	*Nicotiana tomentosiformis*	Yes	No
NtSPL8_2a	Nitab4.5_0001823g0030.1	254	312	2	3	Yes	*Nicotiana sylvestris*	Yes	No
NtSPL8_2b	Nitab4.5_0000706g0150.1	267	267	3	3	No	*Nicotiana tomentosiformis*	Yes	No
NtSPL9/15_1a	Nitab4.5_0003572g0010.1	319	390	4	3	Yes	*Nicotiana sylvestris*	Yes	Yes
NtSPL9/15_1b	Nitab4.5_0000016g0300.1	401	386	3	3	Yes	*Nicotiana tomentosiformis*	Yes	Yes

NtSPL9/15_2a	Nitab4.5_0001538g0080.1	366	366	3	3	No	*Nicotiana sylvestris*	Yes	Yes
NtSPL9/15_2b	Nitab4.5_0000991g0020.1	369	369	3	3	No	*Nicotiana tomentosiformis*	Yes	Yes

Table S2-B. *SPL* genes in *N. tabacum* with their characteristics: original name of annotation in Edwards *et al*., 2017, number of exons in the original Edward 2017 model and in the updated model, protein length as number of amino acids in the Edwards 2017 model and in the new model, if they were updated in the new model, presence or absence of SBP domain, the subgenome donor, presence or absence of miRNA156 target site.

GB ID	Description
GB0307	pUPD2
GB1207	tRNA and scaffold for the assembly of GBoligomers for the last position (positon [n]) of a polycistronic tRNA-gRNA
GB1208	tRNA and scaffold for the assembly of GBoligomers for the first position (positon [D1_n-1]) of a polycistronic tRNA-gRNA regulated by the U6-26 or U6-1 promoter
GB1001	U6-26 promoter
GB2630	pUPD2 with sgSPL1.5
GB2631	pUPD2 with sgSPL1.6
GB0017	pDGB3_alpha2
GB0019	pDGB3_omega1
GB2234	Module for the constitutive expression of the nptII, Cas9 and DsRed genes in pDGB3_alpha1
GB3296	U6-26:sgSPL1.5:sgSPL1.6 in alpha2
GB3298	nptII:Cas9:DsRed_U6-26:sgSPL1.5:sgSPL1.6 in omega 1

Table S3. Goldenbraid plasmids used for this work

Primer name	Sequence 5' > 3'
CDP20Apr05_NbSPL13_1a_g1FW	CTTTGTTACTTCGCAATTAGAGCG
CDP20Apr06_NbSPL13_1a_g1RV	GTGGCGTACTAAGGGTCAAGT
CDP20Apr07_NbSPL13_1a_g2FW	AAATGTTCAATCCCTGGACGAC
CDP20Apr08_NbSPL13_1a_g2RV	ACCATGTCGCTGTCCGTTTTG
CDP20Apr09_NbSPL13_1b_g1FW	GGTTTTGATTCTTGGCGTTAGGAC
CDP20Apr10_NbSPL13_1b_g1RV	CGGGCTCTTGCAGAAATGCC
CDP20Apr11_NbSPL13_1b_g2FW	GCCCAGGTTTGAATGCATTAGGG
CDP20Apr11_NbSPL13_1b_g2RV	ACCATGTCGTTGTCCGTTTTC

Table S4. List of primers used for amplification of the targeted sites.

Sample	Species	Raw reads	% Remaining reads	%Clean sequences with average quality Q >=30	%GC in clean reads	Number of bases per sample after cleaning	Total Clean reads mapped and properly paired	% Clean reads mapped and properly paired
J_L5_1	*N. benthamiana*	148976240	74.69%	99.22%	45.50%	16.600.766.356.00	107812466	96.89%
J_L5_2	*N. benthamiana*	112663478	78.70%	99.20%	45.50%	13.221.448.672.00	86403250	97.45%
J_L5_3	*N. benthamiana*	153128010	73.33%	99.26%	45.50%	16.773.423.129.00	109148068	97.20%
P_L5_1	*N. benthamiana*	126432800	79.60%	99.23%	45.50%	15.032.662.317.00	98246582	97.62%
P_L5_2	*N. benthamiana*	109356400	77.71%	99.26%	45.00%	12.699.831.592.00	82678996	97.29%
P_L5_3	*N. benthamiana*	89883490	83.05%	99.18%	45.00%	11.135.826.064.00	73120212	97.96%
F_L5_1	*N. benthamiana*	119924082	70.42%	99.23%	45.00%	12.611.232.299.00	81955376	97.04%
F_L5_2	*N. benthamiana*	101222782	80.35%	99.10%	44.50%	12.055.527.511.00	78942352	97.06%
F_L5_3	*N. benthamiana*	109772218	73.75%	99.23%	45.00%	12.082.942.224.00	78454528	96.91%
J_L5_1	*N. tabacum*	123642084	83.18%	99.26%	44.00%	15.366.709.536.00	99406288	96.66%
J_L5_2	*N. tabacum*	119703504	81.74%	99.27%	44.50%	14.626.943.244.00	94209828	96.29%
J_L5_3	*N. tabacum*	126513526	74.43%	99.13%	44.00%	14.047.776.285.00	90454752	96.06%
P_L5_1	*N. tabacum*	111114358	77.09%	99.16%	43.50%	12.774.234.042.00	82828534	96.70%
P_L5_2	*N. tabacum*	92974442	82.12%	99.18%	43.50%	11.395.955.921.00	73994396	96.91%
P_L5_3	*N. tabacum*	113446684	85.17%	99.23%	43.50%	14.435.744.260.00	93879062	97.16%
F_L5_1	*N. tabacum*	141179282	75.16%	99.26%	43.50%	15.854.938.037.00	102338588	96.44%
F_L5_2	*N. tabacum*	108202142	81.61%	99.17%	43.00%	13.181.345.493.00	85953364	97.34%
F_L5_3	*N. tabacum*	127348004	75.48%	99.24%	43.00%	14.357.358.429.00	92977396	96.73%

Table S5. RNA-seq sequencing and mapping stats.

Chapter 2. Extension of vegetative phase in *Nicotiana benthamiana* using genome editing of *FLOWERING LOCUS T (FT)* and *SQUAMOSA PROMOTER BINDING-LIKE (SPL)* genes.

Abstract

Plant Molecular farming (PMF) is the production of industrially relevant and commercially valuable proteins in plants. Plants exhibit many advantages as bioreactors such as scalability, cost-effectiveness, and the ability to produce complex proteins. Species of the *Nicotiana* genus, especially tobacco and *Nicotiana benthamiana*, have become increasingly important as production platforms for PMF due to their advantages such as high biomass yield, ease of transformation, robust protein expression, and non-food crop status. Nevertheless, these species are not yet ideal production platforms - breeding goals such as flowering delay or abolition to enhance plant biomass could improve *N. benthamiana* as a prime chassis for molecular farming. In this chapter, our focus was the knockout of key genes for flowering, such as members of the FLOWERING LOCUS T (FT) proteins family. The flowering inducers *NbFT4* and the homeologous pair *NbFT5_1a/NbFT5_1b* together with *NbSPL13_1a*, member of the SQUAMOSA PROMOTER BINDING-LIKE (SPL) transcription factors family, were targets of CRISPR/Cas9 editing. The lines that exhibited biallelic mutations for these genes, alone and in combination, showed delayed flowering and a remarkable increase in biomass, height and branching. These characteristics could be the foundation for the improvement of *N. benthamiana* as a molecular farming production platform.

My contribution to this chapter was essential. I searched for the *FT* genes in *N. benthamiana*, I generated the edited lines presented in this chapter, genotyped, phenotyped them, and assayed their expression potential for eGFP. I wrote and corrected the text of the chapter and generated its figures.

Introduction

Plant Molecular Farming (PMF) is the production of industrially relevant and commercially important proteins in plants (Eidenberger *et al*., 2023). The objective of PMF is to offer safe and cost-effective means to produce bioproducts at a large scale. This research field shows a great potential for the efficient and cost-effective production of therapeutic proteins, antibodies, vaccines and other bioactive molecules. Plants exhibit many advantages as bioreactors over more conventional platforms for the production of recombinant proteins, such as scalability, cost-effectiveness, and the capacity to produce complex proteins (Ma *et al*., 2005; Stoger *et al*., 2014; Twyman *et al*., 2003).

Although various plant species have been employed as chassis for PMF approaches, in recent years *Nicotiana benthamiana* has become increasingly important as a production platform for PMF mainly for its suitability for *Agrobacterium*-mediated transient gene expression mediated by agrobacterium. A non-exhaustive list of recombinant products produced using this platform includes antibodies against Ebola virus (Qiu *et al*., 2014), human immunodeficiency virus (HIV) (Hamorsky *et al*., 2013), Zika virus (Diamos *et al*., 2020) and enterotoxigenic *Escherichia coli* (Teh *et al*., 2021), vaccines for non-Hodgkin's lymphoma (Bendandi *et al*., 2010), follicular lymphoma (Tusé *et al*., 2015), influenza (D'Aoust *et al*., 2008, 2010), and COVID-19 disease (Maharjan & Choe, 2021; Ruocco & Strasser, 2022). Moreover, *N. benthamiana* has been a powerful chassis for the production of metabolites, such as taxadiene (Hasan *et al*., 2014), casbene (Forestier *et al*., 2023), crocins (L. Xie *et al*., 2023), moth sex pheromones for sustainable pest control (Mateos-Fernández *et al*., 2021), and enzymes, as glucocerebrosidase (Limkul *et al*., 2015), glucose oxidase (Talens-Perales *et al*., 2023), a bacterial laccase (van Eerde *et al*., 2022) and fungal enzymes for the degradation of lignin (Khlystov *et al*., 2021). Despite the advantages and examples mentioned above, *N. benthamiana* is not yet an ideal platform for molecular farming. Some breeding efforts have been made in *N. benthamiana*, and incidentally also in *Nicotiana tabacum*, towards maximizing the yield of recombinant protein. For instance, co-expressing folding helpers and chaperons from the same species of the recombinant protein can be a promising strategy for this objective: the expression of human calreticulin in *N. benthamiana* has been shown to effectively increase the accumulation of human viral glycoproteins (Margolin *et al*., 2020). Another problem in recombinant protein production can be the activity of host endogenous proteases, that could degrade the product. To overcome this issue, the silencing of proteases through RNA interference (RNAi) has been proven successful to increase recombinant protein production in tobacco (Duwadi *et al*., 2015; Mandal *et al*., 2014). In *N. benthamiana,* the co-expression of the tomato protease inhibitor SlCYS8 enhanced antibody yield (Jutras *et al*., 2016) and protease inhibitors from *N. benthamiana* and human boosted the accumulation of a monoclonal antibody, erythropoietin and α-galactosidase (Grosse-Holz *et al*., 2018). Another issue in PMF is the non-human glycosylation pattern on secreted proteins that could be immunogenic for humans (Singh *et al*., 2021). Jansing *et al*. (2019) utilized CRISPR/Cas9 in *N. benthamiana* for the knockout of glycosyltransferases gene, obtaining lines that produced "humanized"

non-immunogenic glycoproteins, lacking β-1,2-xylose and core α-1,3-fucose (Jansing *et al*., 2019). Moreover, *N. benthamiana* lacks tyrosylprotein sulfotransferase (TPST) gene, that catalyses tyrosine sulfation, essential for the functionality of various antibodies targeting HIV (Singh *et al*., 2021). The transient co-expression of human TPST allowed the production in *N. benthamiana* of antibodies with proper tyrosine sulfation, that guaranteed their correct functionality against HIV (Singh *et al*., 2020).

One of the main disadvantages of *N. benthamiana* as compared with other PMF chassis is its relative low productivity in terms of biomass. As Alam *et al*. (2018) reported for the production of griffithsin at industrial scale, *N. benthamiana* plants are grown for 24 days from germination, inoculated with the construct of interest, and collected 7- 14 days after, with a total duration of 38 days for the upstream phase (Alam *et al*., 2018). Typically, *N. benthamiana* productive phase is limited to five-seven weeks post-germination, a period that correspond to the juvenile stage, since the capacity for recombinant protein production is severely reduced after flowering (Sheludko *et al*., 2007). In this context, delayed flowering could be considered as an interesting breeding goal for enhancing plant biomass. To this end, the underlying molecular mechanisms of reproductive phase transition need to be understood so that the key genes regulating this process could be eventually engineered.

Reproductive phase change in plants is regulated by various factors, one of the most important being the FLOWERING LOCUS T (FT) proteins, a member of the phosphatidylethanolamine-binding protein (PEBP) family. These proteins play a crucial role in floral transition (Chardon & Damerval, 2005; Hedman *et al*., 2009; Karlgren *et al*., 2011). In the model plant *Arabidopsis thaliana*, FT is a floral promoter and is primarily expressed in leaf phloem companion cells (Samach *et al*., 2000; Valverde *et al*., 2004). Under long-day conditions, FT protein migrates through the sieve elements from the leaves to the shoot apical meristem (SAM). Once in the SAM, FT interacts with FLOWERING LOCUS D (FD) transcription factor, thereby initiating a cascade of events that activate downstream target genes like SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) and APETALA 1 (AP1), a floral meristem identity gene. This network ultimately triggers floral development (Abe *et al*., 2005; Hanano & Goto, 2011; Wigge *et al*., 2005).

FT-like proteins have been identified in various plant species (Cao *et al*., 2016; Chardon & Damerval, 2005; Faure *et al*., 2007; Izawa *et al*., 2002; Meng *et al*., 2011; Yan *et al*., 2006) and their specific functions can differ. In tobacco, four FT paralogs were identified: NtFT1, NtFT2, NtFT3, and NtFT4. NtFT1, NtFT2, and NtFT3 were found to act as floral repressors. On the other hand, NtFT4 appeared to be a floral inducer, promoting the transition to flowering (Harig *et al*., 2012). Further investigations led to the discovery of an additional FT-like protein in tobacco, which was named NtFT5 (G. Wang *et al*., 2018). The overexpression of this gene in tobacco plants resulted in early flowering, highlighting its role as a crucial floral activator (G. Wang *et al*., 2018). Later, Schmidt *et al*. (2020) successfully knocked out NtFT5 in tobacco plants employing CRISPR/Cas technology. Under long-day conditions, the edited tobacco plants were completely unable to flower (F. J. Schmidt *et al*., 2020), providing evidence for the

critical role of NtFT5 in the flowering process of this species and highlighting the potential of NtFT5 as a target for crop improvement strategies aiming at the manipulation of flowering time in tobacco.

In addition to FT genes, also SQUAMOSA PROMOTER BINDING-LIKE (SPL) genes are essential players in plant physiology, controlling processes such as vegetative phase transition (Xu *et al*., 2016), leaf initiation rate and shoot and inflorescence branching (Schwarz *et al*., 2008; Wu & Poethig, 2006), flowering time (Gandikota *et al*., 2007; Xu *et al*., 2016), floral organ development and fertility (Xing *et al*., 2010), fruit development and ripening (Ferreira e Silva *et al*., 2014), the development of pollen sac (Unte *et al*., 2003), root (Yamasaki *et al*., 2009; Yu *et al*., 2015) and trichomes (Yu *et al*., 2010). These genes were originally discovered in *Antirrhinum majus* and earned their name due to their interaction with the gene SQUAMOSA, a floral meristem identity gene (Klein *et al*., 1996). The SPL transcription factor family is characterized by the SBP domain, a DNA binding domain containing zinc-finger motifs and a nuclear localization signal (Birkenbihl *et al*., 2005; Cardon *et al*., 1999). Some of the genes belonging to this family are regulated by microRNA 156, which impacts plant phase transitions (H. Wang & Wang, 2015). In *Arabidopsis thaliana*, 16 SPL genes were identified, each with distinct functions (Preston & Hileman, 2013; Xu *et al*., 2016). Similar gene family analyses have been performed in various plant species, including cotton (Cai *et al*., 2018), Moso Bamboo (Pan *et al*., 2017), rice (Z. Yang *et al*., 2008), citrus (Zeng *et al*., 2019), maize (Mao *et al*., 2016), tomato (Salinas *et al*., 2012), Petunia (Zhou *et al*., 2018), Chrysanthemum (Song *et al*., 2016), Tartary Buckwheat (M. Liu *et al*., 2019), and more recently in tobacco and *N. benthamiana* (Chapter 1, (De Paola *et al*., 2023)).

This chapter focuses on the edition in *N. benthamiana* of the orthologous genes of the tobacco flowering promoters *NtFT4* and *NtFT5* (*NbFT4*, *NbFT5_1a/1b*) either alone or in combination with the edition of *NbSPL13_1a*. After the edition, the phenotype of the generated lines and their potential as platforms for heterologous protein production was evaluated.

Materials and methods

Protein identification and phylogenetic tree construction

FT protein sequences from tobacco were retrieved by BLAST at NCBI using the FT sequence from *Arabidopsis* as reference sequence. These were then used to perform a BLAST search against the *N. benthamiana* genome (H2020 version, https://www.nbenth.com/) (Ranawaka *et al.*, 2023). The retrieved protein sequences together with the ones from *N. tabacum* were used to perform a multiple sequence alignment (MSA) using CLUSTAL v.1.2.4 (Sievers *et al.*, 2011). A phylogenetic tree was constructed with iqtree v1.6.12 (Nguyen *et al.*, 2015) using model JTTDCMut+G4 and a bootstrap value of 1000.

Gene structure

Coding sequences (CDS) and 3' untranslated regions (3'UTRs) of *NbSPL13_1a, NbFT5_1a, NbFT5_1b, NbFT4* were retrieved from *N. benthamiana* LAB330 v3.02 gene models. A browser extensible data (BED) file containing gene models was created with an in-house python script. This file was graphically represented using GSDS (Gene Structure Display Server) (http://gsds.cbi.pku.edu.cn/).

Plasmid assembly

Constructs used for transformations were assembled using GoldenBraid (Vazquez-Vilar *et al.*, 2020, 2021). For the assembly of guide RNAs on level 0, two partially complementary primers were designed at https://gbcloning.upv.es/do/crispr/multi_cas9_gRNA_domesticator_1 using the target sequences of Table S6 as input. The primers were used in a BsmBI restriction–ligation reaction together with pUPD2 and the corresponding level −1 tRNA-scaffold plasmid (GB1208 for sgSPL1.5 and sgFT4.2, GB1207 for sgSPL1.6, sgFT4.1 and sgFT6, GB1205 for sgFT5, GB1206 for sgFT6.1). Later, multipartite BsaI restriction–ligation reactions from level 0 parts and binary BsaI or BsmBI restriction–ligation reactions were performed to obtain all the level ≥1 assemblies. All plasmids were validated by restriction enzyme (RE) analysis. The sequences of all level ≥1 constructs can be found entering their IDs (displayed in Table S10) at https://gbcloning.upv.es/search/features/.

Plant material and stable transformation

The WT *N. benthamiana* LAB strain was used in the stable transformations for the knockout of *NbSPL13_1a, NbFT5_1a* and *NbFT5_1b*, while plants with the genetic backgrounds FT5 and FT5 SPL13 were used for the ones targeting *NbFT4* (see Results). The *Agrobacterium tumefaciens*-mediated transformations followed a standard protocol (Horsch *et al.*, 1985). Briefly, fully expanded leaves of WT plants were sterilized with 5% commercial bleach for 10 minutes followed by four consecutive washing steps with sterile _deionised water. Leaf discs (d= 0.8 cm) were cut with a cork borer and incubated

overnight in co-culture plates (4.9 g/L MS supplemented with vitamins (Duchefa, The Netherlands https://www.duchefa-biochemie.com/), 3% sucrose (Sigma-Aldrich, USA, https://www.sigmaaldrich.com/), 0.8% Phytoagar (Duchefa, The Netherlands), 1 mg/L BAP (Sigma-Aldrich, USA), 0.1 mg/L NAA (Sigma-Aldrich, USA), pH=5.7). Leaf discs were incubated for 15 minutes with a culture of *A. tumefaciens* LBA4404 harboring plasmids GB3298, GB3299, GB3301 or GB3699 (OD600=0.3). Discs were returned to the co-cultivation plates and incubated for two days in darkness. Next, discs were transferred to selection medium (4.9 g/L MS supplemented with vitamins (Duchefa, The Netherlands), 3% sucrose (Sigma-Aldrich, USA), 0.8% Phytoagar (Duchefa, The Netherlands), 1 mg/L BAP (Sigma-Aldrich, USA), 0.1 mg/L NAA (Sigma-Aldrich, USA), 500 mg/L carbenicillin, 100 mg/L kanamycin, pH=5.7). Discs were transferred to fresh medium every seven days until shoots appeared (four-six weeks). Shoots were cut and transferred to rooting medium (4.9 g/L MS supplemented with vitamins (Duchefa, The Netherlands), 3% sucrose (Sigma-Aldrich, USA), 0.8% Phytoagar (Duchefa, The Netherlands), 500 mg/L carbenicillin, 100 mg/L kanamycin, pH=5.7) until roots appeared. Screening for T-DNA segregation in T1 generations was achieved through DsRed fluorescence detection.

Transient expression and fluorescence detection

Transient expression assays were performed as described in Moreno-Giménez *et al*. (2022) with minor differences. Five- and nine-weeks old *N. benthamiana* plants of each genotype (WT, SPL13, FT5, FT5 SPL13, FT4 FT5, FT4 FT5 SPL13, see Results) were used for Agroinfiltration. Bacterial suspensions were adjusted to an optical density of 0.05 at 600 nm (OD600). For enhanced GFP (eGFP) expression with the TMV-based expression system ICON (Giritch *et al*., 2006), the bacterial suspensions harboring the MagnICON® Integrase (pICH14011), the MagnICON® 5' module (pICH17388) and the 3' eGFP module (GB4294, eGFP cloned in a vector adapted for BsaI cloning from MagnICON® pICH7410 (Diego-Martin *et al*., 2020)) were mixed in equal volumes. For eGFP expression with a geminiviral replicon system based on the Bean Yellow Dwarf Virus (BeYDV) (Dahan-Meir *et al*., 2018), equal volumes of bacterial suspensions harbouring plasmids GB3598 and GB4312 were mixed. Bacterial suspension harbouring vector GB4279 was used for eGFP expression under 35S promoter. As control, bacterial suspension harbouring empty vector GB0107 was used. Three leaves per plant were agroinfiltrated. Leaf samples were collected at four and seven days post infiltration (dpi) from the five-weeks old plants and at five and seven dpi from the nine-weeks old plants. To determine fluorescence, one 0.5 cm diameter disc was excised from each Agroinfiltrated leaf and discs were transferred to a black 96-well microplate. Subsequently, enhanced GFP (eGFP) fluorescence was determined using microplate reader Victor™ X5 (Perkin Elmer, USA) using excitation filter 480/30 nm (HH35000902) and emission filter 530/30 nm (HH35000903), following the manufacturer instructions.

Vacuum infiltration and fluorescence detection

Vacuum infiltration assays were performed as in Diego-Martin *et al.*, (2020) with minor differences. Bacterial suspensions for the expression of eGFP with the ICON system (Giritch *et al.*, 2006) were prepared as above. Agrobacterium was delivered to plant cells through vacuum infiltration using a vacuum degassing chamber (model DP118, Applied Vacuum Engineering, UK) equipped with a 30 L infiltration tank. The upper portion of the entire plants (eight plants simultaneously) was immersed in the Agrobacterium infiltration solution. Vacuum was applied for two minutes at a pressure of 0.8 bar, followed by gradual release. After seven days plants were harvested, weighed and ground in liquid nitrogen. To obtain total soluble protein extract, a small fraction of the ground tissue was collected, mixed with PBS buffer (20 mM NaH_2PO_4, 80 mM $Na_2HPO_4.7H_2O$, 100 mM NaCl, pH 7.4) in a 1:3 w/v ratio and centrifuged at 13,000 rpm for 15 min at 4°C. These were placed in a black 96-well microplate and enhanced GFP (eGFP) fluorescence was determined using microplate reader Victor™ X5 (Perkin Elmer, USA) using excitation filter 480/30 nm (HH35000902) and emission filter 530/30 nm (HH35000903), following the manufacturer instructions.

Genomic DNA extraction and editing efficiency evaluation

150 mg of leaf material was used for genomic DNA extraction using the CTAB (cetyl trimethylammonium bromide) method (Murray & Thompson, 1980). The genomic regions flanking the nuclease target sites were PCR amplified using MyTaq™ DNA Polymerase (Bioline, https://www.bioline.com/) and primers listed on Table S9. The PCR amplicons were confirmed on a 1% agarose gel electrophoresis and purified with ExoSAP-IT™ PCR Product Cleanup Reagent (ThermoFisher Scientific, USA, https://www.thermofisher.com) following the manufacturer's instructions prior to Sanger sequencing. Chromatograms of Cas9-edited genomic DNA were analyzed using Inference of CRISPR Edits (ICE) v2 tool from Synthego (https://ice.synthego.com/). All analyses were manually curated.

Phenotyping

For each line, six plants were grown in 17 cm diameter pots under a 14-h light (21.5°C)/10-h dark (18°C) regime in greenhouse. At flowering time, determined as the time from sowing to anthesis, each plant was phenotyped for the following parameters: number of leaves beneath the first flower bud, total height, number of lateral branches, stem diameter, fresh weight and total number and area of leaves. Leaves under 5 cm of length were not considered in the counting and in the total area determination.

Results

Generation of FT5 and SPL13 knock-out lines

Our previous results showed a small delay in flowering time when knocking out *NbSPL13_1a* (Chapter 1, (De Paola *et al.*, 2023)). Considering the relevance of *NtFT5* in tobacco flowering (F. J. Schmidt *et al.*, 2020), our first objective was the editing of *NbSPL13_1a* and the homeologous pair *NbFT5_1a* and *NbFT5_1b*, as first attempts to significantly delay flowering in *N. benthamiana*. The identification of the *FT* genes in *N. benthamiana* was possible thanks to a homology search with their tobacco orthologs in the *N. benthamiana* genome (version LAB3.30) (Ranawaka *et al.*, 2023), followed by the creation of a phylogenetic tree together with their tobacco counterparts (Fig. S4). The IDs of the genes are reported in Table 1.

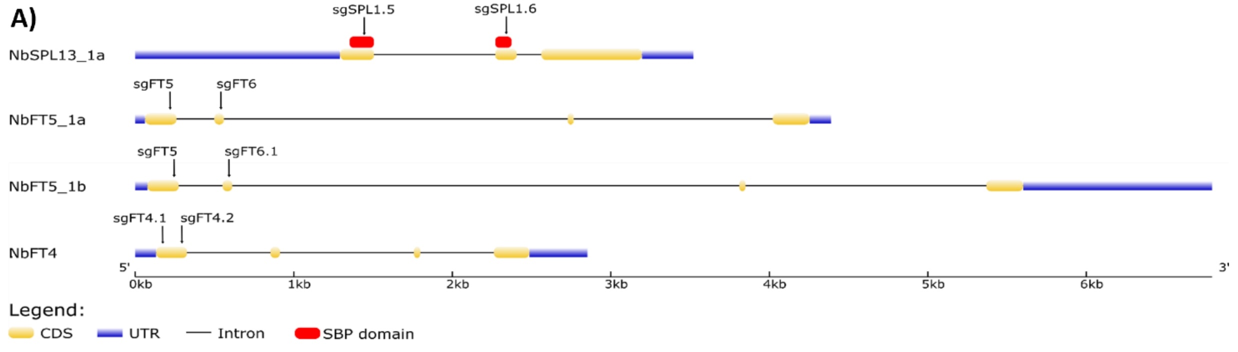
Gene	ID (LAB3.30)	Editing sgRNA	Position	Mutation	Genomic sequence	Lines
NbFT5_1a	gene.47270.0.4.p1	sgFT6	nc257	+A	CCTCCAATTGGTT	FT5 E2-3, FT5 SPL13 22E-5, FT5 FT4 4-11, FT5 FT4 SPL13 40-1
NbFT5_1b	gene.45744.0.4.p1	sgFT6.1	nc257	+A	ACCTCCAATTGGT	
NbSPL13_1a	gene.4997.0.0.p1	sgSPL1.6	c288	+T	AGGCGTTCGAAGG	FT5 SPL13 22E-5
				-ACGC	CCTTCG/CTGTTA	SPL13-3
NbFT4	gene.36576.0.3.p1	sgFT4-1	nc60	+A	CCATTCAACAAGA	FT5 FT4 SPL13 40-1
				-ACAAG	CCATTC/ATCTGT	FT5 FT4 4-11

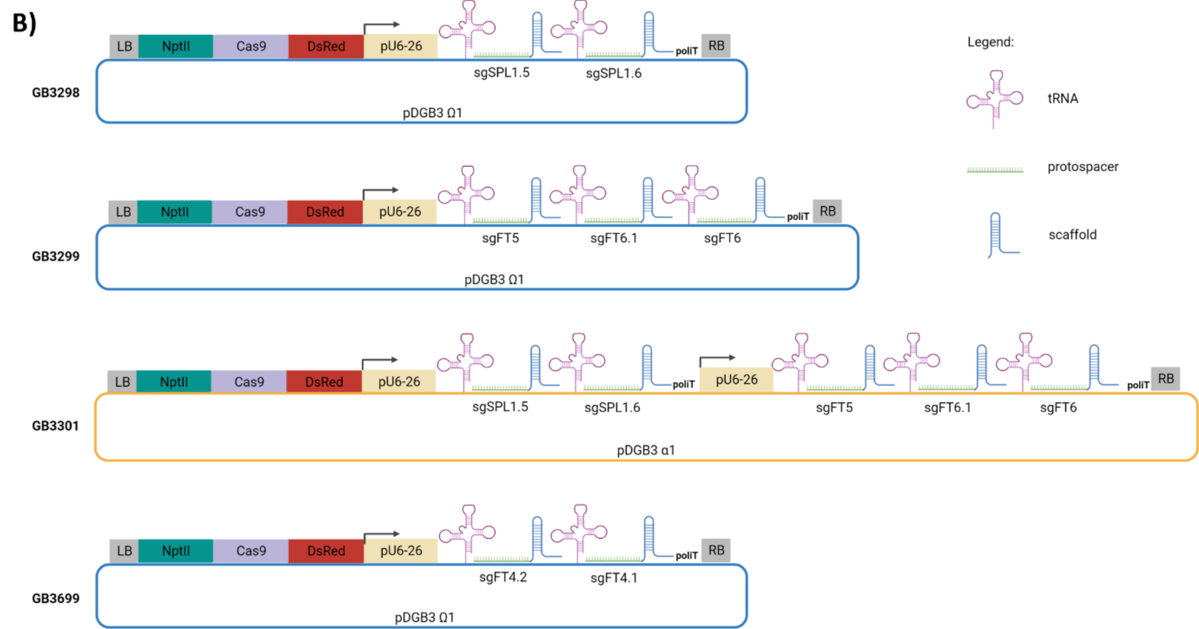
Table 1. List of target genes, sgRNAs and mutations.

The genes chosen as targets of CRISPR/Cas experiments are listed together with their IDs in the LAB3.30 *N. benthamiana* genome version. The sgRNAs that caused a frameshift mutation in their sequence are also listed. Their position is defined as the distance of the Cas9 cutting site from the ATG in their coding sequences. "c" indicates that the sgRNA is designed on the coding strand and "nc" that it is designed on the non-coding strand. The table also reports the type of mutation, with + representing an insertion and – standing for deletion, the resulting genomic sequence, and the T1 lines in which the mutation is present.

In order to edit these genes, sgRNAs were designed to target the CDS at two different positions, one in the first exon and one in the second exon (Fig. 13A). For *NbSPL13_1a*, we focused on targeting its SBP domain, which is important for its role as a transcription factor. The sequences of the chosen sgRNAs are listed in Table S6. After design, these sgRNAs were assembled following a tRNA-spaced polycistronic strategy (K. Xie *et al.*, 2015) using Goldenbraid (Vazquez-Vilar *et al.*, 2021). The final vectors used for *N. benthamiana* transformation were made up of the sgRNA expression cassette(s), the transcriptional units for the expression of the Cas9 protein, the red fluorescent protein DsRed, and the kanamycin resistance gene nptII (Fig. 13B). As depicted in figure 13C, three independent transformations were performed on WT plants: one for editing both *NbFT5* genes, one for knocking out *NbSPL13_1a* gene

specifically, and a third transformation for editing *NbFT5* and *NbSPL13_1a* genes simultaneously. The T0 progenies arising from these transformations were genotyped to search for explants carrying biallelic mutations (i.e., containing both alleles mutated but with different mutations), heterozygous mutations (containing a WT allele and a mutated allele) or homozygous mutations (with both alleles carrying the same mutation).


A)
sgSPL1.5
sgSPL1.6
NbSPL13_1a
sgFT5
sgFT6
NbFT5_1a
sgFT5
sgFT6.1
NbFT5_1b
sgFT4.1
sgFT4.2
NbFT4
5'
3'
0kb
1kb
2kb
3kb
4kb
5kb
6kb
Legend:
CDS
UTR
Intron
SBP domain


B)
GB3298
LB
NptII
Cas9
DsRed
pU6-26
sgSPL1.5
sgSPL1.6
poliT
RB
pDGB3 Ω1
GB3299
sgFT5
sgFT6.1
sgFT6
pDGB3 Ω1
GB3301
pDGB3 α1
GB3699
sgFT4.2
sgFT4.1
pDGB3 Ω1
Legend:
tRNA
protospacer
scaffold

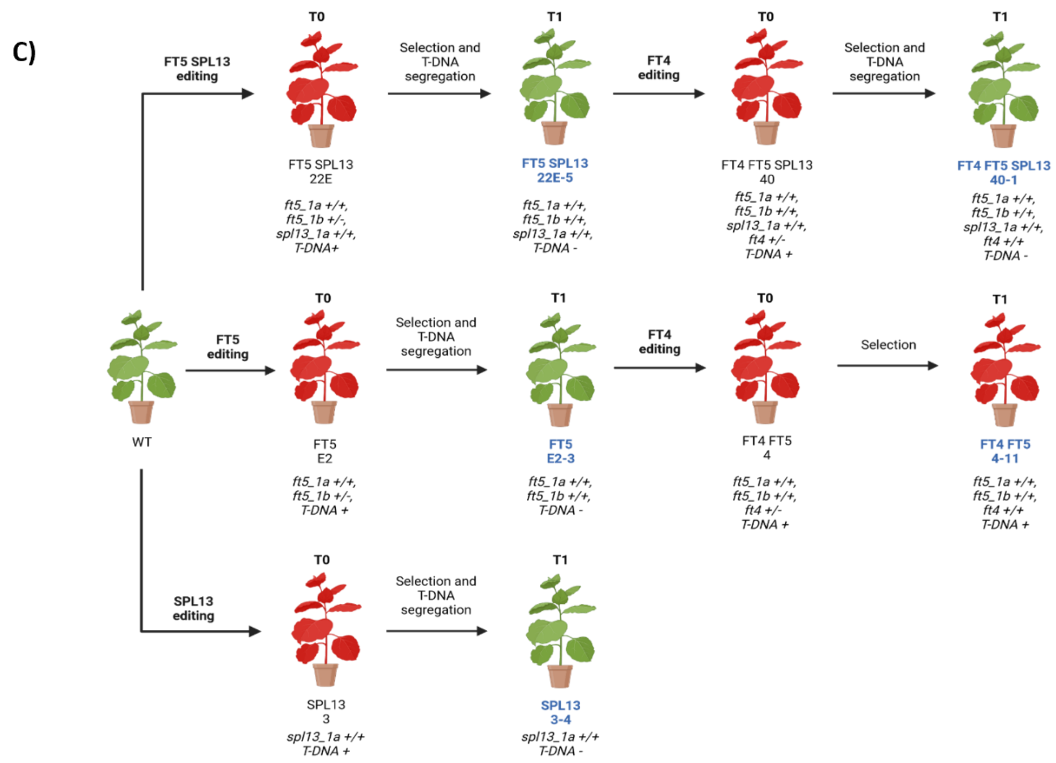

C)
FT5 SPL13 editing
T0
FT5 SPL13 22E
ft5_1a +/+, ft5_1b +/-, spl13_1a +/+, T-DNA+
Selection and T-DNA segregation
T1
FT5 SPL13 22E-5
ft5_1a +/+, ft5_1b +/+, spl13_1a +/+, T-DNA -
FT4 editing
T0
FT4 FT5 SPL13 40
ft5_1a +/+, ft5_1b +/+, spl13_1a +/+, ft4 +/- T-DNA +
Selection and T-DNA segregation
T1
FT4 FT5 SPL13 40-1
ft5_1a +/+, ft5_1b +/+, spl13_1a +/+, ft4 +/+ T-DNA -
WT
FT5 editing
T0
FT5 E2
ft5_1a +/+, ft5_1b +/-, T-DNA +
Selection and T-DNA segregation
T1
FT5 E2-3
ft5_1a +/+, ft5_1b +/+, T-DNA -
FT4 editing
T0
FT4 FT5 4
ft5_1a +/+, ft5_1b +/+, ft4 +/- T-DNA +
Selection
T1
FT4 FT5 4-11
ft5_1a +/+, ft5_1b +/+, ft4 +/+ T-DNA +
SPL13 editing
T0
SPL13 3
spl13_1a +/+ T-DNA +
Selection and T-DNA segregation
T1
SPL13 3-4
spl13_1a +/+ T-DNA -

Figure 13. Edition of the genes *NbSPL13_1a, NbFT5_1a/1b* and *NbFT4* in *N. benthamiana*.

A) Structure of target genes. The arrows represent the target sites of the sgRNAs used for editing. Red boxes represent the SBP domain of *NbSPL13_1a* gene, blue boxes represent UTRs, yellow boxes represent CDS and black lines represent introns. B) Goldenbraid vectors used for editing the above-mentioned genes. Each one of constructs carries *NptII* (resistance to kanamycin), *Cas9* and *DsRed* genes, for visual selection of transformed plants. The transcription of gRNAs is under the control of U6-26 promoter. gRNAs were assembled using the tRNA strategy (Xie *et al*., 2015). Vector GB3298 targets *NbSPL13_1a*, GB3299 targets *NbFT5* homeologous pair, GB3301 targets *NbSPL13_1a* and *NbFT5* homeologous pair and GB3699 targets *NbFT4*. LB: left border, RB: right border, poliT: poliT tail. C) Schematic representation of the generation of the edited N. benthamiana lines. From a first round of three transformations, T1 lines that were biallelic for *NbFT5* (E2-3), *NbSPL13* (3-4) and the combination of both (22 E-5) were selected. Later, a new round of supertransformation for *NbFT4* editing was perfomed on FT5 and FT5 SPL13 background, allowing to obtain T1 lines that were biallelic for *NbFT4* and *NbFT5* (4-11) and *NbFT4*, *NbFT5* and *NbSPL13_1a* (40-1). Red plants represent the ones still carrying the T-DNA with *DsRed*. As it could be observed in the scheme, T-DNA segregation was not possible for line FT4 FT5 4-11. The selected T1 lines, later used for phenotyping, are highlighted in blue.

As it can be observed in Table S7, 23 T0 plants were retrieved and genotyped from FT5/SPL13 transformation, nine from FT5 transformation and eight from SPL13 transformation. According to the percent edition values reported by Synthego (https://ice.synthego.com/). T0 plants were considered as chimeras (percent editing below 40%), carrying heterozygous mutations (percent editing around 50%) or carrying biallelic mutations (>90%). Regarding the *NbFT5* homeologous pair, the edition was more effective for *NbFT5_1a* than for *NbFT5_1b*, overall. In FT5 SPL13 transformation, 19 out of 23 plants (83%) showed edition percent in *NbFT5_1a* above 90%, while the remaining ones exhibited edition values below 40% or failed sequencing. Among the T0 plants of FT5 transformation, seven out of nine (78%) displayed edition values in *NbFT5_1a* above 90%, while the remaining ones had values below 50%. With regards to *NbFT5_1b* gene, in FT5 SPL13 transformation nine out of 23 T0 plants (39%) showed edition values around or above 50%. In FT5 transformation, only one plant out of nine (11%) exhibited an edition value around 50% in *NbFT5_1b*, while the remaining ones were not edited. Considering *NbSPL13_1a*, the edition values of this gene in the FT5 SPL13 transformation were around or above 50% for 12 out of 23 plants (52%) and in the SPL13 transformation in four out of eight plants (50%) (Table S7).

To continue with further characterizations, the most edited plants in this first targeted mutagenesis round were selected: (i) line FT5 SPL13 22E having a biallelic mutation for *NbFT5_1a*, a heterozygous mutation for *NbFT5_1b*, and a biallelic mutation for *NbSPL13_1a*; (ii) line FT5_E2, with a biallelic mutation for *NbFT5_1a* and a heterozygous mutation for *NbFT5_1b*; and (iii) line SPL13_3, containing a biallelic mutation for *NbSPL13_1a*. The three T0 lines were self-pollinated to obtain T1 seeds.

Next, the T1 offspring was further screened to obtain non-transgenic full KO plants. Segregation of Cas9 transgene was followed by DsRed negative selection. Plants that segregated Cas9 and that showed

homozygous mutations for those genes carrying heterozygous mutations in the previous generation were chosen for the following experiments, i.e. lines FT5 SPL13 22E-5, FT5 E2-3 and SPL13 3-4 respectively. The sequence of each specific mutation in the selected lines is depicted in Table 1.

The T2 plants originating from the selected lines were phenotyped for flowering time and number of lateral branches at flowering (Fig. 14). As expected, SPL13 3-4 plants showed a small delay of 2-3 days compared to WT. Interestingly FT5 E2-3 and FT5 SPL13 22E-5 lines exhibited a longer delay of approximately 10 days compared to WT. All mutant lines displayed more lateral branches upon flowering than WT. No other obvious phenotypic changes were observed for these plants.

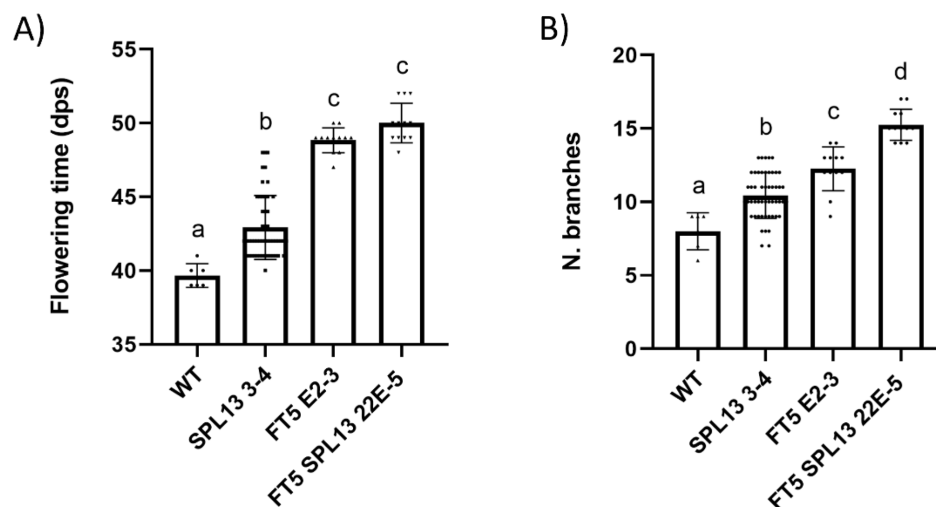


Figure 14. Phenotyping of SPL13 3-4, FT5 E2-3 and FT5 SPL13 22E-5 *N. benthamiana* lines.

A first phenotyping was conducted on T2 progeny coming from the selected T1 lines for flowering time (A) and number of lateral branches at flowering time (B). Each bar represents mean value ± SD. Bars labelled with different letters represent values that are mutually significant. Bars labelled with the same letter are not mutually significant. A one-way Anova test was conducted (WT n=6, SPL13 3-4 n=51, FT5 E2-3 n=12, FT5 SPL13 22E-5 n=12).

Stacking FT4 mutations on FT5 and SPL13 knock-out lines

Since lines FT5 SPL13 22E-5, FT5 E2-3 and SPL13 3-4 showed only a moderate delay in flowering, we decided to knock out also *NbFT4* gene, whose ortholog *NtFT4* in tobacco is also implicated in promoting flowering (Harig *et al.*, 2012). Contrary to *NbFT5*, *NbFT4* has only one homeologous gene in *N. benthamiana*, meaning that only one *FT4* gene is present in the genome. Two sgRNAs were designed to target its first exon, since exons two and three were too short to design sgRNAs with a high predicted on-target score on their sequence (Fig. 13A). T1 lines E2-3 (homozygous for *NbFT5* homeologous pair) and 22E-5 (homozygous for *NbFT5* homeologous pair and *NbSPL13_1a*) (Fig. 13C), were chosen as

background lines for *NbFT4* editing. The vector GB3699 containing the sgRNAs targeting *NbFT4* was used to transform the plants (Fig. 13B).

The genotyping of T0 plants from this experiment can be followed in Table S8. From the transformation in FT5 background, only one plant was recovered carrying a heterozygous mutation in *NbFT4* out of 18 primary transformants, with 17 not edited T0 plants. From the transformation in FT5 SPL13 background, only three plants out of 19 had a heterozygous mutation, with 16 not edited lines. Lines FT4 FT5 4 and FT4 FT5 SPL13 40, respectively, were selected for generating T1 plants, aiming to obtain homozygous mutants for these genes. These plants were self-pollinated and the progeny subsequently genotyped. As expected, approximately ¼ of the progeny resulted in homozygous mutant lines (not shown). Among their T1 progeny, the lines 40-1 and 4-11, homozygous for *FT4* in FT5 SPL13 and FT5 background, respectively, were chosen for the following phenotyping experiments. The selected FT4 FT5 SPL13 40-1 line was DsRed negative, indicating that the T-DNA containing Cas9 was not present in this line. As depicted in figure 13C, the segregation of T-DNA was not possible for FT4 FT5 4 progeny, therefore FT4 FT5 4 -1 line was selected for further analysis despite containing the T-DNA. The sequence of each specific mutation in the selected lines is depicted in Table 1.

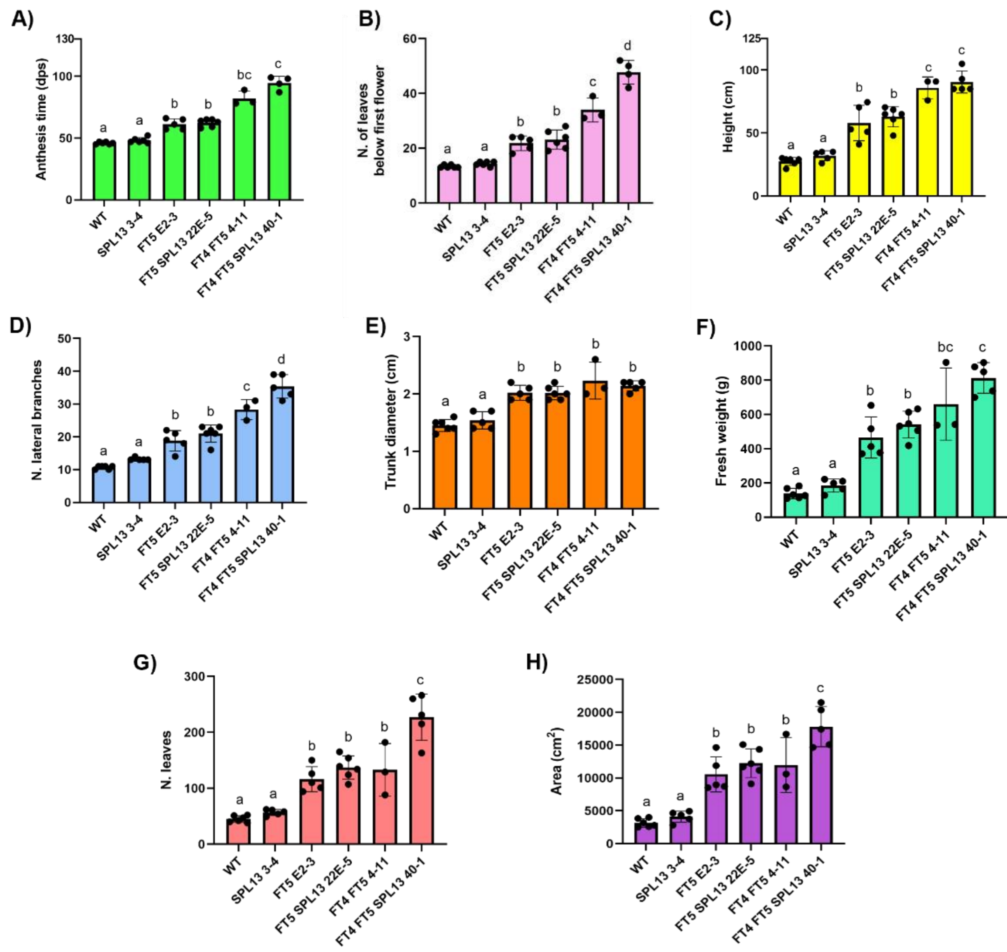
The overall strategy for directed mutagenesis of *FT4*, *FT5* and *SPL13* genes in *N. benthamiana*, as depicted also in figure 13, comprised four T-DNA constructs targeting four different genes and employing eight different gRNAs. It took four generations to reach the final mutagenesis goal. This process could have been shorter, but it should be noted that the decision to stack the editing of *NbFT4* was not in the initial design: this objective was achieved via super-transformation instead of multiplexing. As expected, the efficiency of the different employed sgRNAs varied dramatically, as shown in the graph of figure 15. Each point represents the average of the edition percentage (calculated using Synthego) for each gene of the T0 plants obtained in each transformation in which the referred guide RNA was used. sgFT6 (targeting the second exon of *NbFT5_1a*) had the highest efficiency among all guides (above 80%), while sgFT6.1 (targeting the second exon of *NbFT5_1b*) exhibited an efficiency value around 10%. sgFT5 (targeting the first exon of both *NbFT5* genes) displayed an efficiency close to 0 (Fig. 15). The difference in the efficiency of these sgRNAs resulted in a better editing of *NbFT5_1a* compared to *NbFT5_1b*. The editing of *NbSPL13_1a* gene was mainly due to sgSPL1.6. This guide RNA (targeting the second exon of the gene) exhibited around 40% of efficiency value, while sgSPL1.5 (targeting the first exon) displayed an efficiency close to 0%. With regards to *NbFT4*, the efficiencies of both sgRNAs targeting this gene were close to 0%, with sgFT4.1 efficiency value slightly higher than the one of sgFT4.2 (Fig. 15). It is interesting to note that sgRNAs with the highest efficiency are the ones in the last position of the assembled transcriptional units (TU), while the efficiency dramatically decreases for those in the second- or third-to-last position (Fig. 13B, Fig. 15).

Figure 15. Efficiency of the sgRNAs.

The efficiency of each guide RNA was calculated as the mean of the editing percentages of every T0 plant generated in one transformation. These values are represented by dots in the graph. Bars represent the mean of the efficiencies of every transformation in which the guide RNA was used.

Comparative analysis of flowering time and biomass-related parameters in FT4, FT5 and SPL13 mutant combinations.

For the phenotyping of the generated mutant lines, the T2 progeny derived from each chosen T1 line was utilized, since the homozygous mutations sought in our target genes were stabilized in this generation. Our objective was the evaluation of the differences in flowering time and biomass of the different lines. For each line under analysis, six plants were grown in 17 cm diameter pots. Flowering time was determined for each plant as the time elapsed from sowing until the anthesis of the first flower. Then, at the stage of the first flower anthesis, each plant was phenotyped for the following parameters: number of leaves beneath the first flower bud, total height, number of lateral branches, stem diameter, total fresh weight, total number of leaves and total area of leaves. Data were collected and analysed to produce the graphs of figure 16.


A)
Anthesis time (dps)
B)
N. of leaves below first flower
C)
Height (cm)
D)
N. lateral branches
E)
Trunk diameter (cm)
F)
Fresh weight (g)
G)
N. leaves
H)
Area (cm²)
WT
SPL13 3-4
FT5 E2-3
FT5 SPL13 22E-5
FT4 FT5 4-11
FT4 FT5 SPL13 40-1

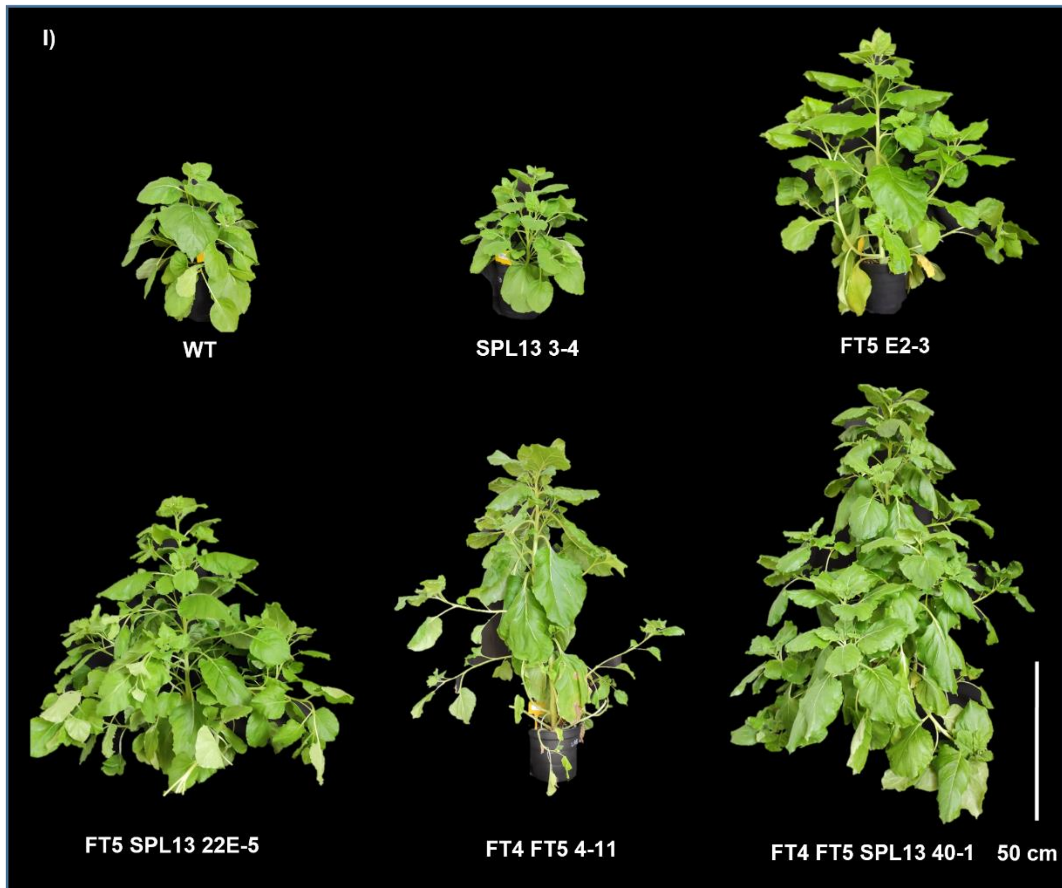

I)
WT
SPL13 3-4
FT5 E2-3
FT5 SPL13 22E-5
FT4 FT5 4-11
FT4 FT5 SPL13 40-1
50 cm

Figure 16. Phenotyping of *N. benthamiana* mutant lines.

A) First anthesis time. This time point was recorded when the first flower entered anthesis. All remaining parameters were measured at first anthesis time for each line . B) Number of leaves on the main axis below the first flower . C) Height D) Branching. This parameter indicates the number of lateral branches. E) Stem diameter. F) Fresh weight. G) Total number of leaves. H) Total area of leaves. For each plant, every leaf measuring more than 5 cm in length was harvested and its area calculated. These values were then summed to obtain the total leaf area for each plant. Each bar represents mean value ± SD. For every parameter, bars labelled with different letters represent values that are mutually significant. Bars labelled with the same letter are not mutually significant. A one-way Anova test was conducted. I) Images of plants representative for each line. Plants were recollected and photographed at first flower anthesis.

As it can be observed in figure 16A, the flowering time increased as more genes were edited: SPL13 3-4 line showed a small average delay of two days compared to WT, a difference that was not statistically significant with the number of plants analysed. In contrast to *SPL13*, the knocking out of *FT* genes had a marked effect in *N. benthamiana* flowering time. The lines FT5 E2-3 and FT5 SPL13 22E-5 showed a similar delay compared to WT (15 and 16 days, respectively). Interestingly, the combination of *FT4* and *FT5* KOs had a dramatic effect in delaying flowering time: line FT4 FT5 4-11 showed a remarkable delay of 36 days in the appearance of the first flower in anthesis. Finally, although *SPL13* mutation alone had little effect on delaying flowering time in a WT background, its combination with the *FT4/FT5* background clearly strengthened its late flowering effect. As shown in figure 16A, the FT4 FT5 SPL13 40-1 line showed the most substantial delay in reaching first anthesis, amounting to 48 days, doubling the flowering time of WT plants in the assayed growth conditions.

Another analysed parameter related to flowering time is the number of leaves below the first flower, this being a convenient indication of the delays occurring in the developmental flowering program. *N. benthamiana* shows and indeterminate growth; when the juvenile phase ends, single-flower inflorescences appear in the main axes, each new flower appearing after two adult leaves (see Fig. 17). The same architecture is repeated in lateral branches (S. J. Park *et al.*, 2014).

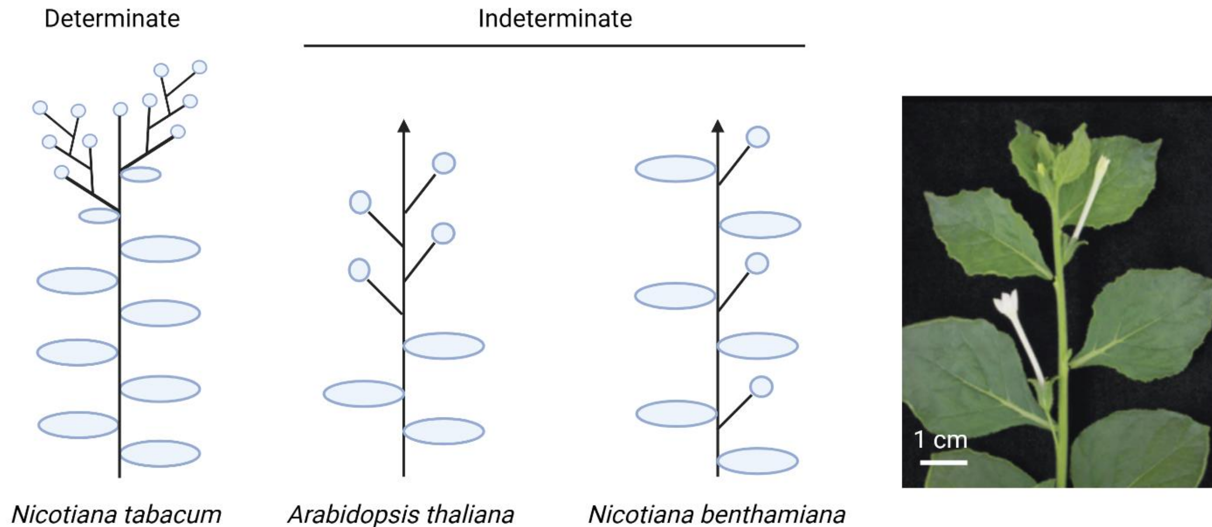


Figure 17. Scheme of inflorescence architecture in *N. tabacum, A. thaliana* and *N. benthamiana*.

Tobacco presents a determinate growth, meaning that this plant forms terminal flowers. *A. thaliana* and *N. benthamiana* exhibit an indeterminate type of growth, in which the main axis grows indefinitely, generating lateral flowers. Ovals represent leaves, circles represent flowers. Schemes adapted from Benlloch *et al.*, 2007 and Park *et al.*, 2014. On the right, a picture of a *N. benthamiana* plant with flowers from Park *et al.*, 2014.

In the assayed conditions, wild type plants typically start producing single flowers after 15-16 young leaves. As shown in figure 16B, this parameter showed no significant changes in single *SPL13* mutants. However, it increases significantly when *FT5* is mutated, and it is almost doubled in *FT4 FT5* KO plants. Similarly, as it was observed for flowering time, the *SPL13* mutation results in a dramatic developmental delay only when knocked out in a *FT4 FT5* mutant background.

N. benthamiana is typically agroinfiltrated and/or harvested at pre-flowering stages for maximizing productivity. Taking this into consideration, several biomass parameters were measured in our mutant lines at the start of flowering, to account for the gains in "juvenile" biomass occurring as a consequence of the delay in flowering. Parameters as plant height (Fig. 16C), the number of lateral branches (Fig 16D), fresh weight (Fig. 16F), the total number of leaves (Fig. 16G) and the total area of leaves (Fig. 16H), all followed a similar trend. The diameter of the stem does not follow this general trend: while line SPL13 3-4 stem diameter value is not significant compared to WT as expected, lines FT5 E2-3, FT5 SPL13 22E-5, FT4 FT5 4-11 and FT4 FT5 SPL13 40-1 values are significant compared to WT and SPL13 3-4 but not among them. The most outstanding differences regarding all parameters were the ones obtained in line FT4 FT5 SPL13 40-1 in comparison to WT. The doubling in anthesis time, from 46 ± 0.9 days to 94.5 ± 5.4 days had consequences on the overall biomass: fresh weight increased approximately eight times, from 138.4 ± 30.3 g to 813.2 ± 89.8 g, height triplicated from 27.4 ± 3.2 cm to 90.6 ± 8.8 cm, the number of leaves raised from 45 ± 5.8 to 227 ± 41.4 and the total area of leaves from 3160.1 ± 645.9 cm^2 to

17814 ± 3062.9 cm^2, an increase of 560%. Finally, also the number of lateral branches increased from 10.7 ± 0.5 to 35.4 ± 3.6.

Figure 16I shows images of plants representing each line. Plants were photographed at anthesis. The increasing height and dimensions can be appreciated as more genes are edited, being FT4 FT5 SPL13 40-1 the biggest one, compared to WT.

To confirm that the observed phenotypes were conserved in different growth conditions, a second experiment was conducted, in which plants were grown at high density, a condition in which flowering time is often accelerated. For this, 24 plants for each line were sown in small pots and divided in three blocks of eight plants each. Blocks were organized in random positions and separated to avoid negative effects on growth due to over-crowding. When 50% of the plants of each line had floral bud, they were collected and weighed. Figure 18A shows the flowering time of each genotype, expressed as days post sowing. As in the previous experiment, it can be noticed that this parameter increases with the edition of more genes. Figure 18B and C depict fresh and dry weight of the blocks, normalized for the number of plants. These parameters follow a trend that is globally like the previous experiment with low density conditions. In this case however, lines FT5 E2-3 and FT5 SPL13 22E-5 showed not significant changes compared to WT and SPL13 3-4 lines, with line FT5 E2-3 weighing slightly more than line FT5 SPL13 22E-5. Line FT4 FT5 SPL13 40-1 exhibited the highest biomass values as expected, but for fresh weight was not significant compared to FT4 FT5 4-11 line. Figure 18D shows pictures of blocks of eight plants for each line, taken at flowering time.

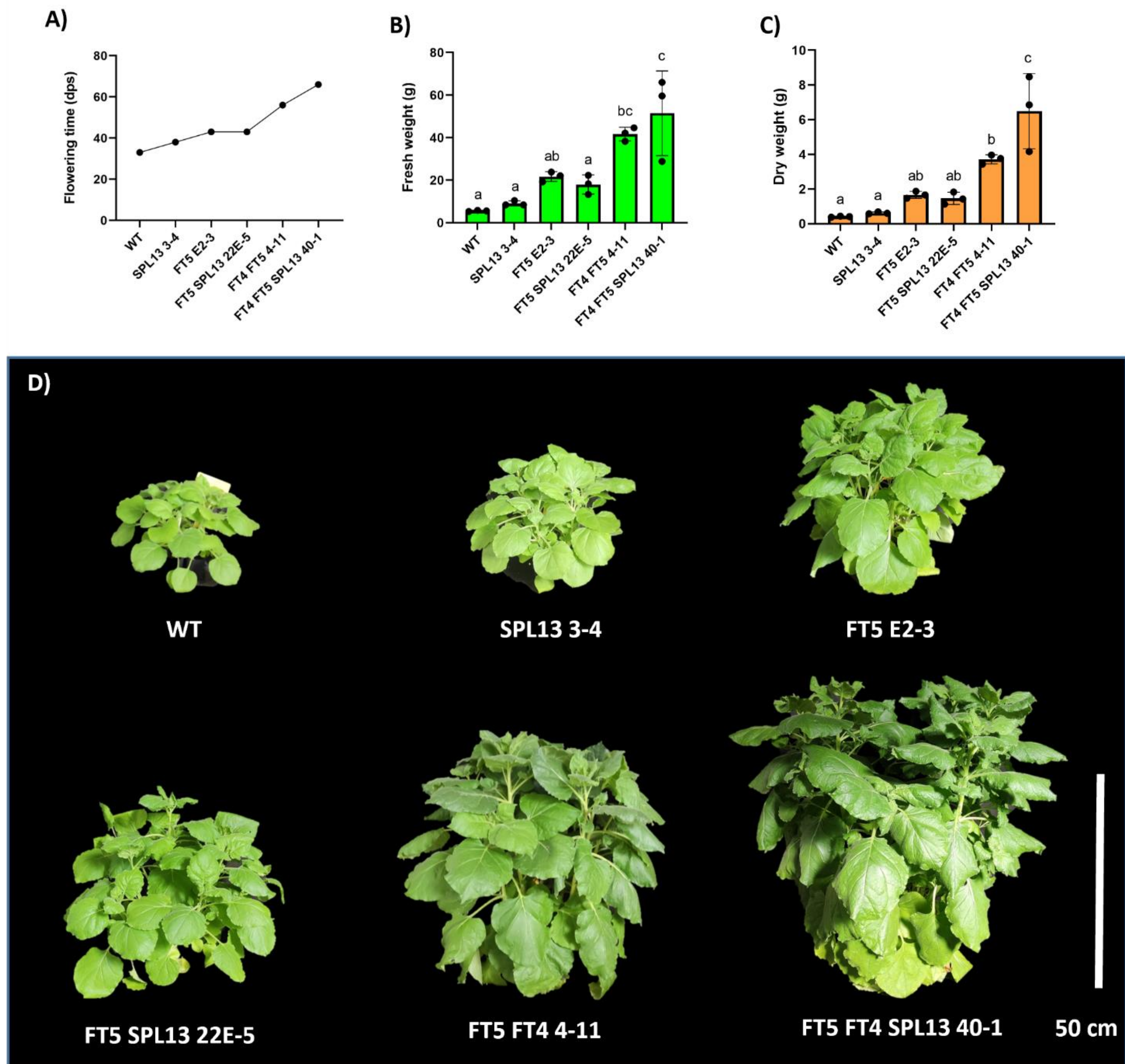


Figure 18. Phenotyping of *N. benthamiana* plants grown in high density.

A) Flowering time for each genotype was recorded when 50% of the plants had a floral bud. B) Fresh weight at the same time point. C) Dry weight at the same time point. Plants of each genotype were weighed together when harvested. Each point in the graphs represents values of a block of eight plants, normalized by the number of plants. Each bar represents mean value ± SD. Bars labelled with different letters represent values that are mutually significant. Bars labelled with the same letter are not mutually significant. A one-way Anova test was conducted (n=3 for each line). D) Images of the lines grown in small pots taken at flowering time. Each picture represents a block of eight plants grown together in small pots.

Analysis of recombinant protein expression in mutant *N. benthamiana* lines

In addition to their phenotype, we were also interested in exploring the recombinant protein expression potential of the generated lines. For this experiment, six plants were grown for each line, which were

agroinfiltrated on the highest fully expanded leaf and on the leaf below using a syringe with three alternative expression systems, namely a non-replicative system, a geminivirus-based DNA replicative system (BeYDV), and a Tobacco Mosaic Virus (TMV)-based vector (ICON). All three systems were used to express an enhanced green fluorescent protein (eGFP). An empty binary vector (EV) was agroinfiltrated as a negative control. Contrary to the previous phenotyping experiment, in which plants at the same growth stage (flowering) were taken in consideration, here all plants were infiltrated at the same time. Plants were first infiltrated at five weeks of age, the typical infiltration time for WT plants, and leaf disks were collected from the infiltration area to measure their fluorescence at four days post infiltration (dpi) and at seven dpi. A second infiltration was performed when plants were nine-weeks old, and samples were collected at five dpi and seven dpi respectively. As expected, the non-replicative 35S-eGFP construct resulted in low fluorescence values, while both replicative systems (BeYDV-eGFP and ICON-eGFP) reached higher fluorescence values.

For five-weeks old plants, four dpi, BeYDV-eGFP fluorescence values hover around $6x10^6$ arbitrary units (a. u.) for all lines, except for line FT4 FT5 4-11 that displays a lower fluorescence value (Fig. 19A). All these values decreased at around $5x10^6$ a. u. at seven dpi (Fig. 19B). ICON-eGFP fluorescence values followed an opposite trend, increasing from around $1x10^6$ a. u. at four dpi to around $5x10^6$ a. u. at seven dpi for all constructs.

In the experiment in which agroinfiltration was performed on nine-weeks old plants, fluorescence values of all constructs decreased with respect to those performed at week five. In figure 19C it can be noticed that at five dpi BeYDV-eGFP fluorescence values ranged between 2 and $3x10^6$ a. u., with FT4 FT5 SPL13 40-1 line having the highest fluorescence value. At seven dpi these values followed the same trend, but no significance could be found between samples. Also, ICON-eGFP fluorescence values were lower in nine-weeks old plants, ranging between 1 and $7x10^5$ a. u. at five dpi and around $1x10^6$ a. u. at seven dpi, being the fluorescence value of line FT4 FT5 SPL13 40-1 the highest one and the only sample showing a significant difference.

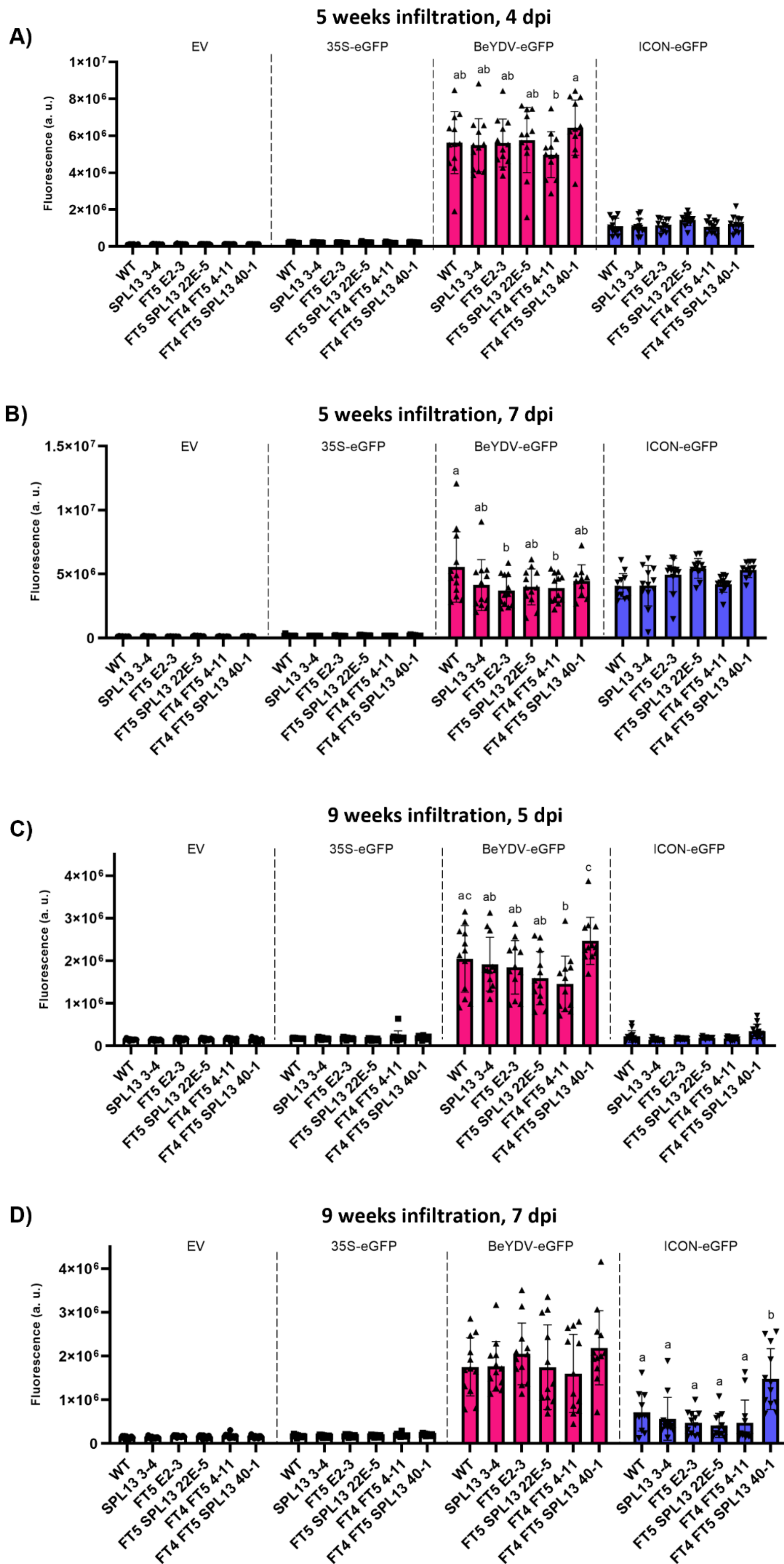

A)
5 weeks infiltration, 4 dpi
EV
35S-eGFP
BeYDV-eGFP
ICON-eGFP
Fluorescence (a. u.)
WT
SPL13 3-4
FT5 E2-3
FT5 SPL13 22E-5
FT4 FT5 4-11
FT4 FT5 SPL13 40-1
B)
5 weeks infiltration, 7 dpi
C)
9 weeks infiltration, 5 dpi
D)
9 weeks infiltration, 7 dpi

Figure 19. GFP expression in different *N. benthamiana* genotypes.

Upon infiltration with different constructs on five-weeks old plants, fluorescence was detected after four (A) and seven days (B). Another round of infiltration was performed when plants were nine-weeks old, and fluorescence was measured after five (C) and seven days post infiltration (D). The contructs employed were, from left to right: empty vector, 35S-eGFP, BeYDV-eGFP, ICON-eGFP. Bars represent mean ± SD. Within the same construct, bars labelled with different letters represent values that are mutually significant. Bars labelled with the same letter or not labelled are not mutually significant. A one-way Anova test was conducted (n=12 for each line).

Since it was observed a significant increase of recombinant protein production associated to line FT4 FT5 SPL13 40-1 using ICON-eGFP in nine-weeks old plants (Fig. 19D), we decided to perform a vacuum infiltration experiment with this line in high density conditions to confirm these results. Three blocks of eight plants were grown in small pots for both WT and FT4 FT5 SPL13 40-1 line. As the previous infiltration experiment, plants of the two genotypes were infiltrated not at the same stage (for instance flowering), but at the same time. Therefore, at nine weeks post sowing plants were vacuum infiltrated with ICON-eGFP, and at seven dpi they were harvested. The fresh weight of the plants was measured as "leaves", and "stem plus leaves", and the obtained values were normalized for the number of plants in each block. As it can be observed in figure 20A, the weight of FT4 FT5 SPL13 40-1 plants is slightly above the weight of WT plants when considering both "leaves" weight and "stem plus leaves" weight, but without a significant difference. After weighting, each sample was ground in liquid nitrogen and the tissue was homogenized in phosphate buffered saline (PBS) to obtain protein extract. Then, the fluorescence of the protein extract coming from each sample was measured. In figure 20B, the fluorescence values of PBS buffer and the protein extract of a non-infiltrated WT plant are represented as controls, together with the fluorescence values of the protein extracts of the infiltrated WT plants and the infiltrated FT4 FT5 SPL13 40-1 plants. Contrary to our previous results, no significant difference between WT and FT4 FT5 SPL13 40-1 fluorescence values was found. Moreover, the fluorescence of WT protein extract is even slightly higher than FT4 FT5 SPL13 40-1 protein extract fluorescence.

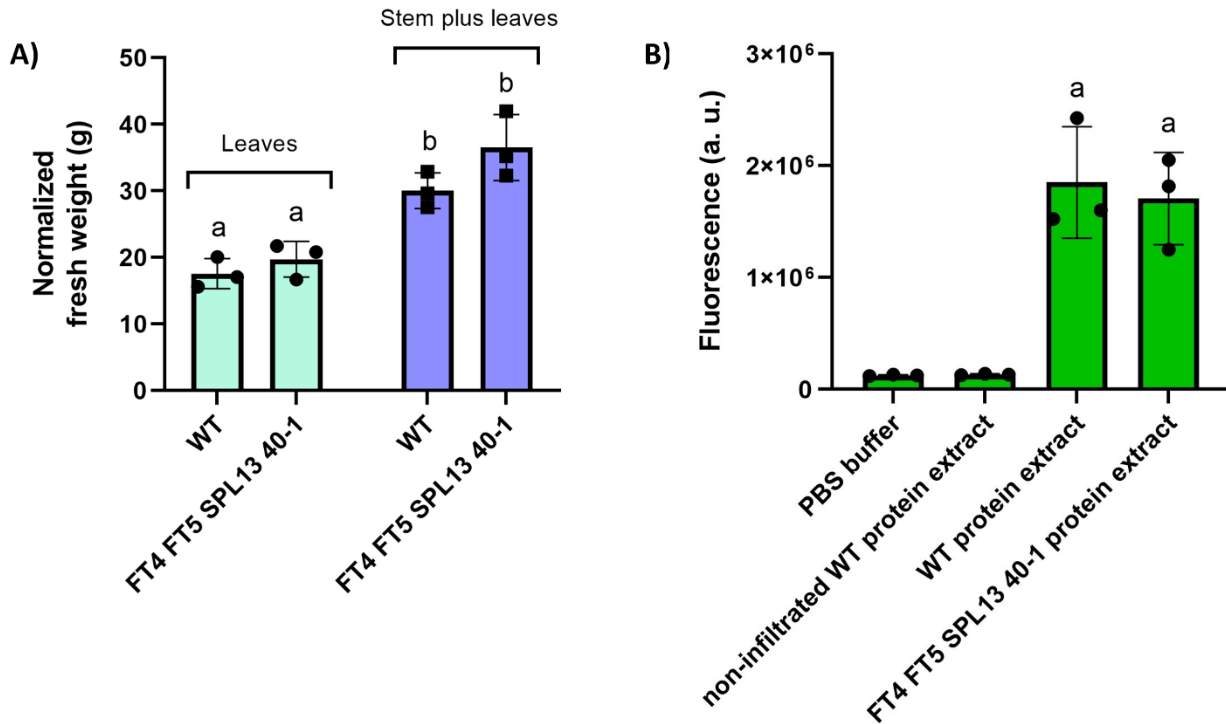


Figure 20. Recombinant GFP production in *N. benthamiana* plants grown in high density.

A) Fresh weight of WT and FT4 FT5 SPL13 40-1 *N. benthamiana* plants grown in high density, harvested seven days post vacuum infiltration, measured as weight of only leaves (left) and of stem plus leaves (right). For each genotype, three replicates were made. B) Fluorescence of protein extracts of vacuum-infiltrated plants. Three replicates were made for the measurement. Bars represent mean ± SD. Bars labelled with the same letter or not labelled are not mutually significant. A one-way Anova test was conducted.

Discussion

CRISPR/Cas has been widely used to edit plants to enhance productivity or to obtain desired traits. This is facilitated by the ability to deliver various guide RNAs at the same time, a feature called multiplexing, that enables the targeting of different genes simultaneously. Several studies already made use of CRISPR multiplexing to improve crops. The editing of six genes important for productivity and yield, namely, *SELF-PRUNING (SP), OVATE (O), FASCIATED (FAS), FRUIT WEIGHT 2.2 (FW2.2), MULTIFLORA (MULT)* and *LYCOPENE BETA CYCLASE (CycB*), allowed *de novo* domestication of wild tomato (*Solanum pimpinellifolium*), improving this species in terms of fruit number, fruit size and nutritional value (Zsögön *et al.*, 2018). Also in tomato, Kwon et al. (2020) edited *SP*, its paralog *SELF-PRUNING 5G (SP5G),* and *ERECTA (ER),* genes related to plant architecture and yield, obtaining tomato plants reduced in size that produce precocious fruits and are better adapted to urban agriculture (Kwon *et al.*, 2020). The KO of genes *Gna1*, *DEP1* and *GS3* resulted in rice mutants with a dense erect panicles, increased grain number and grain size (M. Li *et al*., 2016). Semi-dwarf rapeseed with increased branching was obtained by knocking out the two homologs of the gene *BnaMAX1*, improving silique yield (Zheng *et al*., 2020); *N. benthamiana* was glyco-engineered by knocking out two xylosyl and four fucosyl transferase genes (Jansing *et al*., 2019), and resistance to powdery mildew was achieved in tomato and wheat through the knock-out of *MLO* alleles (Nekrasov *et al*., 2017; Y. Wang *et al*., 2014).

Multiplexing is particularly powerful for polyploid species as *N. benthamiana*, an ancient allotetraploid. This tool was used to manipulate growth of *N. benthamiana,* aiming at the improvement of its performance as a molecular farming platform. Our primary objective was to abolish or at least to significantly delay flowering to prolong the juvenile phase, so that the expression capability of heterologous proteins would be at its maximum potential for longer periods, and together with it, increase biomass to have a higher production volume. These two combined features would have resulted in increased production yield. To achieve this objective, as a first strategy sgRNAs targeting the CDS of *NbSPL13_*1a and *NbFT5* homeologous pair were designed. Among the T1 lines obtained in which these genes were edited, it was not observed abolition of flowering, but only a moderate delay. This allowed us to acknowledge the difference in flowering activation between *N. tabacum* and *N. benthamiana*: in the former, the knock-out of *NtFT5* homologous pair is sufficient for the complete abolition of flowering (F. J. Schmidt *et al*., 2020), while in the latter the transition to reproductive phase clearly depends on more complex mechanisms, despite being two species that are very much related.

Due to this phenomenon, the knockout of NbFT*4* was stacked onto *NbFT5* and *NbSPL13_1a* knockout lines. New sgRNAs targeting *NbFT4* CDS were designed, cloned in vectors and used to perform transformations in FT5 and FT5 SPL13 backgrounds. Even though most of the sgRNAs that were used had a very low efficiency, it was possible to generate all the homozygous mutations we were aiming for. Our observations indicate that sgRNAs at the end of the polycystronic transcriptional unit (TU) in the assembled vectors had the highest efficiency in editing their target gene, while efficiency decreased for

the guides in the previous positions. As a matter of fact, the guides with the highest efficiency were, in decreasing order, sgFT6 and sgSPL1.6, both situated at the end of the TU. Then, guide sgFT6.1 which was situated on the second-to-last position. Guides sgFT5 and sgSPL1.5 had efficiencies close to 0, being positioned in the first position of the TU. As for the *NbFT4* guide RNAs, they both had very low efficiencies, but it is true that sgFT4.1 exhibited a slightly higher efficiency value, being situated in the last position of the TU. Vazquez-Vilar *et al*. (2021) already observed that the position of the guide RNAs in the TU influenced their editing efficiency, with the last position conferring the highest efficiency (Vazquez-Vilar *et al*., 2021). These differences in the editing efficiency could be explained by the fact that the last position in the assembled TU is flanked by a single tRNA, while the previous positions need to release 5'and 3' tRNAs to produce a functional gRNA. Another explanation could be that the presence of the polyA tail at the 3' end makes the last sgRNA more stable than the other ones in the TU (Vazquez-Vilar *et al*., 2021).

The segregation of T-DNA among the T1 progeny is often desirable to avoid both the presence of the transgene and continued Cas editing. In the case of FT4 FT5 4 progeny, T-DNA segregation was not possible, probably due to a high copy number of T-DNA. Therefore, for this line we had to genotype and phenotype T1 plants without T-DNA segregation, assuming that no other mutations were produced in the genome.

Up to now, our primary objective of complete flowering abolition was not achieved. Knockouts of other genes known as floral inducers, such as *TWIN SISTER OF FT (TSF)* or *MOTHER OF FT AND TFL1 (MFT)* (Yamaguchi *et al*., 2005; S. Y. Yoo *et al*., 2004) could be considered in order to reach this final objective. However, the editing of *NbSPL13_1a*, *NbFT5* homeologous pair together with *NbFT4* allowed us to obtain *N. benthamiana* lines with a considerably delayed flowering time, almost doubling the duration of the pre-flowering phase in the wild type. This was reflected on various aspects of plant growth, that was phenotyped through different parameters. For all of them, the mean value increased as more genes were edited.

The most evident effect of the delayed flowering was the increase in biomass at flowering time, that was measured in terms of weight but also in terms of number of lateral branches. Previous studies already focused on increasing biomass for heterologous protein production. Light intensity and photoperiod length are the main factors positively impacting leaf biomass (Dorais & Gosselin, 2002). In accordance to this, Stevens *et al*. (2000) showed that high-light conditions increased biomass and in turn had a positive influence on the production of mouse antibody MGR48 in tobacco plants (Stevens *et al*., 2000). Shang *et al*. (2018) reported that the combination of CO_2 enrichment and LED inter-lighting increased leaf biomass in *N. benthamiana*, which consequently increased influenza virus hemagglutinin H1 expression (Shang *et al*., 2018). This can be explained by the shift of the carboxylation-oxygenation equilibrium of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) towards CO_2 fixation. Various studies reported that CO_2 enrichment and increased light enhance photosynthetic rate and in

turn the biomass of leaves (Drake *et al*., 1997; Kitaya *et al*., 1998, 2003). Goulet *et al*. (2019) showed that in *N. benthamiana* the primary contributors to the yield of agroinfiltrated recombinant proteins are the young leaves of the axillary stems. The spraying of the cytokinin 6-benzylaminopurine (6-BAP) increased branching, and prolonging the photoperiod from 16 to 24 hours resulted in a higher biomass and a considerably higher titre of influenza virus hemagglutinin H1 (Goulet *et al*., 2019). The supply of phyto-hormones, the prolongation of photoperiod or the increasing of CO_2 concentration involve consumption of additional economic resources. Therefore, the generation of *N. benthamiana* lines that already possess an increased branching and biomass is desirable for molecular farming purposes. In addition, Gao *et al*. (2018) demonstrated that the silencing of *SPL13* in alfalfa caused a delay in flowering time and an increased number of lateral branches (R. Gao *et al*., 2018). Therefore, the editing of the homolog of this gene in *N. benthamiana* could be a valid knockout target for our purpose (in fact it had similar effects on our lines). Another approach that was used to obtain increased branching in rapeseed was the knockout through gene editing of the two genes *BnaMAX1*, homologs of *Arabidopsis* gene *MAX1*. Knocked-out lines showed increased branching, that resulted in increased seed productivity (Zheng *et al*., 2020). Therefore, the editing of the homologs of these genes in *N. benthamiana* could be considered to obtain an effect on branching even more pronounced than the one observed on our lines.

The increase in pre-flower biomass per planting cycle is with no doubt an advantage derived from delayed flowering. Less effort was required to obtain equivalent biomass with FT4 FT5 SPL13 40-1 plants compared to wild type plants. It was shown that mutant lines accumulate almost six times more biomass in terms of fresh weigh (from 138.4 g to 813.2 g on average, see fig. 15F) in only twice the time (from 46 days to 94.5 days on average, see fig. 15A). This also reduces the planting efforts, saving seeds, pots and substrate. However, a drawback of this aspect is that the time in which plants are exposed to diseases is also extended, increasing the risks of batch failure. This is particularly problematic for *N. benthamiana*, a species with high susceptibility to viral infections. Therefore, precautions must be maximized to avoid contamination of these *N. benthamiana* lines with pathogens. It should be noted that all the considerations discussed above apply for plant batches harvested before the flowering stage, which is the usual practice in *N. benthamiana* biofactory approaches. It is commonly accepted that productivity is highly reduced when plants reach the reproductive phase. However, comparisons with flowering plants of the same age should be also performed.

After phenotyping, the recombinant protein expression capacity of the mutated lines was evaluated at two different time points. We expected that delay in flowering could be accompanied by the maintenance of similar capacity for recombinant protein expression at early (week five, the usual infiltration time for WT plants) and late (week nine) stages, but this result was not observed in any of the vectors employed for expressing eGFP. In the case of FT4 FT5 SPL13 40-1 plants, the replicative system based in TMV vector was significantly more productive at the later stage than in the rest of genotypes when infiltrated in low density plants, indicating that, at least in this mutant line, some of the juvenile expression capacity

was retained at nine-weeks old plants. However, this effect could not be reproduced when plants were grown at high density. The age-associated decrease in the expression potential of *N. benthamiana* has been related to the ending of juvenility after flowering (Sheludko *et al*., 2007). Apparently, the editing of *NbFT5* homeologous pair and *NbFT4* was clearly not sufficient to maintain juvenility and therefore the expression capability of *N. benthamiana.* Only the concomitant presence of *NbSPL13_1a* seems to have a positive effect in this direction.

As a future perspective, the edition of additional genes of the SPL family could prolong leaf juvenility, helping to maintain the expression capability of *N. benthamiana* even after week five-six. This task could be tackled by editing genes such as *NbSPL9/15* or *NbSPL3/4/5*, or other floral inducers, such as *TWIN SISTER OF FT (TSF)* or *MOTHER OF FT AND TFL1 (MFT)* (Yamaguchi *et al*., 2005; S. Y. Yoo *et al*., 2004). Extending juvenility-associated productivity next to flower initiation would be of extreme utility in plants used for molecular farming, since they would considerably increase their biomass while maintaining their expression potential for longer times, thus resulting in a considerable increase in heterologous protein production.

Although the decrease in the expression capacity observed in nine-weeks old non-flowered plants limits the applicability of the flowering delay approach, still the remarkable increase in biomass could be sufficient to provide economic advantages to the use of these new lines in molecular farming, both for transient expression as well as in stable transformation approaches. Furthermore, the potential of CRISPR/Cas as a powerful tool for editing genes was demonstrated: despite sgRNAs with low efficiencies, especially in the case of the ones having *NbFT4* as target, all the desired mutations were obtained in different combinations.

Supplementary material

Figure S4: Phylogenetic tree of *FT* genes from tobacco and *N. benthamiana*.
The tree was constructed using the maximum-likelihood method with a bootstrap value of n = 1000 iterations. Bootstrap support values are shown at the nodes. The genes ID are the ones from *N. benthamiana* genome version H2020. *NbFT4* is indicated in violet and *NbFT5_1a/1b* are indicated in blue.

sgRNA name	Targeted genes and position	Protospacer sequence	PAM
sgSPL1.5	NbSPL13_1a-nc134	GGACCTCACAAACTTTATGG	CGG
sgSPL1.6	NbSPL13_1a-c288	ATGGACATAACAGGCGTCGA	AGG
sgFT5	NbFT5_1a-nc182/NbFT5_1b-nc182	CCAGAGTGTAAAAGGTACGA	CGG
sgFT6	NbFT5_1a-nc257	AGGACGATACTAACCAATGG	AGG
sgFT6.1	NbFT5_1b-nc257	AAAGAGATGCTAACCAATGG	AGG
sgFT4.1	NbFT4-nc60	AGGTCAACAGATCTTGTGAA	TGG
sgFT4.2	NbFT4-nc132	GGTTGCTTAACAATTTGAGA	AGG

Table S6. Guide RNA sequences used for editing. sgRNA targeted positions is defined as the distance of the Cas9 cutting site from the ATG in their coding sequences. "c" indicates that the gRNA is designed on the coding strand and "nc" that gRNA is designed on the non-coding strand.

	NbFT5_1a		NbFT5_1b		NbSPL13_1a sgSPL1.5		NbSPL13_1a sgSPL1.6	
Plant	**% editing**	**KO score**	**% editing**	**KO score**	**% editing**	**KO score**	**% editing**	**KO score**
FT5 SPL13 1	99	46	4	3	0	0	14	14
FT5 SPL13 2	99	99	0	0	0	0	0	0
FT5 SPL13 5	84	79	0	0	0	0	0	0
FT5 SPL13 6A	87	83	16	16	0	0	29	29
FT5 SPL13 6B	100	100	33	33	0	0	failed	failed
FT5 SPL13 6C	98	56	0	0	0	0	14	14
FT5 SPL13 14A	99	99	97	4	0	0	42	33
FT5 SPL13 14B	99	99	0	0	0	0	failed	failed
FT5 SPL13 18A	99	99	0	0	0	0	95	68
FT5 SPL13 18B	99	99	0	0	0	0	95	66
FT5 SPL13 20	0	0	0	0	0	0	0	0
FT5 SPL13 22A	100	100	50	50	0	0	46	46
FT5 SPL13 22B	99	99	62	62	0	0	47	47
FT5 SPL13 22C	failed		50	50	0	0	49	49
FT5 SPL13 22D	99	99	64	64	0	0	46	46
FT5 SPL13 22E	100	100	50	50	0	0	99	99
FT5 SPL13 22F	99	99	44	44	0	0	98	98
FT5 SPL13 22G	100	100	41	41	0	0	99	99
FT5 SPL13 22H	100	100	49	49	0	0	52	52
FT5 SPL13 23	98	98	39	39	0	0	46	46
FT5 SPL13 25	99	99	31	31	0	0	31	28
FT5 SPL13 28	37	37	0	0	0	0	0	0
FT5 SPL13 32	failed		failed		0	0	0	0
FT5 A1	34	34	0	0				
FT5 A2	100	100	0	0				

FT5 C	99	99	0	0				
FT5 E1	100	100	0	0				
FT5 E2	99	99	48	48				
FT5 F	99	99	4	4				
FT5 R	99	99	2	2				
FT5 S1	18	18	0	0				
FT5 S2	97	94	0	0				
SPL13 1A					0	0	48	48
SPL13 3					0	0	92	92
SPL13 6					20	0	19	19
SPL13 7					5	4	0	0
SPL13 8					0	0	53	53
SPL13 9					failed	failed	34	22
SPL13 10					0	0	15	15
SPL13 11					0	0	48	48

Table S7. Genotype of T0 lines from FT5 SPL13, FT5 and SPL13 transformations. Plants from the performed transformations are listed. For each plant the editing efficiency values and the knockout values reported by Synthego are reported. The sgRNA used for editing are listed on the top of the table. High editing values are highlighted in green, low editing values are highlighted in yellow. The most edited lines that were selected to generate subsequent T1 generation are highlighted in orange. “Failed” indicates that the sequencing of a target genomic region failed.

	NbFT4	
Plant	**% editing**	**KO score**
FT4 FT5 1	0	0
FT4 FT5 2A	14	13
FT4 FT5 2B	1	1
FT4 FT5 3	0	0
FT4 FT5 4	46	46
FT4 FT5 6	0	0
FT4 FT5 7A	0	0
FT4 FT5 8A	0	0
FT4 FT5 5	0	0
FT4 FT5 9	0	0
FT4 FT5 13	0	0
FT4 FT5 14	0	0
FT4 FT5 15	0	0
FT4 FT5 16	0	0
FT4 FT5 17	0	0
FT4 FT5 18	0	0
FT4 FT5 19A	0	0
FT4 FT5 SPL13 1	0	0
FT4 FT5 SPL13 2	0	0
FT4 FT5 SPL13 5	0	0
FT4 FT5 SPL13 9	0	0
FT4 FT5 SPL13 36	0	0
FT4 FT5 SPL13 40	33	33
FT4 FT5 SPL13 44A	0	0
FT4 FT5 SPL13 44B	0	0
FT4 FT5 SPL13 49	0	0
FT4 FT5 SPL13 51	0	0
FT4 FT5 SPL13 61	0	0
FT4 FT5 SPL13 34	0	0
FT4 FT5 SPL13 43	0	0
FT4 FT5 SPL13 60	50	50
FT4 FT5 SPL13 62	54	54
FT4 FT5 SPL13 63	0	0

Table S8. Genotype of T0 lines from FT4 transformations. Plants from the performed transformations are listed. For each plant the editing efficiency values and the knockout values reported by Synthego are reported. High edition values are highlighted in green, low edition values are highlighted in yellow. The most edited lines that were selected to generate subsequent T1 generation are highlighted in orange.

Primer name	Sequence 5' > 3'
CDP20Apr01_NbFT5_1a_FW	GCCAAGAGAACGTGAACCTCTA
CDP20Apr02_NbFT5_1a_RV	CGGGAAAAGGTCGGATCACA
CDP20Apr03_NbFT5_1b_FW	GCCAAGAGAACGTGAACCTCTG
CDP20Apr04_NbFT5_1b_RV	CAAAATTCGTTAGTATCGAGACTGGC
CDP20Apr05_NbSPL13_1a_g1FW	CTTTGTTACTTCGCAATTAGAGCG
CDP20Apr06_NbSPL13_1a_g1RV	GTGGCGTACTAAGGGTCAAGT
CDP20Apr07_NbSPL13_1a_g2FW	AAATGTTCAATCCCTGGACGAC
CDP20Apr08_NbSPL13_1a_g2RV	ACCATGTCGCTGTCCGTTTTG
CDP21Apr01_FT4 FW3	CTGGCATCAACCAGAATCGGA
CDP21Apr02_FT4 RV3	CAGGTTAGGGTTGCTTGGGC

Table S9. List of primers used for amplification of the targeted sites.

GB ID	Description
GB0307	pUPD2
GB1205	tRNA and scaffold for the assembly of GBoligomers for the first position (positon [D1_2]) of a polycistronic tRNA-gRNA regulated by the U6-26 or U6-1 promoter
GB1206	tRNA and scaffold for the assembly of GBoligomers for the intermediate position (positon [2_n-1]) of a polycistronic tRNA-gRNA
GB1207	tRNA and scaffold for the assembly of GBoligomers for the last position (positon [n]) of a polycistronic tRNA-gRNA
GB1208	tRNA and scaffold for the assembly of GBoligomers for the first position (positon [D1_n-1]) of a polycistronic tRNA-gRNA regulated by the U6-26 or U6-1 promoter
GB1001	U6-26 promoter
GB2630	pUPD2 with sgSPL1.5
GB2631	pUPD2 with sgSPL1.6
GB0017	pDGB3_alpha2
GB0019	pDGB3_omega1
GB2234	Module for the constitutive expression of the nptII, Cas9 and DsRed genes in pDGB3_alpha1
GB3296	U6-26:sgSPL1.5:sgSPL1.6 in alpha2
GB3298	nptII:Cas9:DsRed_U6-26:sgSPL1.5:sgSPL1.6 in omega 1
GB3293	sgFT5 in pUPD2
GB3294	sgFT6.1 in pUPD2
GB2625	sgFT6 in pUPD2
GB3297	U6-26:sgFT5:sgFT6.1:sgFT6 in alpha2
GB3301	nptII:Cas:DsRed-U6-26:sgSPL1.5:sgSPL1.6-U6-26:sgFT5:sgFT6.1:sgFT6 in alpha 1
GB3696	sgFT4.1 in pUPD2
GB3697	sgFT4.2 in pUPD2
GB3698	U6-26:sgFT4.2:sgFT4.1 in alpha 2
GB3699	nptII:Cas9:DsRed-U6-26:sgFT4.2:sgFT4.1 in omega 1

Table S10. Goldenbraid plasmids used in this work.

General discussion

The objective of this thesis was the improvement of *N. benthamiana* as a chassis for molecular farming. First, a genome-wide analysis of SPL genes was performed in this species and its close relative *N. tabacum*, identifying 49 *SPL* genes in *N. tabacum* cv. K326 and 43 *SPL* genes in *N. benthamiana* LAB strain. *SPL* genes of the two species were clustered into eight phylogenetic groups according to the *SPL* classification of Arabidopsis. Highly conserved exon-intron gene structure and SBP domains were found in homeologues and orthologues. Thirty of the *NbSPL* genes and 33 of the *NtSPL* genes were found to be targets of microRNA156. RNA-seq analysed the expression of *SPL* genes in leaves at three different stages, revealing that genes constitutively expressed at high levels were generally not under miR156 control, whereas miR156-regulated genes showed lower expression levels, often developmentally regulated. Based on this new data, the *N. benthamiana SPL13_1a* gene was selected as target for a CRISPR/Cas9 KO experiment. The full knock out of this single gene lead to a significant delay in flowering time of 4-5 days and increased branching. Other than *NbSPL13_1a* lines, more CRISPR/Cas9 knock outs were performed in *N. benthamiana* with the objective of obtaining flowering abolition. Floral activators *NbFT5_1a/1b* were knocked out alone and in combination with *NbSPL13_1a*. The biallelic mutations in these genes did not confer flowering abolition, but also in this case a moderate delay of approximately 10 days was obtained. For this reason, the knockout of *NbFT4*, another floral inducer, was stacked onto FT5 and FT5 SPL13 backgrounds. In the most edited line FT4 FT5 SPL13 40-1 flowering time was doubled in comparison to WT plants, but also in this case flowering abolition was not achieved. The delay of 4-5 days in flowering time of *NbSPL13* edited plants shown in chapter 1 is discrepant with the 2 days of delay of the same line shown in chapter 2. This could be due to differences in growth conditions, since the comparison of *NbSPL13* edited plants against WT plants alone was conducted in growth chamber, with a minimum temperature of 20 °C , a maximum of 24 °C, and 16 hours of light per day, while the general phenotyping of all lines (also the ones edited in *NbFT4* and *NbFT5*) was performed in greenhouse with a minimum temperature of 18 °C, a maximum of 21.5 °C, and 14 hours of light per day. The delayed flowering had consequences on various aspects of plant growth, that were quantified through various parameters: the lines that possessed more edited genes had increased biomass, height, number of leaves and total leaves area compared to the ones edited in fewer genes and WT at flowering time. Moreover, the generated lines were assayed for their potential of expression of heterologous proteins. They were infiltrated with different viral constructs harbouring the coding sequence of enhanced GFP (eGFP) at five and nine weeks after sowing. We expected that they could be capable of maintaining high expression levels even after week five (the usual infiltration time of wild type plants), but this result was not observed. Expression levels dropped at week nine for every line, with FT4 FT5 SPL13 line displaying the least decrease. Future work with the same objective of flowering abolition in *N. benthamiana* include stacking our lines with knockouts of other genes, such as *NbSPL9/15* and *NbSPL3/4/5*, which are other important players in juvenile-adult transition and flower initiation (Jung

et al., 2012). We hope that the edition of this genes could also subsequently maintain high expression levels of heterologous proteins for longer periods of time than the usual infiltration window of five-six weeks.

Delayed flowering also has its drawbacks. One of them is the increased possibility of catching diseases before flowering, which could endanger production. If a plant showed signs of pathogen infection, it should be immediately removed to avoid the spreading of the disease to the rest of the batch and save future production. Moreover, the number of cycles (generations) per year are reduced in late flowering varieties, which means lower seeds production (for propagation means) as compared to WT plants. The problems arising from delayed flowering would also impact stable transgenics strategies, and not just plant lines destined for transient expression. Nevertheless, the increased biomass of late flowering lines can be a desirable feature when considering stable transgenics: recombinant genes stably transformed into late flowering lines could guarantee protein production at much higher scale.

Genome editing, for the generation of our plant lines, is a very efficient technique. Nowadays the production of the desired mutations relies on the use of *in vitro* tissue cultures. However, this method has disadvantages, such as the significant time and equipment required, and its limited applicability to certain species (Altpeter *et al.*, 2016). To bypass tissue culture, there have been efforts to achieve transformation by directly introducing transgenes to meristems or egg cells. Nonetheless, the success of this approach was limited to *Arabidopsis thaliana* and its close relatives (Clough & Bent, 1998; Hamada *et al.*, 2017). Therefore, more efficient methods for editing plants are being developed. Up to now, various groups have used a technique called virus-induced genome editing (VIGE), that relies on viruses to deliver sgRNAs to transgenic plants already expressing Cas9. VIGE often results in high editing frequencies in somatic cells but also in low recovery of mutant progeny (Ali *et al.*, 2018; Cody *et al.*, 2017; Q. Gao *et al.*, 2019). Higher mutation frequencies can be obtained if sgRNAs have better access to the germline. For this purpose, *FT* gene can be helpful, since its transcription takes place in leaf vascular tissue, and subsequently, its mRNA is transported to the shoot apical meristem (SAM) to trigger the flowering process (Jackson & Hong, 2012; Notaguchi *et al.*, 2015). Ellison *et al.* (2020) suggested that the fusion of *FT* mRNA to a sgRNA could facilitate its entry to the SAM, potentially leading to heritable genetic mutations. Therefore, the *A. thaliana FT* coding sequence was fused to the 3′-end of the sgRNAs targeting the genes of interest. The FT sequence joined to the sgRNAs was assembled in a tobacco rattle virus (TRV) vector and introduced through agroinfiltration into a transgenic *N. benthamiana* plant expressing Cas9. The recovered progeny was mutated at high frequencies ranging from 65 to 100% (Ellison *et al.*, 2020). Therefore, efficient gene editing can be achieved without the need of performing *in vitro* tissue culture: a plant already expressing Cas9 can be infected with a viral vector containing the desired sgRNAs fused to *FT* coding sequence, that will move to the SAM, creating mutations in the progeny.

This method was also employed by our group for generating plants edited in *NbFT4*, as an alternative to the traditional *in vitro* transformation method. Relatively high editing efficiencies in *NbFT4* were recovered in somatic tissues few days after infiltration, ranging from 50 to 70% of indels. Unfortunately, in our hands these mutations were not inherited to the progeny of the infected plants, therefore this approach was abandoned. Despite the great advantage of avoiding *in vitro* transformation, VIGE is bound to the limit of few sgRNAs that can be cloned in the viral vector, limiting the possibility of editing large gene families, such as *SPL* genes. Therefore, this strategy cannot completely substitute *in vitro* transformation, but depending on the desired target it could be a valid alternative.

The lines described in this thesis would surely have to face the present legal situation about genetically modified organisms (GMOs) for their commercial use, that could have very different outcomes depending on the countries where they would be adopted. For over two decades, biotech crops, genetically engineered for various traits, have been grown. Between 1996 and 2018, global economic benefits derived from biotech crops reached USD 224.9 billion, benefiting nearly 16–17 million farmers, especially in developing countries (Marone *et al*., 2023). The USA continue to lead the commercialization and development of genetically modified (GM) crops. The global cultivation of GM crops raised from 1.7 million hectares to 190.4 million hectares between 1996 and 2019, spanning 29 countries. Key players in GM crop cultivation include the USA, Canada, Argentina, Brazil and India. Notably, China and India stand out as the world's greatest cotton producers, with 95% of their cotton being of Bt cotton varieties (Marone *et al*., 2023). The USA is leader in the number of approved GM crops, followed by Japan, Canada, Brazil, and South Korea. Maize shows the highest count of approved events, with herbicide tolerance and insect resistance being the most prevalent traits (Marone *et al*., 2023).

Despite their impact on global agriculture in the last two decades, GMOs are still controversial, and their commercialization still face important drawbacks, particularly in Europe. With the emergence of genome editing, new scenarios on the regulation of this particular type of GMOs are emerging. The perspectives for development and cultivation of edited crops depend on international regulations and whether they are subject to the same strict laws about GMOs (Jenkins *et al*., 2021). Only a few genome-edited crop traits have been approved for commercialization so far, with herbicide-tolerant canola being one of the earliest (Marone *et al*., 2023). Some genome-edited crops, like high-oleic soybean oil and nutritionally enhanced tomatoes, are already on the market in the USA, Canada, and Japan (Demorest *et al*., 2016; Nonaka *et al*., 2017; Waltz, 2021). In 2022, Argentina ruled that specific genome-edited *Camelina* varieties are exempt from pre-market authorization. This decision was based on the absence of foreign DNA in genome-edited lines and their similarity to conventional varieties of Camelina (Marone *et al*., 2023). The regulatory landscape for transgenic and genome-edited crops is different across countries, influencing aspects such as research, cultivation, commercialization and utilization as food and feed.

The European Union has the most rigid regulation for GMO presence in food and feed. It follows a precautionary principle, which may cause problems for products with indistinguishable mutations imported from countries without regulations (Wolt *et al*., 2016). The EU court has established that crop varieties coming from new site-directed mutagenesis techniques are GMOs, but a new legal framework would be needed to regulate products derived from novel genomic techniques (Marone *et al*., 2023). The scientific community is actively working to persuade the public and policymakers that genome-edited plants are similar to the varieties obtained by conventional breeding methods or classical mutagenesis. Recently, a new proposal has been made by the European Commission to modify the existing regulatory framework about GMOs (Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on Plants Obtained by Certain New Genomic Techniques and Their Food and Feed, and Amending Regulation (EU) 2017/625, 2023). According to this proposal, biotech varieties obtained by new genomic techniques (NGT) would be classified in three categories:

- "Category 1 NGT plants" would be the ones that could have arisen naturally or that could have been generated by conventional breeding techniques. In this category would fall all varieties resulting from targeted mutagenesis or cisgenesis (introduction of genetic material from a sexually compatible species), that comply with certain criteria (e.g. not exceeding a maximum number of modifications). They will have to comply with the rest of the European standards that apply to any new variety to guarantee food safety and environmental protection; a public register of these products will be created, and information will be included in the labelling of seeds and plant propagating material, as well as in commercial registers and catalogues.
- "Category 2 NGT plants" would be other NGT varieties not meeting the criteria of the previous category but containing traits that could contribute to the overall performance of varieties as regards sustainability (e.g. drought tolerance or resilience to climate change). Category 2 NGT plants and products would remain subject to traceability and labelling requirements in the Union's GMO legislation. However, the authorization requirements for these varieties containing sustainable traits will be substantially reduced. For instance, monitoring plan for environmental effects should not be required if the category 2 NGT plant is unlikely to pose risks that need monitoring, such as indirect, delayed or unforeseen effects on human health or on the environment. Moreover, the possibility for Member States to restrict or prohibit cultivation in their territory will not apply to such category 2 NGT plants.
- All remaining NGT plants (not falling under the two previous categories), would be considered as GMOs and evaluated following the current GMO regulation.

In this new regulatory context, our CRISPRed *N. benthamiana* varieties would fall into the first category. Hopefully, CRISPR mutations in these lines would be considered as mutations that could also occur naturally, such as the frameshift mutation that the LAB strain of *N. benthamiana* already possess in its

RDR1 gene, making it more susceptible to infections (S.-J. Yang *et al*., 2004). This would be a considerable advantage since they would not be subject to the present GMO regulation. Consequently, the CRISPR-improved biofactory lines would not require expensive "GMO-ready" contained facilities for cultivation prior to Agroinfiltration. This could reduce costs in the handling and multiplication of plants during upstream operations and facilitate the upscale bioproduction.

Nevertheless, the impact of CRISPR deregulation on *N. benthamiana* protein biofactories would be relatively mild, since they normally require indoor (GMO contained) growth conditions, minimally after the infiltration phase. On the contrary, the new proposal from the European Commission would have a greater impact for large-scale molecular farming platforms, such as genome edited tobacco. Recently, through the EU-funded project Newcotiana (www.newcotiana.org), open field trials of tobacco plants with increased biomass as a result of edited *FT* and *SPL* genes have been carried out in Extremadura, Spain. Moreover, some of these tobacco lines were further CRISPR-engineered to accumulate high levels of anatabine, an alkaloid metabolite useful in the pharmaceutical field (Ruiz Castro *et al*., 2020). The new proposal about NGT plants in the EU would surely facilitate the development of new large-scale biofactory varieties such as non-flowering, high-biomass anatabine-rich tobacco varieties, and others to come in the future.

Conclusions

1. A genome-wide analysis of the *SPL* gene family in representative species of the *Nicotiana* genus identified 49 members in *Nicotiana tabacum* and 43 members in *Nicotiana benthamiana*. The whole *SPL* gene collection was clustered in eight subfamilies, which exhibited a conserved exon-intron structure between the two species. Other features as the presence of microRNA 156 target sites, as well as the temporal expression patterns, were also conserved between subfamilies in the two plant species.

2. The knockout of the *N. benthamiana NbSPL13_1a* gene through CRISPR/Cas9 led to a small delay in flowering time of four-five days, and an increase in branching with respect to WT plants.

3. The combined knockout of the floral activators *NbFT5_1a* and *NbFT5_1b* in *N. benthamiana* produced a considerable delay in flowering time of approximately 10 days. When combined with the mutation of the *NbSPL13_1a* gene, the flowering time was only extended one extra day on average.

4. The stacking of knockout mutations in all three *N. benthamiana* flowering activators of the FT family, namely *NbFT5_1a*, *NbFT5_1b* and *NbFT4,* led to a delay of approximately 36 days in flowering time in the growth conditions assayed.

5. *N. benthamiana* plants with knockout mutations in *NbFT5_1a, NbFT5_1b* and *NbFT4* in combination with *NbSPL13_1a,* led to maximum delays of flowering time of 48 days in the conditions assayed. At flowering time, quadruple mutants showed also increased branching and eight times more biomass in terms of fresh weigh with respect to WT.

6. Minor differences were observed in the recombinant eGFP production capacity of quadruple mutant lines when compared with WT plants employing Tobacco Mosaic Virus-based replicative systems, although such differences need to be substantiated at a larger production scale.

Bibliography

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., & Araki, T. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science (New York, N.Y.)*, *309*(5737), 1052-1056. https://doi.org/10.1126/science.1115983

Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., Shmakov, S., Makarova, K. S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E. S., Koonin, E. V., & Zhang, F. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science (New York, N.Y.)*, *353*(6299), aaf5573. https://doi.org/10.1126/science.aaf5573

Alam, A., Jiang, L., Kittleson, G. A., Steadman, K. D., Nandi, S., Fuqua, J. L., Palmer, K. E., Tusé, D., & McDonald, K. A. (2018). Technoeconomic Modeling of Plant-Based Griffithsin Manufacturing. *Frontiers in Bioengineering and Biotechnology*, *6*. https://www.frontiersin.org/articles/10.3389/fbioe.2018.00102

Ali, Z., Eid, A., Ali, S., & Mahfouz, M. M. (2018). Pea early-browning virus-mediated genome editing via the CRISPR/Cas9 system in Nicotiana benthamiana and Arabidopsis. *Virus Research*, *244*, 333-337. https://doi.org/10.1016/j.virusres.2017.10.009

Alkanaimsh, S., Corbin, J. M., Kailemia, M. J., Karuppanan, K., Rodriguez, R. L., Lebrilla, C. B., McDonald, K. A., & Nandi, S. (2019). Purification and site-specific N-glycosylation analysis of human recombinant butyrylcholinesterase from Nicotiana benthamiana. *Biochemical Engineering Journal*, *142*, 58-67. https://doi.org/10.1016/j.bej.2018.11.004

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., & Stewart, C. N., Jr. (2016). Advancing Crop Transformation in the Era of Genome Editing. *The Plant Cell*, *28*(7), 1510-1520. https://doi.org/10.1105/tpc.16.00196

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

Arcalis, E., Ibl, V., Peters, J., Melnik, S., & Stoger, E. (2014). The dynamic behavior of storage organelles in developing cereal seeds and its impact on the production of recombinant proteins. *Frontiers in Plant Science*, *5*. https://www.frontiersin.org/articles/10.3389/fpls.2014.00439

Arshad, M., Feyissa, B. A., Amyot, L., Aung, B., & Hannoufa, A. (2017). MicroRNA156 improves drought stress tolerance in alfalfa (Medicago sativa) by silencing SPL13. *Plant Science*, *258*, 122-136. https://doi.org/10.1016/j.plantsci.2017.01.018

Arzola, L., Chen, J., Rattanaporn, K., Maclean, J. M., & McDonald, K. A. (2011). Transient co-expression of post-transcriptional gene silencing suppressors for increased in planta expression of a recombinant anthrax receptor fusion protein. *International Journal of Molecular Sciences*, *12*(8), 4975-4990. https://doi.org/10.3390/ijms12084975

Balanzà, V., Martínez-Fernández, I., & Ferrándiz, C. (2014). Sequential action of FRUITFULL as a modulator of the activity of the floral regulators SVP and SOC1. *Journal of Experimental Botany*, *65*(4), 1193-1203. https://doi.org/10.1093/jxb/ert482

Bally, J., Jung, H., Mortimer, C., Naim, F., Philips, J. G., Hellens, R., Bombarely, A., Goodin, M. M., & Waterhouse, P. M. (2018). The Rise and Rise of *Nicotiana benthamiana*: A Plant for All Reasons. *Annual Review of Phytopathology*, *56*(1), 405-426. https://doi.org/10.1146/annurev-phyto-080417-050141

Bally, J., Nakasugi, K., Jia, F., Jung, H., Ho, S. Y. W., Wong, M., Paul, C. M., Naim, F., Wood, C. C., Crowhurst, R. N., Hellens, R. P., Dale, J. L., & Waterhouse, P. M. (2015). The extremophile Nicotiana benthamiana has traded viral defence for early vigour. *Nature Plants*, *1*(11), Article 11. https://doi.org/10.1038/nplants.2015.165

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., & Nekrasov, V. (2013). Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods*, *9*(1), 39. https://doi.org/10.1186/1746-4811-9-39

Bendandi, M., Marillonnet, S., Kandzia, R., Thieme, F., Nickstadt, A., Herz, S., Fröde, R., Inogés, S., Lòpez-Dìaz de Cerio, A., Soria, E., Villanueva, H., Vancanneyt, G., McCormick, A., Tusé, D., Lenz, J., Butler-Ransohoff, J.-E., Klimyuk, V., & Gleba, Y. (2010). Rapid, high-yield production in plants of individualized idiotype vaccines for non-Hodgkin's lymphoma. *Annals of Oncology*, *21*(12), 2420-2427. https://doi.org/10.1093/annonc/mdq256

Benlloch, R., Berbel, A., Serrano-Mislata, A., & Madueño, F. (2007). Floral Initiation and Inflorescence Architecture: A Comparative View. *Annals of Botany*, *100*(3), 659-676. https://doi.org/10.1093/aob/mcm146

Birkenbihl, R. P., Jach, G., Saedler, H., & Huijser, P. (2005). Functional dissection of the plant-specific SBP-domain: Overlap of the DNA-binding and nuclear localization domains. *Journal of Molecular Biology*, *352*(3), 585-596. https://doi.org/10.1016/j.jmb.2005.07.013

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Bornscheuer, U., Buchholz, K., & Seibel, J. (2014). Enzymatic Degradation of (Ligno)cellulose. *Angewandte Chemie International Edition*, *53*(41), 10876-10893. https://doi.org/10.1002/anie.201309953

Bortesi, L., & Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, *33*(1), 41-52. https://doi.org/10.1016/j.biotechadv.2014.12.006

Bortesi, L., Zhu, C., Zischewski, J., Perez, L., Bassié, L., Nadi, R., Forni, G., Lade, S. B., Soto, E., Jin, X., Medina, V., Villorbina, G., Muñoz, P., Farré, G., Fischer, R., Twyman, R. M., Capell, T., Christou, P., & Schillberg, S. (2016). Patterns of CRISPR/Cas9 activity in plants, animals and microbes. *Plant Biotechnology Journal*, *14*(12), 2203-2216. https://doi.org/10.1111/pbi.12634

Breyer, D., De Schrijver, A., Goossens, M., Pauwels, K., & Herman, P. (2012). Biosafety of Molecular Farming in Genetically Modified Plants. En A. Wang & S. Ma (Eds.), *Molecular Farming in Plants: Recent Advances and Future Prospects* (pp. 259-274). Springer Netherlands. https://doi.org/10.1007/978-94-007-2217-0_12

Buyel, J. F. (2019). Plant Molecular Farming – Integration and Exploitation of Side Streams to Achieve Sustainable Biomanufacturing. *Frontiers in Plant Science*, *9*, 1893. https://doi.org/10.3389/fpls.2018.01893

Buyel, J. F., Hubbuch, J., & Fischer, R. (2016). Comparison of Tobacco Host Cell Protein Removal Methods by Blanching Intact Plants or by Heat Treatment of Extracts. *JoVE (Journal of Visualized Experiments)*, *114*, e54343. https://doi.org/10.3791/54343

Buyel, J. F., Stöger, E., & Bortesi, L. (2021). Targeted genome editing of plants and plant cells for biomanufacturing. *Transgenic Research*, *30*(4), 401-426. https://doi.org/10.1007/s11248-021-00236-z

Buyel, J. F., Twyman, R. M., & Fischer, R. (2015). Extraction and downstream processing of plant-derived recombinant proteins. *Biotechnology Advances*, *33*(6, Part 1), 902-913. https://doi.org/10.1016/j.biotechadv.2015.04.010

Cai, C., Guo, W., & Zhang, B. (2018). Genome-wide identification and characterization of SPL transcription factor family and their evolution and expression profiling analysis in cotton. *Scientific Reports*, *8*(1), 762. https://doi.org/10.1038/s41598-017-18673-4

Cao, K., Cui, L., Zhou, X., Ye, L., Zou, Z., & Deng, S. (2016). Four Tomato FLOWERING LOCUS T-Like Proteins Act Antagonistically to Regulate Floral Initiation. *Frontiers in Plant Science*, *6*. https://www.frontiersin.org/articles/10.3389/fpls.2015.01213

Cardon, G., Höhmann, S., Klein, J., Nettesheim, K., Saedler, H., & Huijser, P. (1999). Molecular characterisation of the Arabidopsis SBP-box genes. *Gene*, *237*(1), 91-104. https://doi.org/10.1016/s0378-1119(99)00308-x

Chardon, F., & Damerval, C. (2005). Phylogenomic analysis of the PEBP gene family in cereals. *Journal of Molecular Evolution*, *61*(5), 579-590. https://doi.org/10.1007/s00239-004-0179-4

Chase, M. W., Knapp, S., Cox, A. V., Clarkson, J. J., Butsko, Y., Joseph, J., Savolainen, V., & Parokonny, A. S. (2003). Molecular systematics, GISH and the origin of hybrid taxa in Nicotiana (Solanaceae). *Annals of Botany*, *92*(1), 107-127. https://doi.org/10.1093/aob/mcg087

Clarkson, J. J., Knapp, S., Garcia, V. F., Olmstead, R. G., Leitch, A. R., & Chase, M. W. (2004). Phylogenetic relationships in Nicotiana (Solanaceae) inferred from multiple plastid DNA regions. *Molecular Phylogenetics and Evolution*, *33*(1), 75-90. https://doi.org/10.1016/j.ympev.2004.05.002

Clough, S. J., & Bent, A. F. (1998). Floral dip: A simplified method for Agrobacterium -mediated transformation of Arabidopsis thaliana. *The Plant Journal*, *16*(6), 735-743. https://doi.org/10.1046/j.1365-313x.1998.00343.x

Cody, W. B., Scholthof, H. B., & Mirkov, T. E. (2017). Multiplexed Gene Editing and Protein Overexpression Using a Tobacco mosaic virus Viral Vector. *Plant Physiology*, *175*(1), 23-35. https://doi.org/10.1104/pp.17.00411

Commandeur, U., Twyman, R. M., & Fischer, R. (2003). The biosafety of molecular farming in plants. *AgBiotechNet*, *5*(ABN 110), 1-9.

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., & Coupland, G. (2007). FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. *Science*, *316*(5827), 1030-1033. https://doi.org/10.1126/science.1141752

Craddock, C. P., Adams, N., Bryant, F. M., Kurup, S., & Eastmond, P. J. (2015). Regulation of endomembrane biogenesis in arabidopsis by phospatidic acid hydrolase. *Plant Signaling & Behavior*, *10*(10), e1065367. https://doi.org/10.1080/15592324.2015.1065367

Dahan-Meir, T., Filler-Hayut, S., Melamed-Bessudo, C., Bocobza, S., Czosnek, H., Aharoni, A., & Levy, A. A. (2018). Efficient in planta gene targeting in tomato using geminiviral replicons and the CRISPR/Cas9 system. *The Plant Journal*, *95*(1), 5-16. https://doi.org/10.1111/tpj.13932

Dai, X., Zhuang, Z., & Zhao, P. X. (2018). psRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Research*, *46*(W1), W49-W54. https://doi.org/10.1093/nar/gky316

Damasceno, L. M., Anderson, K. A., Ritter, G., Cregg, J. M., Old, L. J., & Batt, C. A. (2007). Cooverexpression of chaperones for enhanced secretion of a single-chain antibody fragment in Pichia pastoris. *Applied Microbiology and Biotechnology*, *74*(2), 381-389. https://doi.org/10.1007/s00253-006-0652-7

D'Aoust, M.-A., Couture, M. M.-J., Charland, N., Trépanier, S., Landry, N., Ors, F., & Vézina, L.-P. (2010). The production of hemagglutinin-based virus-like particles in plants: A rapid, efficient and safe response to pandemic influenza. *Plant Biotechnology Journal*, *8*(5), 607-619. https://doi.org/10.1111/j.1467-7652.2009.00496.x

D'Aoust, M.-A., Lavoie, P.-O., Couture, M. M.-J., Trépanier, S., Guay, J.-M., Dargis, M., Mongrand, S., Landry, N., Ward, B. J., & Vézina, L.-P. (2008). Influenza virus-like particles produced by transient expression in Nicotiana benthamiana induce a protective immune response against a lethal viral challenge in mice. *Plant Biotechnology Journal*, *6*(9), 930-940. https://doi.org/10.1111/j.1467-7652.2008.00384.x

Darko, E., Heydarizadeh, P., Schoefs, B., & Sabzalian, M. R. (2014). Photosynthesis under artificial light: The shift in primary and secondary metabolism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1640), 20130243. https://doi.org/10.1098/rstb.2013.0243

De Paola, C., Garcia-Carpintero, V., Vazquez-Vilar, M., Kaminski, K., Fernandez-del-Carmen, A., Sierro, N., Ivanov, N. V., Giuliano, G., Waterhouse, P., & Orzaez, D. (2023). Comparative analysis of the Squamosa Promoter Binding-Like (SPL) gene family in Nicotiana benthamiana and Nicotiana tabacum. *Plant Science*, *335*, 111797. https://doi.org/10.1016/j.plantsci.2023.111797

de Ruijter, J. C., Koskela, E. V., & Frey, A. D. (2016). Enhancing antibody folding and secretion by tailoring the Saccharomyces cerevisiae endoplasmic reticulum. *Microbial Cell Factories*, *15*(1), 87. https://doi.org/10.1186/s12934-016-0488-5

Demorest, Z. L., Coffman, A., Baltes, N. J., Stoddard, T. J., Clasen, B. M., Luo, S., Retterath, A., Yabandith, A., Gamo, M. E., Bissen, J., Mathis, L., Voytas, D. F., & Zhang, F. (2016). Direct stacking of sequence-specific nuclease-induced mutations to produce high oleic and low linolenic soybean oil. *BMC Plant Biology*, *16*(1), 225. https://doi.org/10.1186/s12870-016-0906-1

Derevnina, L., Kamoun, S., & Wu, C.-H. (2019). Dude, where is my mutant? Nicotiana benthamiana meets forward genetics. *The New Phytologist*, *221*(2), 607-610. https://doi.org/10.1111/nph.15521

Diamos, A. G., Hunter, J. G. L., Pardhe, M. D., Rosenthal, S. H., Sun, H., Foster, B. C., DiPalma, M. P., Chen, Q., & Mason, H. S. (2020). High Level Production of Monoclonal Antibodies Using an Optimized Plant Expression System. *Frontiers in Bioengineering and Biotechnology*, *7*, 472. https://doi.org/10.3389/fbioe.2019.00472

Diego-Martin, B., González, B., Vazquez-Vilar, M., Selma, S., Mateos-Fernández, R., Gianoglio, S., Fernández-del-Carmen, A., & Orzáez, D. (2020). Pilot Production of SARS-CoV-2 Related Proteins in Plants: A Proof of Concept for Rapid Repurposing of Indoor Farms Into Biomanufacturing Facilities. *Frontiers in Plant Science*, *11*. https://www.frontiersin.org/articles/10.3389/fpls.2020.612781

Donini, M., Lombardi, R., Lonoce, C., Di Carli, M., Marusic, C., Morea, V., & Di Micco, P. (2015). Antibody proteolysis: A common picture emerging from plants. *Bioengineered*, *6*(5), 299-302. https://doi.org/10.1080/21655979.2015.1067740

Dorais, M., & Gosselin, A. (2002). PHYSIOLOGICAL RESPONSE OF GREENHOUSE VEGETABLE CROPS TO SUPPLEMENTAL LIGHTING. *Acta Horticulturae*, *580*, 59-67. https://doi.org/10.17660/ActaHortic.2002.580.6

Doudna, J. A., & Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science (New York, N.Y.)*, *346*(6213), 1258096. https://doi.org/10.1126/science.1258096

Drake, B. G., Gonzalez-Meler, M. A., & Long, S. P. (1997). MORE EFFICIENT PLANTS: A Consequence of Rising Atmospheric CO2? *Annual Review of Plant Physiology and Plant Molecular Biology*, *48*, 609-639. https://doi.org/10.1146/annurev.arplant.48.1.609

Drapal, M., Enfissi, E. M. A., & Fraser, P. D. (2021a). Metabolic changes in leaves of N. tabacum and N. benthamiana during plant development. *Journal of Plant Physiology*, *265*, 153486. https://doi.org/10.1016/j.jplph.2021.153486

Drapal, M., Enfissi, E. M. A., & Fraser, P. D. (2021b). Metabolic effects of agro-infiltration on N. benthamiana accessions. *Transgenic Research*, *30*(3), 303-315. https://doi.org/10.1007/s11248-021-00256-9

Drapal, M., Enfissi, E. M. A., & Fraser, P. D. (2022). The chemotype core collection of genus Nicotiana. *The Plant Journal*, *110*(5), 1516-1528. https://doi.org/10.1111/tpj.15745

Duwadi, K., Chen, L., Menassa, R., & Dhaubhadel, S. (2015). Identification, Characterization and Down-Regulation of Cysteine Protease Genes in Tobacco for Use in Recombinant Protein Production. *PLOS ONE*, *10*(7), e0130556. https://doi.org/10.1371/journal.pone.0130556

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792-1797. https://doi.org/10.1093/nar/gkh340

Edwards, K. D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A. D., Bombarely, A., Allen, F., Hurst, R., White, B., Kernodle, S. P., Bromley, J. R., Sanchez-Tamburrino, J. P., Lewis, R. S., & Mueller, L. A. (2017). A reference genome for Nicotiana tabacum enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics*, *18*(1), 448. https://doi.org/10.1186/s12864-017-3791-6

Eidenberger, L., Kogelmann, B., & Steinkellner, H. (2023). Plant-based biopharmaceutical engineering. *Nature Reviews Bioengineering*, *1*(6), 1-14. https://doi.org/10.1038/s44222-023-00044-6

Ellison, E. E., Nagalakshmi, U., Gamo, M. E., Huang, P., Dinesh-Kumar, S., & Voytas, D. F. (2020). Multiplexed heritable gene editing using RNA viruses and mobile single guide RNAs. *Nature Plants*, *6*(6), Article 6. https://doi.org/10.1038/s41477-020-0670-y

Enfissi, E. M. A., Fraser, P. D., Lois, L.-M., Boronat, A., Schuch, W., & Bramley, P. M. (2005). Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. *Plant Biotechnology Journal*, *3*(1), 17-27. https://doi.org/10.1111/j.1467-7652.2004.00091.x

Faure, S., Higgins, J., Turner, A., & Laurie, D. A. (2007). The FLOWERING LOCUS T-Like Gene Family in Barley (Hordeum vulgare). *Genetics*, *176*(1), 599-609. https://doi.org/10.1534/genetics.106.069500

Feeney, M., Frigerio, L., Cui, Y., & Menassa, R. (2013). Following Vegetative to Embryonic Cellular Changes in Leaves of Arabidopsis Overexpressing LEAFY COTYLEDON2. *Plant Physiology*, *162*(4), 1881-1896. https://doi.org/10.1104/pp.113.220996

Feng, S., Xu, Y., Guo, C., Zheng, J., Zhou, B., Zhang, Y., Ding, Y., Zhang, L., Zhu, Z., Wang, H., & Wu, G. (2016). Modulation of miR156 to identify traits associated with vegetative phase change in tobacco (Nicotiana tabacum). *Journal of Experimental Botany*, *67*(5), 1493-1504. https://doi.org/10.1093/jxb/erv551

Ferreira e Silva, G. F., Silva, E. M., Azevedo, M. da S., Guivin, M. A. C., Ramiro, D. A., Figueiredo, C. R., Carrer, H., Peres, L. E. P., & Nogueira, F. T. S. (2014). microRNA156-targeted SPL/SBP box transcription factors regulate tomato ovary and fruit development. *The Plant Journal: For Cell and Molecular Biology*, *78*(4), 604-618. https://doi.org/10.1111/tpj.12493

Feyissa, B. A., Arshad, M., Gruber, M. Y., Kohalmi, S. E., & Hannoufa, A. (2019). The interplay between miR156/SPL13 and DFR/WD40-1 regulate drought tolerance in alfalfa. *BMC Plant Biology*, *19*(1), 434. https://doi.org/10.1186/s12870-019-2059-5

Fischer, R., & Buyel, J. F. (2020). Molecular farming – The slope of enlightenment. *Biotechnology Advances*, *40*, 107519. https://doi.org/10.1016/j.biotechadv.2020.107519

Fischer, R., Holland, T., Sack, M., Schillberg, S., Stoger, E., Twyman, R. M., & Buyel, J. F. (2021). Glyco-Engineering of Plant-Based Expression Systems. En E. Rapp & U. Reichl (Eds.), *Advances in Glycobiotechnology* (pp. 137-166). Springer International Publishing. https://doi.org/10.1007/10_2018_76

Forestier, E. C. F., Cording, A. C., Loake, G. J., & Graham, I. A. (2023). An Engineered Heat-Inducible Expression System for the Production of Casbene in Nicotiana benthamiana. *International Journal of Molecular Sciences*, *24*(14), Article 14. https://doi.org/10.3390/ijms241411425

Forestier, E. C. F., Czechowski, T., Cording, A. C., Gilday, A. D., King, A. J., Brown, G. D., & Graham, I. A. (2021). Developing a Nicotiana benthamiana transgenic platform for high-value diterpene production and candidate gene evaluation. *Plant Biotechnology Journal*, *19*(8), 1614-1623. https://doi.org/10.1111/pbi.13574

Fu, H., Machado, P. A., Hahm, T. S., Kratochvil, R. J., Wei, C. I., & Lo, Y. M. (2010). Recovery of nicotine-free proteins from tobacco leaves using phosphate buffer system under controlled conditions. *Bioresource Technology*, *101*(6), 2034-2042. https://doi.org/10.1016/j.biortech.2009.10.045

Gandikota, M., Birkenbihl, R. P., Höhmann, S., Cardon, G. H., Saedler, H., & Huijser, P. (2007). The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *The Plant Journal: For Cell and Molecular Biology*, *49*(4), 683-693. https://doi.org/10.1111/j.1365-313X.2006.02983.x

Gao, Q., Xu, W.-Y., Yan, T., Fang, X.-D., Cao, Q., Zhang, Z.-J., Ding, Z.-H., Wang, Y., & Wang, X.-B. (2019). Rescue of a plant cytorhabdovirus as versatile expression platforms for planthopper and cereal genomic studies. *The New Phytologist*, *223*(4), 2120-2133. https://doi.org/10.1111/nph.15889

Gao, R., Gruber, M. Y., Amyot, L., & Hannoufa, A. (2018). SPL13 regulates shoot branching and flowering time in Medicago sativa. *Plant Molecular Biology*, *96*(1-2), 119-133. https://doi.org/10.1007/s11103-017-0683-8

Garabagi, F., Gilbert, E., Loos, A., McLean, M. D., & Hall, J. C. (2012). Utility of the P19 suppressor of gene-silencing protein for production of therapeutic antibodies in Nicotiana expression hosts. *Plant Biotechnology Journal*, *10*(9), 1118-1128. https://doi.org/10.1111/j.1467-7652.2012.00742.x

Gielen, H., Remans, T., Vangronsveld, J., & Cuypers, A. (2016). Toxicity responses of Cu and Cd: The involvement of miRNAs and the transcription factor SPL7. *BMC Plant Biology*, *16*(1), 145. https://doi.org/10.1186/s12870-016-0830-4

Giritch, A., Marillonnet, S., Engler, C., van Eldik, G., Botterman, J., Klimyuk, V., & Gleba, Y. (2006). Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. *Proceedings of the National Academy of Sciences*, *103*(40), 14701-14706. https://doi.org/10.1073/pnas.0606631103

Głowacka, K., Kromdijk, J., Kucera, K., Xie, J., Cavanagh, A. P., Leonelli, L., Leakey, A. D. B., Ort, D. R., Niyogi, K. K., & Long, S. P. (2018). Photosystem II Subunit S overexpression increases the efficiency of water use in a field-grown crop. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-03231-x

González, M. N., Massa, G. A., Andersson, M., Turesson, H., Olsson, N., Fält, A.-S., Storani, L., Décima Oneto, C. A., Hofvander, P., & Feingold, S. E. (2020). Reduced Enzymatic Browning in Potato

Tubers by Specific Editing of a Polyphenol Oxidase Gene via Ribonucleoprotein Complexes Delivery of the CRISPR/Cas9 System. *Frontiers in Plant Science*, *10*. https://www.frontiersin.org/articles/10.3389/fpls.2019.01649

Goodin, M. M., Zaitlin, D., Naidu, R. A., & Lommel, S. A. (2008). Nicotiana benthamiana: Its history and future as a model for plant-pathogen interactions. *Molecular Plant-Microbe Interactions: MPMI*, *21*(8), 1015-1026. https://doi.org/10.1094/MPMI-21-8-1015

Göritzer, K., Grandits, M., Grünwald-Gruber, C., Figl, R., Mercx, S., Navarre, C., Ma, J. K.-C., & Teh, A. Y.-H. (2022). Engineering the N-glycosylation pathway of Nicotiana tabacum for molecular pharming using CRISPR/Cas9. *Frontiers in Plant Science*, *13*. https://www.frontiersin.org/articles/10.3389/fpls.2022.1003065

Goulet, M.-C., Gaudreau, L., Gagné, M., Maltais, A.-M., Laliberté, A.-C., Éthier, G., Bechtold, N., Martel, M., D'Aoust, M.-A., Gosselin, A., Pepin, S., & Michaud, D. (2019). Production of Biopharmaceuticals in Nicotiana benthamiana—Axillary Stem Growth as a Key Determinant of Total Protein Yield. *Frontiers in Plant Science*, *10*. https://www.frontiersin.org/articles/10.3389/fpls.2019.00735

Green, C., & Tibbetts, C. (1980). Targeted deletions of sequences from closed circular DNA. *Proceedings of the National Academy of Sciences*, *77*(5), 2455-2459. https://doi.org/10.1073/pnas.77.5.2455

Grosse-Holz, F., Kelly, S., Blaskowski, S., Kaschani, F., Kaiser, M., & van der Hoorn, R. A. L. (2017). The transcriptome, extracellular proteome and active secretome of agroinfiltrated Nicotiana benthamiana uncover a large, diverse protease repertoire. *Plant Biotechnology Journal*, *16*(5), 1068-1084. https://doi.org/10.1111/pbi.12852

Grosse-Holz, F., Madeira, L., Zahid, M. A., Songer, M., Kourelis, J., Fesenko, M., Ninck, S., Kaschani, F., Kaiser, M., & van der Hoorn, R. A. L. (2018). Three unrelated protease inhibitors enhance accumulation of pharmaceutical recombinant proteins in Nicotiana benthamiana. *Plant Biotechnology Journal*, *16*(10), 1797-1810. https://doi.org/10.1111/pbi.12916

Hamada, H., Linghu, Q., Nagira, Y., Miki, R., Taoka, N., & Imai, R. (2017). An in planta biolistic method for stable wheat transformation. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-11936-0

Hamorsky, K. T., Grooms-Williams, T. W., Husk, A. S., Bennett, L. J., Palmer, K. E., & Matoba, N. (2013). Efficient Single Tobamoviral Vector-Based Bioproduction of Broadly Neutralizing Anti-HIV-1 Monoclonal Antibody VRC01 in Nicotiana benthamiana Plants and Utility of VRC01 in Combination Microbicides. *Antimicrobial Agents and Chemotherapy*, *57*(5), 2076-2086. https://doi.org/10.1128/aac.02588-12

Han, Y.-Y., Ma, Y.-Q., Li, D.-Z., Yao, J.-W., & Xu, Z.-Q. (2016). Characterization and phylogenetic analysis of fifteen NtabSPL genes in Nicotiana tabacum L. cv. Qinyan95. *Development Genes and Evolution*, *226*(1), 1-14. https://doi.org/10.1007/s00427-015-0522-3

Hanania, U., Ariel, T., Tekoah, Y., Fux, L., Sheva, M., Gubbay, Y., Weiss, M., Oz, D., Azulay, Y., Turbovski, A., Forster, Y., & Shaaltiel, Y. (2017). Establishment of a tobacco BY2 cell line devoid of plant-specific xylose and fucose as a platform for the production of biotherapeutic proteins. *Plant Biotechnology Journal*, *15*(9), 1120-1129. https://doi.org/10.1111/pbi.12702

Hanano, S., & Goto, K. (2011). Arabidopsis TERMINAL FLOWER1 Is Involved in the Regulation of Flowering Time and Inflorescence Development through Transcriptional Repression. *The Plant Cell*, *23*(9), 3172-3184. https://doi.org/10.1105/tpc.111.088641

Harig, L., Beinecke, F. A., Oltmanns, J., Muth, J., Müller, O., Rüping, B., Twyman, R. M., Fischer, R., Prüfer, D., & Noll, G. A. (2012). Proteins from the FLOWERING LOCUS T-like subclade of the PEBP family act antagonistically to regulate floral initiation in tobacco. *The Plant Journal*, *72*(6), 908-921. https://doi.org/10.1111/j.1365-313X.2012.05125.x

Hasan, Md. M., Kim, H.-S., Jeon, J.-H., Kim, S. H., Moon, B., Song, J.-Y., Shim, S. H., & Baek, K.-H. (2014). Metabolic engineering of Nicotiana benthamiana for the increased production of taxadiene. *Plant Cell Reports*, *33*(6), 895-904. https://doi.org/10.1007/s00299-014-1568-9

Hayashi, S., Watanabe, M., Kobayashi, M., Tohge, T., Hashimoto, T., & Shoji, T. (2020). Genetic Manipulation of Transcriptional Regulators Alters Nicotine Biosynthesis in Tobacco. *Plant and Cell Physiology*, *61*(6), 1041-1053. https://doi.org/10.1093/pcp/pcaa036

Hedman, H., Källman, T., & Lagercrantz, U. (2009). Early evolution of the MFT-like gene family in plants. *Plant Molecular Biology*, *70*(4), 359-369. https://doi.org/10.1007/s11103-009-9478-x

Herman, E. M., & Larkins, B. A. (1999). Protein Storage Bodies and Vacuoles. *The Plant Cell*, *11*(4), 601-613. https://doi.org/10.1105/tpc.11.4.601

Herpen, T. W. J. M. van, Cankar, K., Nogueira, M., Bosch, D., Bouwmeester, H. J., & Beekwilder, J. (2010). Nicotiana benthamiana as a Production Platform for Artemisinin Precursors. *PLOS ONE*, *5*(12), e14222. https://doi.org/10.1371/journal.pone.0014222

Hiatt, A., Caffferkey, R., & Bowdish, K. (1989). Production of antibodies in transgenic plants. *Nature*, *342*(6245), Article 6245. https://doi.org/10.1038/342076a0

Hoernstein, S. N. W., Fode, B., Wiedemann, G., Lang, D., Niederkrüger, H., Berg, B., Schaaf, A., Frischmuth, T., Schlosser, A., Decker, E. L., & Reski, R. (2018). Host Cell Proteome of Physcomitrella patens Harbors Proteases and Protease Inhibitors under Bioproduction Conditions. *Journal of Proteome Research*, *17*(11), 3749-3760. https://doi.org/10.1021/acs.jproteome.8b00423

Hood, E. E., Kusnadi, A., Nikolov, Z., & Howard, J. A. (1999). Molecular farming of industrial proteins from transgenic maize. *Advances in Experimental Medicine and Biology*, *464*, 127-147. https://doi.org/10.1007/978-1-4615-4729-7_11

Horsch, R. B., Fry, J. E., Hoffmann, N. L., Wallroth, M., Eichholtz, D., Rogers, S. G., & Fraley, R. T. (1985). A Simple and General Method for Transferring Genes into Plants. *Science*, *227*(4691), 1229-1231. https://doi.org/10.1126/science.227.4691.1229

Hundleby, P. A. C., Sack, M., & Twyman, R. M. (2018). Biosafety, Risk Assessment, and Regulation of Molecular Farming. En *Molecular Pharming* (pp. 327-351). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118801512.ch13

Izawa, T., Oikawa, T., Sugiyama, N., Tanisaka, T., Yano, M., & Shimamoto, K. (2002). Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. *Genes & Development*, *16*(15), 2006-2020. https://doi.org/10.1101/gad.999202

Jackson, S., & Hong, Y. (2012). Systemic movement of FT mRNA and a possible role in floral induction. *Frontiers in Plant Science*, *3*. https://www.frontiersin.org/articles/10.3389/fpls.2012.00127

Jansing, J., Sack, M., Augustine, S. M., Fischer, R., & Bortesi, L. (2019). CRISPR/Cas9-mediated knockout of six glycosyltransferase genes in Nicotiana benthamiana for the production of recombinant proteins lacking β-1,2-xylose and core α-1,3-fucose. *Plant Biotechnology Journal*, *17*(2), 350-361. https://doi.org/10.1111/pbi.12981

Jenkins, D., Dobert, R., Atanassova, A., & Pavely, C. (2021). Impacts of the regulatory environment for gene editing on delivering beneficial products. *In Vitro Cellular & Developmental Biology. Plant: Journal of the Tissue Culture Association*, *57*(4), 609-626. https://doi.org/10.1007/s11627-021-10201-4

Jiang, M., He, Y., Chen, X., Zhang, X., Guo, Y., Yang, S., Huang, J., & Traw, M. B. (2020). CRISPR-based assessment of genomic structure in the conserved SQUAMOSA promoter-binding-like gene clusters in rice. *The Plant Journal*, *104*(5), 1301-1314. https://doi.org/10.1111/tpj.15001

Jiao, Y., Wang, Y., Xue, D., Wang, J., Yan, M., Liu, G., Dong, G., Zeng, D., Lu, Z., Zhu, X., Qian, Q., & Li, J. (2010). Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nature Genetics*, *42*(6), 541-544. https://doi.org/10.1038/ng.591

Jung, J.-H., Ju, Y., Seo, P. J., Lee, J.-H., & Park, C.-M. (2012). The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in Arabidopsis. *The Plant Journal*, *69*(4), 577-588. https://doi.org/10.1111/j.1365-313X.2011.04813.x

Jutras, P. V., Marusic, C., Lonoce, C., Deflers, C., Goulet, M.-C., Benvenuto, E., Michaud, D., & Donini, M. (2016). An Accessory Protease Inhibitor to Increase the Yield and Quality of a Tumour-Targeting mAb in Nicotiana benthamiana Leaves. *PLOS ONE*, *11*(11), e0167086. https://doi.org/10.1371/journal.pone.0167086

Kaiser, A., Sell, S., & Hehl, R. (2002). Heterologous Expression of a Bacterial Homospermidine Synthase Gene in Transgenic Tobacco: Effects on the Polyamine Pathway. *Archiv Der Pharmazie*, *335*(4), 143-151. https://doi.org/10.1002/1521-4184(200204)335:4<143::AID-ARDP143>3.0.CO;2-B

Karlgren, A., Gyllenstrand, N., Källman, T., Sundström, J. F., Moore, D., Lascoux, M., & Lagercrantz, U. (2011). Evolution of the PEBP gene family in plants: Functional diversification in seed plant evolution. *Plant Physiology*, *156*(4), 1967-1977. https://doi.org/10.1104/pp.111.176206

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, *20*(4), 1160-1166. https://doi.org/10.1093/bib/bbx108

Khlystov, N. A., Yoshikuni, Y., Deutsch, S., & Sattely, E. S. (2021). A plant host, Nicotiana benthamiana, enables the production and study of fungal lignin-degrading enzymes. *Communications Biology*, *4*(1), Article 1. https://doi.org/10.1038/s42003-021-02464-9

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907-915. https://doi.org/10.1038/s41587-019-0201-4

Kim, N.-S., Kim, T.-G., Kim, O.-H., Ko, E.-M., Jang, Y.-S., Jung, E.-S., Kwon, T.-H., & Yang, M.-S. (2008). Improvement of recombinant hGM-CSF production by suppression of cysteine proteinase gene expression using RNA interference in a transgenic rice culture. *Plant Molecular Biology*, *68*(3), 263-275. https://doi.org/10.1007/s11103-008-9367-8

Kitaya, Y., Niu, G., Kozai, T., & Ohashi, M. (1998). Photosynthetic Photon Flux, Photoperiod, and CO2 Concentration Affect Growth and Morphology of Lettuce Plug Transplants. *HortScience*, *33*(6), 988-991. https://doi.org/10.21273/HORTSCI.33.6.988

Kitaya, Y., Okayama, T., Murakami, K., & Takeuchi, T. (2003). Effects of CO2 concentration and light intensity on photosynthesis of a rootless submerged plant, Ceratophyllum demersum L., used for aquatic food production in bioregenerative life support systems. *Advances in Space Research: The Official Journal of the Committee on Space Research (COSPAR)*, *31*(7), 1743-1749. https://doi.org/10.1016/s0273-1177(03)00113-3

Klabunde, J., Kleebank, S., Piontek, M., Hollenberg, C. P., Hellwig, S., & Degelmann, A. (2007). Increase of calnexin gene dosage boosts the secretion of heterologous proteins by Hansenula polymorpha. *FEMS Yeast Research*, *7*(7), 1168-1180. https://doi.org/10.1111/j.1567-1364.2007.00271.x

Klein, J., Saedler, H., & Huijser, P. (1996). A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA. *Molecular & General Genetics: MGG*, *250*(1), 7-16. https://doi.org/10.1007/BF02191820

Knapp, S., Bohs, L., Nee, M., & Spooner, D. M. (2004). Solanaceae—A model for linking genomics with biodiversity. *Comparative and Functional Genomics*, *5*(3), 285-291. https://doi.org/10.1002/cfg.393

Kromdijk, J., Głowacka, K., Leonelli, L., Gabilly, S. T., Iwai, M., Niyogi, K. K., & Long, S. P. (2016). Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science*, *354*(6314), 857-861. https://doi.org/10.1126/science.aai8878

Kunert, R., & Reinhart, D. (2016). Advances in recombinant antibody manufacturing. *Applied Microbiology and Biotechnology*, *100*(8), 3451-3461. https://doi.org/10.1007/s00253-016-7388-9

Kwon, C.-T., Heo, J., Lemmon, Z. H., Capua, Y., Hutton, S. F., Van Eck, J., Park, S. J., & Lippman, Z. B. (2020). Rapid customization of Solanaceae fruit crops for urban agriculture. *Nature Biotechnology*, *38*(2), 182-188. https://doi.org/10.1038/s41587-019-0361-2

Li, A., Jia, S., Yobi, A., Ge, Z., Sato, S. J., Zhang, C., Angelovici, R., Clemente, T. E., & Holding, D. R. (2018). Editing of an Alpha-Kafirin Gene Family Increases, Digestibility and Protein Quality in Sorghum. *Plant Physiology*, *177*(4), 1425-1438. https://doi.org/10.1104/pp.18.00200

Li, C., & Lu, S. (2014). Molecular characterization of the SPL gene family in Populus trichocarpa. *BMC Plant Biology*, *14*, 131. https://doi.org/10.1186/1471-2229-14-131

Li, J.-F., Zhang, D., & Sheen, J. (2014). Chapter Twenty-One—Cas9-Based Genome Editing in Arabidopsis and Tobacco. En J. A. Doudna & E. J. Sontheimer (Eds.), *Methods in Enzymology* (Vol. 546, pp. 459-472). Academic Press. https://doi.org/10.1016/B978-0-12-801185-0.00022-2

Li, M., Li, X., Zhou, Z., Wu, P., Fang, M., Pan, X., Lin, Q., Luo, W., Wu, G., & Li, H. (2016). Reassessment of the Four Yield-related Genes Gn1a, DEP1, GS3, and IPA1 in Rice Using a CRISPR/Cas9 System. *Frontiers in Plant Science*, *7*. https://www.frontiersin.org/articles/10.3389/fpls.2016.00377

Limkul, J., Misaki, R., Kato, K., & Fujiyama, K. (2015). The combination of plant translational enhancers and terminator increase the expression of human glucocerebrosidase in Nicotiana benthamiana plants. *Plant Science*, *240*, 41-49. https://doi.org/10.1016/j.plantsci.2015.08.018

Liu, M., Sun, W., Ma, Z., Huang, L., Wu, Q., Tang, Z., Bu, T., Li, C., & Chen, H. (2019). Genome-wide identification of the SPL gene family in Tartary Buckwheat (Fagopyrum tataricum) and expression analysis during fruit development stages. *BMC Plant Biology*, *19*(1), 299. https://doi.org/10.1186/s12870-019-1916-6

Liu, W., & Stewart, C. N. (2016). Plant synthetic promoters and transcription factors. *Current Opinion in Biotechnology*, *37*, 36-44. https://doi.org/10.1016/j.copbio.2015.10.001

Long, S. P., Marshall-Colon, A., & Zhu, X.-G. (2015). Meeting the Global Food Demand of the Future by Engineering Crop Photosynthesis and Yield Potential. *Cell*, *161*(1), 56-66. https://doi.org/10.1016/j.cell.2015.03.019

Lowder, L. G., Paul, J. W., & Qi, Y. (2017). Multiplexed Transcriptional Activation or Repression in Plants Using CRISPR-dCas9-Based Systems. En K. Kaufmann & B. Mueller-Roeber (Eds.), *Plant Gene Regulatory Networks: Methods and Protocols* (pp. 167-184). Springer. https://doi.org/10.1007/978-1-4939-7125-1_12

Ludman, M., Burgyán, J., & Fátyol, K. (2017). Crispr/Cas9 Mediated Inactivation of Argonaute 2 Reveals its Differential Involvement in Antiviral Responses. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-01050-6

Lyzenga, W. J., Harrington, M., Bekkaoui, D., Wigness, M., Hegedus, D. D., & Rozwadowski, K. L. (2019). CRISPR/Cas9 editing of three CRUCIFERIN C homoeologues alters the seed protein profile in Camelina sativa. *BMC Plant Biology*, *19*(1), 292. https://doi.org/10.1186/s12870-019-1873-0

Ma, J. K.-C., Barros, E., Bock, R., Christou, P., Dale, P. J., Dix, P. J., Fischer, R., Irwin, J., Mahoney, R., Pezzotti, M., Schillberg, S., Sparrow, P., Stoger, E., & Twyman, R. M. (2005). Molecular farming for new drugs and vaccines. *EMBO reports*, *6*(7), 593-599. https://doi.org/10.1038/sj.embor.7400470

Ma, J. K.-C., Drossard, J., Lewis, D., Altmann, F., Boyle, J., Christou, P., Cole, T., Dale, P., van Dolleweerd, C. J., Isitt, V., Katinger, D., Lobedan, M., Mertens, H., Paul, M. J., Rademacher, T., Sack, M., Hundleby, P. A. C., Stiegler, G., Stoger, E., … Fischer, R. (2015). Regulatory approval and a first-in-human phase I clinical trial of a monoclonal antibody produced in transgenic tobacco plants. *Plant Biotechnology Journal*, *13*(8), 1106-1120. https://doi.org/10.1111/pbi.12416

Magnusdottir, A., Vidarsson, H., Björnsson, J. M., & Örvar, B. L. (2013). Barley grains for the production of endotoxin-free growth factors. *Trends in Biotechnology*, *31*(10), 572-580. https://doi.org/10.1016/j.tibtech.2013.06.002

Mahalik, S., Sharma, A. K., & Mukherjee, K. J. (2014). Genome engineering for improved recombinant protein expression in Escherichia coli. *Microbial Cell Factories*, *13*(1), 177. https://doi.org/10.1186/s12934-014-0177-1

Maharjan, P. M., & Choe, S. (2021). Plant-Based COVID-19 Vaccines: Current Status, Design, and Development Strategies of Candidate Vaccines. *Vaccines*, *9*(9), Article 9. https://doi.org/10.3390/vaccines9090992

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., … Koonin, E. V. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology*, *13*(11), Article 11. https://doi.org/10.1038/nrmicro3569

Mandal, M. K., Ahvari, H., Schillberg, S., & Schiermeyer, A. (2016). Tackling Unwanted Proteolysis in Plant Production Hosts Used for Molecular Farming. *Frontiers in Plant Science*, *7*. https://www.frontiersin.org/articles/10.3389/fpls.2016.00267

Mandal, M. K., Fischer, R., Schillberg, S., & Schiermeyer, A. (2014). Inhibition of protease activity by antisense RNA improves recombinant protein production in Nicotiana tabacum cv. Bright

Yellow 2 (BY-2) suspension cells. *Biotechnology Journal*, *9*(8), 1065-1073. https://doi.org/10.1002/biot.201300424

Mao, H.-D., Yu, L.-J., Li, Z.-J., Yan, Y., Han, R., Liu, H., & Ma, M. (2016). Genome-wide analysis of the SPL family transcription factors and their responses to abiotic stresses in maize. *Plant Gene*, *6*, 1-12. https://doi.org/10.1016/j.plgene.2016.03.003

Margolin, E., Oh, Y. J., Verbeek, M., Naude, J., Ponndorf, D., Meshcheriakova, Y. A., Peyret, H., van Diepen, M. T., Chapman, R., Meyers, A. E., Lomonossoff, G. P., Matoba, N., Williamson, A.-L., & Rybicki, E. P. (2020). Co-expression of human calreticulin significantly improves the production of HIV gp140 and other viral glycoproteins in plants. *Plant Biotechnology Journal*, *18*(10), 2109-2117. https://doi.org/10.1111/pbi.13369

Marin Viegas, V. S., Ocampo, C. G., & Petruccelli, S. (2017). Vacuolar deposition of recombinant proteins in plant vegetative organs as a strategy to increase yields. *Bioengineered*, *8*(3), 203-211. https://doi.org/10.1080/21655979.2016.1222994

Marone, D., Mastrangelo, A. M., & Borrelli, G. M. (2023). From Transgenesis to Genome Editing in Crop Improvement: Applications, Marketing, and Legal Issues. *International Journal of Molecular Sciences*, *24*(8), Article 8. https://doi.org/10.3390/ijms24087122

Marty, F. (1999). Plant Vacuoles. *The Plant Cell*, *11*(4), 587-599. https://doi.org/10.1105/tpc.11.4.587

Mason, H. S., Ball, J. M., Shi, J. J., Jiang, X., Estes, M. K., & Arntzen, C. J. (1996). Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice. *Proceedings of the National Academy of Sciences*, *93*(11), 5335-5340. https://doi.org/10.1073/pnas.93.11.5335

Mateos-Fernández, R., Moreno-Giménez, E., Gianoglio, S., Quijano-Rubio, A., Gavaldá-García, J., Estellés, L., Rubert, A., Rambla, J. L., Vazquez-Vilar, M., Huet, E., Fernández-del-Carmen, A., Espinosa-Ruiz, A., Juteršek, M., Vacas, S., Navarro, I., Navarro-Llopis, V., Primo, J., & Orzáez, D. (2021). Production of Volatile Moth Sex Pheromones in Transgenic Nicotiana benthamiana Plants. *BioDesign Research*, *2021*, 9891082. https://doi.org/10.34133/2021/9891082

Matsuo, K., & Atsumi, G. (2019). CRISPR/Cas9-mediated knockout of the RDR6 gene in Nicotiana benthamiana for efficient transient expression of recombinant proteins. *Planta*, *250*(2), 463-473. https://doi.org/10.1007/s00425-019-03180-9

McNulty, M. J., Gleba, Y., Tusé, D., Hahn-Löbmann, S., Giritch, A., Nandi, S., & McDonald, K. A. (2020). Techno-economic analysis of a plant-based platform for manufacturing antimicrobial proteins for food safety. *Biotechnology Progress*, *36*(1), e2896. https://doi.org/10.1002/btpr.2896

Meng, X., Muszynski, M. G., & Danilevskaya, O. N. (2011). The FT-like ZCN8 Gene Functions as a Floral Activator and Is Involved in Photoperiod Sensitivity in Maize. *The Plant Cell*, *23*(3), 942-960. https://doi.org/10.1105/tpc.110.081406

Mercx, S., Smargiasso, N., Chaumont, F., De Pauw, E., Boutry, M., & Navarre, C. (2017). Inactivation of the β(1,2)-xylosyltransferase and the α(1,3)-fucosyltransferase genes in Nicotiana tabacum BY-2 Cells by a Multiplex CRISPR/Cas9 Strategy Results in Glycoproteins without Plant-Specific Glycans. *Frontiers in Plant Science*, *8*. https://www.frontiersin.org/articles/10.3389/fpls.2017.00403

Mimida, N., Goto, K., Kobayashi, Y., Araki, T., Ahn, J. H., Weigel, D., Murata, M., Motoyoshi, F., & Sakamoto, W. (2001). Functional divergence of the TFL1-like gene family in Arabidopsis revealed by characterization of a novel homologue. *Genes to Cells*, *6*(4), 327-336. https://doi.org/10.1046/j.1365-2443.2001.00425.x

Molina-Hidalgo, F. J., Vazquez-Vilar, M., D'Andrea, L., Demurtas, O. C., Fraser, P., Giuliano, G., Bock, R., Orzáez, D., & Goossens, A. (2021). Engineering Metabolism in Nicotiana Species: A Promising Future. *Trends in Biotechnology*, *39*(9), 901-913. https://doi.org/10.1016/j.tibtech.2020.11.012

Montero-Morales, L., & Steinkellner, H. (2018). Advanced Plant-Based Glycan Engineering. *Frontiers in Bioengineering and Biotechnology*, *6*, 81. https://doi.org/10.3389/fbioe.2018.00081

Moreno, P., Fexova, S., George, N., Manning, J. R., Miao, Z., Mohammed, S., Muñoz-Pomer, A., Fullgrabe, A., Bi, Y., Bush, N., Iqbal, H., Kumbham, U., Solovyev, A., Zhao, L., Prakash, A., García-Seisdedos, D., Kundu, D. J., Wang, S., Walzer, M., … Papatheodorou, I. (2022). Expression Atlas update: Gene

and protein expression in multiple species. *Nucleic Acids Research*, *50*(D1), D129-D140. https://doi.org/10.1093/nar/gkab1030

Moreno-Giménez, E., Selma, S., Calvache, C., & Orzáez, D. (2022). GB_SynP: A Modular dCas9-Regulated Synthetic Promoter Collection for Fine-Tuned Recombinant Gene Expression in Plants. *ACS Synthetic Biology*, *11*(9), 3037-3048. https://doi.org/10.1021/acssynbio.2c00238

Murray, M. G., & Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*, *8*(19). https://doi.org/10.1093/nar/8.19.4321

Nagels, B., Van Damme, E. J. M., Pabst, M., Callewaert, N., & Weterings, K. (2011). Production of Complex Multiantennary N-Glycans in Nicotiana benthamiana Plants. *Plant Physiology*, *155*(3), 1103-1112. https://doi.org/10.1104/pp.110.168773

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., & Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-00578-x

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268-274. https://doi.org/10.1093/molbev/msu300

Nonaka, S., Arai, C., Takayama, M., Matsukura, C., & Ezura, H. (2017). Efficient increase of ɣ-aminobutyric acid (GABA) content in tomato fruits by targeted mutagenesis. *Scientific Reports*, *7*(1), 7057. https://doi.org/10.1038/s41598-017-06400-y

Notaguchi, M., Higashiyama, T., & Suzuki, T. (2015). Identification of mRNAs that Move Over Long Distances Using an RNA-Seq Analysis of Arabidopsis/Nicotiana benthamiana Heterografts. *Plant and Cell Physiology*, *56*(2), 311-321. https://doi.org/10.1093/pcp/pcu210

Ocampo, C. G., Lareu, J. F., Marin Viegas, V. S., Mangano, S., Loos, A., Steinkellner, H., & Petruccelli, S. (2016). Vacuolar targeting of recombinant antibodies in Nicotiana benthamiana. *Plant Biotechnology Journal*, *14*(12), 2265-2275. https://doi.org/10.1111/pbi.12580

Oey, M., Lohse, M., Kreikemeyer, B., & Bock, R. (2009). Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *The Plant Journal*, *57*(3), 436-445. https://doi.org/10.1111/j.1365-313X.2008.03702.x

Opdensteinen, P., Clodt, J. I., Müschen, C. R., Filiz, V., & Buyel, J. F. (2019). A Combined Ultrafiltration/Diafiltration Step Facilitates the Purification of Cyanovirin-N From Transgenic Tobacco Extracts. *Frontiers in Bioengineering and Biotechnology*, *6*. https://www.frontiersin.org/articles/10.3389/fbioe.2018.00206

Padmanabhan, M. S., Ma, S., Burch-Smith, T. M., Czymmek, K., Huijser, P., & Dinesh-Kumar, S. P. (2013). Novel positive regulatory role for the SPL6 transcription factor in the N TIR-NB-LRR receptor-mediated plant innate immunity. *PLoS Pathogens*, *9*(3), e1003235. https://doi.org/10.1371/journal.ppat.1003235

Pan, F., Wang, Y., Liu, H., Wu, M., Chu, W., Chen, D., & Xiang, Y. (2017). Genome-wide identification and expression analysis of SBP-like transcription factor genes in Moso Bamboo (Phyllostachys edulis). *BMC Genomics*, *18*(1), 486. https://doi.org/10.1186/s12864-017-3882-4

Park, S. J., Eshed, Y., & Lippman, Z. B. (2014). Meristem maturation and inflorescence architecture—Lessons from the Solanaceae. *Current Opinion in Plant Biology*, *17*, 70-77. https://doi.org/10.1016/j.pbi.2013.11.006

Park, Y., & Runkle, E. S. (2018). Spectral effects of light-emitting diodes on plant growth, visual color quality, and photosynthetic photon efficacy: White versus blue plus red radiation. *PLOS ONE*, *13*(8), e0202386. https://doi.org/10.1371/journal.pone.0202386

Parsons, J., Altmann, F., Graf, M., Stadlmann, J., Reski, R., & Decker, E. L. (2013). A gene responsible for prolyl-hydroxylation of moss-produced recombinant human erythropoietin. *Scientific Reports*, *3*(1), Article 1. https://doi.org/10.1038/srep03019

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), 290-295. https://doi.org/10.1038/nbt.3122

Pillay, P., Kibido, T., du Plessis, M., van der Vyver, C., Beyene, G., Vorster, B. J., Kunert, K. J., & Schlüter, U. (2012). Use of Transgenic Oryzacystatin-I-Expressing Plants Enhances Recombinant Protein Production. *Applied Biochemistry and Biotechnology*, *168*(6), 1608-1620. https://doi.org/10.1007/s12010-012-9882-6

Podevin, N., Davies, H. V., Hartung, F., Nogué, F., & Casacuberta, J. M. (2013). Site-directed nucleases: A paradigm shift in predictable, knowledge-based plant breeding. *Trends in Biotechnology*, *31*(6), 375-383. https://doi.org/10.1016/j.tibtech.2013.03.004

Poorter, H., Niklas, K. J., Reich, P. B., Oleksyn, J., Poot, P., & Mommer, L. (2012). Biomass allocation to leaves, stems and roots: Meta-analyses of interspecific variation and environmental control. *New Phytologist*, *193*(1), 30-50. https://doi.org/10.1111/j.1469-8137.2011.03952.x

Preston, J. C., & Hileman, L. C. (2013). Functional Evolution in the Plant SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) Gene Family. *Frontiers in Plant Science*, *4*, 80. https://doi.org/10.3389/fpls.2013.00080

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on Plants Obtained by Certain New Genomic Techniques and Their Food and Feed, and Amending Regulation (EU) 2017/625 (2023). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52023PC0411

Puchol Tarazona, A. A., Lobner, E., Taubenschmid, Y., Paireder, M., Torres Acosta, J. A., Göritzer, K., Steinkellner, H., & Mach, L. (2020). Steric Accessibility of the Cleavage Sites Dictates the Proteolytic Vulnerability of the Anti-HIV-1 Antibodies 2F5, 2G12, and PG9 in Plants. *Biotechnology Journal*, *15*(3), 1900308. https://doi.org/10.1002/biot.201900308

Qiu, X., Wong, G., Audet, J., Bello, A., Fernando, L., Alimonti, J. B., Fausther-Bovendo, H., Wei, H., Aviles, J., Hiatt, E., Johnson, A., Morton, J., Swope, K., Bohorov, O., Bohorova, N., Goodman, C., Kim, D., Pauly, M. H., Velasco, J., … Kobinger, G. P. (2014). Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp™. *Nature*, *514*(7520), 47-53. https://doi.org/10.1038/nature13777

Ranawaka, B., An, J., Lorenc, M. T., Jung, H., Sulli, M., Aprea, G., Roden, S., Llaca, V., Hayashi, S., Asadyar, L., LeBlanc, Z., Ahmed, Z., Naim, F., de Campos, S. B., Cooper, T., de Felippes, F. F., Dong, P., Zhong, S., Garcia-Carpintero, V., … Waterhouse, P. M. (2023). A multi-omic Nicotiana benthamiana resource for fundamental research and biotechnology. *Nature Plants*, *9*(9), Article 9. https://doi.org/10.1038/s41477-023-01489-8

Robert, S., Goulet, M.-C., D'Aoust, M.-A., Sainsbury, F., & Michaud, D. (2015). Leaf proteome rebalancing in Nicotiana benthamiana for upstream enrichment of a transiently expressed recombinant protein. *Plant Biotechnology Journal*, *13*(8), 1169-1179. https://doi.org/10.1111/pbi.12452

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, *26*(1), 139-140. https://doi.org/10.1093/bioinformatics/btp616

Rogers, K., & Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. *The Plant Cell*, *25*(7), 2383-2399. https://doi.org/10.1105/tpc.113.113159

Ruiz Castro, P. A., Kogel, U., Lo Sasso, G., Phillips, B. W., Sewer, A., Titz, B., Garcia, L., Kondylis, A., Guedj, E., Peric, D., Bornand, D., Dulize, R., Merg, C., Corciulo, M., Ivanov, N. V., Peitsch, M. C., & Hoeng, J. (2020). Anatabine ameliorates intestinal inflammation and reduces the production of pro-inflammatory factors in a dextran sulfate sodium mouse model of colitis. *Journal of Inflammation*, *17*(1), 29. https://doi.org/10.1186/s12950-020-00260-6

Ruocco, V., & Strasser, R. (2022). Transient Expression of Glycosylated SARS-CoV-2 Antigens in Nicotiana benthamiana. *Plants*, *11*(8), Article 8. https://doi.org/10.3390/plants11081093

Sack, M., Rademacher, T., Spiegel, H., Boes, A., Hellwig, S., Drossard, J., Stoger, E., & Fischer, R. (2015). From gene to harvest: Insights into upstream process development for the GMP production of a monoclonal antibody in transgenic tobacco plants. *Plant Biotechnology Journal*, *13*(8), 1094-1105. https://doi.org/10.1111/pbi.12438

Salinas, M., Xing, S., Höhmann, S., Berndtgen, R., & Huijser, P. (2012). Genomic organization, phylogenetic comparison and differential expression of the SBP-box family of transcription factors in tomato. *Planta*, *235*(6), 1171-1184. https://doi.org/10.1007/s00425-011-1565-y

Samach, A., Onouchi, H., Gold, S. E., Ditta, G. S., Schwarz-Sommer, Z., Yanofsky, M. F., & Coupland, G. (2000). Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. *Science (New York, N.Y.)*, *288*(5471), 1613-1616. https://doi.org/10.1126/science.288.5471.1613

Sánchez-León, S., Gil-Humanes, J., Ozuna, C. V., Giménez, M. J., Sousa, C., Voytas, D. F., & Barro, F. (2018). Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnology Journal*, *16*(4), 902-910. https://doi.org/10.1111/pbi.12837

Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., … Sherry, S. T. (2022). Database resources of the national center for biotechnology information. *Nucleic Acids Research*, *50*(D1), D20-D26. https://doi.org/10.1093/nar/gkab1112

Schachtsiek, J., & Stehle, F. (2019). Nicotine-free, nontransgenic tobacco (Nicotiana tabacum l.) edited by CRISPR-Cas9. *Plant Biotechnology Journal*, *17*(12), 2228-2230. https://doi.org/10.1111/pbi.13193

Scheben, A., Wolter, F., Batley, J., Puchta, H., & Edwards, D. (2017). Towards CRISPR/Cas crops – bringing together genomics and genome editing. *New Phytologist*, *216*(3), 682-698. https://doi.org/10.1111/nph.14702

Schiavinato, M., Marcet-Houben, M., Dohm, J. C., Gabaldón, T., & Himmelbauer, H. (2020). Parental origin of the allotetraploid tobacco Nicotiana benthamiana. *The Plant Journal: For Cell and Molecular Biology*, *102*(3), 541-554. https://doi.org/10.1111/tpj.14648

Schmidt, F. J., Zimmermann, M. M., Wiedmann, D. R., Lichtenauer, S., Grundmann, L., Muth, J., Twyman, R. M., Prüfer, D., & Noll, G. A. (2020). The Major Floral Promoter NtFT5 in Tobacco (Nicotiana tabacum) Is a Promising Target for Crop Improvement. *Frontiers in Plant Science*, *10*. https://www.frontiersin.org/articles/10.3389/fpls.2019.01666

Schmidt, J. A., McGrath, J. M., Hanson, M. R., Long, S. P., & Ahner, B. A. (2019). Field-grown tobacco plants maintain robust growth while accumulating large quantities of a bacterial cellulase in chloroplasts. *Nature Plants*, *5*(7), Article 7. https://doi.org/10.1038/s41477-019-0467-z

Schoberer, J., & Strasser, R. (2018). Plant glyco-biotechnology. *Seminars in Cell & Developmental Biology*, *80*, 133-141. https://doi.org/10.1016/j.semcdb.2017.07.005

Schuck, S., Prinz, W. A., Thorn, K. S., Voss, C., & Walter, P. (2009). Membrane expansion alleviates endoplasmic reticulum stress independently of the unfolded protein response. *Journal of Cell Biology*, *187*(4), 525-536. https://doi.org/10.1083/jcb.200907074

Schwarz, S., Grande, A. V., Bujdoso, N., Saedler, H., & Huijser, P. (2008). The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis. *Plant Molecular Biology*, *67*(1), 183-195. https://doi.org/10.1007/s11103-008-9310-z

Shaaltiel, Y., Bartfeld, D., Hashmueli, S., Baum, G., Brill-Almon, E., Galili, G., Dym, O., Boldin-Adamsky, S. A., Silman, I., Sussman, J. L., Futerman, A. H., & Aviezer, D. (2007). Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. *Plant Biotechnology Journal*, *5*(5), 579-590. https://doi.org/10.1111/j.1467-7652.2007.00263.x

Shang, L., Gaudreau, L., Martel, M., Michaud, D., Pepin, S., & Gosselin, A. (2018). Effects of CO2 enrichment, LED inter-lighting, and high plant density on growth of Nicotiana benthamiana used as a host to express influenza virus hemagglutinin H1. *Horticulture, Environment, and Biotechnology*, *59*(5), 637-648. https://doi.org/10.1007/s13580-018-0085-0

Sharma, A. K., Shukla, E., Janoti, D. S., Mukherjee, K. J., & Shiloach, J. (2020). A novel knock out strategy to enhance recombinant protein expression in Escherichia coli. *Microbial Cell Factories*, *19*(1), 148. https://doi.org/10.1186/s12934-020-01407-z

Sheen, S. J. (1983). Biomass and Chemical Composition of Tobacco Plants Under High Density Growth. *Contributions to Tobacco & Nicotine Research*, *12*(1), 35-42. https://doi.org/10.2478/cttr-2013-0523

Sheludko, Y. v., Sindarovska, Y. r., Gerasymenko, I. m., Bannikova, M. a., & Kuchuk, N. v. (2007). Comparison of several Nicotiana species as hosts for high-scale Agrobacterium-mediated transient expression. *Biotechnology and Bioengineering*, *96*(3), 608-614. https://doi.org/10.1002/bit.21075

Shoji, Y., Farrance, C. E., Bautista, J., Bi, H., Musiychuk, K., Horsey, A., Park, H., Jaje, J., Green, B. J., Shamloul, M., Sharma, S., Chichester, J. A., Mett, V., & Yusibov, V. (2012). A plant-based system for rapid production of influenza vaccine antigens. *Influenza and Other Respiratory Viruses*, *6*(3), 204-210. https://doi.org/10.1111/j.1750-2659.2011.00295.x

Sierro, N., Battey, J. N. D., Ouadi, S., Bakaher, N., Bovet, L., Willig, A., Goepfert, S., Peitsch, M. C., & Ivanov, N. V. (2014). The tobacco genome sequence and its comparison with those of tomato and potato. *Nature Communications*, *5*, 3833. https://doi.org/10.1038/ncomms4833

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*(1), 539. https://doi.org/10.1038/msb.2011.75

Singh, A. A., Pillay, P., & Tsekoa, T. L. (2021). Engineering Approaches in Plant Molecular Farming for Global Health. *Vaccines*, *9*(11), 1270. https://doi.org/10.3390/vaccines9111270

Singh, A. A., Pooe, O., Kwezi, L., Lotter-Stark, T., Stoychev, S. H., Alexandra, K., Gerber, I., Bhiman, J. N., Vorster, J., Pauly, M., Zeitlin, L., Whaley, K., Mach, L., Steinkellner, H., Morris, L., Tsekoa, T. L., & Chikwamba, R. (2020). Plant-based production of highly potent anti-HIV antibodies with engineered posttranslational modifications. *Scientific Reports*, *10*(1), 6201. https://doi.org/10.1038/s41598-020-63052-1

Slater, G. S. C., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, *6*, 31. https://doi.org/10.1186/1471-2105-6-31

Song, A., Gao, T., Wu, D., Xin, J., Chen, S., Guan, Z., Wang, H., Jin, L., & Chen, F. (2016). Transcriptome-wide identification and expression analysis of chrysanthemum SBP-like transcription factors. *Plant Physiology and Biochemistry: PPB*, *102*, 10-16. https://doi.org/10.1016/j.plaphy.2016.02.009

South, P. F., Cavanagh, A. P., Liu, H. W., & Ort, D. R. (2019). Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science*, *363*(6422), eaat9077. https://doi.org/10.1126/science.aat9077

Stevens, L. H., Stoopen, G. M., Elbers, I. J., Molthoff, J. W., Bakker, H. A., Lommen, A., Bosch, D., & Jordi, W. (2000). Effect of climate conditions and plant developmental stage on the stability of antibodies expressed in transgenic tobacco. *Plant Physiology*, *124*(1), 173-182. https://doi.org/10.1104/pp.124.1.173

Stoger, E., Fischer, R., Moloney, M., & Ma, J. K.-C. (2014). Plant Molecular Pharming for the Treatment of Chronic and Infectious Diseases. *Annual Review of Plant Biology*, *65*(1), 743-768. https://doi.org/10.1146/annurev-arplant-050213-035850

Stone, J. M., Liang, X., Nekl, E. R., & Stiers, J. J. (2005). Arabidopsis AtSPL14, a plant-specific SBP-domain transcription factor, participates in plant development and sensitivity to fumonisin B1. *The Plant Journal: For Cell and Molecular Biology*, *41*(5), 744-754. https://doi.org/10.1111/j.1365-313X.2005.02334.x

Strasser, R. (2016). Plant protein glycosylation. *Glycobiology*, *26*(9), 926-939. https://doi.org/10.1093/glycob/cww023

Strasser, R. (2018). Protein Quality Control in the Endoplasmic Reticulum of Plants. *Annual Review of Plant Biology*, *69*(1), 147-172. https://doi.org/10.1146/annurev-arplant-042817-040331

Takaiwa, F. (2013). Increasing the production yield of recombinant protein in transgenic seeds by expanding the deposition space within the intracellular compartment. *Bioengineered*, *4*(3), 136-139. https://doi.org/10.4161/bioe.24187

Takaiwa, F., Wakasa, Y., Hayashi, S., & Kawakatsu, T. (2017). An overview on the strategies to exploit rice endosperm as production platform for biopharmaceuticals. *Plant Science*, *263*, 201-209. https://doi.org/10.1016/j.plantsci.2017.07.016

Talens-Perales, D., Nicolau-Sanus, M., Marín-Navarro, J., Polaina, J., & Daròs, J.-A. (2023). Production in Nicotiana benthamiana of a thermotolerant glucose oxidase that shows enzybiotic activity

against Escherichia coli and Staphylococcus aureus. *Current Research in Biotechnology*, *6*, 100148. https://doi.org/10.1016/j.crbiot.2023.100148

Teh, A. Y.-H., Cavacini, L., Hu, Y., Kumru, O. S., Xiong, J., Bolick, D. T., Joshi, S. B., Grünwald-Gruber, C., Altmann, F., Klempner, M., Guerrant, R. L., Volkin, D. B., Wang, Y., & Ma, J. K.-C. (2021). Investigation of a monoclonal antibody against enterotoxigenic Escherichia coli, expressed as secretory IgA1 and IgA2 in plants. *Gut Microbes*, *13*(1), 1-14. https://doi.org/10.1080/19490976.2020.1859813

Tewolde, F. T., Shiina, K., Maruo, T., Takagaki, M., Kozai, T., & Yamori, W. (2018). Supplemental LED inter-lighting compensates for a shortage of light for plant growth and yield under the lack of sunshine. *PLOS ONE*, *13*(11), e0206592. https://doi.org/10.1371/journal.pone.0206592

Tremblay, R., Wang, D., Jevnikar, A. M., & Ma, S. (2010). Tobacco, a highly efficient green bioreactor for production of therapeutic proteins. *Biotechnology Advances*, *28*(2), 214-221. https://doi.org/10.1016/j.biotechadv.2009.11.008

Tripathi, R. K., Bregitzer, P., & Singh, J. (2018). Genome-wide analysis of the SPL/miR156 module and its interaction with the AP2/miR172 unit in barley. *Scientific Reports*, *8*(1), 7085. https://doi.org/10.1038/s41598-018-25349-0

Tusé, D., Ku, N., Bendandi, M., Becerra, C., Collins, R., Langford, N., Sancho, S. I., López-Díaz de Cerio, A., Pastor, F., Kandzia, R., Thieme, F., Jarczowski, F., Krause, D., Ma, J. K.-C., Pandya, S., Klimyuk, V., Gleba, Y., & Butler-Ransohoff, J. E. (2015). Clinical Safety and Immunogenicity of Tumor-Targeted, Plant-Made Id-KLH Conjugate Vaccines for Follicular Lymphoma. *BioMed Research International*, *2015*, 648143. https://doi.org/10.1155/2015/648143

Tusé, D., Nandi, S., McDonald, K. A., & Buyel, J. F. (2020). The Emergency Response Capacity of Plant-Based Biopharmaceutical Manufacturing-What It Is and What It Could Be. *Frontiers in Plant Science*, *11*. https://www.frontiersin.org/articles/10.3389/fpls.2020.594019

Twyman, R. M., Stoger, E., Schillberg, S., Christou, P., & Fischer, R. (2003). Molecular farming in plants: Host systems and expression technology. *Trends in Biotechnology*, *21*(12), 570-578. https://doi.org/10.1016/j.tibtech.2003.10.002

Unte, U. S., Sorensen, A.-M., Pesaresi, P., Gandikota, M., Leister, D., Saedler, H., & Huijser, P. (2003). SPL8, an SBP-box gene that affects pollen sac development in Arabidopsis. *The Plant Cell*, *15*(4), 1009-1019. https://doi.org/10.1105/tpc.010678

Uranga, M., Aragonés, V., Selma, S., Vázquez-Vilar, M., Orzáez, D., & Daròs, J.-A. (2021). Efficient Cas9 multiplex editing using unspaced sgRNA arrays engineering in a Potato virus X vector. *The Plant Journal*, *106*(2), 555-565. https://doi.org/10.1111/tpj.15164

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., & Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science (New York, N.Y.)*, *303*(5660), 1003-1006. https://doi.org/10.1126/science.1091761

van der Veen, S. J., Hollak, C. E. M., van Kuilenburg, A. B. P., & Langeveld, M. (2020). Developments in the treatment of Fabry disease. *Journal of Inherited Metabolic Disease*, *43*(5), 908-921. https://doi.org/10.1002/jimd.12228

van Eerde, A., Várnai, A., Wang, Y., Paruch, L., Jameson, J.-K., Qiao, F., Eiken, H. G., Su, H., Eijsink, V. G. H., & Clarke, J. L. (2022). Successful Production and Ligninolytic Activity of a Bacterial Laccase, Lac51, Made in Nicotiana benthamiana via Transient Expression. *Frontiers in Plant Science*, *13*. https://www.frontiersin.org/articles/10.3389/fpls.2022.912293

Vazquez-Vilar, M., Gandía, M., García-Carpintero, V., Marqués, E., Sarrion-Perdigones, A., Yenush, L., Polaina, J., Manzanares, P., Marcos, J. F., & Orzaez, D. (2020). Multigene Engineering by GoldenBraid Cloning: From Plants to Filamentous Fungi and Beyond. *Current Protocols in Molecular Biology*, *130*(1), e116. https://doi.org/10.1002/cpmb.116

Vazquez-Vilar, M., Garcia-Carpintero, V., Selma, S., Bernabé-Orts, J. M., Sanchez-Vicente, J., Salazar-Sarasua, B., Ressa, A., de Paola, C., Ajenjo, M., Quintela, J. C., Fernández-del-Carmen, A., Granell, A., & Orzáez, D. (2021). The GB4.0 Platform, an All-In-One Tool for CRISPR/Cas-Based Multiplex Genome Engineering in Plants. *Frontiers in Plant Science*, *12*. https://www.frontiersin.org/articles/10.3389/fpls.2021.689937

Verdú-Navarro, F., Moreno-Cid, J. A., Weiss, J., & Egea-Cortines, M. (2023). The advent of plant cells in bioreactors. *Frontiers in Plant Science*, *14*, 1310405. https://doi.org/10.3389/fpls.2023.1310405

Vicuna Requesens, D., Gonzalez Romero, M. E., Devaiah, S. P., Chang, Y.-K., Flory, A., Streatfield, S., Ring, R., Phillips, C., Hood, N. C., Marbaniang, C. D., Howard, J. A., & Hood, E. E. (2019). The maize α-zein promoter can be utilized as a strong inducer of cellulase enzyme expression in maize kernels. *Transgenic Research*, *28*(5), 537-547. https://doi.org/10.1007/s11248-019-00162-1

Wakasa, Y., Yasuda, H., Oono, Y., Kawakatsu, T., Hirose, S., Takahashi, H., Hayashi, S., Yang, L., & Takaiwa, F. (2011). Expression of ER quality control-related genes in response to changes in BiP1 levels in developing rice endosperm. *The Plant Journal*, *65*(5), 675-689. https://doi.org/10.1111/j.1365-313X.2010.04453.x

Waltz, E. (2021). GABA-enriched tomato is first CRISPR-edited food to enter market. *Nature Biotechnology*, *40*(1), 9-11. https://doi.org/10.1038/d41587-021-00026-2

Wang, G., Wang, P., Gao, Y., Li, Y., Wu, L., Gao, J., Zhao, M., & Xia, Q. (2018). Isolation and functional characterization of a novel FLOWERING LOCUS T homolog (NtFT5) in Nicotiana tabacum. *Journal of Plant Physiology*, *231*, 393-401. https://doi.org/10.1016/j.jplph.2018.10.021

Wang, H., & Wang, H. (2015). The miR156/SPL Module, a Regulatory Hub and Versatile Toolbox, Gears up Crops for Enhanced Agronomic Traits. *Molecular Plant*, *8*(5), 677-688. https://doi.org/10.1016/j.molp.2015.01.008

Wang, J., Yu, H., Xiong, G., Lu, Z., Jiao, Y., Meng, X., Liu, G., Chen, X., Wang, Y., & Li, J. (2017). Tissue-Specific Ubiquitination by IPA1 INTERACTING PROTEIN1 Modulates IPA1 Protein Levels to Regulate Plant Architecture in Rice. *The Plant Cell*, *29*(4), 697-707. https://doi.org/10.1105/tpc.16.00879

Wang, X., Ye, L., Lyu, M., Ursache, R., Löytynoja, A., & Mähönen, A. P. (2020). An inducible genome editing system for plants. *Nature Plants*, *6*(7), Article 7. https://doi.org/10.1038/s41477-020-0695-2

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., & Qiu, J.-L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology*, *32*(9), Article 9. https://doi.org/10.1038/nbt.2969

Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., & Weigel, D. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. *Science (New York, N.Y.)*, *309*(5737), 1056-1059. https://doi.org/10.1126/science.1114358

Wolt, J. D., Wang, K., & Yang, B. (2016). The Regulatory Status of Genome-edited Crops. *Plant Biotechnology Journal*, *14*(2), 510-518. https://doi.org/10.1111/pbi.12444

Wu, G., & Poethig, R. S. (2006). Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. *Development (Cambridge, England)*, *133*(18), 3539-3547. https://doi.org/10.1242/dev.02521

Xie, K., Minkenberg, B., & Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(11), 3570-3575. https://doi.org/10.1073/pnas.1420294112

Xie, L., Luo, Z., Jia, X., Mo, C., Huang, X., Suo, Y., Cui, S., Zang, Y., Liao, J., & Ma, X. (2023). Synthesis of Crocin I and Crocin II by Multigene Stacking in Nicotiana benthamiana. *International Journal of Molecular Sciences*, *24*(18), Article 18. https://doi.org/10.3390/ijms241814139

Xing, S., Salinas, M., Höhmann, S., Berndtgen, R., & Huijser, P. (2010). miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in Arabidopsis. *The Plant Cell*, *22*(12), 3935-3950. https://doi.org/10.1105/tpc.110.079343

Xu, M., Hu, T., Zhao, J., Park, M.-Y., Earley, K. W., Wu, G., Yang, L., & Poethig, R. S. (2016). Developmental Functions of miR156-Regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) Genes in Arabidopsis thaliana. *PLoS Genetics*, *12*(8), e1006263. https://doi.org/10.1371/journal.pgen.1006263

Yamaguchi, A., Kobayashi, Y., Goto, K., Abe, M., & Araki, T. (2005). TWIN SISTER OF FT (TSF) Acts as a Floral Pathway Integrator Redundantly with FT. *Plant and Cell Physiology*, *46*(8), 1175-1189. https://doi.org/10.1093/pcp/pci151

Yamamoto, T., Hoshikawa, K., Ezura, K., Okazawa, R., Fujita, S., Takaoka, M., Mason, H. S., Ezura, H., & Miura, K. (2018). Improvement of the transient expression system for production of recombinant proteins in plants. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-23024-y

Yamasaki, H., Hayashi, M., Fukazawa, M., Kobayashi, Y., & Shikanai, T. (2009). SQUAMOSA Promoter Binding Protein-Like7 Is a Central Regulator for Copper Homeostasis in Arabidopsis. *The Plant Cell*, *21*(1), 347-361. https://doi.org/10.1105/tpc.108.060137

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S., & Dubcovsky, J. (2006). The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(51), 19581-19586. https://doi.org/10.1073/pnas.0607142103

Yang, S., Overlander-Chen, M., Carlson, C. H., & Fiedler, J. D. (2022). A SQUAMOSA promoter binding protein-like transcription factor controls crop ideotype for high productivity in barley. *Plant Direct*, *6*(9), e450. https://doi.org/10.1002/pld3.450

Yang, S.-J., Carter, S. A., Cole, A. B., Cheng, N.-H., & Nelson, R. S. (2004). A natural variant of a host RNA-dependent RNA polymerase is associated with increased susceptibility to viruses by Nicotiana benthamiana. *Proceedings of the National Academy of Sciences*, *101*(16), 6297-6302. https://doi.org/10.1073/pnas.0304346101

Yang, W. C., Minkler, D. F., Kshirsagar, R., Ryll, T., & Huang, Y.-M. (2016). Concentrated fed-batch cell culture increases manufacturing capacity without additional volumetric capacity. *Journal of Biotechnology*, *217*, 1-11. https://doi.org/10.1016/j.jbiotec.2015.10.009

Yang, Z., Wang, X., Gu, S., Hu, Z., Xu, H., & Xu, C. (2008). Comparative study of SBP-box gene family in Arabidopsis and rice. *Gene*, *407*(1-2), 1-11. https://doi.org/10.1016/j.gene.2007.02.034

Yoo, S. J., Chung, K. S., Jung, S. H., Yoo, S. Y., Lee, J. S., & Ahn, J. H. (2010). BROTHER OF FT AND TFL1 (BFT) has TFL1-like activity and functions redundantly with TFL1 in inflorescence meristem development in Arabidopsis. *The Plant Journal*, *63*(2), 241-253. https://doi.org/10.1111/j.1365-313X.2010.04234.x

Yoo, S. Y., Kardailsky, I., Lee, J. S., Weigel, D., & Ahn, J. H. (2004). Acceleration of flowering by overexpression of MFT (MOTHER OF FT AND TFL1). *Molecules and Cells*, *17*(1), 95-101.

Yu, N., Cai, W.-J., Wang, S., Shan, C.-M., Wang, L.-J., & Chen, X.-Y. (2010). Temporal control of trichome distribution by microRNA156-targeted SPL genes in Arabidopsis thaliana. *The Plant Cell*, *22*(7), 2322-2335. https://doi.org/10.1105/tpc.109.072579

Yu, N., Niu, Q.-W., Ng, K.-H., & Chua, N.-H. (2015). The role of miR156/SPLs modules in Arabidopsis lateral root development. *The Plant Journal: For Cell and Molecular Biology*, *83*(4), 673-685. https://doi.org/10.1111/tpj.12919

Zeng, R.-F., Zhou, J.-J., Liu, S.-R., Gan, Z.-M., Zhang, J.-Z., & Hu, C.-G. (2019). Genome-Wide Identification and Characterization of SQUAMOSA-Promoter-Binding Protein (SBP) Genes Involved in the Flowering Development of Citrus Clementina. *Biomolecules*, *9*(2), E66. https://doi.org/10.3390/biom9020066

Zhang, T., Wang, J., & Zhou, C. (2015). The role of miR156 in developmental transitions in Nicotiana tabacum. *Science China. Life Sciences*, *58*(3), 253-260. https://doi.org/10.1007/s11427-015-4808-5

Zhang, Y., Pribil, M., Palmgren, M., & Gao, C. (2020). A CRISPR way for accelerating improvement of food crops. *Nature Food*, *1*(4), Article 4. https://doi.org/10.1038/s43016-020-0051-8

Zhang, Y., Schwarz, S., Saedler, H., & Huijser, P. (2007). SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in Arabidopsis. *Plant Molecular Biology*, *63*(3), 429-439. https://doi.org/10.1007/s11103-006-9099-6

Zheng, M., Zhang, L., Tang, M., Liu, J., Liu, H., Yang, H., Fan, S., Terzaghi, W., Wang, H., & Hua, W. (2020). Knockout of two *Bna MAX 1* homologs by CRISPR /Cas9-targeted mutagenesis improves plant architecture and increases yield in rapeseed (*Brassica napus* L.). *Plant Biotechnology Journal*, *18*(3), 644-654. https://doi.org/10.1111/pbi.13228

Zhou, Q., Zhang, S., Chen, F., Liu, B., Wu, L., Li, F., Zhang, J., Bao, M., & Liu, G. (2018). Genome-wide identification and characterization of the SBP-box gene family in Petunia. *BMC Genomics*, *19*(1), 193. https://doi.org/10.1186/s12864-018-4537-9

Zhu, H., Bhatt, B., Sivaprakasam, S., Cai, Y., Liu, S., Kodeboyina, S. K., Patel, N., Savage, N. M., Sharma, A., Kaufman, R. J., Li, H., & Singh, N. (2019). Ufbp1 promotes plasma cell development and ER expansion by modulating distinct branches of UPR. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-08908-5

Zimran, A., Gonzalez-Rodriguez, D. E., Abrahamov, A., Cooper, P. A., Varughese, S., Giraldo, P., Petakov, M., Tan, E. S., & Chertkoff, R. (2018). Long-term safety and efficacy of taliglucerase alfa in pediatric Gaucher disease patients who were treatment-naïve or previously treated with imiglucerase. *Blood Cells, Molecules, and Diseases*, *68*, 163-172. https://doi.org/10.1016/j.bcmd.2016.10.005

Zischewski, J., Sack, M., & Fischer, R. (2016). Overcoming low yields of plant-made antibodies by a protein engineering approach. *Biotechnology Journal*, *11*(1), 107-116. https://doi.org/10.1002/biot.201500255

Zsögön, A., Čermák, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., Freschi, L., Voytas, D. F., Kudla, J., & Peres, L. E. P. (2018). De novo domestication of wild tomato using genome editing. *Nature Biotechnology*, *36*(12), 1211-1216. https://doi.org/10.1038/nbt.4272