Universitat Autònoma de Barcelona

PLANT CELL BIOREACTORS FOR PEPTIDE PRODUCTION

Tarik Ruiz Medina

UNIVERSITAT AUTÒNOMA DE BARCELONA

Facultat de Biociències

Dept. Biologia Animal, Biologia Vegetal i Ecologia

PLANT CELL BIOREACTORS FOR PEPTIDE PRODUCTION

Tarik Ruiz Medina

Barcelona, March 2020

UNIVERSITAT AUTÒNOMA DE BARCELONA

Facultat de Biociències

Dept. Biologia Animal, Biologia Vegetal i Ecologia

Doctorat en Biologia i Biotecnologia Vegetal

PhD thesis

PLANT CELL BIOREACTORS FOR PEPTIDE PRODUCTION

Dissertation presented by Tarik Ruiz Medina for the degree of Doctor in Biology and Plant Biotechnology by Universistat Autònoma de Barcelona

This work was carried out in Center for Research in Agricultural Genomics (CRAG), Cerdanyola del Vallès (Barcelona)

Dr. Juan José López-Moya	**Dr. María Coca**	**Dr. Òscar Expósito**	**Tarik Ruiz Medina**
Thesis director	Thesis director Tutor	Company tutor	Author

Barcelona, March 2020

Mira siempre hacia dónde vas, pero nunca olvides de dónde has salido.

ACKNOWLEDGMENTS

Me gustaría terminar esta etapa dando las gracias a las personas que han estado a mi lado, de uno u otro modo, durante estos últimos años.

En primer lugar, mi más sincero agradecimiento a Òscar Expósito y Albert Jané por ofrecerme la oportunidad de dedicar mis esfuerzos como trabajador de Vytrus al desarrollo de esta tesis. Siempre estaré agradecido por haber apostado por mí, por vuestro respeto y colaboración incondicional en los momentos más complejos, así como haberme tenido siempre en cuenta. Vuestra actitud os honra, muchas gracias.

En segundo lugar, y no menos importante, querría mostrar mi agradecimiento a mis directores de tesis: María Coca y Juan José López-Moya, por brindarme la oportunidad de colaborar con vosotros. Os agradezco de todo corazón vuestra dedicación, tiempo y formación. El desarrollo de esta tesis a vuestro lado me ha enseñado mucho tanto como científico como a nivel personal. Ha sido un placer poder formar parte de vuestros equipos y conoceros.

Los cuatro, de una o otra forma, me habéis ayudado a fomentar y moldear mi creatividad, permitiéndome tener ideas innovadoras y minimizando la divagación, os lo agradezco.

También me gustaría agradecer a nuestros colaboradores de Alemania: Stefan Schillberg, Stefan Rasche, Holger Spiegel y Helga Schinkel por acogerme en vuestro laboratorio con los brazos abiertos y haberme hecho la estancia en Alemania más llevadera.

Querría agradecer también a los compañeros que me han estado ayudando, soportando mis quejas en el día a día: a los que están (Ornela, Inma, Adrià, Xiaoqing y Eric) y los que han pasado (Marisa, Bia, Laia y Jhony). De los compañeros del CRAG también me gustaría agradecer a todo el equipo de invernaderos, secuenciación, microscopia e imagen. Especialmente me gustaría dar mi agradecimiento a Pilar Fontanet y Montse Amenós, ambas me habéis brindado vuestra ayuda siempre que la he necesitado, gran parte de esta tesis no habría sido posible sin vuestra ayuda.

A los compañeros de Vytrus, gracias por vuestra colaboración, comprensión y ayuda cuando ha sido necesario. Me gustaría agradecer especialmente a Pau Riera y Ana Gallego por su ayuda en este proyecto, así como a Manel López y Adrián Pérez por vuestra ayuda con las hormonas y demás materiales.

Y por último, y más importante, a mi familia: mamá, papá y Andrea, siempre habéis estado a mi lado y me habéis dado amor y comprensión. Sin vosotros no sería la persona que soy y estaría donde estoy. Siempre os tengo presentes esté donde esté. También a Carlos, Montse y Anna, por ser una parte indispensable de mi vida.

Y, como no, a la persona que me ha apoyado incondicionalmente durante toda mi vida adulta, Annabel, gracias por estar a mi lado, soportarme y no arrojarme por la ventana pese a que quizá sea necesario en algunos momentos.

Gracias a todos.

INDEX OF CONTENTS

INDEX OF FIGURES

INTRODUCTION

RESULTS

DISCUSSION

INDEX OF TABLES

INTRODUCTION

MATERIALS AND METHODS

RESULTS

ABSTRACT

The production of proteins in plant cell cultures and whole plants represents great opportunities to develop products for commercial use. The main objective of this industrial thesis was to develop economic and efficient plant production systems to bring proteins of interest to the market. We explored two different systems, *Daucus carota* cell cultures and *Nicotiana benthamiana* leaves, each having advantages and drawbacks depending on the intended use of the products. As a proof of concept, both systems were applied in the production of the human insulin-like growth factor 1 (IGF1), a high value peptide for the cosmetic and therapeutic industries. Innovative strategies to enhance gene expression and to facilitate product purification were used to improve yields and to reduce costs. Moreover, the biological activity of the produced IGF1 and derivatives was evaluated and compared to the chemically synthesized peptides to demonstrate the usefulness of production systems.

Our first approach to enhance gene expression and improve peptide yields was with RNA silencing suppressors (RSSs). Using transient expression assays and the green fluorescent protein (GFP) as reporter, we selected the P1b from the *Cucumber vein yellowing virus* (CVYV) Ipomovirus as the RSSs to enhance gene expression in carrot cell cultures. Our results demonstrated that transgenic lines overexpressing IGF1 or the derivative CPP-IGF1 (a variant tailored to enhance the delivery to human cells) reached up to 4-fold higher peptide yields in combination with P1b than without. The IGF1 or CPP-IGF1 was targeted to the culture media being easily purified by simple clarification of suspensions. Moreover, we found that the media containing the produced IGF1 or CPP-IGF1 stimulated the division of human fibroblasts. A cryopreservation process was applied to the transgenic lines to avoid the reduction in peptide production found over successive propagation cycles. This allowed us to recover the original yields, opening up the possibility of establishing master cell banks.

We also developed a transient production system of IGF1 and CPP-IGF1 using *N. benthamiana* leaves and a derived tobacco mosaic virus vector. This system resulted in similar yields of active peptides to cell cultures with the main advantage of shortening production times, although requiring more complex downstream purification.

Our innovative strategy to facilitate the purification of IGF1 from plant matrices was the use of oleosin fusion technology for lipid droplet (LDs) targeting. This technology has been previously used in LD-rich seeds, but unexplored in plant cell cultures or LD-poor tissues such as leaves. Our work showed that model

cell cultures from *Nicotiana tabacum* or *Arabidopsis thaliana* were an exception, as many other plant cell cultures, including *D. carota* cells, do contain a large number of LDs and are susceptible to produce oleosin fusion proteins. However, as the stable expression of oleosin fusions severely affected callus cell growth, we tested the technology in transient expression in leaves. Due to the low level of LDs in leaves, oleosin fusion proteins production was in combination with triacylglycerol (TAG) induction to increase LD content simultaneously. For this purpose, key components of the TAG biosynthetic pathway, *A. thaliana* derived elements such as the enzyme *DGAT1* and the regulatory factor *WRI1* were co-expressed with the IGF1 oleosin fusion proteins in *N. benthamiana* leaves. Using this strategy, we obtained yields up to 1 µg/g of IGF1 bound to LDs, easily purified and fully active.

Our work provides evidence of the potential of plant matrices to produce valuable peptides. Also, the oleosin technology, the use of RSSs and viral vectors explored will serve to overcome some of the known limitations of plant systems to produce active products of industrial interest.

RESUMEN

La producción de proteínas recombinantes en plantas representa una oportunidad para su obtención y uso comercial. El objetivo principal de esta tesis industrial ha sido el desarrollo de sistemas vegetales de producción de proteínas, eficientes y competitivos a nivel económico, con posibilidades de llevarlas al mercado. Para ello hemos explorado dos sistemas: los cultivos celulares de *Daucus carota* y las hojas de *Nicotiana benthamiana*, cada uno con sus ventajas y limitaciones. Como prueba de concepto, ambos sistemas fueron utilizados para la producción de "'insulin-like growth factor 1" (IGF1), un péptido de alto valor añadido para las industrias cosmética y farmacéutica. Se ensayaron varias estrategias innovadoras para mejorar los rendimientos de producción aumentando la expresión génica y para reducir costes de purificación del producto. Además, la actividad biológica de IGF1 y sus derivados producidos en plantas se evaluó en comparación con péptidos sintéticos.

Como primera estrategia se ensayaron supresores del silenciamiento de ARN de origen viral para incrementar la expresión génica. En ensayos de expresión transitoria con la proteína verde fluorescente como marcadora, seleccionamos la proteína P1b del ipomovirus *Cucumber vein yellowing virus* (CVYV). Nuestros resultados con líneas celulares de zanahoria sobreexpresoras de IGF1 o su péptido derivado CPP-IGF1 (variante diseñada para mejorar su penetración en células humanas) mostraron que en combinación con P1b alcanzaban rendimientos de producción 4 veces mayores que las líneas sin el supresor del silencing. Además, los péptidos fueron dirigidos al medio de cultivo para facilitar su aislamiento por simple clarificación. En ensayos de actividad, las fracciones obtenidas confirmaron ser capaces de incrementar la división de fibroblastos humanos. En relación a la estabilidad de la producción, se observó una reducción cercana al 33% después de veintiún ciclos de propagación sucesivos, por lo que se implementó la criopreservación de las líneas transgénicas para mantener los rendimientos de producción originales, y así establecer bancos de líneas celulares para usos futuros.

También se desarrolló un sistema de producción transitoria de IGF1 y CPP-IGF1 en hojas de *N. benthamiana* utilizando un vector derivado del virus del mosaico del tabaco, *Tobacco mosaic virus* (TMV). Este sistema permitió reducir el tiempo de obtención del péptido activo, aunque en comparación con la producción en líneas celulares la obtención del producto no fue tan sencilla.

Con el fin de facilitar la purificación de IGF1 desde matrices vegetales, aplicamos una estrategia innovadora basada en fusiones a oleosina para dirigir la producción a cuerpos lipídicos. Esta tecnología ya

había sido utilizada en semillas, pero no en cultivos celulares, y escasamente en hojas. Nuestras observaciones mostraron la presencia de abundantes cuerpos lipídicos en numerosos cultivos celulares, incluyendo los de *D. carota*, con la excepción de las dos especies modelo analizadas, *Nicotiana tabacum* y *Arabidopsis thaliana*. Desafortunadamente, la expresión estable de fusiones a oleosina pareció afectar gravemente el crecimiento de los callos celulares, por lo que se exploró la alternativa de su aplicación a la producción en hojas. Para aumentar la cantidad de cuerpos lipídicos, la producción de las fusiones a oleosina se realizó simultáneamente con inductores de la acumulación de triacilgliceroles, usando elementos clave de su ruta biosintética en *A. thaliana*: la enzima DGAT1 y el factor de transcripción WRI1. Cuando ambos inductores fueron co-expresados en combinación con fusiones de oleosina e IGF1 en plantas de *N. benthamiana*, se obtuvo hasta 1 µg/g de IGF1 unida a los cuerpos lipídicos, fácilmente aislable y activo.

Nuestro trabajo proporciona evidencias de que la utilización de supresores del silenciamiento de ARN, los vectores virales y la tecnología de oleosinas contribuyen al potencial de las matrices vegetales para la producción de proteínas de interés.

RESUM

La producció de proteïnes recombinants en plantes representa una oportunitat per a la seva obtenció i utilització comercial. L'objectiu principal d'aquesta tesi industrial ha estat el desenvolupament de sistemes vegetals de producció de proteïnes, eficients i competitius econòmicament, amb possibilitats de portar-les al mercat. Per fer-ho hem explorat dos sistemes: els cultius cel·lulars de *Daucus carota* i les fulles de *Nicotiana benthamiana*, cadascun d'ells amb els seus avantatges i limitacions. Com a prova de concepte, ambdós sistemes van ser utilitzats per a la producció d'"insulin-like growth factor 1" (IGF1), un pèptid d'alt valor afegit per a les indústries cosmètica i farmacèutica. S'assajaren vàries estratègies innovadores per a millorar els rendiments de producció augmentant l'expressió gènica i reduir costos de purificació del producte. A més, l'activitat biològica de l'IGF1 i els seus derivats produïts en planta va ser avaluada en comparació amb pèptids sintètics.

Com a primera estratègia s'assajaren supressors del silenciament de l'ARN d'origen viral per tal d'incrementar l'expressió gènica. En assajos d'expressió transitòria amb la proteïna verda fluorescent com a marcadora, seleccionàrem la proteïna P1b del ipomovirus *Cucumber vein yellowing virus* (CVYV). Els nostres resultats amb línies cel·lulars de pastanaga sobreexpressores de l'IGF1 o el seu pèptid derivat CPP-IGF1 (variant dissenyada per a millorar la seva penetració en cèl·lules humanes) mostraren que en combinació amb P1b s'arribava a rendiments de producció 4 vegades majors que les línies sense el supressor del silenciament. A més, els pèptids foren dirigits al medi de cultiu per facilitar el seu aïllament mitjançant una simple clarificació. En assajos d'activitat, les fraccions obtingudes confirmaren ser capaces d'incrementar la divisió de fibroblasts humans. En relació amb l'estabilitat de la producció, s'observà una reducció propera al 33% després de vint-i-un cicles de propagació successius, per la qual cosa s'implementà la criopreservació de les línies transgèniques per mantenir els rendiments de producció originals, i així establir bancs cel·lulars per usos futurs.

Alhora, es desenvolupà un sistema de producció transitòria de l'IGF1 i el CPP-IGF1 en fulles de *N. benthamiana* utilitzant un vector derivat del virus del mosaic del tabac, *Tobacco mosaic virus* (TMV). Aquest sistema va permetre reduir el temps d'obtenció del pèptid actiu, encara que en comparació amb la producció a les línies cel·lulars l'obtenció del producte no fou tan senzilla.

Per tal de facilitar la purificació de l'IGF1 a partir de les matrius vegetals, aplicàrem una estratègia innovadora basada en fusions a oleosina per dirigir la producció a cossos lipídics. Aquesta tecnologia ja

havia estat utilitzada en llavors, però no en cultius cel·lulars i escassament en fulles. Les nostres observacions mostraren la presència d'abundants cossos lipídics en nombrosos cultius cel·lulars incloent-hi els de *D. carota* amb l'excepció de les dues espècies model analitzades, *Nicotiana tabacum* i *Arabidopsis thaliana*. Desafortunadament, l'expressió estable de fusions a l'oleosina sembla que va afectar greument al creixement dels calls cel·lulars, pel que s'exploraren alternatives de la seva aplicació a la producció en fulles. Per tal d'augmentar la quantitat de cossos lipídics, la producció de les fusions a l'oleosina es realitzà simultàniament amb la d'inductors de l'acumulació de triacilglicerols utilitzant elements clau de la seva ruta biosintètica en *A. thaliana*: l'enzim DGAT1 i el factor de transcripció WRI1. Quan ambdós inductors foren co-expressats en combinació amb fusions a oleosina i l'IGF1 en plantes de *N. benthamiana*, es va obtenir fins 1 µg/g d'IGF1 unit als cossos lipídics, fàcilment aïllable i actiu.

El nostre treball proporciona evidències que la utilització de supressors del silenciament de l'ARN, els vectors virals i la tecnologia de les oleosines contribueixen al potencial de les matrius vegetals per a la producció de proteïnes d'interès.

ABBREVIATIONS USED

35S:	*Cauliflower mosaic virus* 35S promoter
Aa:	Aminoacid
ABRC:	The Arabidopsis Biological Resource Center
AGO:	Argonaute protein
AKT:	Protein Kinase B
AtOle2:	Arabidopsis thaliana Oleosin 2
BEVS:	Baculovirus Expression Vector System
BY2:	*Nicotiana tabacum* Bright Yellow 2 cell line
CFG:	Consortium of Functional Glycomics
CFP:	Cian Fluorescent Protein
CHO:	Chinese Hamster Ovary
CP:	Coat Protein
CPP:	Cell Penetrating Peptide
CVYV:	*Cucumber vein yellowing mosaic virus*
Da:	Daltons
DCL:	Dicer like protein
DGAT1:	Diacylglycerol O-acyltransferase 1
DNA:	Deoxyribonucleic Acid
dsRNA:	Double stranded RNA
ELISA:	Enzyme-Linked ImmunoSorbent Assay
ER:	Endoplasmic Reticulum
EV:	Empty vector
FA:	Fatty Acid
FDA:	U.S. Food and Drug Administration
FW:	Fresh Weight

GA: Golgi Aparatus

GFP: Green Fluorescent Protein

GMO: Genetically Modified Organism

GMP: Good Manufacturing Practices

HCPro: Helper Component Protease

hGH: Human Growth Hormone

ICTV: International Committee on Taxonomy of Viruses

IGF1: Insulin-Like Growth Factor 1

IR: Insulin Receptor

kDa: Kilo Daltons

LD: Lipid Droplet

MS: Murashige and Skoog medium

MSAt: Murashige and Skoog *Arabidopsis thaliana* medium

MSBY2: Murashige and Skoog BY2 medium

MSDc: Murashige and Skoog *Daucus carota* medium

ORF: Open Reading Frame

PAMPS: Pathogen-Associated Molecular Patterns

P1b: Protein P1b

PCR: Polymerase Chain Reaction

PSV: Protein Storage Vacuole

PTM: Post-Translational Modification

RdRp: RNA-dependent RNA polymerase

rhLF: Human lactoferrin

RISC: TNA induced silencing complexes

RNA: Ribonucleic acid

RSS: viral RNA Silencing Suppressors

RT: Reverse transcription

SAR: Scaffold Attachment Region

siRNA:	Small interfering RNA
SDS-PAGE:	Polyacrylamide gel electrophoresis in presence of SDS
SFPMV:	*Sweet potato feathery mottle virus*
SPFMVs:	*Sweet potato feathery mottle virus* protease cleavage site
Ta:	Exhaustion time
TAG:	Triacylglycerol
TAT:	Trans-Activator of Transcription
TEV:	*Tobacco etch virus*
TEVs:	*Tobacco etch virus* protease cleavage site
TMV:	*Tobacco mosaic virus*
UV:	Ultraviolet
VEC:	Cell Packing Volume
Vexp:	Exponential growth velocity
vsiRNAs:	viral small interferent RNAs
WMV:	*Watermelon mosaic virus*
WRI1:	WRINKLED1
WT:	Wild Type
YFP:	Yellow Fluorescent Protein
ºB:	Brix degrees

INTRODUCTION

INTRODUCTION

HETEROLOGOUS PROTEIN PRODUCTION

Proteins are macromolecules formed by amino acids as building blocks. In some cases, proteins can be associated with other molecules. They are characterized by an extremely wide range of diversity and specificity and play essential roles in nearly all biological processes in known living entities. Their activities range from key structural functions to the catalysis of biochemical reactions. Their importance for most biological reactions has made them key targets for treatments against animal and human diseases, triggering the development of protein-derived therapeutics. A proposed classification for these is based on their therapeutic activity: proteins with enzymatic or regulatory activity, proteins with special targeting properties, vaccines, proteins with application in diagnostics (Leader *et al.*, 2008). In 2017, the THPdb database of FDA approved peptides and proteins as therapeutics included 239 molecules corresponding to 380 drug variants (http://crdd.osdd.net/raghava/thpdb/). THPdb also shows that the main applications of these proteins are as treatments for metabolic or immunological disorders (Usmani *et al.*, 2017). Molecular biology and biotechnology have paved the way for producing proteins of interest in organisms or systems different to their natural origin. These are known in general as "recombinant proteins". Since the approval of the first recombinant protein, human insulin (Goeddel *et al.*, 1979), to the 380 drug variants on the market in 2017, recombinant therapeutic proteins have garnered a lot of attention from business and science. From a business point of view, the market for these recombinant proteins has experienced rapid growth in the last decades, and as they are expected to represent more than 10% of the world's pharmaceutical market in the next few years (Craik *et al.*, 2013; Usmani *et al.*, 2017). From the standpoint of scientific professionals, these recombinant proteins are essential elements in research activities for numerous reasons. As examples of their high scientific value, we can mention the target specificity of the developed therapies based on the use of proteins, and the wide range of possibilities for using the proteins and peptides as drug conjugates or molecular functionalizers (Arlotta and Owen, 2019; Jauset and Beaulieu, 2019).

Recombinant proteins can be produced in a huge variety of organisms. Among the easiest and fastest systems to implement are those based on the use of bacterial cells, such as *Escherichia coli*. However, complex, multimeric or post-translationally modified proteins from eukaryotic origin might not be correctly produced, processed or modified in prokaryotic organisms or in systems derived from them.

Technology approaches have been developed to overcome these limitations, for instance performing the last steps of the protein production *in vitro*, but these methods often increase the production costs to unaffordable levels for industrial processes (Rosano and Ceccarelli, 2014). Consequently, many alternatives to *E.coli* production have been explored and evaluated. Whole organisms can be used, such as plants, invertebrates like insects, and even animals including mammals, as well as the corresponding cell lines derived from them when available for *in vitro* culture systems. A summary of both the merits and limitations of different systems used for recombinant protein production is provided in the next table (Table I1), which is based on recent reviews (Baghban *et al.*, 2019; Chambers *et al*., 2018; Mei *et al.*, 2019; Mojzita *et al.*, 2019; Zhu, 2012).

Because human and animal health products require a high level of safety and quality, while the markets require low cost and high yields, industry is looking for alternatives to the classical expression host. Protein expression in prokaryotic hosts has the previously mentioned problem with post-translational modifications (PTMs), but is also affected negatively by the production of endotoxins, which can be difficult to eliminate from the final products even using cell lines with low endotoxin levels (Mamat *et al.*, 2015). For other industrial choices, such as mammalian-derived cells, costs may be a problem, as well as safety concerns caused by possible contamination with mammalian viruses and prions. An example of the risks of these systems is illustrated by the protein production in Chinese hamster ovary (CHO) cell systems in 2009, which caused a crisis in Genzyme Corp (Weisman, 2009), forcing the company to stop the production of two drugs for rare genetic disorders due to virus contamination of the equipment. In this particular case, the supply to 8000 patients was compromised, affecting their treatments. Additional problems of CHO cell production systems are associated to cell line instability when producing recombinant proteins (Kim *et al.*, 2011).

Alternative expression hosts could be unicellular eukaryotic organisms like yeasts, that have the advantage of fast and easy growth, and lack the risk of prion and mammalian viral contamination. However, other problems occur, such as the frequent hyperglycosylation that can affect the proteins produced in this host. Hyperglycosylation could produce loss of activity, or generate an undesired immune response in the animals or patients treated with this recombinant protein (Baghban *et al.*, 2019). In the case of filamentous fungi, the lack of knowledge about their metabolism, or downstream problems have restricted their use to produce specific proteins, mainly fungal proteins.

Alternatives have been explored when working with insect cells; stable or transient relying on virus-based strategies such as the baculovirus expression vector systems (BEVS). In the case of stable expression, the

yields have been substantially lower than that obtained using transfection with BEVS (Chambers, *et al.* 2018). However, BEVS is still problematic, mainly because of the time required to generate recombinant virus and the instability for certain target proteins (Scholz and Suppmann, 2017).

Table I1. Comparison of common systems of recombinant protein production

Host system	Merits	Limitations
E. coli	Easy to transform and work with Extensively studied metabolism Economical production costs Rapid growth rates Capacity of continuous fermentations Good production rates	Unable to removing introns from transcripts Codon usage bias Can recognize foreign gene regions as cryptic promoters and produce artifacts Lack of PTMs Possibility of production of proteins in insoluble forms or inclusion bodies Production of endotoxins High protein degradation rates
Yeast systems	Rapid growth rates in low cost medium Appropriate post-translational modifications (PTMs) No endotoxins produced Good production rates Metabolism of some species studied	Possible protein hyperglycosylation Codon usage bias Inefficient in secreting the proteins, leading to intracellular retention
Filamentous fungi	High-level of expression Can produce anti-microbial peptides	Complex upstream and downstream processes Lack of knowledge in metabolism
Baculovirus/ Insect systems	High-level of expression Appropriate PTMs Safe for vertebrate proteins Good tool for recombinant glycoprotein production	Continuous expression is not possible More demanding culture conditions Complex upstream and downstream processes
Mammalian cells/ systems	Proper protein folding Appropriate PTMs Proper glycosylation patterns	High cost Complex technologies Risk of contamination with animal viruses
Transgenic plants	Easy scaling up at low cost Proteins can be localized in different organs depending of the growth stage High yields Transient expression technologies, avoiding GMO Appropriate PTMs in most cases	Expression levels are target dependent Risk of endotoxins in agrobacterium based transient strategies
Plant cell cultures	Easy scaling up at low cost No need of complex growth/production media High yields Transient expression technologies, avoiding GMO Production in bioreactors Appropriate PTMs in most cases Safety	Expression levels are target dependent Risk of endotoxins in agrobacterium based transient strategies
Transgenic animals	Appropriate PTMs and product assembly Proper glycosylation	High cost Relatively longer production period Low yield Risk of contamination with animal viruses

Adapted from Gomes, 2016.

Finally, plant expression strategies would appear to be a robust alternative to all these systems, avoiding most of the problems associated with other production platforms, and achieving most of the goals.

PLANT-BASED PROTEIN PRODUCTION

Plant expression systems offer several advantages over traditionally used protein production platforms, including safety and costs that are key factors from an industrial perspective. In contrast to prokaryotic systems, they lack endotoxins, harmful molecules that are difficult to eliminate even during downstream purification processes and that seriously affect the health of target organisms. When comparing with yeasts, hyperglycosylation, that could affect the protein activities and/or induce the immune response in target organisms, does not occur in plant produced proteins. Nor do plant systems carry mammalian pathogens such as viruses, a serious concern when using mammalian-derived systems. Although plants can carry their own viruses and pathogens, mammalian organisms are constantly exposed to them, without risk, and there are no reports that plant viruses might induce diseases in humans or other mammals (Fischer *et al.*, 2012). The presence of certain secondary metabolites such as nicotine, especially if using species from the *Nicotiana* genus, can be a challenge in downstream stages (Fischer *et al.*, 2012). However, edible plants such as lettuce (van Eerde *et al.*, 2019), rice (Takaiwa *et al.*, 2015) and carrot (Permyakova *et al.*, 2015) can be used as safer production systems, alleviating the problems associated with secondary metabolites.

In contrast to prokaryotic systems, plants are capable of processing complex molecules and most post-translational modifications. These modifications are important for many biological functions, pharmacokinetic properties, immunogenicity, solubility, folding and protein-protein interactions (Heilmann *et al.*, 2012; Moremen *et al.*, 2012). Plant systems add complex glycan structures in a manner very similar to mammalian ones (Bosch *et al.*, 2013; Strasser, 2016). However, while modifications such as N-glycosylations are similar, they are not identical, and therefore some precautions are needed. N-glycosylation of proteins is considered the most common co- and post-translational modification of proteins entering the secretory pathway. The process starts simultaneously with translation in the endoplasmic reticulum (ER) where an oligosaccharide precursor is attached to the protein, then a series of reactions trigger the maturation of the precursor in the final glycoside of the glycoproteins in the Golgi apparatus (GA). In plants, the N-glycan structures contain 1,3-fucose, while in mammalians they contain

1,6-fucose, and mammalian proteins carry 1,4-galactose and terminal sialic acid residues while plants have 1,2-xylose residues (Fig. I1).

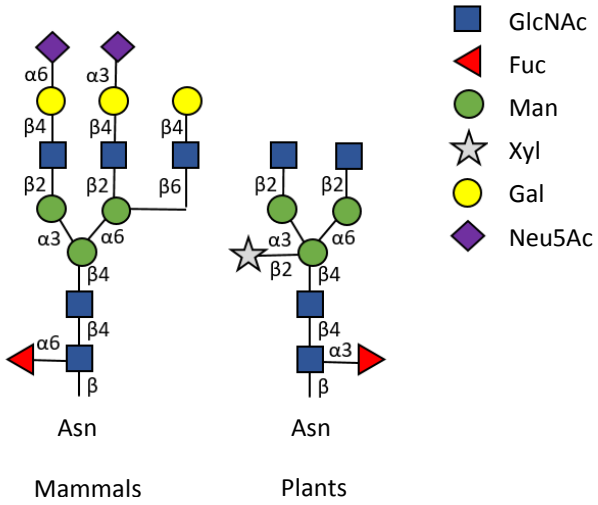


Figure I1. Typical N-glycan structures from mammal and plant proteins. The monosaccharides are represented as symbols following the recommendations from the Consortium for Functional Glycomics (CFG).

Strategies for production in whole plants

The use of whole plants for the production of recombinant proteins was proposed as an option for high yields at low cost, potentially surpassing the limits imposed by the capacity of fermenters, and reducing the cost of treatments required to supply a large amount of recombinant proteins in emergency situations within a short timeframe. An example of a health emergency was the application of the plant produced antibody cocktail ZMapp™ for the treatment of Ebola infections (Hiatt *et al.*, 2015; Qiu *et al.*, 2014). In the case of recombinant proteins for topical applications, nutraceuticals or even for stimulation of the immune system, the plant strategy could be a particularly good solution due to the simple downstream processing needed to reach correct activity (Fischer *et al.*, 2013; Ruiz *et al.*, 2018).

The plants can be grown in contained environments (greenhouses) or in open fields as "regular" crops (Fig I2). However, the current legislation of transgenic use and the bad perception of transgenic plants, especially in occidental Europe, make it difficult to cultivate them outdoors (Fischer *et al.*, 2013; Wunderlich and Gatto, 2015). There are recent examples of accidents that have increased this lack of trust in the security of growing transgenics in open fields: the security breach of Prodigene Inc. in 2002

that lead to the pollinization of fields of non-transgenic maize with the pollen of corn designed to produce a pig vaccine, and the growth of transgenic maize in a soybean field previously used for transgenic experimentation (Ellstrand, 2003). However, other examples indicate the value of the strategy, such as Ventria Bioscience currently produce human lactoferrin (rhLF) in rice seeds from open field grown plants (Bethell and Huang, 2004). Nevertheless, the production in open fields have a key limitation regarding certifications: in general, good manufacturing practices (GMP) cannot be easily achieved due to the variability of production conditions. GMP certification is required for the commercialization of pharmaceutical products.

Figure I2. Images showing *N. benthamiana* culture in greenhouses for the production of recombinant proteins (Palmer and Cooper, 2014).

The use of contained environments is gaining relevance because it allows greater control of the production and reduces the risk of security breaks. Several processes under confinement have been awarded the GMP certification (Takeyama *et al.*, 2015) and clinical trials have even been developed using plant derived products (Landry *et al.*, 2010). For alternative applications which do not require GMP certification, as in the case of cosmetics, companies such as Plantaderma Biolab SL. and Orf Liftaekni Hf are commercializing plant produced recombinant proteins

Plant cell culture strategies

Growing plant cells *in vitro* was first experimentally demonstrated by Gottlieb Haberlandt in 1958 when he grew carrot cells as callus (Fehér, 2015). During the next decades, the industry pursued this technology to develop a series of products from cell cultures, mainly secondary metabolites with uses in cosmetics, food and others (Hara *et al.*, 2004; Hibino and Ushiyama, 1999; Sato and Yamada, 1984). One of the most

important milestones was reached in 2000, when an anticancer compound was produced in plant cell cultures and approved by the FDA. This product was paclitaxel that was commercialized by Phyton Biotech. Taxol, the active compound of paclitaxel, was produced using *Taxus brevifolia* cell cultures in 75,000 L stirred tank bioreactors (Wilson and Roberts, 2012). The approval of paclitaxel stimulated the study and development of other pharmacologically valuable compounds produced in cell cultures such as resveratrol, podophyllotoxin, genistein and lithospermic acid B (Routh and Nandagopal, 2017) . In parallel, other cell culture technologies were developed, including hairy roots, moss cultures, microalgae and duckweeds.

Recombinant proteins have also been produced in plant cell culture matrixes (Xu *et al.*, 2012), plant cell suspensions being the most used. These are induced from a starting plant material through different steps (Fig I3). First, usually after a sterilization procedure to eliminate microbial contaminants, fresh biomass (such as explants) is plated on culture media with different plant hormone combinations to induce callus formation. Different plants and different plant tissues require different hormone combinations for optimal callus induction. In addition, the same plant species can develop different callus types (organogenic or friable) depending on the hormone combinations. Besides the hormones, the media used for plant cell cultures includes macro- and micro- nutrients, and a carbon source such as sucrose or glucose. Most often the micro and macro nutrients are included in Murashige and Skoog media (MS) (Murashige and Skoog, 1962). The media can include other elements, but in most cases, they contain no animal derived components.

Transgenic plant cell culture suspensions can be obtained from two different sources: from callogenesis using transgenic plant material, or directly transforming cell suspensions. In the first case, the callogenesis is a fast process but the resulting cell lines need adaptation to liquid media and usually further optimization to achieve growing dynamics compatible with the industrial use. In the second case, previously optimized cell lines are transformed, and therefore the transformed cells adapt more easily to liquid media. The suspensions can be transformed by different methods, the most common being bombardment of the nucleus (biolistic delivery of genetic material) and Agrobacterium-mediated transformation. When using Agrobacterium for transformation, the number and position of gene insertions occur randomly in the host genome so requiring screening to identify the most adequate lines (Kirchhoff *et al.*, 2012; Magy *et al.*, 2014). This variability can be a problem because usually cells do not grow isolated but in small clusters, and even every cluster can contain different transformants with diverse production levels. Given that cell clusters are difficult to separate during propagation, the resulting

suspensions can be heterogeneous, containing a mosaic of cells producing different amounts of recombinant proteins and having different growth rates. Cells with high growth rates but low production levels can gradually dominate the population leading to a reduction of recombinant protein yield (James and Lee, 2006). Strategies to mitigate this problem are currently being studied, including isolation of transformed cells by protoplast and cell line regeneration (James and Lee, 2006) and the use of elements to target the insertion of the transgenes (Allen *et al.*, 2000).

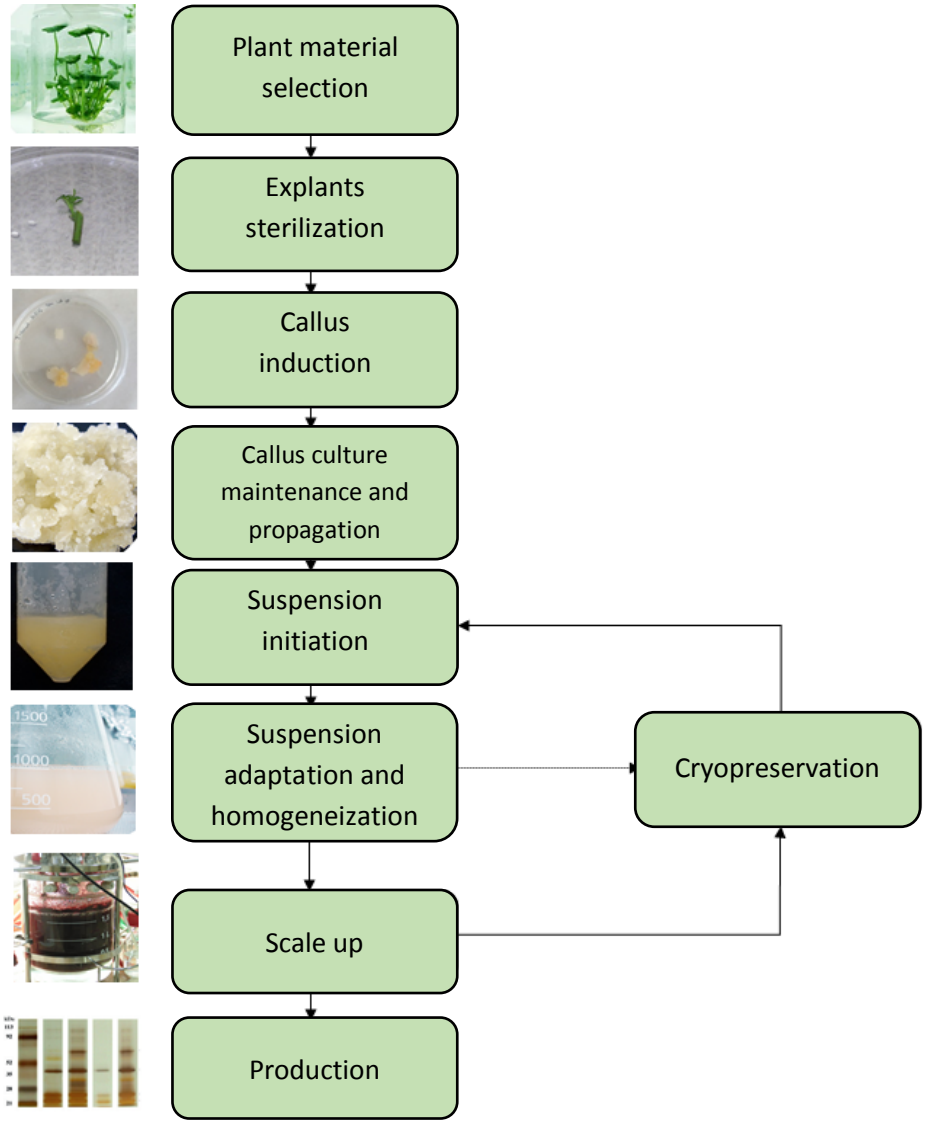


Figure I3. Schematic representation of the procedure for plant cell culture suspension establishment. Arrows indicate progression of the process.

Once the transgenic cell lines have been established and evaluated, the next step of the industrial process is to scale up. Plant cell cultures can be grown in systems including wavebags, orbital shaking vessels, and standard stirred tank bioreactors (Ferrari, 2010). There are also tailor-made options such as the Protalix BioTherapeutics cultivation system: air mixed polyethylene bags for the culture of *Daucus carota* cells demonstrate the versatility and robustness of plant cell culture growth (Tekoah *et al.*, 2015). Although using stirred tank bioreactors is a well-established standard in the industry, the use of single use bioreactors, especially wavebags, is an emerging trend. These bioreactors use pre-sterilized culture bags for up to 500 l working volume (Wave bioreactor 500/1000, GE Healthcare™) and do not require expensive investment in equipment. Companies such Vytrus Biotech S.L. have demonstrated the compatibility of wavebag bioreactors in the culture of diverse plant cell lines at an industrial scale.

During the scale-up process, growth parameters need to be optimized. The parameters that could be limiting are aeration, mixing and, in some cases, media composition (Gao and Lee, 1992). For each cell line, the parameters have to be optimized experimentally, and even two cell lines derived from the same plant material can have different requirements. Furthermore, cell lines can change their optimal growing needs and characteristics over time. To avoid variation in growth and production dynamics during successive propagation cycles, the most common strategy is cryopreservation. However, and unlike the case of microbial cells, cryopreservation in plant cell cultures is a complex process with very low success rates. Morphological characteristics of plant cell cultures, including the properties of the plant cell wall, and the high-water content seem to be responsible for the complexity of the process. Different cryopreservation protocols have been developed and tested, but most are only valid for model cell lines, with suboptimal results in other plant species (Mustafa *et al.*, 2011; Ogawa *et al.*, 2012). Protocols can be based on slow freezing, vitrification, and encapsulation or dehydration approaches. The gold standard for cell culture preservation is the approach based on controlled-rate slow freezing.

Transient and stable recombinant protein production

Regardless of the plant matrix used for recombinant protein production in plants, two expression strategies are currently available: transient or stable. In transient expression, the transgene of interest is delivered into somatic tissues of the plant host and expressed during a certain timeframe. Then the plant material is processed, and the host eliminated. The transgene is integrated in the plant germline, so no

transgenic plant line derives from the transformed material, and it is considered that no transgenic plants or cell cultures are involved in transient expression, leaving this strategy outside the umbrella for regulation of transgenics. The main method used in transient strategy is agrobacterium-mediated transformation, successfully used for the expression of recombinant proteins in plants (Kopertekh and Schiemann, 2019) and plant cell cultures (Rademacher *et al.*, 2019). On the other hand, with strategies based on stable expression in plants and plant cell cultures the transgene of interest is delivered into the host with the main objective of propagating it in the progeny. This strategy allows the development of plants with the transgene integrated into their genome, that can be expressed constitutively in all the tissues and throughout the life cycle of the plant, or during a certain timeframe in certain tissues if using inducible promoters. Stable expression has important legal limitations, and also often has a bad perception in society (Fischer *et al.*, 2013; Wunderlich and Gatto, 2015). Other limitations can derive from the toxicity or undesired side effects of the recombinant proteins in key processes, such as host development. As with transient expression strategies, the transgenes can be delivered via agrobacterium, by bombardment or alternative methods.

Plasmids for recombinant protein production

The DNA material for plant transformation needs different characteristics depending on the method selected for transgene delivery. In bombardment, the transgene cassette or the plasmid carrying it is coated onto washed, disaggregated tungsten or gold particles for biolistic transformation (Taylor and Fauquet, 2003). Additional co-adjuvants are used in the reaction. In agrobacterium-mediated transformation the bacteria from the *Agrobacterium* genus contains a plasmid, known as the Ti plasmid, which encodes the information to transfer the DNA sequences from the T-DNA fragment to the host cell. In nature, Agrobacterium enters through natural plant wounds and delivers the T-DNA of the Ti plasmid into the plant cell, which is inserted into the plant genome and provokes the development of a tumor in the plant host. The tumor produces a certain environment and substances that the bacteria uses to proliferate (Chilton *et al.*, 1977; Tzfira and Citovsky, 2003). When using agrobacterium for transgene insertion, much of the T-DNA region is substituted by the sequences of interest to be inserted in the plant. The Agrobacterium strains that are used for plant transformation carry a modified T-DNA without the tumor-controlling genes, referred to as “disarmed” strains (Bahramnejad *et al.*, 2019; Bourras *et al.*, 2015; Krenek *et al.*, 2015). The plasmids used by Agrobacterium for delivery of transgenes are called binary

plasmids because the virulence genes required for the transfer of T-DNA are dissociated in a helper plasmid (Komori *et al.*, 2007).

Plant viral vectors for the production of recombinant proteins

Various plant virus genomes and sequences derived from them can be modified and used to develop binary plasmids for the production of recombinant proteins, taking advantage of the highly efficient replication capacities of viruses to drive the expression of foreign genes. These plasmids have been found to significantly improve the yields of recombinant protein yields showing their potential for large-scale production (Gleba *et al.*, 2007; J. A. Lindbo, 2007; Shi *et al.*, 2019), and are currently used to produce vaccine protein antigens, antibodies and recombinant proteins (Hefferon, 2017).

Plant virus-derived expression vectors can be based on different viruses, including those with single stranded DNA or positive-sense RNA genomes (Gleba *et al.*, 2004, 2007; John A. Lindbo, 2007; Shi *et al.*, 2019). Initially the expression vectors used carried full-length virus genomes, with the transgene of interest inserted in the viral sequence an extra gene product. In these approaches, the transgenes can be expressed as fusion proteins with other viral proteins or cloned independently using for instance an additional strong subgenomic promoter. These vectors trigger the production of virions that cannot carry large transgene sequences, theoretically could infect a whole plant, and could spread to close plants if no confinement measures are adopted. TMV-derived vectors were developed later using DNA modules delivered by *Agrobacterium tumefaciens* and assembled in vivo by a site-specific recombinase to produce the TMV-derived vector (Marillonnet *et al.*, 2004). This de-constructed system was improved by introducing silent nucleotide substitutions and multiple introns into the TMV coding sequences, in order to increase infectivity (Marillonnet *et al.*, 2005). Another series of TMV-derived vectors developed were aimed at limiting the risks of virus spread and the need to use different DNA modules leading to the development of “disarmed” viral vectors by removing viral genes that were not strictly necessary for the production of the recombinant protein, such as those coding for the viral coat protein (CP) or movement proteins (MP), and replacing them with one or more transgene sequences (Peyret and Lomonossoff, 2015). An important step forward in the development of this disarmed virus-based expression system was the use of Agrobacterium for initiating infections. In addition, the elimination of elements necessary for the movement and assembly of the virions reduced the risk of spread and limited the expansion of the infection in the infiltrated plants. Additional developments now allow the use of viral vectors for the

production of virions as drug carriers, as in the vectors derived from *Cowpea mosaic virus* (Bunney *et al.*, 2017).

The main viral vectors derive from members of four different genera of plant viruses: *Tobamoviruses, Potexviruses, Comoviruses* and *Geminiviruses*, although there are descriptions of other groups.

The *Tobamovirus* genus includes the first virus used for the deconstructed vector system, *Tobacco mosaic virus* (TMV) (Gleba *et al.*, 2007). This virus has been used to develop multiple viral vectors, for instance after its coat protein (CP) has been totally replaced by an open reading frame (ORF) of the foreign gene (John A. Lindbo, 2007), or partially maintaining key regions for the function of subgenomic viral replication (Shi *et al.*, 2019) (Fig I4) . Several proteins have been produced using vectors derived from this virus, including vaccines (Cummings *et al.*, 2014; Mansour *et al.*, 2018), anti-microbial peptides (Shi *et al.*, 2019) and tools to fight epidemics such as Ebola (Lam *et al.*, 2016).

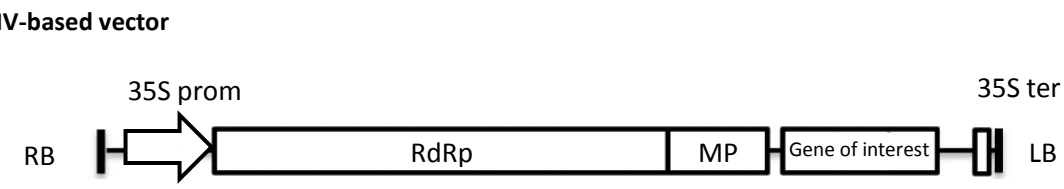


Figure I4. Schematic representation of a TMV-based binary vector. The expression of viral proteins and transgene are driven by *Cauliflower mosaic virus* (CaM) 35S promoter and terminator. The transgene replaces the CP sequence.

In the case of the *Potexvirus* genus, *Potato Virus X* (PVX) is the member most widely used. Diverse proteins have been produced using PVX derived vectors, including epitopes and vaccines (Mardanova *et al.*, 2017), and fusion proteins for the treatment of multiples diseases (Mardanova *et al.*, 2015; Uhde-Holzem *et al.*, 2010).

Comoviruses, with *Cowpea mosaic virus* as the main representative, have been extensively used in the production of virus-like particles (VLPs) and other antigen elements, deconstructed vectors (Montague *et al.*, 2011; Saunders *et al.*, 2009) and nanoparticles for cancer investigation and treatment (Beatty and Lewis, 2019; Bunney *et al.*, 2017).

The family *Geminivirus* includes many single-stranded circular DNA viruses with a broad host range, making them interesting for the development of expression vectors, although some limitations exist.

Viruses of genera in the family, such as the *Bean yellow dwarf virus* (BeYMV) and *Beet Curly top virus* (BCTV) have been used to produce vaccines and antibodies (Chung *et al.*, 2008; Huang, 2018).

Downstream processing

Downstream processes used for the purification of recombinant protein are rather costly steps in the chain of production, mostly regardless of the host used for protein expression. In most cases the main differences in purification of products from different matrices are only in the first steps of the processes, while later key and expensive stages, such as chromatographic purification, are not dependent on the production platform used (Fischer *et al.*, 2012). On average, downstream processes are estimated to account for up to 80% of the costs of manufacturing recombinant proteins (Menkhaus *et al.*, 2004; Raven *et al.*, 2015), so cheaper alternatives are constantly being explored. In the case of productions using plants, the recombinant proteins can be accumulated in the cytoplasm, in cell compartments such as the ER, or exported to the extracellular fluid (ECL). If the recombinant protein is retained in the cells, the biomass requires processing, followed by extraction and exhaustive purification steps, that increase total costs. On the other hand, the secretion of the produced proteins to the ECL could allow for easier protein purification and reduce production times and costly steps, also without the need for major investment in equipment. When recombinant proteins are produced in plant cell lines, they can be secreted to the media culture or retained in the cell. Extraction and purification are easier if the product is secreted in the media, even allowing continuous manufacturing in some cases. If the proteins are retained in the cells, different steps are needed to purify them: biomass separation, disruption of cell walls, clarification and elimination of plant contaminants. Disruption of cell walls can be achieved using enzymatic digestion, sonication, wet milling and pressure homogenization, although all the strategies have limitations when a scale-up is needed (Xu *et al.*, 2011). Additional costs are associated with measures to prevent protein modification and degradation such as the addition of protease inhibitors and/or changes in temperature to slow down the breakdown. In some cases, plant cells accumulating the recombinant proteins are used as the delivery system, as in the case of Protalix BioTherapeutics that uses the cells to deliver the protein glucocerebrosidase (GCD) for the treatment of Gaucher disease (Tekoah *et al.*, 2015).

LIPID DROPLETS

Plants need to store energy to be used during certain steps of their development, as in seed and pollen germination. Plants use different molecules as energy deposits: carbohydrates, proteins and lipids, due to their high energetic value and anhydrous profiles being the most convenient molecules for energy storage. Their content varies depending on the tissues and ranges from ~ 40% of dry weight in oilseeds (Mansour *et al.*, 2014) to approximately 0.05% in plant leaves (Xu and Shanklin, 2016). Lipids are typically accumulated in the form of triacylglycerols (TAGs). In higher plants, TAG synthesis begins mainly in plastids, where fatty acids (FA) are synthesized from acetyl-Coenzyme A (acetyl-CoA). The fatty acids are exported and modified in the endoplasmic reticulum (ER), where they are used as precursors for the synthesis of TAGs, phospholipids, galactolipids, and other lipid derived components (Baud and Lepiniec, 2010; Cagliari *et al.*, 2011; Dyer *et al.*, 2008).

TAG synthesis is a highly regulated process that involves a number of transcription factors and enzymes. Among them, WRINKLED 1 (WRI1), a transcriptional factor that regulates the derivation of carbon sources into FA synthesis (Cernac and Benning, 2004), and diacylglycerol O-Acyltransferase 1 (DGAT1), a protein that catalyzes the conversion of diacylglycerol and fatty acyl CoA into triacylglycerol, are of great importance (Vanhercke *et al.*, 2017). In the ER the TAGs are stored in the acyl moiety region between the bilayers of phospholipids (Yen *et al.*, 2008), from where they bud as a single membrane vesicles named oil bodies, oleosomes or lipid droplets (LDs) (Fig I5a). LDs are storage compartments found mostly in seeds and senescing leaves, formed by TAGs, phospholipids and structural proteins. These structural proteins are grouped in three families: oleosins, caleosins, and steroleosins (Himada and Ishimura, 2010; Lin *et al.*, 2002; Tzen *et al.*, 1990). Classically, LDs have been considered as storage organelles, however they have emerged in recent years as fully entitled organelles with key functions in cell signaling and lipid homeostasis (van der Schoot *et al.*, 2011; Shimada and Hara-Nishimura, 2015).

Other non-plant organisms, as yeast and mammals, also accumulate LDs. However, the lipid and protein composition of these LDs are different from those produced in plants. While the plant LDs lipid core is formed by TAGs, in yeast and mammals steryl esters are also accumulated in the form of LDs. Mammals and yeast do not express oleosins but have other structural proteins on the surface of the LDs (Goodman, 2008; Martin and Parton, 2006; Rajakumari *et al.*, 2008). These surface proteins, unlikely oleosins, do not have a hydrophobic domain, suggesting that they do not penetrate the lipid matrix.

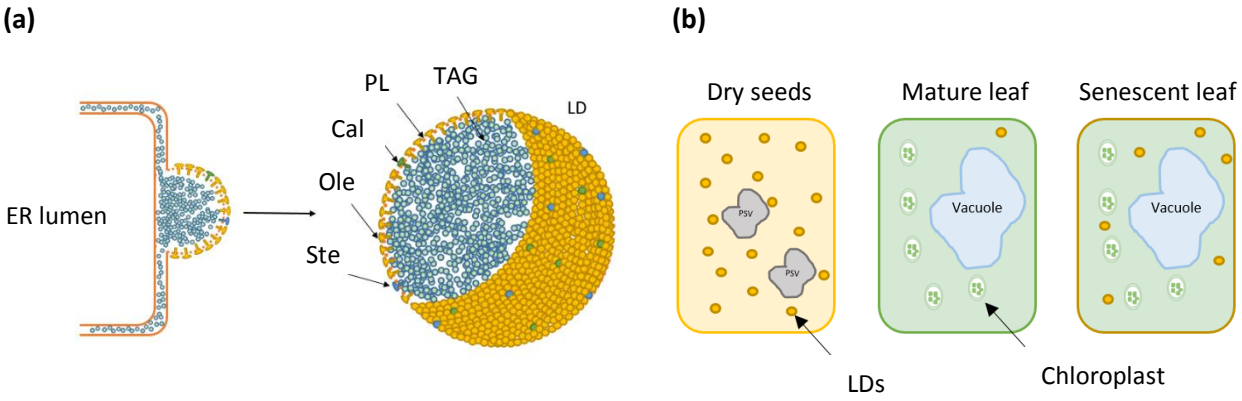


Figure I5. Schematic representation of LD formation and accumulation during development. **(a)** Budding LD on rough ER during LD formation. The structure includes the ER lumen, two phospholipid layers (PL, orange), TAGs (blue), oleosin (Ole, yellow configuration), caleosin (Cal, green configuration) and steroleosin (Ste, blue configuration). The constituents are not drawn to scale. **(b)** LDs are abundant in dry seed cell, practically non-existent in mature leaf cells, and relatively abundant in senescent leaf cells. Vacuole, protein storage vacuole (PSV), chloroplasts and LDs are represented.

LDs are highly accumulated in dry seeds. During seed germination, they are actively degraded through the mobilization of TAGs to peroxisomes where FAs are used to obtain energy and structural components through β-oxidation. In senescent leaves, an increase in LD biogenesis occurs that leads to its accumulation (Fig I5b) (Shimada *et al.*, 2015). LD accumulation also occurs in plant cells in response to biotic and abiotic stresses (Gidda *et al.*, 2016).

Oleosins

Oleosins are the main structural proteins of LDs. They have only been localized in LDs and ER cisterns. Their molecular weights range between 15 and 26 KDa (Huang and Huang, 2015). Structurally, oleosins have three domains: the N- and C-terminal amphipathic domains that reside on the LD surface in contact with the cytoplasm, and a central hydrophobic domain anchored inside the LDs (Fig I6) (Abell *et al.*, 1997; Chuang *et al.*, 1996; Li *et al.*, 2002). The central domain is formed by ~70-80 residues with a hairpin structure formed by two different regions of ~30 amino acids. The presence of oleosins in LDs stabilizes these organelles via steric hindrance and electronegative repulsion, preventing them from coalescence (Huang, 2018). They have been used as oil stabilizers (Wijesundera *et al.*, 2013) and for the production

and purification of heterologous proteins in bacteria (Peng *et al.*, 2004), yeasts (Popa *et al.*, 2019) and plants.

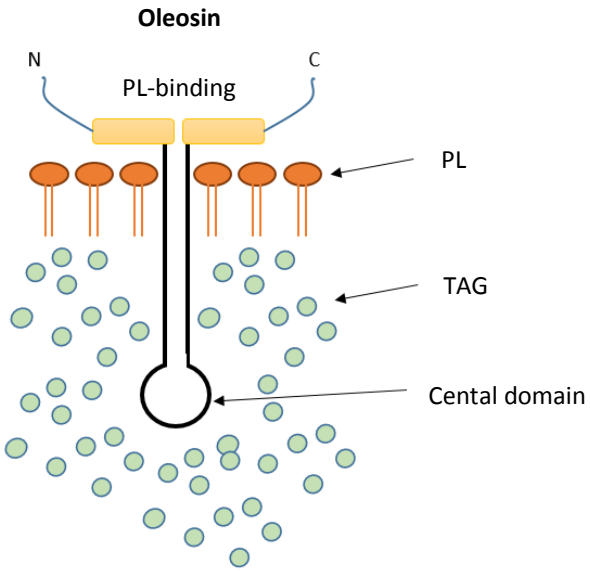


Figure I6. Schematic representation of plant oleosin structure. The representation includes the phospholipids layer (PL, orange), TAGs (blue), oleosin N- and C-terminal, phospholipids binding regions and central domain. The constituents are not drawn to scale.

Production of proteins as oleosin fusions in plants

As previously mentioned, up to 80% of the cost of recombinant protein production is for product purification (Menkhaus *et al.*, 2004; Raven *et al.*, 2015). Therefore, alternatives to the classical protein accumulation in cellular cytoplasm and compartments are highly in demand to reduce production steps and costs.

The physicochemical properties of oleosins and their association with LDs have made them great candidates for the production and purification of recombinant proteins. It is well known that oleosins function as carriers to LDs when proteins are produced as oleosin fusion proteins. Once incorporated into LDs, they can be easily purified by flotation on density gradients (Van Rooijen and Moloney, 1995). The potential of oleosins as recombinant protein carriers was first demonstrated in the production of β-glucuronidase enzyme in *Brassica napus* seeds (Van Rooijen and Moloney, 1995). Further studies proved that the oleosin technology was very efficient for the production of recombinant proteins with industrial interest: hirudin in oilseed rape (Parmenter *et al.*, 1995); apolipoprotein AI in safflower seeds (Nykiforuk *et al.*, 2011); the human growth hormone (hGH) and insulin in *A. thaliana* seeds (Boothe *et al.*, 2010; Nykiforuk *et al.*, 2006); antimicrobial peptides in rice seeds (Bundó *et al.*, 2019; Montesinos *et al.*, 2016),

and other recombinant proteins in different plants, and even yeasts and bacteria (Table I2). In most cases the proteins of interest were linked through proteolytic cleavage sites that allow the release of the recombinant proteins, although in some cases the LDs carrying recombinant fusions have been used as the functional final products (Cai *et al.*, 2018).

Table I2. Current hosts for recombinant protein production using oleosin technology

Host	Foreign protein	Reference
Brassica napus	Maize 18 kDa oleosin	(Lee *et al.*, 1991)
	Hirudin	(Parmenter *et al.*, 1995)
	β-Glucoronidase	(Van Rooijen and Moloney, 1995).
	Xylanase	(Liu et al., 1997)
	Soybean 24 kDa oleosin	(Sarmiento *et al.*, 1997)
	Human erythropoietin	(Newman, 2011)
	Xylanase (XynC)	(Liu *et al.*, 1997)
Brassica carinata	Hirudin	(Chaudhary *et al.*, 1998)
	β-Glucoronidase	
Nicotiana tabacum	Green fluorescent protein	(Wahlroos *et al.*, 2003a)
Arabidopsis thaliana	Insulin	(Nykiforuk et al., 2006)
	Human growth hormone	(Boothe *et al.*, 2010; Nykiforuk *et al.*, 2006)
	Green fluorescent protein	(W. Li et al., 2012)
	di-or tri-meric fusions with GFP	(Banilas *et al.*, 2011)
	Hyaluronidase hPH-20	(Li et al., 2014)
	Human acidic fibroblast growth factor	(Yang *et al.*, 2015)
	Mice fibroblast growth factor 9	(Yi et al., 2015)
	Trimeric version of Arabidopsis thaliana polyoleosin 18kDa (three head-to-tail repeats of oleosin)	(Winichayakul *et al.*, 2012)
	Single-chain variable fragments (scFv) against epitopes from infective Trichostrongylus colubriformis L3	
	Monomeric red fluorescent protein (mCherry)	(Heilmann *et al.*, 2012)
	Mouse fatty acyl-CoA reductase	
	Mouse wax synthase	

Carthamus tinctorius	Apolipoprotein AI	(Nykiforuk et al., 2011)
	Human fibroblast growth factor 9	(Cai *et al.*, 2018)
	Human fibroblast growth factor 10	(Li et al., 2017)
Oryza sativa	cecropin A	(Montesinos *et al.*, 2016)
	PAF102	(Bundó et al., 2019)
	Sesame oleosin 15 kDa	(Lee *et al.*, 2006)
	Sesame 2S albumin	
	Green fluorescent protein	(Shimada *et al.*, 2011)
Arachis hypogaea	Human pro-insulin	(Ling et al., 2016)
Escherichia coli	Nattakinase	(Chiang *et al.*, 2005)
	Hydantoinase	(Chiang et al., 2007)
	Sesame oleosin 15 kDa	(Hsieh *et al.*, 2017) (C. C. Peng *et al.*, 2004)(C.-C. Peng *et al.*, 2004)
	Mastoparan B	(Hsieh et al., 2017)
	Sesame cystatin	(C. C. Peng *et al.*, 2004)
	Green fluorescent protein	(C.-C. Peng *et al.*, 2004)
	Royalisin	(Tseng *et al.*, 2011)
	D-Psicose 3-epimerase	(Tseng et al., 2014)
Saccharomyces cerevisiae	Maize 18 kDa oleosin	(Ting *et al.*, 1997)
	Sunflower oleosin 1	(Beaudoin et al., 2000)
	Peanut oleosin 3	(Parthibane *et al.*, 2012)
	Mouse-ear cress oleosin 3	(Vindigni et al., 2013)
	Monomeric red fluorescent protein (mCherry)	(Heilmann *et al.*, 2012)
	Mouse fatty acyl-CoA reductase	
	Mouse wax synthase	
	Green Fluorescent protein	(Vindigni et al., 2013)
Pichia pastoris	Rice 18 kDa oleosin	(Popa *et al.*, 2019)
	Green fluorescent protein	
	PAF102	
Yarrowia lipolytica	Sesame oleosin 15 kDa	(Han et al., 2013)
	Monomeric red fluorescent protein (mCherry)	
	Scaffoldin CipC fragments	
	Green fluorescent protein	

GENE SILENCING

Plants have different mechanisms to defend themselves against pathogen aggression. One of the first layers of defense are physical barriers to impede the entrance and colonization of pathogens. These physical barriers include exterior surfaces with waxy cuticles and preformed antimicrobials that prevent the entry of pathogenic entities. Plant cell walls also act a defense against microbes to impede pathogen penetration. However, the plant can detect those pathogens that overcome these barriers and activate its immune response. Plant immunity can be divided into two components operating on different time scales. The basal defense system appears early in pathogen interaction, while the resistance (*R*) gene-mediated defense operates on the time scale of hours. Basal defense response is mediated by PAMPs (pathogen-associated molecular patterns), which include lipopolysaccharide, peptidoglycan, bacterial flagellin, and mannans of yeast, and others (Nürnberger *et al.*, 2004), that trigger the production of a series of plant compounds to fight against infection. Gene-mediated defense is inherited and is specific to a particular pathogen or pathogenic group. Plants have dominant *R* genes whose products can recognize those of the pathogen's complementary *Avr* alleles. Avr proteins are effector proteins secreted into the plant cell to promote pathogen virulence and to overcome host defenses (Grennan, 2006). Once the pathogens are detected, plants express a series of factors to impair or eliminate the invaders. Among gene strategies to fight infections, plants have developed gene silencing as a defense mechanism to fight against virus infections and control transposable elements, mediated by small RNAs as a key component.

Gene silencing is a process conserved in all eukaryotic organisms from fungi (Nakayashiki, 2005), to insects (Vogel et al., 2019), plants (Pumplin and Voinnet, 2013) and mammalian cells (Schuster et al., 2019). In plants, in addition to its importance for pathogen defense, this mechanism is involved in several biological processes, such as development and adaptation to stress conditions.

Short interfering RNAs (siRNA) are small RNAs involved in the silencing process that can act at two expression levels: during transcription ("Transcriptional Gene Silencing", TGS) or in post-transcriptional nucleic acids processing ("Post-Transcriptional Gene Silencing", PTGS). Viral replication generates certain amounts of double stranded RNAs (dsRNAs) on infecting cells in different situations: during dsRNA viruses infection (Martin *et al.*, 2011; Sabanadzovic *et al.*, 2009, 2010); with single-stranded RNA viruses as a result of secondary structures, or replication intermediates (Lacomme *et al.*, 2005; Lu *et al.*, 2003) and by overlapping RNA transcripts derived from the transcription of DNA viruses (Ding and Voinnet, 2007). The dsRNAs induce the production of virus-derived small interfering RNAs (vsRNAs or vsiRNAs) (Ding and

Voinnet, 2007), limiting viral infection or modulating host gene expression through the interaction with plant DNAs and RNAs (Ruiz-Ferrer and Voinnet, 2009).

The silencing process starts with the recognition of dsRNAs by DICER-LIKE (DCL), a type III RNAse. This recognition triggers the processing of the RNA molecule in small fragments (vsRNAs or vsiRNAs), around 20-25 bases. The vsiRNAs are recognized by Argonaute (AGO) protein and loaded onto an RNA-induced silencing complex (RISC), acting as guide molecules to target homologous RNA for degradation (Fig I7). The vsiRNAs are also amplified using dsRNAs synthesized by plant RpRD proteins. The amplified vsiRNAs are transported through the phloem conferring systemic silencing to the virus infection that triggered the response (Choudhary *et al.*, 2019; Voinnet, 2008).

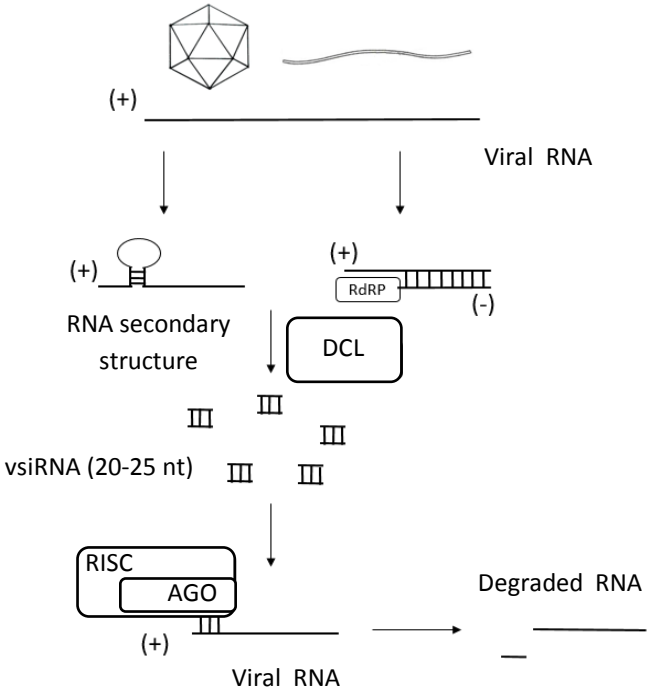


Figure I7. Schematic representation of RNA virus-induced gene silencing in plants. The abbreviations correspond to: RdRP, RNA dependent RNA polymerase; DCL, Dicer-like protein; RISC, RNA-induced silencing complex; AGO, Argonaute protein.

Viral silencing suppressors

During their coevolution as pathogen and hosts, plant viruses have developed strategies to avoid plant silencing processes. The main strategy is based on blocking RNA silencing mechanisms through the expression of RNA silencing suppressors (RSS). These viral proteins are divided into categories based on how they operate: (I) binding to the triggering dsRNAs; (II) interference with the binding of DCLs to dsRNAs; (III) sequestration of siRNAs; and (IV) inhibition of activity of AGOs and/or RISC assembly (Li and

Wang, 2019) . RSSs have diverse sequences and share no structural similarities between major plant virus groups. They are usually multifunctional gene products, and can be key components of virus activities or helpers playing essential roles during infection, in addition to its adopted role of RSSs (Csorba *et al.*, 2015; Yang and Li, 2018). In some cases, the silencing suppressor activity is not restricted to a single virus protein, with different proteins from the same virus being able to act against different silencing-related processes (Díaz-Pendón and Ding, 2008; Giner *et al.*, 2010; Pumplin and Voinnet, 2013).

Potyviruses as source of RSS and biotechnological tools

The *Potyviridae* family currently contains almost 200 members according to the International Committee on Taxonomy of Viruses (ICTV). It is the second largest taxonomic group of plant-infecting viruses only after the family *Geminiviridae* (DNA genomes), and is considered the largest family of plant RNA viruses (Scholthof *et al.*, 2011; Wylie *et al.*, 2017). The *Potyviridae* family includes at least ten different genera, with *Potyvirus* the most highly represented. The genera are divided according to their genome organization, sequence similarity and natural transmission vector. Over 170 members of the *Potyvirus* genus have been described while there are less than 10 in the other genera (Table I3).

Table I3. Table including the different *Potyvirus* genus genera

Genus[1]	Vector type	Genome	Number of described members
Bevemovirus	To be defined	Monopartite	1
Brambyvirus	To be defined. Potential aerial vector	Monopartite	1
Bymovirus	Fungoid protist	Bipartite	6
Ipomovirus	Whiteflies	Monopartite	8
Macluravirus	Aphids	Monopartite	9
Poacevirus	Mites	Monopartite	3
Potyvirus	Aphids	Monopartite	>170
Roymovirus	To be defined	Monopartite	1
Rymovirus	Mites	Monopartite	3
Tritimovirus	Mites	Monopartite	6

[1] Data from ICTV, 2018b Release. (https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/potyviridae)

All the members of the *Potyviridae* family have flexuous, filamentous viral particles around 680-900 nm length and 11-14nm diameter (Kendall *et al.*, 2008; McDonald *et al.*, 2010) (Fig I8a). A characteristic structural trait of these viral particles are the formation of cylindrical crystal inclusions in the cytoplasm of the infected cells (Fig I8b), known as "pinwheels" due to their morphology (Hollings, 1981; Rubio Huertos, 1966).

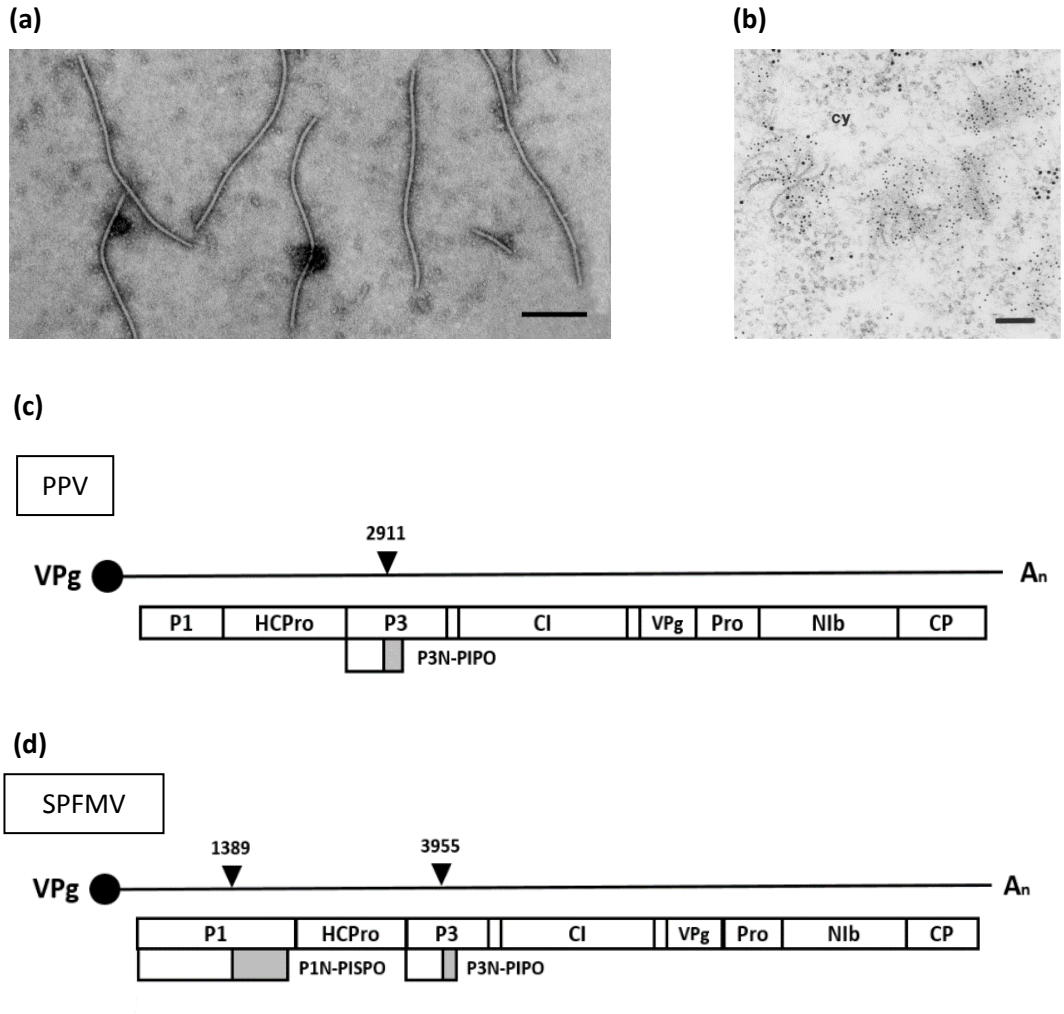


Figure I8 Viral particles and cytological alternations in the host of members of the *Potyviridae* family, and genome organization of potyviruses. **(a)** Negative staining electron micrograph of particles of an isolate of *Plum pox virus*, stained with 1% PTA, pH 6.0. Bar = 200 nm. Obtained from ICTV website 2019 (https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/potyviridae). **(b)** Immunodetection of CI proteins in pinwheels detached from the cell wall (from Rodriguez-Cerezo et al. 1997). Bar = 100 nm; cy = cytoplasm. **(c-d)** The genomes of the *Plum pox virus* (PPV) **(c)**, and *Sweet potato feathery mottle virus* (SPFMV) **(d)** are shown schematically. Processed gene products are shown as boxes. Additional ORFs corresponding to out-of-frame PIPO and PISPO regions are also indicated. Black triangles show the polymerase slippage position.

The genome organization of *Potyviridae* family members, is based on encoding a single major polyprotein in their viral RNA. *Bymovirus* encodes two polyproteins in its bipartite genome. The polyprotein undergoes post-translational processing by proteolytic cleavages caused by dedicated viral proteases encoded in its own polyprotein. The proteolytic cleavage potyvirids generate at least ten different protein products, from the N to C-terminus: Protein 1 (P1), Helper Component-Protease (Hc-Pro), Protein 3 (P3), 6-kDa peptide 1 (6K1), Cylindrical Inclusion protein (CI), 6-kDa peptide 2 (6K2), Viral Protein genome-linked (VPg), Nuclear inclusion A-protease (NIa-Pro), Nuclear inclusion B (NIb) and Coat protein (CP). An additional viral protein is expressed by all the *Potyviridae* family members P3N-PIPO. In contrast to the other viral proteins processed from the major polyprotein, P3N-PIPO expression is the result of a polymerase slippage mechanism that provokes a change in the reading frame of the P3 protein (Chung *et al.*, 2008; Olspert *et al.*, 2015) (Fig I8c). The slippage occurs at the conserved motif $G_{1\text{-}2}$ $A_{6\text{-}7}$, upstream of the conserved PIPO region. An additional protein specific to sweet potato-infecting potyviruses is also produced by polymerase slippage, P1N-PISPO (Li *et al.*, 2012; Rodamilans *et al.*, 2015; White, 2015) (Fig I8d). Figure I8 c-d show a schematic representation of the genome of two members of the *Potyviridae* family: *Plum pox virus* (PPV) and *Sweet potato feathery mottle virus* (SPFMV). The expression of P3N-PIPO is represented for PPV and that of P1N-PISPO and P3N-PIPO for SPFMV.

Each viral product has its own activity, although some of them are helpers and additional activities are currently being discovered. Generally speaking, we consider that:

- P1 (Protein 1): It might have significant roles in viral infection, for instance by stimulation of the gene silencing suppressor HC-Pro. It contains a serine protease domain that autocatalytically cleaves its C-terminus to liberate the P1 product from the polyprotein, typically at Tyr/Phe-Ser (Valli *et al.*, 2007).
- HC-Pro (Helper Component-Protease): The HC-Pro protein has many roles in viral plant-to-plant transmission, polyprotein maturation and RNA silencing suppression (Valli *et al.*, 2018). A cysteine protease domain towards the C-terminus cleaves itself from the rest of the polyprotein, usually at a Gly-Gly site.
- P3 (Protein 3): It is essential for the viral replication complex, as demonstrated with *Soybean mosaic virus* (SMV) (Cui *et al.*, 2018).
- 6K1 (6-kDa peptide 1): It is present in the replication complex, and it is considered to have a specific function of anchoring elements to replication sites.

- CI (Cylindrical Inclusion protein): It is a key protein in pinwheel formation, and is involved in replication as a helicase, and in cell-to-cell and long-distance movements, among other activities (Sorel *et al.*, 2014).
- 6K2 (6-kDa peptide 2): This protein is involved in the replication complex structure probably attaching the complex to the ER.
- VPg (Viral Protein genome-linked): It is attached to the 5'-terminus of the genome (Murphy *et al.*, 1991), interacting with one or several isoforms of the eIF4E translation initiation factor crucial for viral infection (Charron *et al.*, 2008; Leonard *et al.*, 2000).
- NIa-Pro (Nuclear inclusion A-protease): It is responsible for polyprotein processing, as a serine-like cysteine protease, cleaving typically at Gln/Glu-Ser/Gly/Ala (Adams *et al.*, 2005).
- CP (Coat protein): A viral capsid protein involved in virus movement and transmission. It can be used to produce virus-like particles (VLPs) with nanobiotechnological interest for peptide presentation (González-Gamboa *et al.*, 2017).
- P3N-PIPO (Pretty Interesting Potyvirus ORF): It has been shown to be a key player in virus intracellular movement (Vijayapalani *et al.*, 2012).
- P1N-PISPO: It has viral RNA silencing suppressor activity (RSS) in the sweet potato-infecting virus analyzed (Mingot *et al.*, 2016; Rodamilans *et al.*, 2015; Untiveros *et al.*, 2016).

Diverse proteins derived from members of the Potyvirus genus have been used as biotechnological tools in research or for industrial purposes. Among them, the *Tobacco etch virus* (TEV) protease has been widely used and is a key component of recombinant protein purification processes thanks to its specificity (Shih, 2005). Also the use of Potyvirus-derived silencing suppressors are quite common and they are being used to enhance recombinant protein expression in plants (Arzola *et al.*, 2011).

Watermelon mosaic virus

Watermelon mosaic virus (WMV) is a Potyvirus with a wide range of plant hosts - more than 170 plant species from 27 different families (Mayo, 1995), including the cucurbitaceous, leguminous, malvaceous and chenopodiaceous families (Purcifull, 1979) generating huge production losses. Other hosts include wild grass species, that act as a reservoir for the virus (Lecoq H, 1992). As a Potyvirus, its genome includes the PN3-PIPO sequence. It was one of the first Potyvirus described as a cucurbit pathogen (Webb RE, 1965) and research is currently focused on its importance in mixed infections with other plant viruses.

Sweet potato feathery mottle virus

Sweet potato feathery mottle virus (SPFMV) is a Potyvirus that infects members of the Convolculaceae family (Tugume *et al.*, 2008). It infects sweet potato plants in most areas where they are cultivated (Clark *et al.*, 2012b), being a major problem in Africa with 40% of the cultivation worldwide. The SPFMV genome includes P3N-PIPO and PN1-PISPO with silencing suppressor activity (Mingot *et al.*, 2016; Rodamilans *et al.*, 2015; Untiveros *et al.*, 2016).

Cucumber vein yellowing virus

Cucumber vein yellowing virus (CVYV) is an Ipomovirus with a narrow host range, restricted to some cucurbitaceous plants and transmitted by the whitefly *Bemisia tabachi* (A. Mansour A. Al-Musa, 1993). This virus lacks the HC-Pro region, having a duplicated P1 sequence after polyprotein processing in two gene products, P1a and P1b, the second having a potent RNA silencing suppressor activity (Valli *et al.*, 2006a). P1b protein has the ability to functionally replace HCPro in potyviral infections of *Nicotiana* plants (Carbonell *et al.*, 2012).

PEPTIDE HORMONES AND DELIVERY SYSTEMS

Insulin-like growth factor 1 (IGF1)

The Insulin-Like Growth Factor 1 (IGF1) is a peptidic hormone having structural similarity with insulin, being encoded by the *IGF1* gene in humans (gene ID: 3479). The *IGF1* gene appeared by duplication of the pro-insulin gene before the evolution of the first vertebrates. It encodes a small peptide of 70 amino acids with three intramolecular disulfide bonds and a molecular weight of 7647 Da (Fig I9) (Rinderknecht and Humbel, 1978). It is mainly synthetized in the liver or peripheral tissues. In liver, IGF1 is secreted in response to growth hormone (GH) signaling that represents 75% of the peptide in blood. IGF1 availability is regulated through its binding to the insulin-like growth factor binding proteins (IGF1BP). These binding proteins affect the half-life of IGF1 and its interaction with the insulin-like growth factor 1 receptor (IGF-1R). Also, they can target IGF1 to a specific location or even modulate the action of IGF1 by direct affecting IGF-1R and other receptors involved in the hormone signaling. (Rosenfeld *et al.*, 1999). In the peripheral tissues, IGF1 secretion serves autocrine or paracrine purposes. Different cell types produce IGF1 in these tissues and its control is regulated by GH and other factors. Due to its close relation to GH, IGF1 is

considered the major mediator of GH-stimulated somatic growth. Regarding the receptors, IGF1 mainly binds IGF-1R however it can act on the insulin receptor (IR) among others (Andersen *et al.*, 2017; Boucher *et al.*, 2010). All its receptors have tyrosine kinase activity and are activators of the Akt pathway (Chitnis *et al.*, 2008; Singh *et al.*, 2014). IGF1 is implicated in several metabolic processes: growth and development; proliferation; lipid metabolism; ageing; inflammation; anabolism, and protection of mitochondria against oxidative damage (Aguirre *et al.*, 2016). IGF1 has been used for cosmetic and sanitary applications, including reduction of skin wrinkles and hair growth (Noordam *et al.*, 2013; Trüeb, 2018).

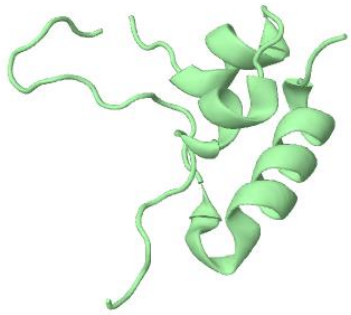


Figure I9. 3D representation of IGF1. The image was designed according to Protein Data Bank (PDB) code.

Cell penetrating peptides (CPPs)

The delivery of complex molecules into organisms for research and industrial purposes needs to surpass biological barriers in order to interact with the desired targets. The demand for tools for delivery systems has stimulated the interest in peptides with penetration activity in different cell lines. The first identified peptides/proteins with the ability to penetrate biological membranes and deliver a wide variety of cargos are the "Trans-Activator of Transcription" (TAT) from the *Human Immunodeficiency Virus* (HIV) and the penetratin protein derived from the homeodomain of Antennapedia (Prochiantz, 2000). Since then, cell penetrating peptides (CPPs) and protein transduction domains (PTDs) have been intensively studied. Their desired characteristics include a high penetration efficiency, being short and having low toxicity in most cases. The developed CPPs are classified in two categories according to their origin or their physicochemical properties (Milletti, 2012). Based on their origin, CPPs can be subdivided as synthetic, protein derived and chimeric; and according to their physicochemical properties, they are subdivided as cationic, amphipathic or hydrophobic peptides. They have different penetration strategies including direct

penetration, via the endocytosis pathway or translocation through the formation of a transitory structure (Fig I10). Depending on the CPP, their cargoes can include peptides, DNA, RNA and other organic or inorganic elements (Guidotti *et al.*, 2017; Kalafatovic and Giralt, 2017; Wang and Duan, 2019). This means that the use of CPPs as peptidic carriers in chimeric proteins could allow the increase in biodisponibility, delivery and even therapeutic potential of the fused proteins.

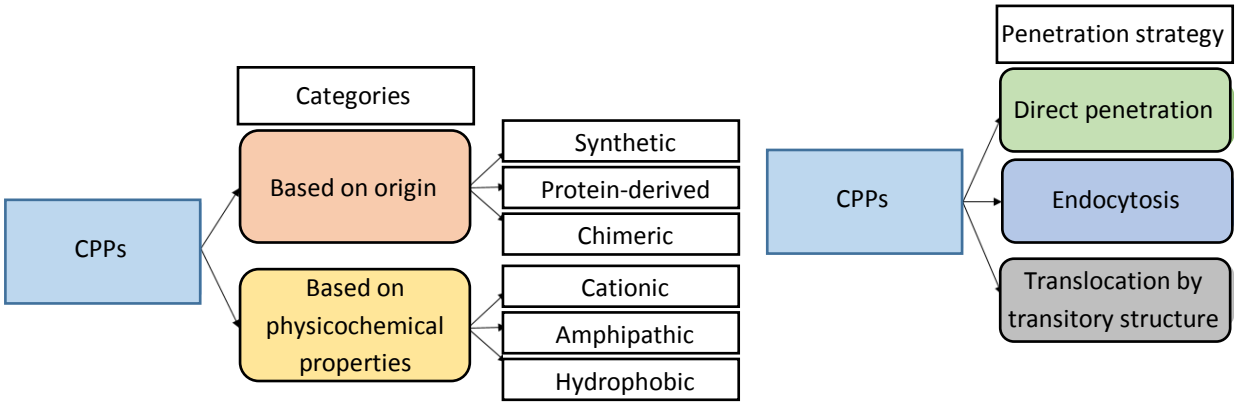


Figure I10. Classifications of CPPs according to origin, physicochemical properties or penetration strategy.

OBJECTIVES

OBJECTIVES

The general objective of this industrial thesis is to develop economic and efficient plant production systems for bringing proteins of interest to the market using the IGF1 as the target peptide. To accomplish this main objective, the following specific objectives were raised:

1. To develop biotechnological tools for enhancing gene expression or assisting product purification in plant production systems, including the selection of a viral RNA silencing suppressor and the preparation of gene constructs for oleosin fusion technology, as well as the establishment of a carrot cell line.

2. To produce the IGF1 and derivatives in carrot cell cultures and to establish and cryopreserve the producer cell line.

3. To produce IGF1 and derivatives in *N. benthamiana* plants using viral expression vectors.

4. To adapt the oleosin fusion technology to plant leaves for the efficient production of IGF1 in LDs.

5. To evaluate the biological activity of plant produced IGF1 and derivatives.

MATERIALS & METHODS

1. Biological material

1.1 Plant material

Different plant cell cultures were kindly provided by Vytrus Biotech S.L. (Vytrus) or by CRAG *in vitro* service (CRAG) (Table M1). The Vytrus cell lines were sampled at the end of the growing stage, according to Vytrus records, and delivered into CRAG facilities. *A. thaliana* and BY2, were propagated in two different culture media (Table M2 and M3), considering 1/4 dilution rate, with an agitation of 110 rpm and 22-24ºC growing temperature. BY2 cultures were propagated into darkness while *A. thaliana* cultures were propagated in a growing chamber with 16h light/ 8h darkness photoperiod. Both cell cultures were sampled at the end of their growing stage (when degree brix (ºB) levels were around 0).

Table M1. Plant cell lines provided by Vytrus.

Plant cell line	Origin
Arnica montana	Vytrus
Curcuma longa	Vytrus
Gossypium herbaceum	Vytrus
Cannabis sativa	Vytrus
Daucus carota	Vytrus
Sarcocapnos crassifolia	Vytrus
Nicotiana tabacum (BY2)	CRAG
Arabidopsis thaliana	CRAG

Table M2. *A. thaliana* cell culture media (MSAt)

	MSAt
Murashige and Skoog (MS) basal medium with vitamins (Duchefa™)	4.3 g/l
Sucrose (Duchefa™)	30 g/l
2-(N-morpholino)ethanesulfonic acid (MES) (Sigma Aldrich™)	0.4 g/l
Naphthaleneacetic acid (NAA) (Duchefa™)	0.5 mg/l
pH*	5.7
Autoclaving	1 atm; 121ºC, 20min

*Adjusted with KOH solution.

Table M3. BY2 cell culture media (MSBY2)

	MSBY2
Murashige and Skoog (MS) basal medium with vitamins (Duchefa™)	4.3 g/l
Sucrose (Duchefa™)	30 g/l
KH2PO4 (Merck™)	0.2 g/l
2,4-Dichlorophenoxyacetic acid (Duchefa™)*	100 µl/l
Thiamina/HCl (Sigma Aldrich™)**	250 µl /l
Myo-inositol (Duchefa™)	0.1 g/l
pH***	5.7
Autoclaving	1 atm; 121ºC, 20min

*2,4 D was dissolved at 0.04 g/10 ml in 1M NaOH **Stock 4mg/l in ultrapure water *** Adjusted with KOH solution.

Commercial carrots (*D*. carota) roots were acquired in a local food market in Terrassa. The carrots were produced by Verdura Masclans, SCP in Caldes de Montbui. The crop was cultured under the "organic" certification issued by the Registre d'operadors del Consell Català de la Producció Agrària Ecológica with the register number CT001617. The cultivation fields were inspected under the report number LIMV180143.

N. benthamiana plant seeds overexpressing *A. thaliana WRI*, *A. thaliana DGAT1* transgenes (*N. benthamiana* LD+) were kindly provided by Dr Petrie, CSIRO, Australia (Vanhercke *et al.*, 2014). Seeds were germinated and planted on soil to be cultivated in CRAG plant growth facilities under the appropriate containment conditions. Seeds stock were produced and stored in order to maintain the plant line.

N. benthamiana wild type plants were used in the assays. Wild type and transgenic *N. benthamiana* plants, were maintained in greenhouse at 22-28 ºC and 16 h light/8 h darkness during the propagation and protein expression assays.

Sweet potato (*Ipomea batatas*) leaves infected with SPFMV isolate AM-MB2 were sampled from the propagated plants at CRAG plant growth facilities. The plants had been acquired in a local food market in Barcelona, and were planted on soil to produce fresh aerial tissue (stems and leaves) that was further propagated (Mingot, 2016).

1.2 Human cell lines

Human pulmonary fibroblasts MRC-5 were acquired from ATCC collection (ATCC® CCL-171™) and used to evaluate the activity of plant produced recombinant proteins by Servei de Cultius Cel·lulars, Producció d'Anticossos i Citometria (SCAC, UAB).

1.3 Bacteria

E. coli and A. tumefaciens

The multiplication and purification of plasmids with the gene constructs was performed using *E. coli* strains TOP10 or DB3 (permissive for empty Gateway™ cloning plasmids). *E. coli* cells were grown at 37ºC in LB media. *A. tumefaciens* strains, EHA105 (plant expression plasmids) and GV3101 carrying the helper plasmid pSoup (Shi *et al.*, 2019), were used for transferring gene constructs into plant cells via infiltration. Agrobacterium was grown in YEB medium at 28ºC.

Bacterial strains were maintained at 4ºC during short period of times and stored at 30% glycerol stock at -80ºC for long times.

LB medium: 10 g/l triptone, 5g/l yeast extract and 10g/l NaCl. Agar was added at 15g/l for solid media.

YEB medium: 5 g/l beef extract, 1 g/l yeast extract, 5 g/l peptone sucrose 5 g/l and 0.48 g/l $MgSO_4$. Agar was added at 15g/l for solid media.

Media were supplemented with the corresponding selection antibiotics.

1.4 Virus RNA

SPFMV isolate AM-MB2 RNA was extracted and purified from sweet potato leaves from plants propagated in CRAG greenhouses.

1.5 Plasmids, cloning and viral vectors

D-TOPO plasmids carrying the codon optimized IGF1 and CPP sequences (pD-TOPO#IGF1, pD-TOPO#CPP-IGF1) were synthesized by Synbio Technologies™.

Plasmids containing *AtWRI* (6530638541), *AtDGAT1* (3510696031) *AtOle2* (U14610) cDNAs were obtained from The Arabidopsis biological resource center (ABRC).

Gateway™ (Invitrogen™) technology was used for the cloning in plant binary plasmids. pENTR™/D-TOPO® was used as the entry plasmid and different pDEST were used as destination plasmids (Table M4). Destination plasmids were previously described in (Nakagawa *et al.*, 2007, 2009; Tanaka *et al.*, 2011).

Table M4. Gateway™ pDEST plasmids used for cloning

pDEST	Expression cassette	GenBank num.
pGWB2	p35S:R1-Cm^R-ccdB-R2:Nost	AB289765.1
pGWB5	p35S:R1-Cm^R-ccdB-R2–**GFP**:Nost	AB289768.2
pGWB6	p35S:**GFP**–R1-Cm^R-ccdB-R2:Nost	AB289769.1
pGWB24	p35S:**GST**–R1-Cm^R-ccdB-R2:Nost	AB289787.1
pGWB402	p35S:promoter–R1-Cm^R-ccdB-R2:Nost	AB294426.1
pGWB709	p35S:**6xHIS**–R1-Cm^R-ccdB-R2:Nost	AB608275.1

p35S, 35S promoter from CaMV; R1, Gateway cassette recombination site 1; Cm^R, chloramphenicol resistance; ccdB, bacterial toxin affecting DNA gyrase; R2, Gateway cassette recombination site 2; GFP, green fluorescent protein; GST, glutathione S-transferase; YFP, yellow fluorescent protein; CFP, cyan fluorescent protein; and 6xHIS; six histidine tandem; Nost, nopaline synthase terminator.

The Table M5 indicates the binary plasmids developed in the project with the expression cassettes, as well as the entry and destination plasmids used for their preparation.

The plasmids pGWB402#p35s:AtOle2-TEVs-IGF1 and pGWB402#p35s:IGF1-TEVs-AtOle2, were engineered using recombinant PCR and Gateway™ cloning technology. First, AtOle2 sequences were amplified from D-TOPO#AtOle2 plasmid by PCR using the primer combinations 20-2 or 18-4 (Table M12). The primers 2 and 18, introduced the sequence for *Tobacco etch virus* protease cleavage site (TEVs) in the cDNAs. Then IGF1 was amplified from the pD-TOPO#IGF1 using primers 13-12 or 17-16 (Table M12). Finally, the cDNA fragments were used in a PCR to generate recombinant products from mixed homologous template populations; AtOle2 (20-2) + IGF1 (13-12) and IGF1 (18-4) + AtOle2 (17-16),

generating the AtOle2-IGF1 and IGF1-AtOle2 cDNAs respectively. The resulting products were cloned in pD-TOPO plasmids and recombined in a pDest pGWB402.

Table M5. Binary plasmids developed using Gateway™ cloning

Plasmid	Expression cassette	pENTRY	pDEST
pGWB5#p35S:AtOle2-GFP	p35S:**AtOle2-TEVs-GFP**:Nost	pD-TOPO#AtOle2-TEVs	pGWB5
pGWB6#p35S:GFP-AtOle2	p35S:**GFP-TEVs-AtOle2**:Nost	pD-TOPO#TEVs-AtOle2	pGWB6
pGWB2#p35S:AP24-IGF1	p35S:**AP24-IGF1**:Nost	pD-TOPO#AP24-IGF1	pGWB2
pGWB2#p35S:AP24-CPP-IGF1	p35S:**AP24-CPP-IGF1**:Nost	pD-TOPO#AP24-CPP-IGF1	pGWB2
pGWB24#p35s:NIa-Pro SPFMV	p35S:**GST-NIa-Pro**:Nost	pD-TOPO#NIa-Pro	pGWB24
pGWB709#p35s:NIa-Pro SPFMV	p35S:**6xHIS-NIa-Pro**:Nost	pD-TOPO#NIa-Pro	pGWB709
pGWB5#p35s:AtOle-SPFMVs-GFP	p35S:**AtOle2-SPFMVs-GFP**:Nost	pD-TOPO#AtOle2-SPFMVs	pGWB5
pGWB2#p35s:AtDGAT1	p35S:**AtDGAT1**:Nost	pD-TOPO#AtDGAT1	pGWB2
pGWB2#p35s:AtWRI	p35S:**AtWRI**:Nost	pD-TOPO#AtWRI	pGWB2
pGWB402#p35s:AtOle2	p35S:**AtOle2**:Nost	pD-TOPO#AtOle2	pGWB402
pGWB402#p35s:AtOle-TEVs-IGF1	p35S:**AtOle2-TEVs-IGF1**:Nost	pD-TOPO#AtOle2-TEVs-IGF1	pGWB402
pGWB402#p35s:IGF1-TEVs-AtOle2	p35S:**IGF1-TEVs-AtOle2**:Nost	pD-TOPO#IGF1-TEVs-AtOle2	pGWB402

p35S, 35S promoter from CaMV; *AtOle2*, *A. thaliana oleosin 2* gene; TEVs, *Tobacco etch virus* protease cleavage site; AP24, *N. tabacum* AP24 protein exporting signaling peptide; CPP, cell penetrating peptide hPP3; IGF1; human Insulin-like growth factor 1 cDNA; NIa-Pro, SPFMV protease; AtDGAT1, *A. thaliana* Diacylglycerol O-acyltransferase 1; AtWRI; Wrinkled transcription factor; GFP, green fluorescent protein; GST, glutathione S-transferase; YFP, yellow fluorescent protein; CFP, cyan fluorescent protein; and 6xHIS; six histidine tandem; Nost, nopaline synthase terminator.

The engineered plasmids used in SAR evaluation were developed using a pTRAK plasmid (Sack *et al.*, 2007) as backbone and digesting it with different combinations of restriction enzymes (Table M6), treating with T4 DNA Polymerase (Invitrogen™) for terminal ends blunting, and ligating the plasmids using T4 DNA Ligase (Invitrogen™). In the case of pTRAK-RFP2 and pTRAK4-RFP, after digestion and blunting, the

plasmids were re-ligated, while in the case of pTRAK3-RFP and pTRAK5-RFP, two independent fragments corresponding to the plasmid and the RFP expression cassettes were ligated.

Table M6. Restriction enzymes used for pTRAK-RFP plasmids cloning

Plasmid	Restriction enzyme 1	Restriction enzyme 2
pTRAK-RFP2	*Sph*I	*Nhe*I
pTRAK-RFP3 (plasmid)	*Sph*I	*Nhe*I
pTRAK-RFP3 (RFP cassette)	*Fse*I	*Pme*I
pTRAK-RFP4	*Sph*I	*Asc*I
pTRAK-RFP5 (plasmid)	*Sph*I	*Pme*I
pTRAK-RFP5 (RFP cassette)	*Asc*I	*Fse*I

pTAREK_GW was developed using pTRAK-RFP as backbone and the Gateway™ cloning cassette derived from pGWB402. First the pTRAK-RFP and pGWB402 were digested with *Pme*I and *Asc*I restriction enzymes, and the plasmid backbone and recombination cassette were purified. Then both plasmid fragments were ligated using T4 DNA Ligase (Invitrogen™).

pTAREK_P1b was developed adding the *P1b* expression cassette to the pTAREK_GW. *P1b* expression cassette was amplified by PCR using the primers 27 and 28 listed in Table M12 and was assembled to the pTAREK_GW backbone using Gibson *Assembly* ® Cloning technology. This required pTAREK_GW be digested with *Pme*I and *Mfe*I restriction enzymes.

pTAREK_GW_GFP and pTAREK_P1b_GFP were engineered using Gateway™ cloning and the p-TOPO#GFP plasmid as pEntry plasmid. The pTRAK derived plasmids developed are listed in Table M7.

Table M7. pTRAK derived plasmids

Plasmid	Plasmid T-DNA
pTRAK-RFP	LB-pNos:Kan^r:Nost-**SAR**-p35S:DsRed:35St-**SAR**-RB
pTRAK2-RFP	LB-**SAR**-p35S:DsRed:35St-**SAR**-RB
pTRAK3-RFP	LB-**SAR**-p35S:DsRed:35St-RB
pTRAK4-RFP	LB-p35S:DsRed:35St-**SAR**-RB
pTRAK5-RFP	LB-p35S:DsRed:35St-RB
pTAREK_GW	LB-pNos:Kan^r:Nost-**SAR-p35S:R1-Cm^R-ccdB-R2**:Nost-RB
pTAREK_P1b	LB-pNos:Kan^r:Nost-**SAR-p35S:R1-Cm^R-ccdB-R2**:**Nost-p35S:P1b**:**Nost**-RB
pTAREK_GW_GFP*	LB-pNos:Kan^r:Nost-**SAR-p35S:GFP:Nost**-RB
pTAREK_P1b_GFP*	LB-pNos:Kan^r:Nost-**SAR-p35S:GFP**:**Nost-p35S:P1b**:**Nost**-RB

LB, T-DNA left border; pNos, nopaline synthase promoter; Kan^r, kanamycin resitance; Nost, nopaline synthase terminator; SAR, Scaffold attachment region: p35S, 35S promoter from CaMV; dsRed, Red fluorescent protein isolated from Discosoma genus; 35St, 35S terminator from CaMV; RB, T-DNA right border; R1, Gateway cassette recombination site 1; Cm^R, chloramphenicol resistance; ccdB, bacterial toxin affecting DNA gyrase; R2, Gateway cassette recombination site 2; P1b, CVYV P1b silencing suppressor; GFP, green fluorescent protein. * p-TOPO#GFP was used for the cloning.

Viral vectors derived from TMV used for the production of IGF1 and CPP-IGF1 in *N. benthamiana* leaves were available in the research group, and were previously described (Shi *et al.*, 2019). Gibson *Assembly* ® Cloning technology was used for cloning into viral vectors using the primers 11 to 16 listed in Table M12. The plasmids were cloned following the protocol described by (Shi *et al.*, 2019). Briefly, an intermediate plasmid (pMTMVn) was digested with *Age*I and *Xho*I and the different cDNAs were assembled in it. Once the cDNA for the recombinant protein was inserted into pMTMVn, the manipulated TMV genome fragment was transferred into the TMV infectious clone (pGTMVa) using a second Gibson assembly reaction. This requires that pGTMV be digested with restriction enzymes *Nco*I and *Pfl*23II, which had unique recognition sites in this plasmid. Table M8 list the plasmids developed.

Table M8. TMV-derived viral vectors

Plasmid	Plasmid T-DNA
pGTMVa#AP24-IGF1	LB-p35S:RdRp-MP-**AP24-IGF1**:35St-RB
pGTMVa#AP24-CPP-IGF1	LB-p35S:RdRp-MP-**AP24-CPP-IGF1**:35St-RB
pGTMVa#GFP	LB-p35S:RdRp-MP-**GFP**:35St-RB

LB, T-DNA left border; p35S, 35S promoter from CaMV; RdRp, *Tobacco mosaic virus* (TMV) polymerase; MP, TMV movement protein; AP24, apoplast exporting signaling peptide; CPP, cell penetrating peptide hPP3; IGF1, Insulin-like growth factor 1; GFP, green fluorescent protein; 35St, CaMV 35S terminator; RB, T-DNA right border.

2. Callus induction and characterization

2.1 *D. carota* callus induction

D. carota callus induction was performed using the cambial region of carrot roots. The plant material was washed profusely with water to eliminate soil residues. Roots were peeled and cut in square pieces. Then root pieces were disinfected by immersion in an ethanol 70% solution for 1 min, and 30 min in a solution containing sodium hypochlorite (20%) and Tween®20 (Sigma-Aldrich™) (0.1%). In sterility, the disinfected explants were washed using PBS and their external layers eliminated with scalpel. Explants were then plated in induction media (Table M9) at different hormone concentrations. Basically, different concentrations of the auxine 2,4 diclorophenoxiacetic (2,4-D) were used (Table M10) When solidified, plates were stored at 4ºC for a maximum of 1 month.

PBS buffer: 8 g/ l of NaCl, 200 mg/l of KCl, 1.44 g/l of Na2HPO4, 240 mg/l KH2PO4. pH 7.4.

Table M9. Media composition used for *D. carota* callus induction (MSI)

	MSI
Murashige and Skoog (MS) basal medium with vitamins (Duchefa™)	4.3 g/l
Sucrose (Duchefa™)	30 g/l
2,4-Dichlorophenoxyacetic acid (Duchefa™)	0.25-2.5 mg/l
pH*	5.8
Autoclaving	1 atm; 121ºC, 20min

*Adjusted with KOH solution.

Table M10. 2,4-D concentrations used in the different MSI media

	2,4-D (mg/l)
MSI1	0.25
MSI2	0.5
MSI3	0.75
MSI4	1
MSI5	1.25
MSI6	1.5
MSI7	1.75
MSI8	2
MSI9	2.25
MSI10	2.5

2.2 Plant cell growth

Carrot cells were grown in MSDc liquid medium, BY2 in MSBY2 and Arabidopsis cells in MSAt media.

Carrot cell growth dynamics were evaluated using suspensions propagated in liquid media for different number of cycles as indicateuhy Cell suspensions were initiated from 40 g of calli, in 200 ml of MSDc media (Table M11), in 500 ml flasks and several parameters were determined, including reduction in brix degrees (ºB), pH, cell packing volume (VEC) and fresh weight (FW) at different times (0, 3, 5, 7, 11 and 15 days after propagation).

ºB were evaluated using a digital refractometer, by dispensing a media drop in the optical channel of the equipment (667-2159, Hanna™).

VEC relates the percentage of cellular volume in the total volume of a suspension. VEC values were determined using 5 ml samples taken using wide mouth serological pipettes of 10 ml (Sarstedt™). Samples were centrifuged for 5 min at 99 g. The VEC was measured using a caliper.

FW was determined by filtering suspension samples using a vacuum pump until biomass completely dry, washing cells with PBS and weighting them using precision scale. Then, cells were washed with PBS and filtered again. Finally, the biomass was weight in a lab scale.

The characterization of the growing curves was performed using the parameters: growth velocity in exponential phase (Vexp), maximum VEC (VECmax), and exhaustion time (Ta). Vexp was defined as the increment of biomass in the linear interval of the exponential curve (Vexp= Δ VEC / Δ t), VECmax was consider the VEC value in the upper threshold of the curve while Ta was considered the time when lack of resources starts to provoke cell death or stress.

Table M11. *D. carota* cell culture media (MSDc)

	MSDc
Murashige and Skoog (MS) basal medium (Duchefa™)	4.3 g/l
Sucrose (Duchefa™)	30 g/l
Casein hydrolysate (Sigma-Aldrich™)	250 mg/l
Vitamin mix 1 *	1 ml/l
Vitamin mix 2 **	1 ml/l
pH***	5.7
Autoclaving	1 atm; 121ºC, 20min

* Vit 1: 1mg/l of calcium pentatonate, 100mg/ml of myo-inositol, 0.01mg/ml biotin, 1mg/ml nicotinic acid, 1mg/ml tiamine and 1mg/ml pyridoxine (Duchefa™). ** Vit 2: 1mg/ml tiamine and 2mg/ml of nicotinic acid (Duchefa™). ***Adjusted with KOH solution.

3. Nucleic acid manipulation

3.1 RNA extraction from plant material

Plant total RNA was extracted from sweet potato leaves material using TRIzol reagent (Invitrogen™) and following manufacturer instructions. The RNA was used to amplify the NIa-pro from SPFMV. Finally, pellet was washed with ethanol 70%, dried and resuspended in ultrapure water.

Genomic DNA was extracted from transgenic callus using Maxwell® RSC Plant DNA Kit (Promega™) and following manufacturer instructions. 25mg of callus culture was used in each extraction procedure.

RNA and DNA concentration on was evaluated using spectrophotometer (Nanodrop™, Thermo Scientific™).

3.2 Reverse transcription

High-Capacity cDNA Reverse Transcription kit (Applied Byosistems™) was used for cDNA amplification using 1 µg of RNA as template. The protocol recommended by the Applied Byosistems™ and random primers were used for the reaction.

3.3 PCR amplification

The amplification of DNA fragments was performed by polymerase chain reaction (PCR) using as template plasmids carrying the sequences of interest or cDNAs obtained by reverse transcription. Phusion® high-fidelity DNA polymerase (New England Biolabs™) was used for the amplification following the manufacturer recommendations. Temperatures and cycles used for the different PCR reactions vary according to Tm of primers and amplicon length. The primers used for the PCR reactions are listed in Table M12.

4. Transformation of microrganisms and plant cells

4.1 *E. coli* and *A. tumefaciens*

Chemical competent *E. coli* and *A. tumefaciens* cells were prepared following two different protocols. For *E.coli*, overnight cultures were chilled for 10 min on ice, spun 10 min at 6000g in 50ml tubes, the pellets were resuspended in 10 ml ice-cold CC buffer per 100ml culture, incubate the solution in ice for 10 min, spun for 10 min at 4000g, suspend in 5ml of ml ice-cold CC buffer and add 1.4 ml DMSO and Distribute the cell suspension in 50-100 µl aliquots in 1.5 ml pre-chilled tubes. In the case of chemocompetent *A. tumefaciens* preparation, overnight cultures around $1OD_{600}$ were chilled on ice for 15 min, spun at 4500g for 10 min at 4ºC, the pellets were resuspended in 1ml of buffer CA per 50ml culture, and dispensed 100 µl aliquots aliquots in 1.5 ml pre-chilled tubes.

CC buffer: 10 mM Hepes 2.38 g/L, 15 mM CaCl2 2.21 g/L, 55 mM MnCl2.4H2O and 10.89 g/L 250 mM KCl 18.64 g/L.

CA buffer: 20 mM CaCl2 (ice cold).

Table M12. Primers used for cloning in this work

	Primer name	Dir	Sequence	Tm	Length	PCR template	Final construct
1	TR_AtOle2	Fw	CACCATGGCGGATACACACCGTGT	66	24	AtOle2 plasmid	D-TOPO#AtOle2-TEVs*
2	TR-AtOle2-TEVs	Rv	ACTCTGAAAATAAAGATTCTCTGCAGCCGTCGTCCTC	66	37	AtOle2 plasmid	D-TOPO#AtOle2-TEVs*
3	TEVs_AtOle2	Fw	CACCGAGAATCTTTATTTTCAGAGTATGGCGGATACACACCGTGT	65.9	45	AtOle2 plasmid	D-TOPO#TEVs-AtOle*
4	TR_TEV-AtOle2	Rv	CTATTGAGAATAAACCAATTCAGACATC	67	28	AtOle2 plasmid	D-TOPO#TEVs-AtOle*
5	TR_AtOle2	Fw	CACCATGGCGGATACACACCGTGT	66	24	AtOle2 plasmid	D-TOPO#AtOle2-SPFMVs*
6	TR_AtOle2	Rv	AGATTGATGATAAACTTCAAGTGCAGCCGTCGTCCTCCCCTG	77	42	AtOle2 plasmid	D-TOPO#AtOle2-SPFMVs*
7	TR_NIaPro	Fw	CACCATGAGTAAGTCACTATTTAGAGG	27	65	SPFMV RT	D-TOPO#NIa-pro SPFMV*
8	TR_NIaPro	Rv	CTATTGAGCATACACTTGTATGCC	24	62	SPFMV RT	D-TOPO#NIa-pro SPFMV*
9	TR-IGF1-Met	Fw	CACCATGGGGCCTGAGACTCTCTGTGGC	76	28	D-TOPO:CPP-IGF1	D-TOPO#IGF1*
10	TR-IGF1-Stop	Rv	CTATGCAGACTTGGCAGGTTTTAAGGGAGC	72	30	D-TOPO:CPP-IGF1	D-TOPO#IGF1*
11	AP24	Fw	CTCAGTTCGTGTTCTTGTCAATGTCCAACAACATGGGC	65.3	38	D-TOPO:CPP-IGF1	pMTMVn#IGF1**
12	IGF1	Rv	CTACCTCAAGTTGCAGGACCTCATGCAGACTTGGCAGG	68	38	D-TOPO:CPP-IGF1	pMTMVn#IGF1**
13	AGPA-IGF1	Fw	AGGTCCCGCAGGGCCTGAGACTCTCTGTG	69.1	29	D-TOPO:CPP-IGF1	pMTMVn#CPP-IGF**
14	AGPA-IGF1	Rv	TCTCAGGCCCTGCGGGACCTGCTCGTTT	70	28	D-TOPO:CPP-IGF1	pMTMVn#CPP-IGF1**
15	AP24-IGF1	Fw	GTGACCTACACTTATGCAGGGCCTGAGACTCTC	64.8	33	D-TOPO:CPP-IGF1	pMTMVn#CPP-IGF1**
16	AP24-IGF1	Rv	GAGAGTCTCAGGCCCTGCATAAGTGTAGGTCAC	64.8	33	D-TOPO:CPP-IGF1	pMTMVn#CPP-IGF1**
17	TR-IGF1-Met	Fw	CACCATGGGGCCTGAGACTCTCTGTGGC	76	28	D-TOPO:CPP-IGF1	D-TOPO#IGF1-AtOle2*
18	TR-TEV-AtOle2	Fw	GAGAATCTTTATTTTCAGAGTATGGCGGATACACACCGTGT	74	41	AtOle2 plasmid	D-TOPO#IGF1-AtOle2*
19	TR-AtOle2	Rv	CTATGCAGCCGTCGTCCTCC	65	20	AtOle2 plasmid	D-TOPO#IGF1-AtOle2*
20	AtOle2	Fw	CACCATGGCGGATACACACCGTGT	66	24	AtOle2 plasmid	D-TOPO#AtOle2-IGF1*
21	TR-AtOle-TEVs	Rv	ACTCTGAAAATAAAGATTCTCTGCAGCCGTCGTCCTC	66	37	AtOle2 plasmid	D-TOPO#AtOle2-IGF1*
22	TR-AtOle-IGF1	Fw	GAGAATCTTTATTTTCAGAGTGGGCCTGAGACTCTCTGTGGCG	68	43	D-TOPO:CPP-IGF1	D-TOPO#AtOle2-IGF1*
23	TR-GFP-Met	Fw	CACCATGGTGAGCAAGGGCGAGGAG	72	25	pGWB6	D-TOPO#GFP***
24	TR-GFP-Stop	Rv	CCGCTTTACTTGTACAGCTCGTCCATG	70	27	pGWB6	D-TOPO#GFP***
25	TR_AtOle2	Fw	CACCATGGCGGATACACACCGTGT	66	24	AtOle2 plasmid	D-TOPO#AtOle2*
26	TR_TEV-AtOle2	Rv	CTATTGAGAATAAACCAATTCAGACATC	67	28	AtOle2 plasmid	D-TOPO#AtOle2*
27	TR_P1bcas	Fw	TAAAACGACGGCCAGTGCCATGAGACTTTTCAACAAAGG	77	39	pGWB402-P1bCVYV	pTAREK_P1b
28	TR_P1bcas	Rv	ACCCTTTGTTGAAAAGTCTCGATCTAGTAACATAGATGACAC	75	42	pGWB402-P1bCVYV	pTAREK_P1b

*Cloning in binary plasmids; **Cloning in viral plasmids; ***Cloning in pTAREK_GW and pTAREK_P1b.

4.2 *N. benthamiana* leaf transient transformation

N. benthamiana leaves (fully expanded, from three to four week old plants) were transiently transformed via Agroinoculation as previously described (Valli *et al.*, 2006b). For plant Agroinoculation, overnight cultures of *A. tumefaciens* were diluted in induction medium at the appropriate optical density of 600 nm (OD_{600}), incubated for 3 h at room temperature, and infiltrated using a needle-less syringe (0.5 OD_{600}). Cultures should have an OD_{600} lower than 1 for good transformation efficiency. In the transient expression assays using viral vectors, Agrobacterium were used at 0.4 OD_{600} while in the co-infiltration assays using the P1b the different Agrobacterium solutions used showed between 0.6-0.8 OD_{600}.

Transient expression was monitored at different time points from 1 day until 7-9 days post-infiltration.

Induction medium: 10 mM MES, 10 mM MgCl2, 200 µM acetosyringone.

4.3 Stable transformation of plant cell cultures

Plant cell lines were stably transformed via *A. tumefaciens* EHA105 strain carrying different plant expression plasmids. For that, plant suspensions at exponential growth were used, as well as *A. tumefaciens* cultures at 0.8-1.2 OD_{600}. The Agrobacterium cultures were induced before the transformation by incubating for 2h in induction media with 200 µM acetosyringone. Plant suspensions and *A. tumefaciens* were mixed by gently agitation and maintained for 60 minutes at room temperature. Liquid media was eliminated by filtering the cells trough a 30 µm nylon monofilament (NMO) mesh. Plant cell biomass was then incubated at 24ºC in darkness for 48h and washed using growth media supplemented with Timentin™ (GoldBio©) and the selection antibiotics. Washes were done in sterile tubes by gently agitation, and filtering through NMO, followed by suspension in culture media. Finally, the plant material was plated in solid growth media supplemented with antibiotics and allowed to grow at 24ºC in darkness until calli appeared. Calli were then transferred to fresh media without Timentin™ for propagation.

When calli were visible they were transfer to fresh media including antibiotics for propagation. Table M13 shows the transgenic calli obtained in the different assays, as well as the plasmids used in transformation.

Table M13. List of transgenic cell cultures including plasmids used for transformation and associated section.

Callus identity	Number of callus	Plasmid 1	Plasmid 2	Results section
GFP-Hygromycin	3	pGWB6	np	*5.2*
GFP-Kanamycin	5	pGWB6	np	*5.2*
EV- Hygromycin	2	pGWB402	np	*5.2*
EV- Kanamycin	3	pGWB402	np	*5.2*
AtOle-GFP- Hygromycin	*3*	pGWB5#p35S:AtOle2-GFP	np	*5.3*
GFP-AtOle- Hygromycin	5	pGWB6#p35S:GFP-AtOle2	np	*5.3*
GFP- Hygromycin	2	pGWB6	np	*5.3*
AtOle-GFP- Kanamycin	2	pGWB5#p35S:AtOle2-GFP	np	*5.3*
GFP-AtOle- Kanamycin	*5*	pGWB5#p35S:AtOle2-GFP	np	*5.3*
GFP- Kanamycin	*3*	pGWB6	np	*5.3*
*IGF1,EV**	*7*	pGWB2#p35S:AP24-IGF1	pGWB402	*6.2*
*CPP-IGF1,EV**	5	pGWB2#p35S:AP24-CPP-IGF1	pGWB402	*6.2*
*IGF1,P1b**	*6*	pGWB2#p35S:AP24-IGF1	pGWB402:P1b	*6.2*
*CPP-IGF1,P1b**	*4*	pGWB2#p35S:AP24-CPP-IGF1	pGWB402:P1b	*6.2*
*P1b**	*4*	pGWB2	pGWB402:P1b	*6.2*

*In the presence of both antibiotics hygromycin and kanamycin.

4.4 Plant cell culture transient transformation

Plant cell cultures were transiently transformed using *A. tumefaciens* strains EHA105 and GV3101 carrying the plasmids for the expression of fluorescent proteins using the method previously reported (Rademacher *et al.*, 2019). One milliliter of plant suspensions at exponential growth phase was dispensed per well in 96-well plates (receiver plate 20µm, Macherey-Nagel™). The plate was subjected to vacuum infiltration (NucleoVac 96 Vacuum Manifold, Macherey-Nagel™) for some seconds to remove the excess of medium and to generate semy-dry plant cell packs (PCPs). Overnight *A. tumefaciens* cultures were induced for 2h in induction media using either of the two different induction media CRAG induction medium (CRAG IM) or IME induction medium (IME IM). Then, Agrobacterium cultures were dispensed in each well and vacuum was applied to infuse into the PCPs. The plate was incubated at room temperature

for different periods of time, ranging from 1 to 60 minutes, and vacuum was applied again to dry the biomass. Plates were then covered with a gas-permeable foil (Macherey-Nager™) and incubated at 24ºC, in darkness and high humidity for the indicated periods of time. When sampling, the gas-permeable foil was retired, and the biomass packs recovered pushing the filters with a needle.

CRAG induction media: 100mM MES, 100mM MgCl2 and 200 µM acetosyringone.

IME induction media: 100 g/L sucrose, 3.98 g/L glucose, 1 g/L Ferty2Mega (Planta Düngemittel GmbH) and 200µM acetosyringone.

5. Confocal visualization

Confocal microscope TCS SP5 confocal II (Leica Microsystems™, Germany) was used for the visualization of plant cell cultures and leaf material. In the case of plant cell cultures, suspensions or disaggregated calli were placed on microscopy slides and culture media were added to maintain the humidity. For plant leaf visualization, leaf patches were placed in microscope slides trying to avoid major veins and with the abaxial side on the top of the sample.

LDs were stained with Nile red (1 ng/ml; Sigma Aldrich™) and visualized using a DPSS laser emitting at 561 nm. GFP signal was detected at 500–550 nm and excited with an Argon ion laser emitting a 488 nm; YFP was detected at 525-575 nm; and CFP was detected at 465-509 nm. Images were taken using 10x, 20x objectives and digital zoom. The images were processed using ImageJ software (National Institutes of Health).

6. Cryopreservation

D. carota cell line was cryopreserved for long term maintenance using a protocol developed by Vytrus Biotech. The protocol was applied to 40 ml cell culture at exponential growth (6 days after subculturing), in which the VEC was determined to be biomass around 30%. The 40ml culture were filtered through 10 µm filters and biomass drained. Then, dried material was introduced into a 100 ml flask and added LPS medium at a ratio 2 volumes of cell culture: 1 volume of LPS medium, considering the VEC value at 30%. Cells and medium were homogenized by gentle pipetting and then incubated for 30 min at room

temperature without agitation. Aliquots of 1.5 ml were then dispensed in cryotubes, which were distributed in a Mr. Frosty pack. It is important to be precise because changes in volume affect dramatically congelation velocities and fill completely the freezing pack with cryotubes. Mr. Frosty pack (Nalgene® Mr. Frosty, Sigma Aldrich™) was introduced in a porexpan (EPS) box, previously tempered to 4ºC, and introduced in a -80ºC freezer for 6 hours. After that, cryotubes were taken out from Mr. Frosty pack and immersed in liquid nitrogen, where they were maintained untill defrosting.

For defrosting process, the crytubes were taken out from liquid nitrogen and introduced in a pot with sterile water at 42ºC for 2 minutes. During the incubation the water was stirred to allow the temperature around the tubes to constant. Then, the cryotubes were centrifuged during 5 minutes at 4400rpms, supernatants removed, and plant cell material placed on solid media plate covered with a sterile filter paper for absorption of toxic substances used during the process.

LPS medium: MS medium + 2M glycerin + 0.4M sucrose + 10g/l proline.

7. Plant extracts preparation for recombinant protein production analysis

Total protein extractions were performed grinding plant material with liquid nitrogen, homogenized in SDS-PAGE loading buffer without 2-mercaptoethanol (400 mg of pulverized material in 1 ml buffer) (Laemmli, 1970), and incubated at 95 C for 10 min. Cell debris and heat-denatured proteins from crude leaf protein extracts were removed by centrifugation (13 000 rpm for 15 min at room temperature).

Extracellular fluids (ECFs) were extracted from fresh leaves by vacuum infiltration using the corresponding buffers (PBS, Asc and Sucrose) supplemented with 0.02% (v/v) Silwet L-77. First, leaves were cleaned to remove soil residues, and introduced into tubes containing the proper buffer. Then, tubes were introduced in a polycarbonate vacuum chamber (Bel-Art™), and vacuum was applied until the extraction buffer penetrates into the leaves. Finally plant leaves were taken out from tubes, surface dried with paper tissues and pressed using a needleless syringe to recover the ECL fraction.

Transgenic plant cell culture media were collected from suspensions at the end of their propagation cycle. The medias were frozen, lyophilized and resuspended in loading buffer and used directly for protein evaluation.

For LD fractionation from plant cell cultures, around 1g of grinded biomass was resuspended in 1 ml of LD isolation buffer and grinded using a mortar and pestle on ice. Then, the mixture was centrifuged at 10000 *x*

g for 20 min. The floating fat pad was collected, resuspended in the same buffer and used for confocal visualization. Integrity of isolated LDs was tested by selective staining with Nile red in all cases. LD proteins were solubilized in loading buffer, separated by SDS-PAGE and analyzed by Western-blot.

PBS buffer: PBS 1x (8 g/ l of NaCl, 200 mg/l of KCl, 1.44 g/l of Na2HPO4, 240 mg/l KH2PO4. pH 7.4), 1 mM ethylenediaminetetraacetic acid (EDTA), 0.02% Silwet-L77.

Asc buffer: 50 mM phosphate buffer (20.214 g of Na2HPO4-7H2O, 3.394 g of NaH2PO4H2O, 1000ml ultrapure water. pH 7.2), 50 mM ascorbic acid, 10 mM EDTA, 0.02% Silwet-L77.

Sucrose buffer: 0.3 M sucrose crystallized, 0.3 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 5mM EDTA.

Loading buffer: 2.5 ml 1 M Tris-HCl pH 6.8, 0.5 ml of ultrapure water, 1.0 g Sodium dodecyl sulfate (SDS), 0.8 ml 0.1% Bromophenol Blue, 4 ml 100% glycerol, 2 ml 14.3 M β-mercaptoethanol.

LD isolation buffer: 2.4 M sucrose, 30 mM sodium phosphate buffer pH 7.5, 1 mM PMSF.

7.1 Characterization of the production kinetics of cell cultures

Plant cell cultures were sampled during 5 months of growth to assess their production kinetics over time. For that, 20ml of cell cultures were sampled every 14 days (at the end of the growing cycle), centrifuged at 99 g for 10 min, and biomass and media frozen independently for subsequent characterization.

8. Identification of recombinant protein production by SDS-PAGE

Protein preparations were separated in Tricine-SDS-PAGE (16.5%) (Schägger and von Jagow, 1987). Gels were stained with Coomassie blue G-250 (Life Technologies™) to detect proteins, or transferred to nitrocellulose membranes (Whatman® Protran 0.2 µm) to immunodetect proteins, as described previously in (Coca *et al.*, 2004). The transfer to nitrocellulose membranes was performed in a XCell SureLock™ (Invitrogen™) at 10V, 4ºC overnight. Nitrocelulose membranes were stained with Ponceau (CAS: 6226-79-5) to confirm a correct protein transfer. IGF1 and CPP-IGF1 were immunodetected using commercial anti-IGF1 antibodies (ab9572, Abcam™) at 1:5000 dilutions. Anti-rabbit polyclonal antibodies tagged with

horseradish peroxidase (HRP) (1:10000) were used as secondary antibodies. Immunoreactions were visualized with Supersignal® maximum sensitivity substrate (Thermo Scientific™) using ChemiDoc™ Touch Imaging System (Bio-Rad™). Synthetic human IGF1 peptide was produced by GeneScript™ and used as a control.

9. Quantification of protein accumulation

The amount of IGF1 and CPP-IGF1 peptide accumulated in the different plant cell matrixes and with the different production strategies was determined by Enzyme-Linked ImmunoSorbent Assay (ELISA) following a modified protocol from Abcam™. Microplates of 96 wells were used for the ELISA assays (Microplate 96 well High binding, Greiner Bio-One™). Coating of microplates was done by adding 50 µl of protein samples in PBS buffer and incubation overnight at 4°C. Then protein samples were removed, and plates washed three times with PBS. Wash solutions were removed by flipping the plate over a sink. The remaining drops were removed by patting the plate on a paper towel. Protein blocking was done incubating the plates at room temperature with 5% non-fat dry milk in PBS. After washing with PBS, plates were incubated with anti-IGF1 (1:2500 dilution) in PBS for 2h at room temperature. Primary antibodies were detected with the secondary antibody Anti-Rabbit IgG H&L (1:5000 dilution) from goat coupled to the alkaline phosphatase (ab97048, Abcam™). After 2h incubation at room temperature, plates were revealed using PNPP substrate (34047, Thermo Scientific™) and evaluated using spectrophotometer at 405nm. Peptide accumulation was estimated using a calibration curve of the synthetic human IGF1.

10. Bioactivity assay of IGF1 peptides

To evaluate the activity of the plant produced peptides a cell proliferation assay was used. The assay was based on the MTT-microculture tetrazolium as earlier described (Mosmann, 1983) and performed using the kit EZ4U (BI-5000, Biomedica™). The MTT is a colorimetric assay for assessing cell metabolic activity, allowing the determination of the cytotoxic or proliferative activity triggered by the molecules tested. Briefly, cells at the exponential growth phase were harvested, centrifuged and incubated in a 5% humidified CO2 incubator at 37°C. The cell number was determined by hemocytometer counting. The recombinant peptides were added to the culture media (10 µl to final concentrations of 5, 20 and 80 ng/ml). In the control groups, 10 µl of the different solutions listed in Table M14 were added. Five

replicates were performed in each case. Cells were collected at 24 and 48 hours. MTT (stock solution 5 mg/ml, PBS (Sigma-Aldrich™)) was added and the plates were again incubated for four hours. Then, the plates were read immediately in a microplate reader (VICTOR Multilabel plate reader, PerkinElmer™) at 540 nm. As exprerimental controls, different samples were included: complete medium alone, extracts obtained from EV,P1b transformed cells, and apoplast extracted from plants infiltred with the EV. As blank control MTT without cells were used.

Table M14. Controls used in MTT assays

Test samples	Negative control	Positive control
dIGF1	EV,P1b medium	EV,P1b medium + sIGF1
dCPP-IGF1	EV,P1b medium	EV,P1b medium + sIGF1
pIGF1	Apoplast (EV)	ECF in PBS buffer + sIGF1
dCPP-IGF1	Apoplast (EV)	ECF in PBS buffer + sIGF1
LDs (AtOle-IGF1)	LDs (AtOle2)	LDs (AtOle2) + sIGF1
LDs (IGF1-AtOle)	LDs (AtOle2)	LDs (AtOle2) + sIGF1

dIGF1, IGF1 produced in *D. carota*; dCPP-IGF1, CPP-IGF1 produced in *D. carota*; pIGF1, *N. benthamiana* plant produced IGF1; pCPP-IGF1 *N. benthamiana* plant produced CPP-IGF1; sIGF1, synthetic IGF1.

RESULTS

1. SELECTION OF A VIRAL RNA SILENCING SUPPRESSOR (RSS)

The co-expression of viral RSS can significantly prevent the transgene-induced gene silencing, and enhance high expression levels of transgenes in plant leaves through Agrobacterium-mediated transient expression processes (Sainsbury and Lomonossoff, 2014). To evaluate whether the presence of viral RSSs can also enhance transgene expression and improve the production of recombinant proteins in plant cell cultures, we decided to assess the effect of viral RSS in the accumulation levels of our recombinant proteins *N. benthamiana*. To select the viral RSSs to be co-expressed in our assays, we first tested the activity of the RSSs derived from two plant viruses of the family *Potyviridae* already in use by our laboratory, the HCPro of WMV and the P1b of CVYV (Fig R1a).

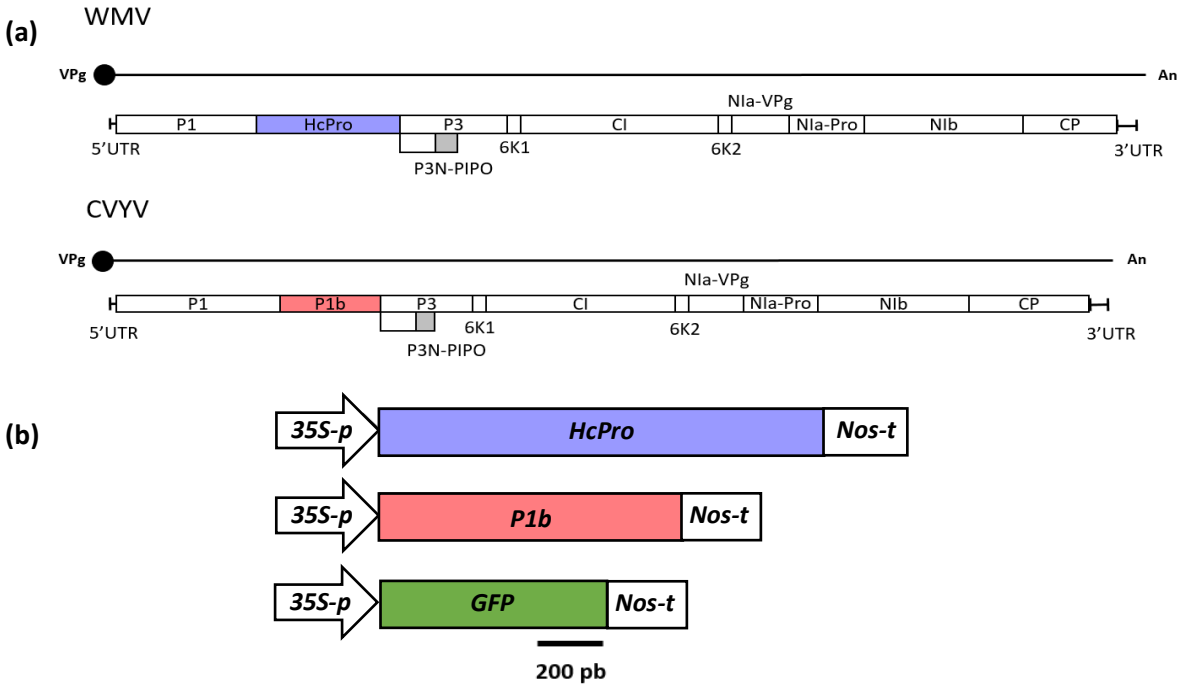


Figure R1. Maps of *Watermelon mosaic virus* (WMV) and *Cucumber vein mosaic virus* (CVYV) genomes, and gene constructs for the expression in plants of RNA silencing suppressors. (a) Schematic representation of the genome maps of WMV and CVYV with boxes corresponding to the mature gene products derived from auto-proteolytic processing of the polyprotein. Also, the partially out-of-frame product P3N-PIPO, deriving from RNA slippage, is indicated. The known RSSs are indicated by purple (HCPro) and salmon (P1b). **(b)** Gene constructs used to produce the HCPro, P1b and the reporter GFP in plants, are schematically represented by boxes, flanked by the CaMV 35S promoter (white arrow) and Nos (nopaline synthase gene) terminator (white box) regulatory elements.

Plant expression vectors containing the translatable cDNAs of the HCPro and P1b RSSs of interest under the control of the CaMV 35S (35S) promoter and the Nos terminator were available in the group (Fig R1b), and they were used to assess their suppression activity. Each plasmid was evaluated in transient gene expression assays in combination with the visually detectable reporter gene encoding the *green fluorescent protein* (*GFP*) (Fig R2). The assays were performed by agroinfiltration of the different

combinations of constructs in patches of *N. benthamiana* leaves. The *GFP* expression was then visualized using UV-light at different time points, and the GFP production was measured by quantitative spectrophotometry.

As shown in Fig R2, the signal corresponding to GFP fluorescence was more intense under UV light in the presence of either HCPro or P1b when compared to GFP agroinfiltrated with the empty vector control. These observations are in agreement with previous results finding RSS activity for HCPro and P1b gene products. The patches where GFP was agroinoculated with P1b were visualized consistently brighter than those with HCPro, suggesting a stronger suppressor activity for P1b. To confirm these observations, the amount of GFP was quantified by measuring the emitted fluorescence of total protein extracts from agroinoculated leaves. Leaf patches from three independently agroinfiltrated leaves were collected, weighted, and total protein extracted in carbonate buffer. A calibration curve using a serial dilution of known amount of GFP was prepared in an equivalent matrix control (protein extract from leaves infiltrated with an empty vector), and fluorescence determined. By comparison with the calibration curve, the amount of GFP in the samples was estimated, and the results are presented in Fig R3. The maximum protein concentration was obtained at 5 days post inoculation (dpi) when a value of 1.073 mg protein/g of fresh leaf was estimated for P1b and 0.899 mg protein/g of fresh leaf for HCPro. Higher values were always measured for P1b at the different time points evaluated, corresponding to values 23.23%, 16.21% and 38.36% higher at 3, 5 and 7 dpi, respectively. The background levels for negative control (GFP without any RSS) where <0.05 mg protein/g of fresh weight at all-time points. Differences were statistically significant between treatments in all the cases (Tukey test, $p \leq 0.05$). These results confirmed the RNA silencing suppression activities of P1b and HCPro and point out to a stronger and durable activity for P1b in *N. benthamiana* leaves. Therefore, the P1b silencing suppressor from CVYV was selected for future studies.

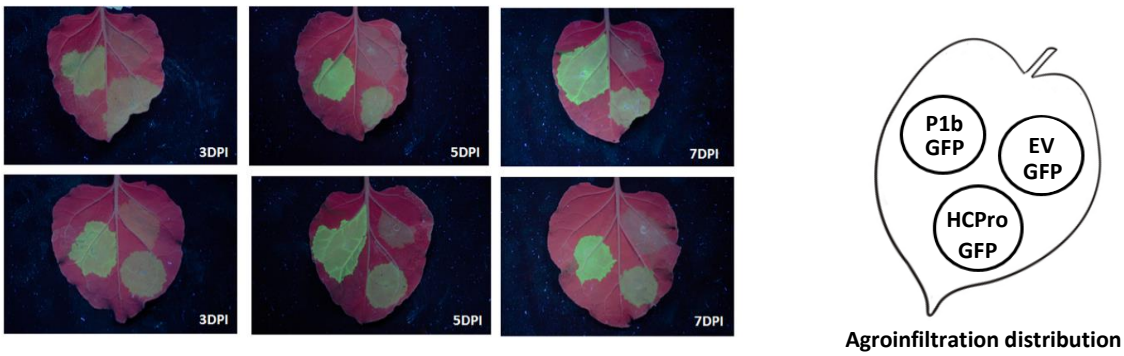


Figure R2. *GFP* expression in agroinfiltrated patches of *N. benthamiana* leaves. Representative images obtained under UV-light of *N. benthamiana* leaves agroinfiltrated with different *A. tumefaciens* combinations at 3, 5 and 7 dpi. The scheme on the right displays the tested combinations of RSSs with GFPs, including the negative control with empty vector (EV). Results of two independent experiments are shown in the upper and lower rows to illustrate the reproducibility of the observed differences.

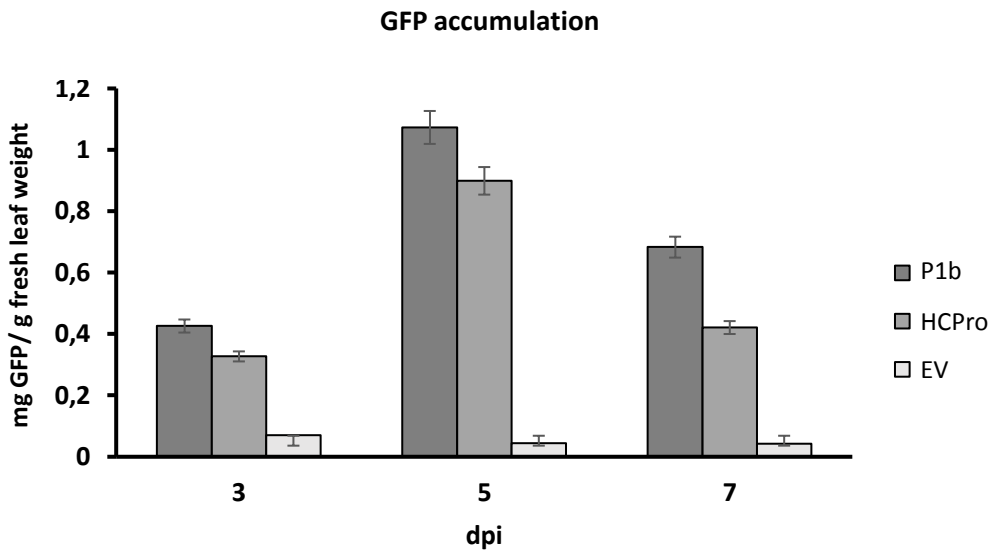


Figure R3. GFP accumulation in leaves agroinfiltrated with the indicated constructs. Quantification of GFP accumulation in protein extracts from leaf patches agroinfiltrated with the *GFP* reporter construct in combination with the indicated RSS gene construct or empty vector (EV). GFP was extracted with carbonate buffer and determined by spectrophotometry analysis at 488nm. The bars represent the average values of 3 leaves at 3, 5 and 7 dpi of 3 independent assays (n =9). The error bars show the ± SD of three independent samples.

2. LIPID BODIES IN PLANT CELL CULTURE SUSPENSIONS OF DIFFERENT PLANT SPECIES

LDs have been identified in diverse plant tissues, however available literature data indicates that they are not abundant in the model plant cell culture BY2 (Zhang *et al.*, 2016), and thus genetic modification or addition of external fatty acids to culture media are needed to achieve viewable LD levels (Gidda *et al.*, 2016; Zhang *et al.*, 2016). To characterize the LD content of other plant cell cultures, different plant cell lines from phylogenetic distant species (Fig R4) and their isolated LDs were observed under confocal fluorescence microscopy using Nile red staining. This staining is specific for neutral lipids, which are the main component of LDs (Greenspan *et al.*, 1985). We examined a battery of cell lines that were derived from *Arnica montana, Curcuma longa, Gossypium herbaceum, Cannabis sativa, Daucus carota, Sarcocapnos crassifolia, Arabidopsis thaliana* and *Nicotiana tabacum* (BY2) plants. They were part of Vytrus Biotech S.L. cell bank and were obtained using different plant tissues for callus induction and line generation (Table R1). All the visualizations were done on samples collected from liquid culture at stationary phase, and the same standard procedure was applied to extract and isolate LDs (see Material and Methods).

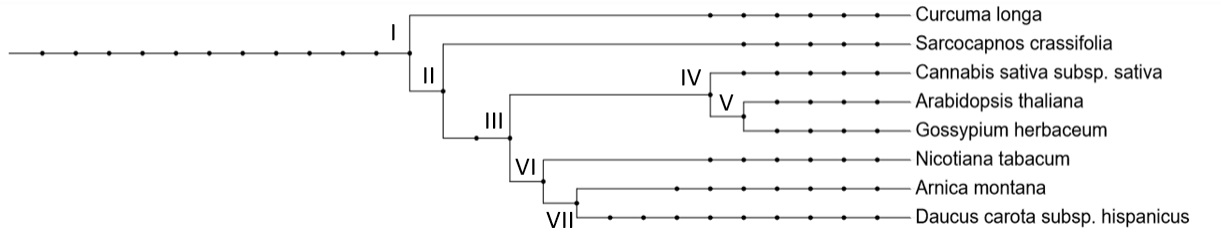


Figure R4. Phylogenetic tree of the plant species of the cell line collection. Species include *Arnica montana* (NCBI: txid436207), *Curcuma longa* (NCBI: txid136217), *Cannabis sativa* (NCBI: txid1678924), *Daucus carota* (NCBI: txid1179240), *Gossypium herbaceum* (NCBI: txid34274) *and Sarcocapnos crassifolia* (NCBI: txid185659). I, corresponds to the liliopsida-eudicotyledons axis, II to early-diverging eudicotyledons-gunneridae, III to rosids-asterids, IV to fabids-malvids, V to brasicales-malvales, VI to lamiids-campanulids and VII to asterales-apiales. The tree was generated with phyloT tool data version 2019.1 (https://phylot.biobyte.de/) using NCBI taxonomy nodes 2 071 652 in the analysis.

Table R1. Plant tissues used to obtain the different cell lines.

Plant cell lines	Tissues used for callus induction
Curcuma longa	Rhizome
Sarcocapnos crassifolia	Seedling
Cannabis sativa	Seedling
Arabidopsis thaliana	Seedling
Gossypium herbaceum	Seedling
Nicotiana tabacum	Seedling
Arnica montana	Seedling
Daucus carota	Root

LDs were visualized as discrete spherical cytoplasmic vesicles in all the cell lines (Fig R5-R7). Cells from *A. montana, C. longa, G. herbaceum, C. sativa, D. carota and S. crassifolia* showed a high number of LDs, (Fig R5-R6) whereas *N. tabacum* and *A. thaliana* cells showed only few LDs (Fig R7). We also observed that LDs distributed homogeneously in the cytoplasm of most of the cell lines with the exception of *C. sativa* and *C. longa* cells in which LDs were aggregated. The number of LDs was measured in 25 cells and statistically significant differences were observed (Fig R8a). The *D. carota* and *G. herbaceum* cells showed the highest number of LDs.

LDs can be easily fractioned from other cell components in density gradient solutions. Using a simple flotation-centrifugation method, LDs were isolated from the different cell lines except *from N. tabacum*, *A. thaliana* and *S. crassifolia* cell lines. These lines were the ones with the lowest LD content

among all the analyzed lines, or the one with the smallest LDs (*S. crassifolia*) (Fig. R7). Isolated LDs were then visualized under confocal microscopy as Nile Red stained spherical organelles with different sizes depending on the species from they were obtained (Fig R8b). Although a high variance in the size of LDs was observed, we detected an opposite trend in LD size and number: those cell lines containing the highest number of LDs showed many of them of small size, whereas those with a low number of LDs contained many large LDs.

Our results show that a high number of LDs are accumulated in the cells of most of the analyzed plant cell lines with the exception of BY2 and *A. thaliana* cells. Our results do not allow us to stablish any correlation between LD profile and phylogenetic relationship or plant tissue origin from cell lines.

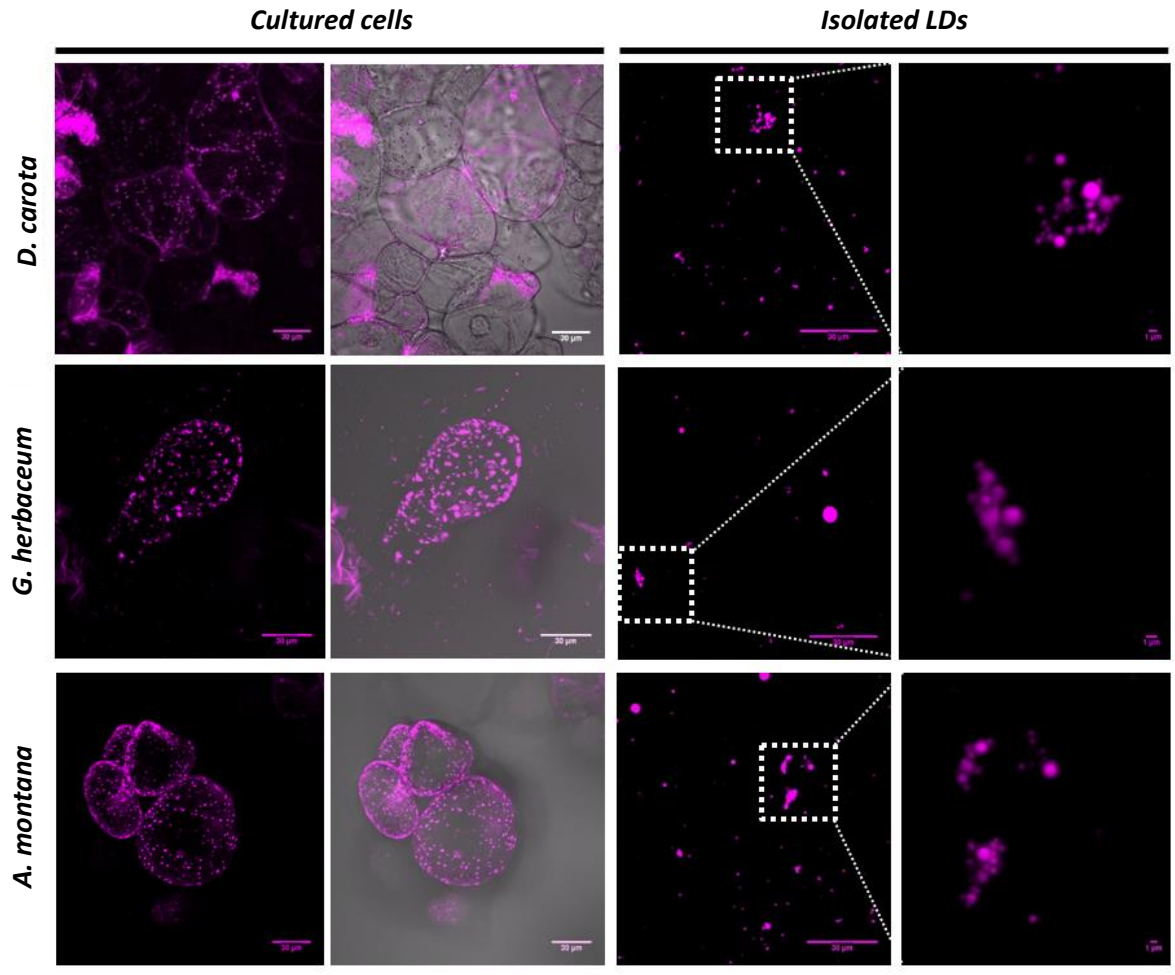


Figure R5. Plant culture cells of certain species contain a significant number of LDs. Representative confocal laser microscopy images of culture cells (left panels) and isolated LDs (right panels) from the species indicated. The samples were stained with Nile Red and visualized under confocal microscopy. The images of plant cell cultures visualization correspond to Nile Red channel or merged with bright field. Amplified images of white squares are shown next to each image. Scale bars correspond to 30 µm and 1 µm as indicated.

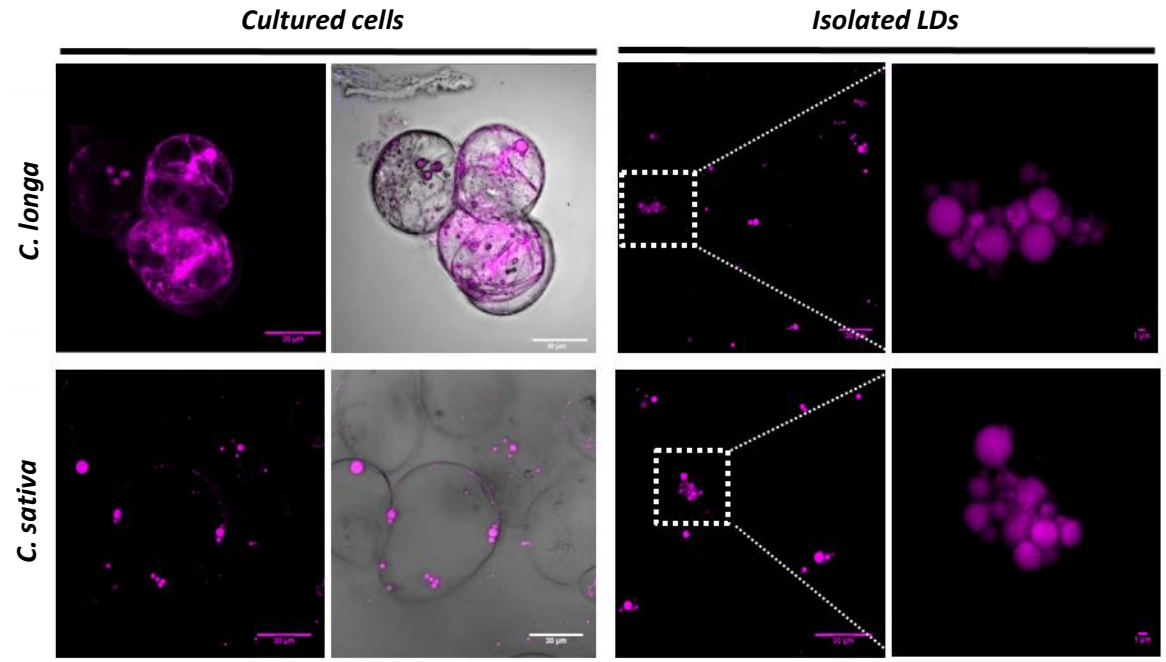


Figure R6. *Curcuma longa* and *Cannabis sativa* culture cells contain large LDs. Representative confocal laser microscopy images of plant culture cell (left panels) and isolated LDs (right panels) from indicated species. The samples were stained with Nile Red and visualized under confocal microscopy. The images of plant cell cultures visualization correspond to Nile Red channel or merged with bright field. White squares are amplified next to each image. Scale bars correspond to 30 µm and 1 µm as indicated.

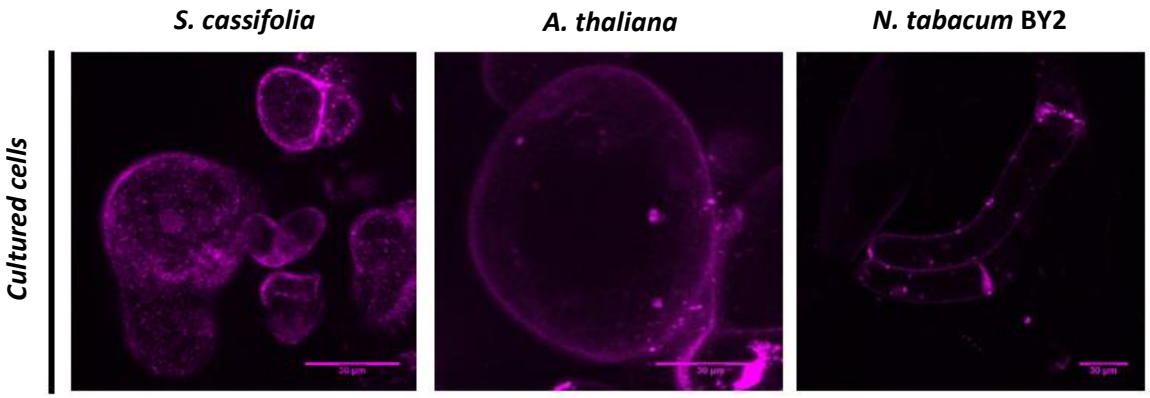
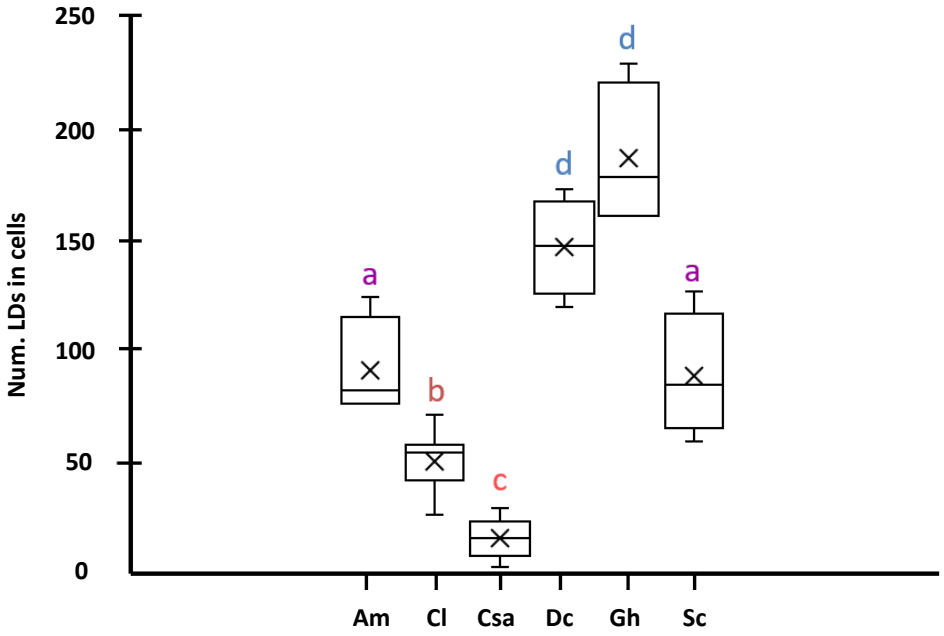


Figure R7. *Sarcocapnos crasifolia*, *Arabidopsis thaliana* and *Nicotiana tabacum* culture cells contain small LDs. Representative confocal laser microscopy images of plant culture cell from indicated species. The samples were stained with Nile Red and visualized under confocal microscopy. The images of plant cell cultures visualization correspond to Nile Red channel or merged with bright field. Scale bars correspond to 30 µm as indicated.

(a)



(b)

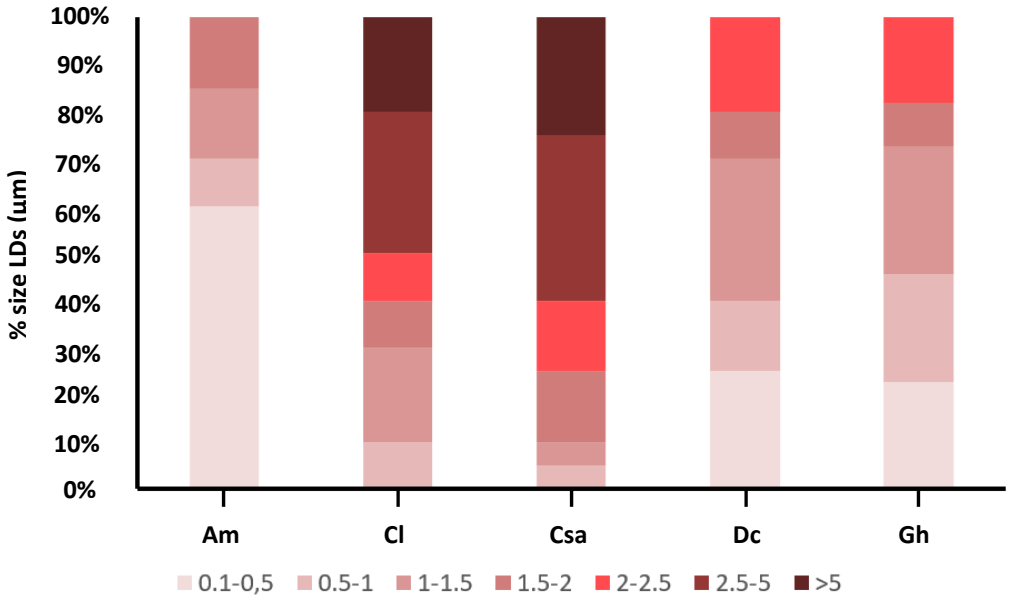


Figure R8. Distribution of LD sizes from plant culture cells. **(a)** Boxplot with average and deviations of the number of LDs per cell of the indicated cell lines: *Arnica montana* (Am), *Curcuma longa* (Cl), *Cannabis sativa* (Cs), *Daucus carota* (Dc), *Gossypium herbaceum* (Gh) and *Sarcocapnos crassifolia* (Sc). LDs were counted in 5 cells per sample and five independent samples (n = 25). Same letters indicate no statistically significant differences according to Tukey test ($p \leq 0.05$). **(b)** Distribution of sizes (categories correspond to diameter in µm) from 20 isolated LDs from the indicated cell lines. LDs could not be isolated in the case of Sc.

3. GENERATION OF A *D. CAROTA* CELL LINE

3.1 *D. carota* callus induction and propagation

Based on our observations of LDs in plant cell suspension cultures, we selected the species *D. carota* for our studies on recombinant protein production using the oleosin fusion strategy. The cell line derived from this species presented a high number of small LDs that can be easily purified, which makes it a suitable matrix for the production of recombinant oleosin fusion proteins. Furthermore, *D. carota* cell growth in solid and liquid media is compatible with industrial productions, and cell lines from this species are currently used for the production of therapeutic proteins (Wolfson, 2013). However, the *D. carota* cell line available from Vytrus cell bank was subjected to restrictions of use. For that reason, we generated a new *D. carota* cell line using carrots as initial material. Carrot roots were acquired and used for callus induction in the appropriate media at different hormone concentration, as indicated in Material and Method. After 26 days of incubation, calli were developed surrounding the carrot explants (Fig R9a), and then transferred to fresh media for propagation for two 4-weeks cycles (Fig R9b). The explants showed whitish color, probably due to discoloration during the sterilization procedure, however the typical orange color of carrot roots remained in the inner part of the explants, and the different individual callus exhibited properties that were evaluated during the selection process.

(a) **(b)**

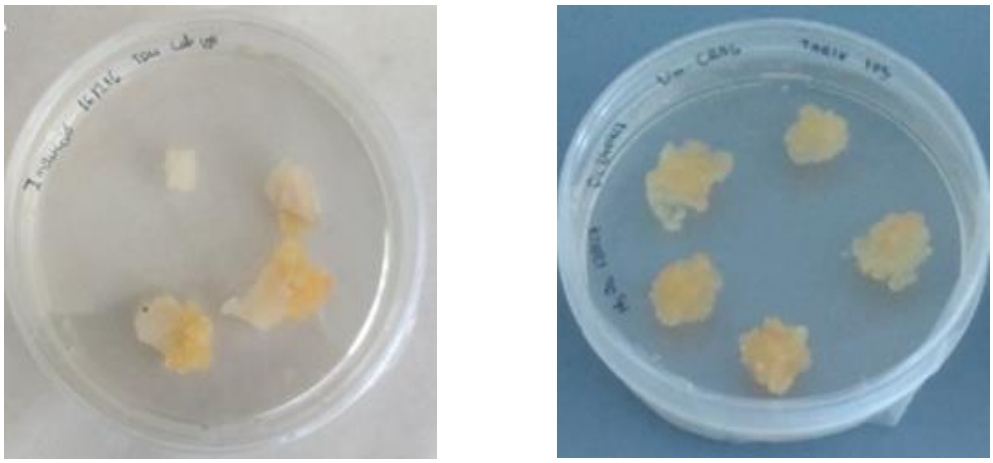


Figure R9. *D. carota* callus induction and propagation. **(a)** Callus induction from carrot root explants in MSI6 media. **(b)** Propagation of the callus obtained from explant plates.

Callus induction rates, fresh weights (FW), cell textures and color were evaluated for callus from different media. Callus formation was observed in all the tested media. No significant differences were observed when callus fresh weights were compared (Tukey test analysis, $p \leq 0.05$). Friable (easy to disperse, soft and with low plant regeneration potential) and organogenic (compact and adequate for plant regeneration) calli were produced using most of the media. Friable calli were selected due to their ability to grow in liquid cultures and potential to propagate in large volumes, while organogenic was discarded. It was noticed that organogenesis was associated to callus with a yellowish coloration, while friable callus tends to show a light orange/orange coloration (Table R2).

Table R2. Effect of 2,4-D in callus induction from *D. carota* explants.

Medium	2,4-D (mg/l)	Number of explants platted	Callus induction %	Callus texture (1)	Colour (2)
MSI1	0,25	4	75	Friables	light orange
MSI2	0,5	4	50	Friables/organogenic	light orange
MSI3	0,75	4	100	Organogenic	yellowish
MSI4	1	4	25	Organogenic	yellowish
MSI5	1,25	4	100	Organogenic	yellowish
MSI6	1,5	4	75	Friables/organogenic	light orange
MSI7	1,75	5	100	Friables/organogenic	light orange
MSI8	2	5	60	Friables/organogenic	light orange
MSI9	2,25	5	80	Friables/organogenic	light orange/ yellowish
MSI10	2,5	5	60	Friables	orange

Quantification at 26 days of *D. carota* callus induction and biomass characteristics in media containing different concentrations of the 2,4D hormone. The percentage of callus induction was calculated using the number of explants and of independent calli per condition. The texture **(1)** and color **(2)** were evaluated visually and using tweezers, respectively.

The selection of the calli to be used for propagation and to initiate a cell suspension was based on friability and the capacity of growth in the industrial compatible media (MSDc). The friable calli derived from MSI6 was selected, then propagated in the same media for three consecutive cycles. Each propagation round comprised a 4-weeks growth period followed by a callus piece selection, and platting in fresh solid media. Organogenic material was discarded during the plating rounds.

Total callus material was divided in two parts. The first part was maintained in induction media (MSI6), and the second was subcultured to MSDc media. This strategy of maintaining half of the material in

MSI6 served as a reservoir in case the adaptation to MSDc media failed. The MSDc media differs with MSI6 in the contents of vitamins and 2,4-D, and therefore callus needs to adapt to MSDc through several propagation rounds. MSDc media had been optimized for *D. carota* cell culture growth, and it is compatible with industrial biomass production. After four rounds of propagation in MSDc, the generated callus continued growing as friable material showing a good adaptation to the media. Once the adaptation process was completed successfully, the backup material in MSI6 media was discarded finally.

3.2 *D. carota* cell suspensions and dynamics of growth

Cell suspensions were initiated using MSDc adapted calli by inoculating 20 ml of MSDc liquid media in 50 ml flasks and incubating them with agitation in a rotary shaker at 110 rpm and 24 °C. Cell suspensions were sub-cultured every fourteen days, with a total of six cycles for the analysis of the dynamics of growth. During every subculture, the biomass was filtered and adjusted to a concentration of 4 g/20ml to initiate the next round of growth. FW was evaluated at day fourteen. The *D. carota* cell line available at Vytrus (Dc-Vytrus) was used for comparative purposes.

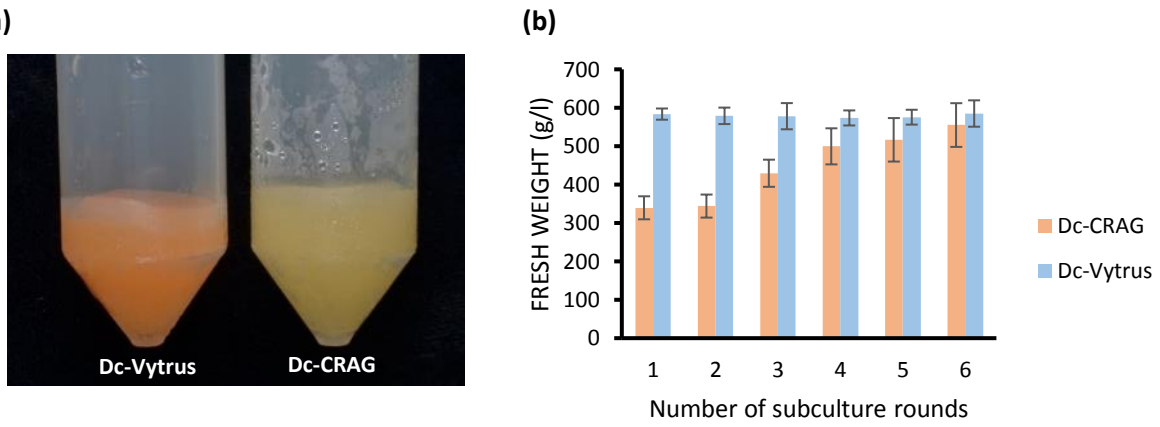


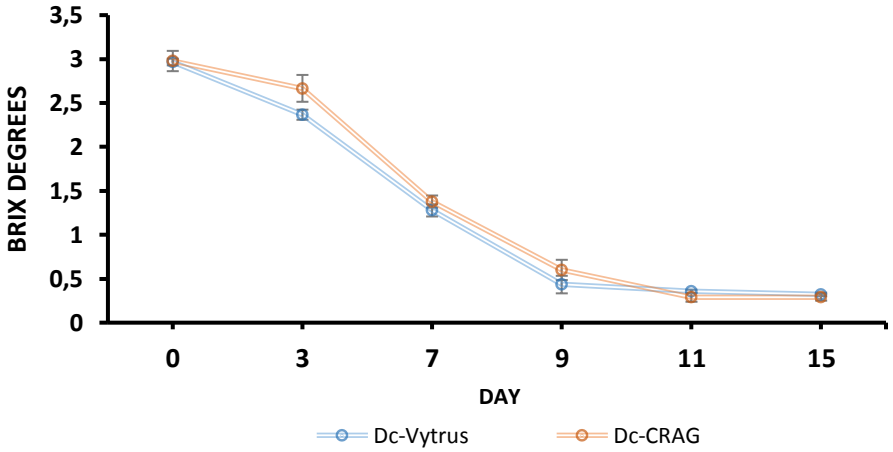
Figure R10. Growth dynamics of *D. carota* cell suspension. **(a)** Appearance of *D. carota* cell culture sediments from the line generated at CRAG (Dc-CRAG) and the one available at Vytrus (Dc-Vytrus) at day 14 after subculture. **(b)** Fresh weight of cell suspension cultures (20 ml) at day 14 after 1 to 6 rounds of subcultures. Values are the mean of three biological replicates ± SD.

The most evident difference between cell lines was the coloration, orange for Dc-Vytrus and yellow for Dc-CRAG (Fig R10a). An improvement in the growth rate was observed along the adaptation process for the Dc-CRAG cell line, that was measured as an increment of the fresh weight in every new subculture round (Fig R10b). Whereas the already established cell line from Vytrus maintained the

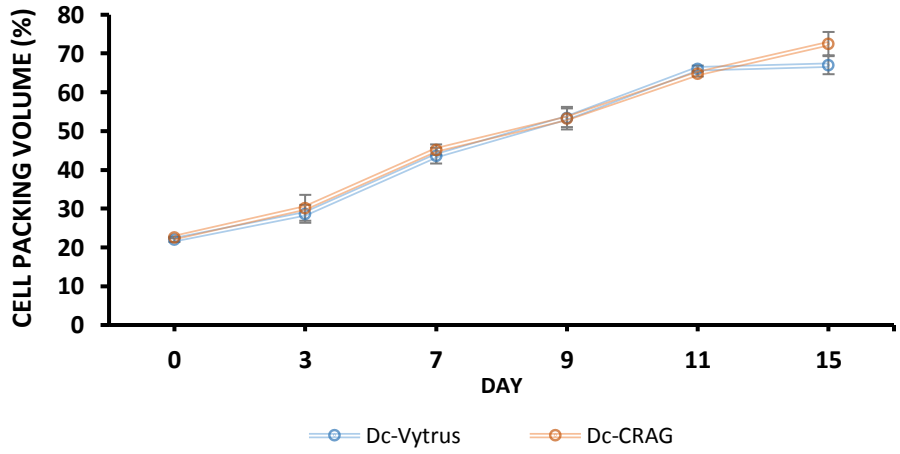
same growth rate along the different subculture rounds (Fig R10b). Similar growth was measured after 6 rounds of subculture for both cell lines reaching growth values around 600 g/l.

Next, we evaluated the growing kinetics of both *D. carota* cell suspension lines at large scale using 0.5L flasks, 200 ml liquid media and 40 g of cell material. Several parameters were scored to have a complete kinetics of growth, including reduction of brix degrees (it is a measure nutrient content of the media, mainly sugars), the pH of the media, the fresh weight (FW) and the cell packing volume (VEC) of the culture. Results are shown in Fig R11. We observed that nutrient consumption dynamics followed a similar behavior for both cell lines, with a more pronounced steep slope during the first 72h for Dc-Vytrus (Fig R11a). The cultures showed a diminution in the pH during the first cultivation days, but later the pH increased reaching similar values to the initial ones in both cell lines. The pH in Dc-CRAG culture along the process was lower than shown by Dc-Vytrus cell cultures (Fig R11b). The biomass increase, measured as VEC and FW, was almost identical for both cell lines (Fig R11c). The three biological replicates showed only minor differences (Fig R11d).

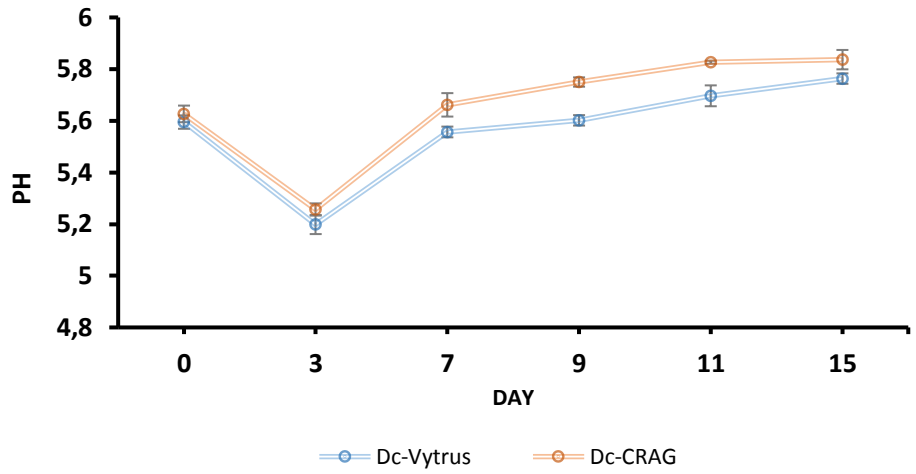
(a)



(b)



(c)



(d)

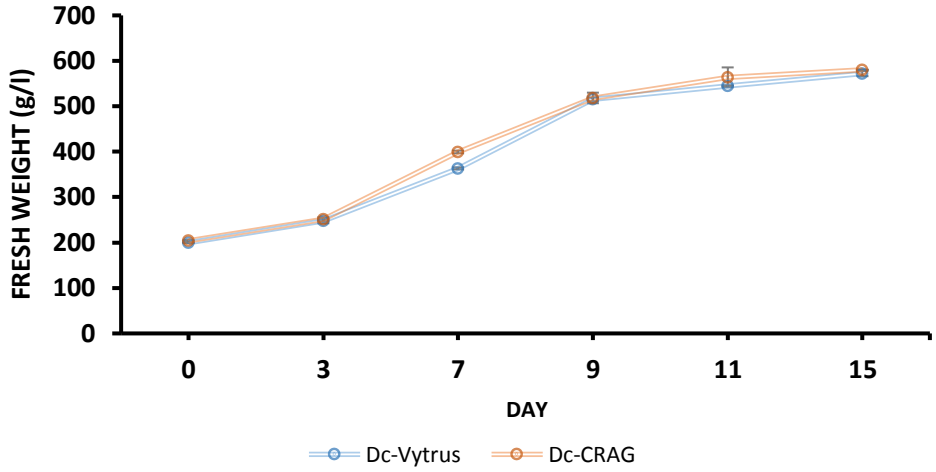


Figure R11. Growth curves of *D. carota* cell suspensions. Graphs represent the mean values ± SD of brix degrees **(a)**, pH **(b)**, cell packing volume expressed (VEC) in % **(c)** and fresh weight (FW) expressed in g/l **(d)**, of Dc-Vytrus and Dc-CRAG cell suspensions measured at indicated time points. Values are the mean of three biological and technical replicates (n=6).

One evaluated the different growth curves of *D. carota* cell cultures, we proceed to evaluate the growth kinetic parameters for both cell lines. Maximum VEC, exhaustion time (Ta) and maximum velocity at the exponential phase (Vexp) were determined according to the equations shown in Materials and Methods. The Fig R12 show the curves used to refer the parameters. Parameter values are listed in Table R3. Regarding the growing phases along the process, our data estimated that the latency phase lasted until 3 days after subculturing, followed by an exponential growth phase extending until day 11, and reaching then a final plateau phase.

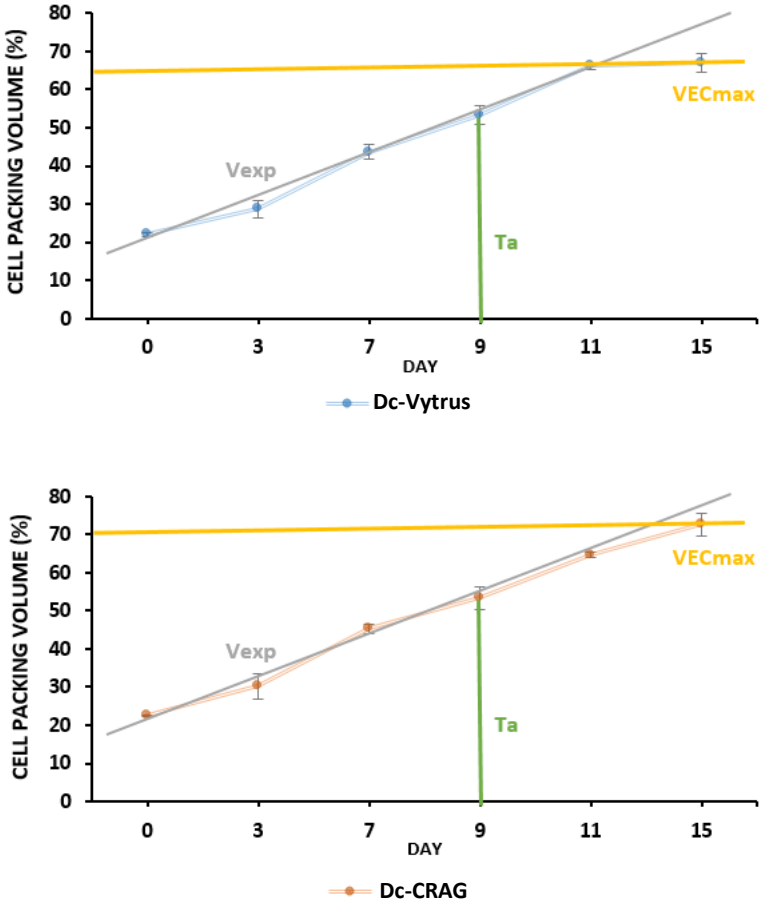


Figure R12. Growth kinetic parameters of Dc-Vytrus and Dc-CRAG cell culture suspensions. VEC curves of Dc-Vytrus (blue) and Dc-CRAG (orange) cell suspensions used for the determination of the growth kinetic, including the VEC max (yellow line), the maximum velocity at the exponential phase (Vexp, gray line) and exhaustion time (Ta, green line). Maximum VEC is represented as yellow line, while Ta in green and V.exp in gray

Table R3. Growth kinetic parameter values of Dc-Vytrus and Dc-CRAG cell suspension lines.

Kinetic value	Dc-Vytrus	Dc-CRAG
VEC$_{max}$ (%)	69	71
Ta (Days)	15	15
V$_{exp}$ (g/h)	0,197	0,187

3.3 LDs content of *D. carota* cell lines

Our main interest on the generated cell line was to use it for the production of recombinant proteins as oleosin fusion proteins, therefore we decided to characterize its LD content. For that, we examined the cultured cells under confocal microscopy at the end of their exponential growth (11 days after

subculture). Cells from the line Dc-Vytrus were visualized in parallel for comparative purposes. As shown in Fig R13a, Dc-Vytrus and Dc-CRAG cells showed LDs that homogenously distributed in the cytoplasm. LDs were easily isolated from cell suspensions showing similar spherical appearance when Nile red stained and visualized under confocal microscopy.

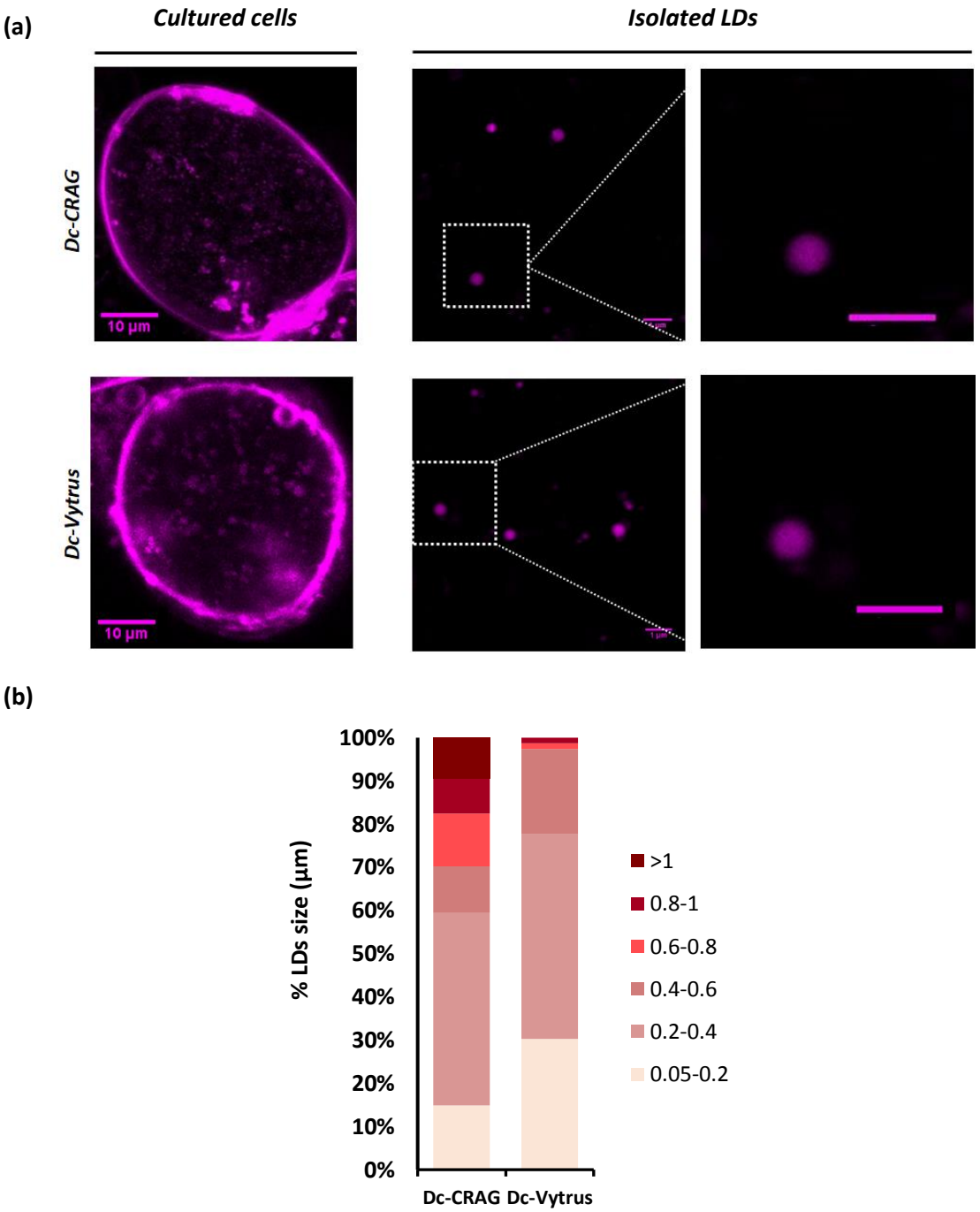


Figure R13. Visualization of LDs from *D carota* cell lines. **(a)** Confocal laser microscopy images of cells from Dc-CRAG and Dc-Vytrus lines (left panels) and their isolated LDs (right panels). The samples were stained with Nile red previous visualization. White square images are amplified next to each image. Scale bars correspond to 10 µm, 1 µm and 0.1 µm as indicated. **(b)** Distribution of range sizes (diameter in µm) measured from a total of 75 isolated LDs.

Differences were observed when compared LDs sizes purified from the different *D. carota* cell lines. The LDs purified from Dc-CRAG exhibited bigger sizes than the purified from Dc-Vytrus. The 85% of the LDs were found to be bigger than 0.2 µm diameter for Dc-CRAG, while only around 69% for Dc-Vytrus (Fig R13b).

4. PRODUCTION OF OLEOSIN FUSION PROTEINS

4.1 Oleosin selection

Genes encoding for different oleosin proteins have been predicted to be present in the genome of *D. carota* (Q43123, A0A164UMJ0). These proteins are not as well characterized as other oleosin proteins from *A. thaliana, Oryza sativa, Brasica napus, Helianthus annus, Zea mays* or *Glycine max* (Chuang *et al.*, 1996; Siloto, 2006; Umate, 2012). Some of these oleosins have already been used as carriers for the production of recombinant proteins in LDs (Markley et al. 2005; Yang et al. 2015; Montesinos et al. 2016). In our assays, we decided to use the AtOle2 protein from *A. thaliana* as the carrier of our recombinant proteins. The selection was based on its size and on the availability of the cDNA in the lab. The AtOle2 is the largest oleosin protein in Arabidopsis plants, with a molecular weight of 21 kDa, which will facilitate the identification of the fusion proteins among the other oleosins by simple differences in protein sizes.

4.2 Plant expression vectors

To evaluate whether oleosin proteins could be used as carrier proteins to LDs in plant cell suspensions, two plant expression vectors were prepared for the production of oleosin fusion proteins to the marker GFP protein. The GFP can be easily visualized under fluorescence microscopy allowing the monitoring of the fusion protein localization and accumulation dynamics. The hypothesis is that AtOle2 protein will carry the GFP to LDs, where AtOle2 will be inserted into the phospholipid layer and the GFP exposed on the LD surface facing the cytoplasm (Fig R14). The fusion proteins contained the GFP fused in N- or C-terminal with the AtOle2 (AtOle2-GFP and GFP-AtOle2) to evaluate the positional effect on LDs characteristics (Fig R15a). The sequence encoding for the *Tobacco etch virus* (TEV) protease recognition site was introduced between the two proteins, the AtOle2 and the GFP (Fig R15b), to allow the proteolytic release of the individual proteins if needed. The chimeric fusion genes were under the control of 35S promoter and the Nos terminator. Gateway technology was used for cloning the

chimeric genes in the plant expression vectors pGWB5 and pGWB6 (GenBank: AB289768.2 and AB289769.1).

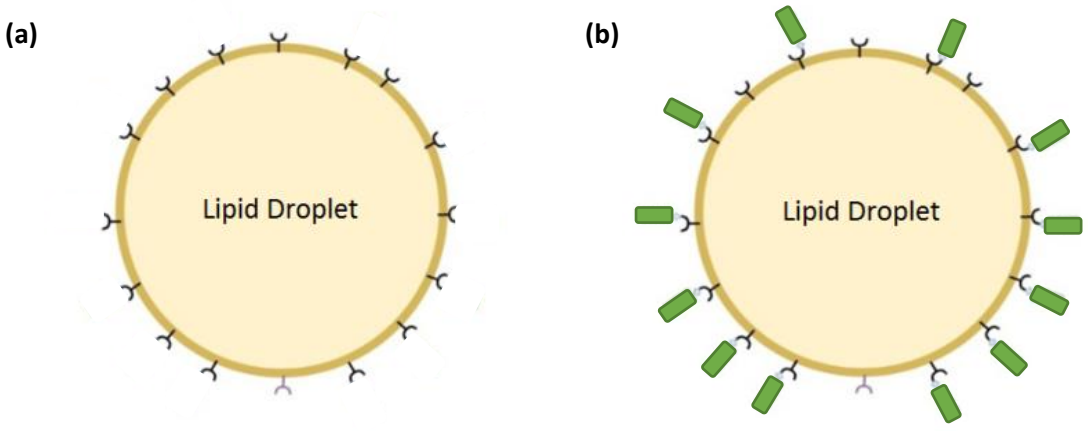
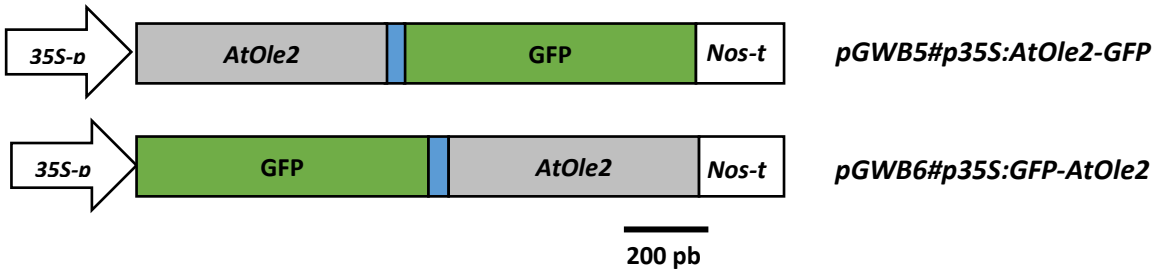


Figure R14. Native and recombinant LDs carrying oleosin-GFP fusion proteins. Schematic representation of a native **(a)** or recombinant **(b)** LD containing oleosin (black fork) or oleosin-GFP (black fork with green box). The lipid layer surrounding the LD is shown in golden color and the LD core of triacylglycerol in yellow.

(a)



(b)

AtOle2-GFP

MADTHRVDRTDRHFQFQSPYEGGRGQGQYEGDRGYGGGGYKSMMPESGPSSTQVLSLLIGVPVVGSL
LALAGLLLAGSVIGLMVALPLFLLFSPVIVPAALTIGLAMTGFLASGMFGLTGLSSISWVMNYLRGTRRTVP
EQLEYAKRRMADAVGYAGQKGKEMGQHVQNKAQDVKQYDISKPHDTTTKGHETQGRTTAAENLYFQ
SKGGRADPAFLYKVVRSRGSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT
GKLPVPWPTLVTTFTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL
VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIG
DGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELY

GFP-AtOle2

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFTYGVQ
CFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGH
KLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS
KDPNEKRDHMVLLEFVTAAGITHGMDELYITSLYKKAGSAAAPFTENLYFQSMADTHRVDRTDRHFQF
QSPYEGGRGQGQYEGDRGYGGGGYKSMMPESGPSSTQVLSLLIGVPVVGSLLALAGLLLAGSVIGLMV
ALPLFLLFSPVIVPAALTIGLAMTGFLASGMFGLTGLSSISWVMNYLRGTRRTVPEQLEYAKRRMADAVG
YAGQKGKEMGQHVQNKAQDVKQYDISKPHDTTTKGHETQGRTTAA

Figure R15. Chimeric genes for the production of AtOle2-GFP and GFP-AtOle2 proteins. **(a)** Schematic diagram of the constructs for the expression of the oleosin fusion genes under the control of the 35S promoter (35S-p) and the Nos terminator (Nos-t). **(b)** Amino acid sequence of the fusion proteins AtOle2-GFP and GFP-AtOle2, containing the GFP (green) and the AtOle2 (gray), linked through TEV protease recognition site (TEVs, blue) and a spacer residual sequence region deriving from Gateway™ cloning site (black). The triangle indicates the TEV protease cleavage site.

4.3 *N. benthamiana* transient expression assays

We evaluated the expression of the chimeric genes in plants using a transient expression assay in *N. benthamina* leaves, as a rapid system to examine intracellular localization of fluorescently-tagged proteins. For that, the *A. tumefaciens* carrying pGWB5#p35S:AtOle2-TEVs-GFP or pGWB6#p35S:GFP-TEVs-AtOle2 plasmids were infiltrated separately or in combination with constructs allowing the expression of the *P1b* RSS into leaf patches. The leaves were visualized by confocal microscopy at 1, 2, 3, 5 and 7 days post infiltration.

At 1 dpi, leaf patches expressing *P1b* and oleosin fusion proteins with the GFP displayed a clear fluorescent signal, whereas those patches without P1b did not show any detectable fluorescent signal, indicating that the presence of an active RSS was needed to allow detectable expression of the oleosin-related products. The intensity of fluorescence was reduced at 2 dpi and disappeared completely at later visualization times. This recombinant protein accumulation dynamic was different from the observed in *N. benthamiana* plants co-expressing *P1b* and *GFP* (in a free form, not fused to oleosin), where fluorescence could be visualized until 7 dpi.

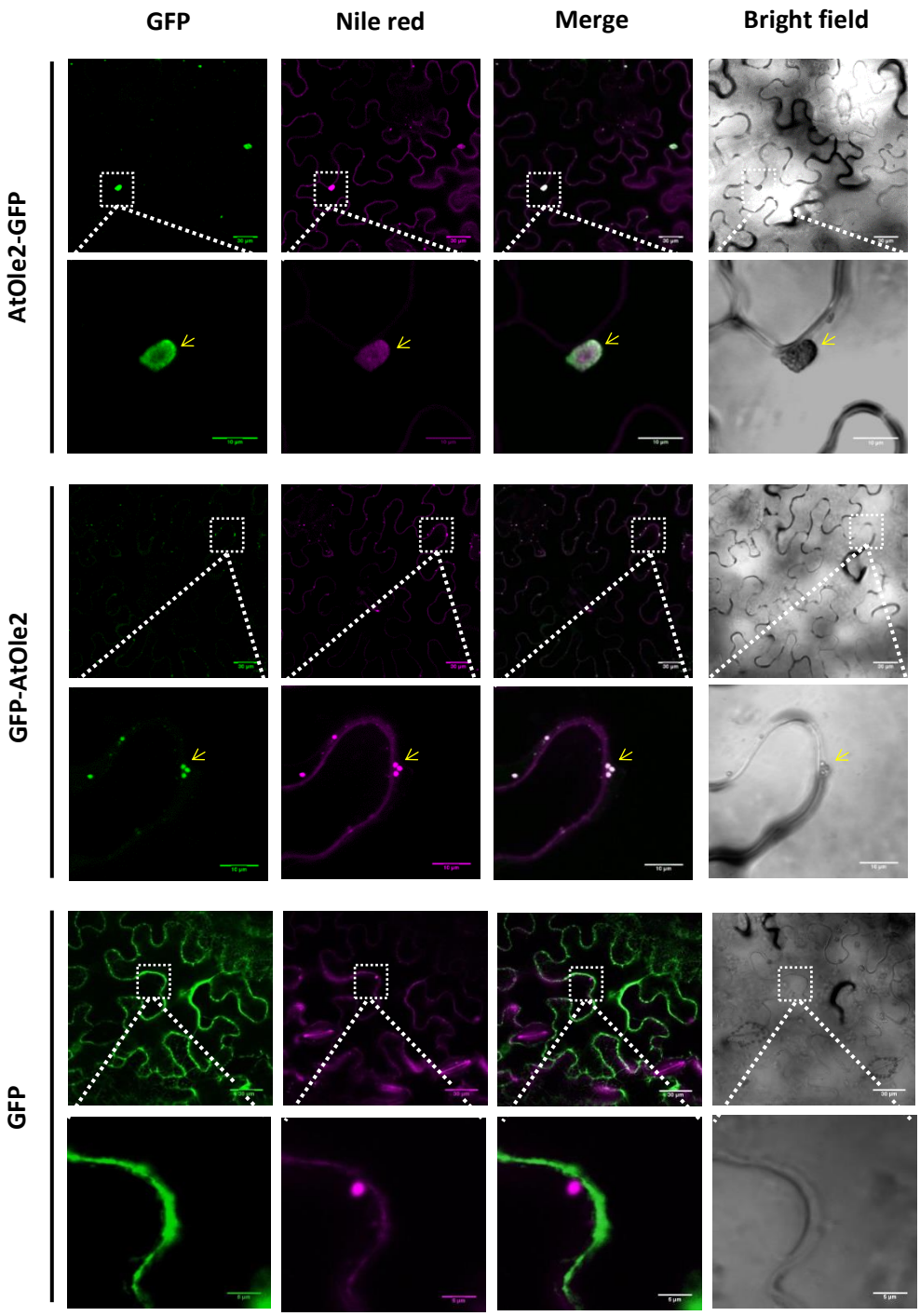


Figure R16. AtOle2-GFP and GFP-AtOle2 proteins are targeted to LDs in *N. benthamiana* leaves. Confocal laser microscopy images of epidermal cells accumulating the AtOle2-GFP (top panels), GFP-AtOle2 (central panels) and GFP (low panels) proteins visualized under fluorescence for GFP, Nile red, and bright field. Merged images between GFP and Nile Red channels are also shown. White squares are amplified in the below images. Arrows point to LDs. Scale bars correspond to 10 µm and 30 µm as indicated.

The *AtOle2-GFP* expression leads to the accumulation of the fusion protein in large cytoplasmic aggregates (around 8-10 µm of diameter) that are also stained with Nile red (Fig R16). These structures seem to be aggregates of tiny LDs containing the recombinant protein, and they were absent in control

leaves infiltrated with GFP. Structures resembling LDs could be intuited as buds in the surface of the aggregates, but they were difficult to be identified properly. On the other hand, the *GFP-AtOle2* expression resulted in the accumulation of the fluorescent protein associated to small, spherical organelles (0,4-0,6 µm of diameter), which were also Nile red stained, and consequently identified as LDs (Fig R16). These LDs were not present in leaves agroinfiltrated with the GFP control. These results suggested that the production of the AtOle2 fusion proteins induced the formation of LDs in *N. benthamiana* leaves. The fusion of the GFP at the C-terminal of AtOle2 appeared to promote the aggregation of the recombinant LDs, whereas the GFP fusion in N-terminal maintained the recombinant LDs as separate structures.

Altogether, our results showed that oleosin fusion to fluorescent proteins are produced *in N. benthamiana* leaves and targeted to LDs. We also observed that the accumulation of the oleosin fusion proteins induces the formation of LDs in the leaves. The recombinant LDs carrying the C-terminal tagged fusions showed a tendency to aggregate, whereas those carrying the N-terminal tagged fusions are maintained as discrete organelles. The fusion proteins are visualized on the periphery of LDs.

5. *D. CAROTA* CELL CULTURE TRANSFORMATION

To obtain calli expressing oleosin fusion genes, we performed a series of transformation assays with Dc-CRAG cell line using *A. tumefaciens*. The protocol for cell culture transformation was previously evaluated using a binary plant expression plasmid for the production of GFP protein whose accumulation can be easily monitored under UV light.

5.1 Antibiotic resistance assay

Previously to the transformation, the tolerance of calli to the selection antibiotics was evaluated. Calli were propagated in MSDc plates supplemented with different kanamycin or hygromycin concentrations (10, 25 and 50 mg/l) and timentin (100 mg/l) in all cases. Timentin is an antibiotic used for *A. tumefaciens* clearance, that is described to have minimum effect on callus growth (Cheng *et al.*, 1998). Calli contours were marked in the lower part of the culture dishes at the beginning of the propagation cycle to visually monitor callus growth. Also, the increase in biomass was evaluated by comparing callus weight at the beginning and at the end of the propagation cycle.

Calli did not show any apparent growth and remained inside the marked area during all the propagation period in the media supplemented with 25 and 50 mg/l kanamycin, whereas in 10 mg/l

kanamycin media grew and surpassed the marked area. Calli presented discoloration in all the plates along the period of cultivation and showed increased levels of apparent hydration and friability. For the hygromycin antibiotic, the calli remained inside the marked area at all tested concentrations and also presented discoloration. The calli platted in media supplemented only with timentin (100 mg/l), grew occupying most of the plate and showing a strong yellowish color and without perceptible hydration or friability changes (Fig R17). Biomass quantification showed that fresh weight did not increase in any of the plates containing kanamycin or hygromicyn antibiotics, with the only exception of the lower concentration of kanamycin (Fig R18). Whereas the calli propagated in the presence of timentin increased their biomass from 0.5 g/callus to 3.7 g/callus. These results indicate that both kanamycin and hygromycin can be used as selectable antibiotics in our transformation experiments, and timentin can be used to eliminate Agrobacterium cells without interfering with callus growth.

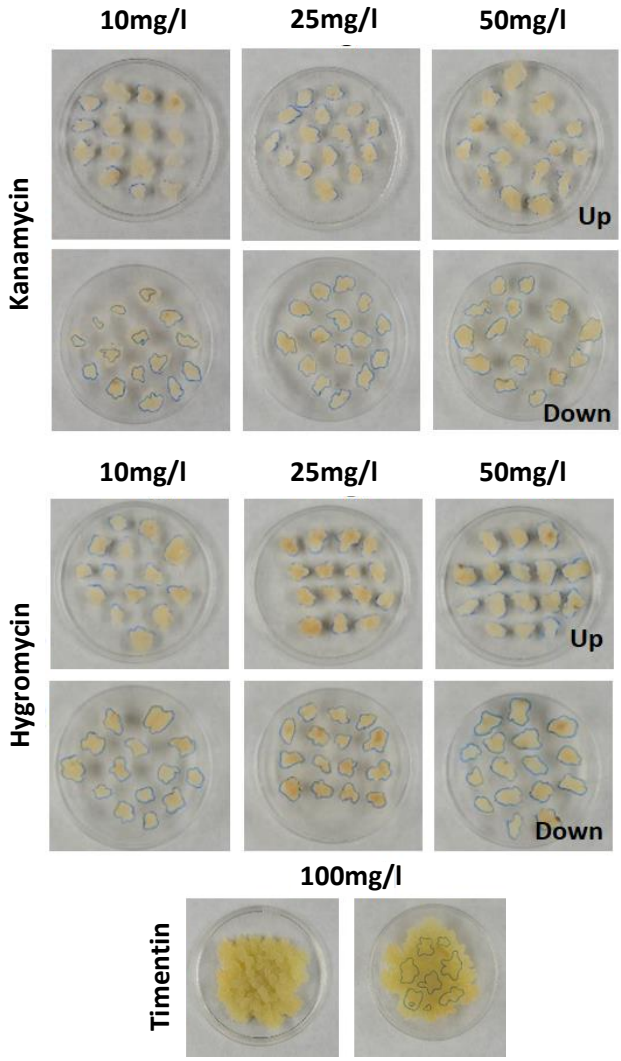


Figure R17. Visualization of *D. carota* calli in media containing the indicated selection antibiotics. Photographs of Dc-CRAG calli propagated for 35 days on MSDc media supplemented with the indicated concentrations of antibiotics. Calli borders were marked at the propagation day. The images are shown from the up and downside of the plate.

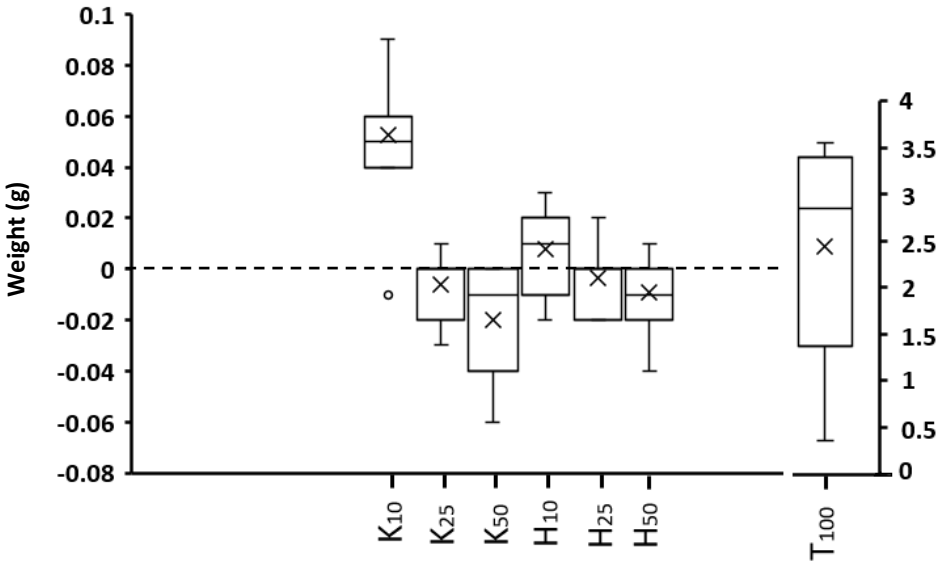


Figure R18. Weight variation of *D. carota* calli in antibiotic containing media. Boxplot of the callus weight difference in a propagation cycle. Callus grown in MSDc plates supplemented with the antibiotics kanamycin (K) or hygromycin (H) at 10, 25 or 50 mg/l, or timentin (T) at 100 mg/l. The boxplots correspond to values of fifteen biological replicates.

5.2 Evaluation of the *D. carota* callus transformation protocol

Dc-CRAG calli were transformed using the plant expression vector pGWB6 that carries the *GFP* expression cassette or an empty vector (pGWB402) via *Agrobacterium tumafaciens*. Hygromycin and kanamycin resistance genes in the T-DNA region were used as selectable markers. After 45 days, 3 growing calli were observed in hygromycin plates and 5 more in kanamycin plates transformed with pGWB6. In the case of the transformation with the empty vector (EV), 2 hygromycin and 3 kanamycin resistant calli were obtained. Therefore, we obtained 8 putative transformed calli carrying the construct for *GFP* expression and 5 more calli carrying the T-DNA of EV.

Antibiotic resistant calli were then transferred to fresh media containing the selection antibiotics for one additional round of propagation. The *GFP* expression was monitored under fluorescence stereoscopy. Strong GFP fluorescence was detected on the calli transformed with the pGWB6 (Fig 5.3). These results confirmed the efficacy of the Agrobacterium-mediated transformation protocol for Dc-CRAG callus, and the suitability of pGWB binary vectors for gene transfer and expression.

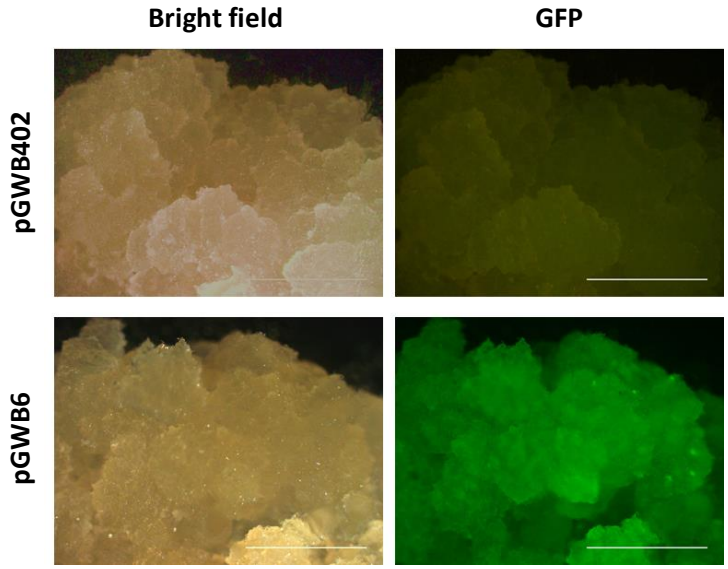


Figure R19. Visualization of GFP fluorescence of *D. carota* transformed calli. Representative images of bright field or fluorescence stereoscopy of *D. carota* calli transformed with the plasmid pGWB402 control (upper panels) or pGWB6 containing the *GFP* gene (lower panels). Scale bars correspond to 2mm.

5.3 Stable transformation of *D. carota* suspensions with oleosin fusion genes

Using the validated transformation protocol, Dc-CRAG suspensions were transformed with the binary vectors carrying the *AtOle2-GFP* and the *GFP-AtOle2* constructs. After incubating with the Agrobacterium suspensions, plant cell cultures were platted in kanamycin selective media (ten plates for each construct). After 45 days any callus growth was observed in the plates. Then, we decided to repeat the assay, that time plating the cell cultures in both kanamycin or hygromycin plates. We decided to use both types of plates to ensure that the lack of growth was not associated with some kind of interaction between the antibiotic resistance and the oleosin fusion produced. Three independent assays were performed, plating ten different plates for each construct and selection antibiotic. In the first two transformation assays, the transformations with the oleosin fusion constructs failed to generate any callus resistant to the selective antibiotics, whereas control transformations with the EV or the vector for *GFP* expression resulted in calli growing on both kanamycin and hygromycin plates. These results suggested that some kind of negative effects might affect callus transformation when using oleosin fusion constructs. In the third assay, we increased the numbers of plates to twenty for each oleosin construct, and some antibiotic resistant calli grew per transformation event. A summary of the number of transformed calli is presented in Table R4.

Table R4. Summary of Dc-CRAG transformed calli.

Antibiotic	Construct	Number of transformed calli	Number of spiked plates	Size of growing calli
Kanamycin	*AtOle2-GFP*	2	20	Small
	GFP-AtOle2	5	20	Small
	GFP	3	10	Large
Hygromycin	*AtOle2-GFP*	3	20	Small
	GFP-AtOle2	5	20	Small
	GFP	2	10	Large

The antibiotic resistant calli transformed with the oleosin fusion constructs were smaller than those transformed with the *GFP* construct. A small piece of the transformed calli was sampled in sterility and visualized under confocal microscopy. We observed GFP fluorescent small spherical vesicles all over the cytoplasm of cells transformed with both oleosin fusion constructs, that were not found in the *GFP* transformed cells. These vesicles were stained with Nile red, that identified them as LDs (Fig R20). These observations show that we obtained transgenic Dc-CRAG calli accumulating the AtOle2-GFP or the GFP-AtOle2 proteins in LDs. Our results indicate that AtOle2 serve as a carrier to LDs in *D. carota* cells.

Antibiotic resistant calli were transferred to fresh media for a second round of selection. Unexpectedly, those calli expressing the oleosin fusion proteins slowed down their growth until imperceptible. After transferring to fresh media for new round of propagation, callus contours were labelled to evaluate their growth. However, calli remained in the marked area without any apparent growth in contrast to GFP transgenic calli that surpassed the delimited area (Fig R21). The calli were weighted before and after the propagation cycle, showing an important increase in biomass for calli expressing *GFP*, while only minor changes in weight were measured for the calli expressing oleosin fusions (Fig R22). The growth increase for transgenic calli expressing *GFP* were statistically significant (Tukey test, $p \leq 0.05$) while insignificant for oleosin fusions.

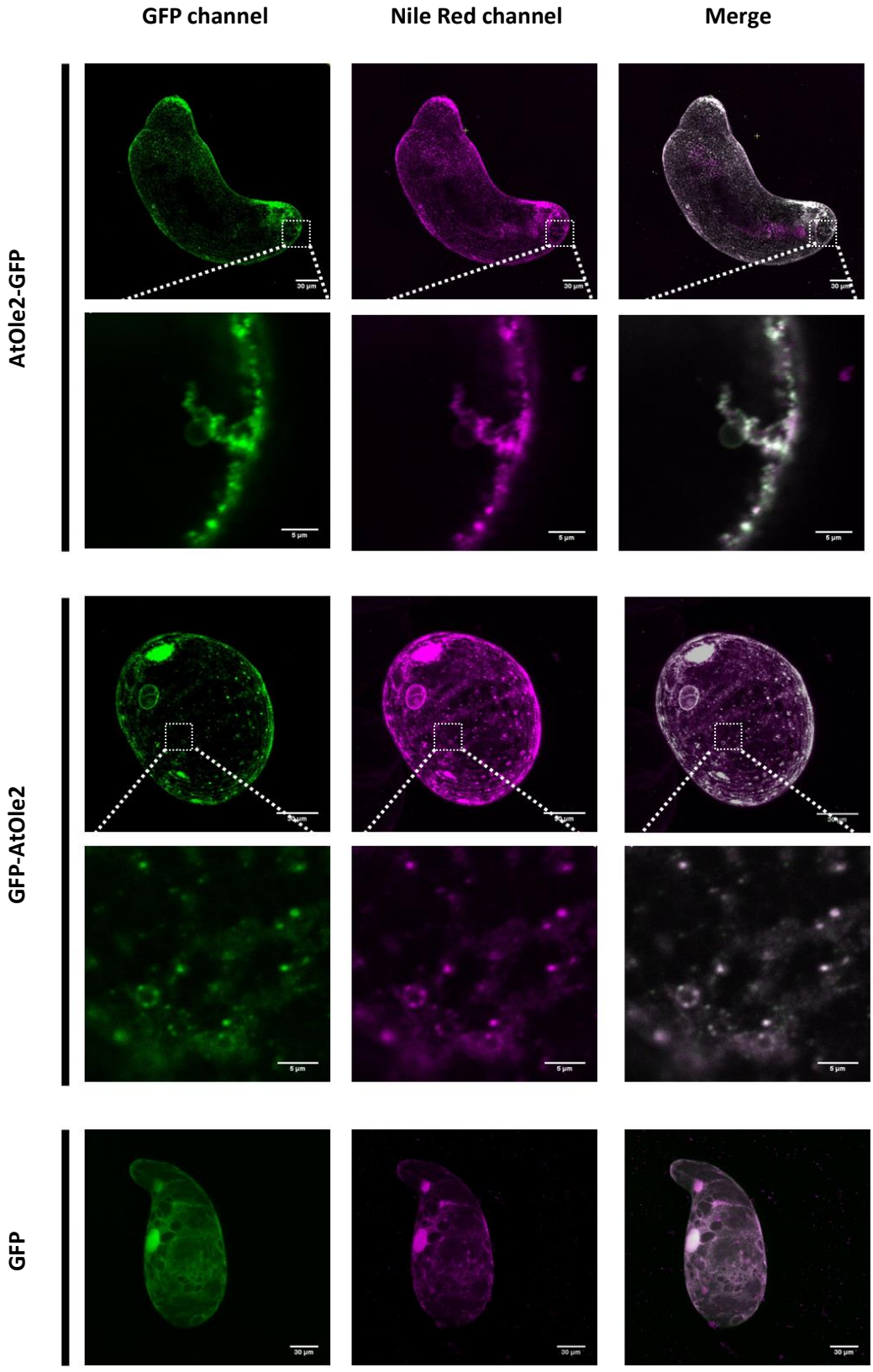
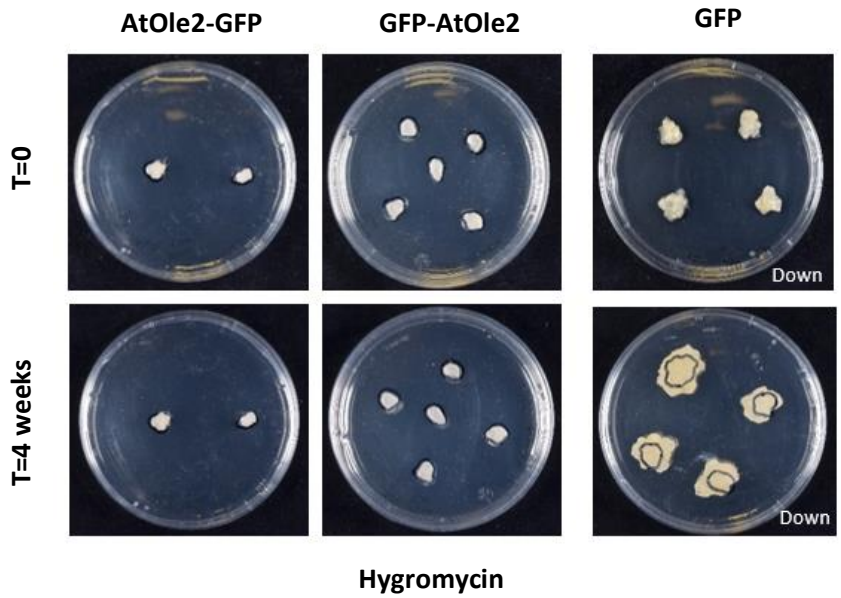


Figure R20. Visualization of *D. carota* cells transformed with the oleosin fusion constructs. Confocal laser microscopy images of Dc-CRAG cells accumulating the indicated AtOle2-GFP, GFP-AtOle2 and GFP proteins. Insets are amplified at the right images. Scale bars correspond to 30 µm and 5 µm as indicated.

(a)



(b)

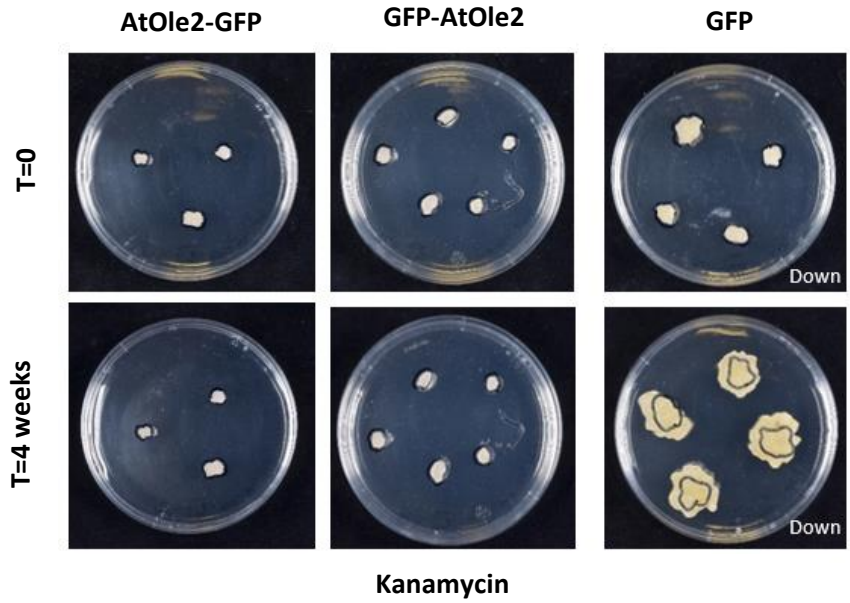


Figure R21. Growth evaluation of *D. carota* calli transformed with the oleosin fusion constructs. Images of transgenic calli accumulating the indicated proteins at the beginning (T=0) or after 4 weeks (T= 4 weeks). Plates are shown from the downside. Callus contour was marked at time 0. Calli were platted in MSDc media supplemented with 50 mg/l kanamycin **(a)** or 50 mg/l hygromycin **(b)** and 100 mg/l timentin in all cases**.**

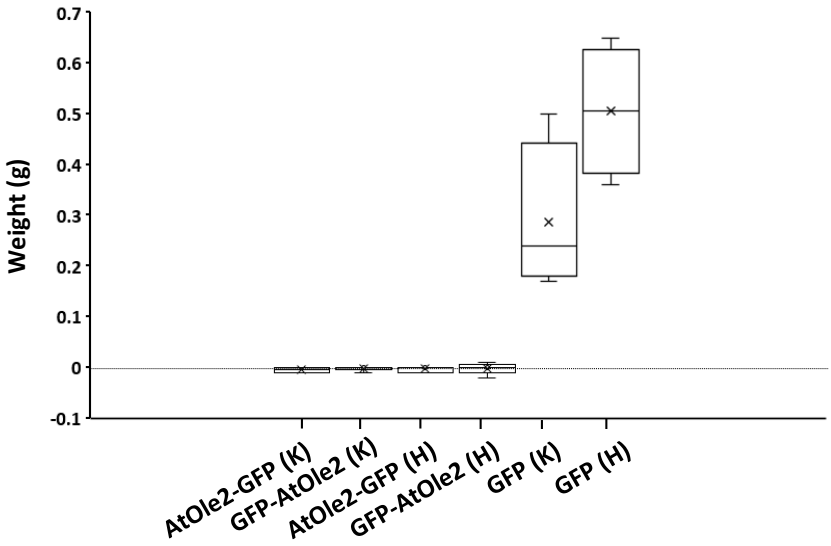


Figure R22. Weight variation in *D. carota* calli transformed with oleosin fusion constructs. Boxplot of the weight variation after four weeks of culture of transgenic calli carrying the oleosin fusion constructs or *GFP* gene. At least 3 calli per construct were plated in media supplemented with 50 mg/l of kanamycin (K) or hygromycin (H).

Altogether, the results show that the *D. carota* cell line generated in this work can be stably transformed via Agrobacterium, and the transformed cells can an accumulate the recombinant proteins. Also, our results indicate that oleosin fusion proteins are accumulated in the Dc-CRAG culture cells and targeted to LDs. However, the accumulation of the AtOle2-GFP or GFP-AtOle2 fusion proteins seems to affect severely the growth of these cells.

6. PRODUCTION OF RECOMBINANT PEPTIDES IN *D. CAROTA* CELL SUSPENSIONS

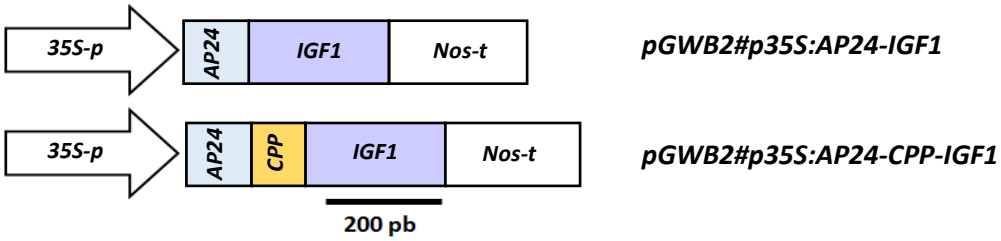
The human growth factor IGF1 has a great industrial interest due to its market potential in cosmetics and pharma sectors. To obtain a stable source of IGF1, we decided to obtain cell lines producing this peptide. For that, we transformed *D. carota* cell lines with gene constructs designed for IGF1 production, we monitored the growth kinetics of the transgenic cell lines and we evaluated the yield and activity of carrot cell produced IGF1.

6.1 Binary vectors preparation

Two plant expression vectors were prepared for the production IGF1 in plant cell cultures (Fig R23). One designed for the production of IGF1 and the other for the production of CPP-IGF1. The peptide CPP-IGF1 is an extension at N-terminal with a cell penetrating peptide (CPP) to facilitate the entrance of IGF1 into mammalian cells. The hPP3 peptide was selected as the CPP to be added due to their

proven capability of protein cargo transporting and penetrating into human fibroblasts (Wang *et al.*, 2016). The constructs carried the codon optimized *IGF1* or *CPP-IGF1* cDNAs extended in N-terminal with a signal peptide encoding sequence for apoplast targeting. The signal peptide corresponds to the one in the tobacco AP24 protein (XP_009782398.1). This signal peptide allows the protein to enter the secretory pathway toward the extracellular space. Given that plant cell cultures would grow in suspension, the proteins would be exported to the culture media and their purification will be facilitated extremely. Gateway™ technology was used for cloning the recombinant genes into the pGWB2 destination plasmid.

(a)



(b) ***IGF1***

ATGTCCAACAACATGGGCAACCTCAGGTCCTCCTTCGTCTTCTTCCTCCTGGCCCTGGTGACCTACACTTATGCAGGG
CCTGAGACTCTCTGTGGCGCCGAATTAGTGGACGCATTGCAATTCGTTTGTGGCGATAGGGGGTTCTATTTCAACAA
GCCTACTGGGTACGGTTCTTCTTCAAGAAGAGCTCCTCAAACTGGCATTGTGGATGAATGTTGTTTCAGATCTTGTGA
TCTCAGGAGACTCGAAATGTACTGTGCTCCCTTAAAACCTGCCAAGTCTGCATGA

IGF1 peptide

MSNNMGNLRSSFVFFLLALVTYTYAGPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLR
RLEMYCAPLKPAKSA*

CPP-IGF1

ATGTCCAACAACATGGGCAACCTCAGGTCCTCCTTCGTCTTCTTCCTCCTGGCCCTGGTGACCTACACTTATGCAATGA
AGATACCGCTCCCCCGTTTTAAATTAAAGTGCATCTTCTGTAAAAAGCGAAGGAAACGAGGGCCTGAGACTCTCTGT
GGCGCCGAATTAGTGGACGCATTGCAATTCGTTTGTGGCGATAGGGGGTTCTATTTCAACAAGCCTACTGGGTACG
GTTCTTCTTCAAGAAGAGCTCCTCAAACTGGCATTGTGGATGAATGTTGTTTCAGATCTTGTGATCTCAGGAGACTCG
AAATGTACTGTGCTCCCTTAAAACCTGCCAAGTCTGCATGA

CPP-IGF1 peptide

MSNNMGNLRSSFVFFLLALVTYTYAMKIPLPRFKLKCIFCKKRRKRAGPAGPETLCGAELVDALQFVCGDRGFYFNKPTGY
GSSSRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPAKSA*

Figure R23.Chimeric genes for the production of IGF1 and CPP-IGF1 peptides in plant cell culture suspensions. **(a)** Schematic diagram of the gene constructs for the expression of *IGF1* and *CPP-IGF1* genes under the control of the *35S* promoter and the *Nos* terminator. **(b)** DNA sequence of recombinant *IGF1* and*CPP-IGF1* genes and the corresponding amino acid sequence are shown, highlighting in blue the signal peptide sequence, in ochre the CPP sequence, in purple the IGF1 and in grey the spacer sequence. Asterisks indicate the stop.

6.2. *D. carota* cell culture transformation

IGF1 and *CPP-IGF1*constructs, in combination with a construct for *P1b* expression or a EV control, were used for the transformation of *D. carota* cell suspensions (Fig R24a). As a control, *P1b* construct was also used for transformation in combination with EV. MSDc plates supplemented with kanamycin and hygromycin were used as selection media. Several calli were obtained for each combination (Fig R24a) and they were propagated for two rounds in solid media. Then, growing calli were sampled and used for DNA extraction. The purified DNAs were used as template for transgene amplification. The primer combination amplifies *IGF1* sequence that is common to all transformations, except for control transformations. All the calli evaluated were PCR positive (Fig R24b). Fig R24b shows representative PCRs of 4 calli per transformation event.

(a)

Transformation	Construct 1	Construct 2	Num. calli
IGF1	*p35S:AP24-IGF1*	*EV*	*7*
CPP-IGF1	*p35S:AP24-CPP-IGF1*	*EV*	*5*
IGF1,P1b	*p35S:AP24-IGF1*	*p35S:P1b*	*6*
CPP-IGF1,P1b	*p35S:AP24-CPP-IGF1*	*p35S:P1b*	*4*
EV,P1b	*p35S:P1b*	*EV*	*4*

(b)

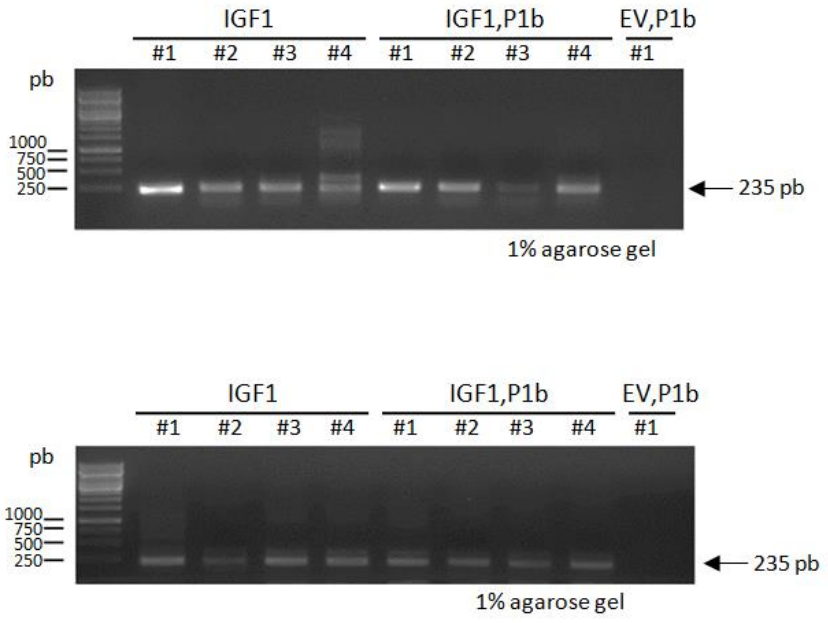
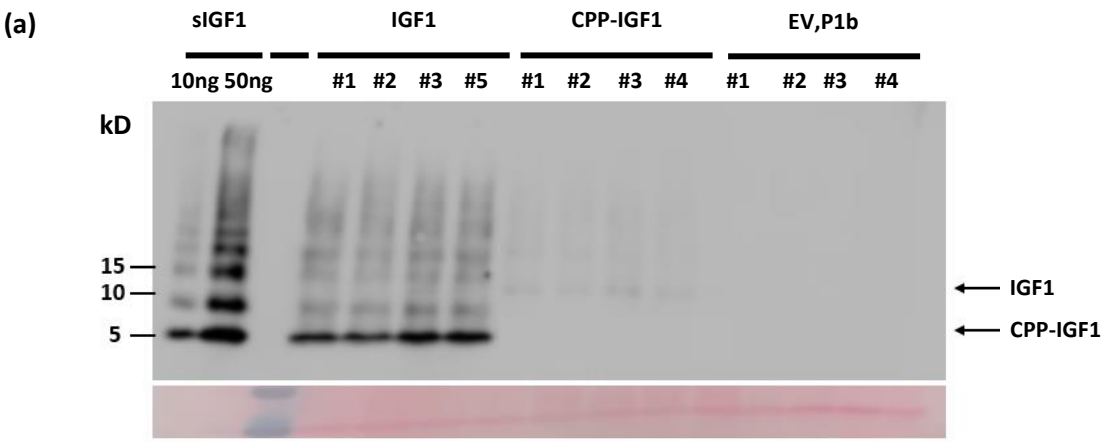


Figure R24. Transformation of *D. carota* cell suspensions with *AP24-IGF1* and *AP24-CPP-IGF1* constructs. (a) Summary of the transformation events. **(b)** PCR analysis on genomic DNA purified from independent calli (#1, #2, #3, and #4) transformed with the indicated constructs. Molecular weight markers are shown on the left.

6.3. Transgenic cell line initiation and peptide production

Once confirmed transgene insertion, calli were propagated for four extra rounds to allow optimal media adaptation. Then, liquid cultures were initiated by inoculating friable callus pieces in 20 ml of media in 50 ml and incubated at 24 °C in a rotary shaker at 110 rpm. These cultures were subcultured every fourteen days for two cycles. The biomass was filtered and standardized to 4g/20ml for each subculture. The depleted culture media of the second round (~20 ml) was collected and lyophilized. Lyophilized products were then resuspended in 1ml protein extraction buffer and evaluated. Samples were analyzed by immunodetection using specific anti-IGF1 antibodies in Western blot. Bands of expected sizes were immunodetected corresponding to IGF1 (theoretical molecular weight 7.7 kDa) and CPP-IGF1 (theoretical molecular weight 10.62 kDa) (Fig R25a). We detected also a ladder of bands corresponding to peptide multimers, also detected with the synthetic peptide, indicating a multimerization tendency for IGF1 peptide. Same results were obtained for calli transformed with *IGF1* constructs in combination with *P1b*. Altogether, these results show that carrot cell lines producing the IGF1 and CPP-IGF1 peptides in the culture media were generated.

Then, we evaluated the peptide accumulation of the different cell lines by ELISA assay using specific anti-IGF1 antibodies. First, we sampled depleted media at the end of the propagation cycle, as well as total culture material. The evaluation of both samples would allow determining if part of the peptide would be retained in the cell and not exported to the media. Also, the samples treatment with different tensoactives (Silwet-L77 and Tween20) was evaluated to study whether the peptide was bond to cells. The results are summarized in Fig R25b and show that the independent cell lines accumulate different peptide levels. The differences in peptide levels will be commented in later sections. Also the results showed that there were minor differences between peptide yields in the media and cell cultures, in the presence or absence of tensoactives, suggesting that practically all the recombinant peptide produced were being exported to the media and they were not bond to cells.



(b)

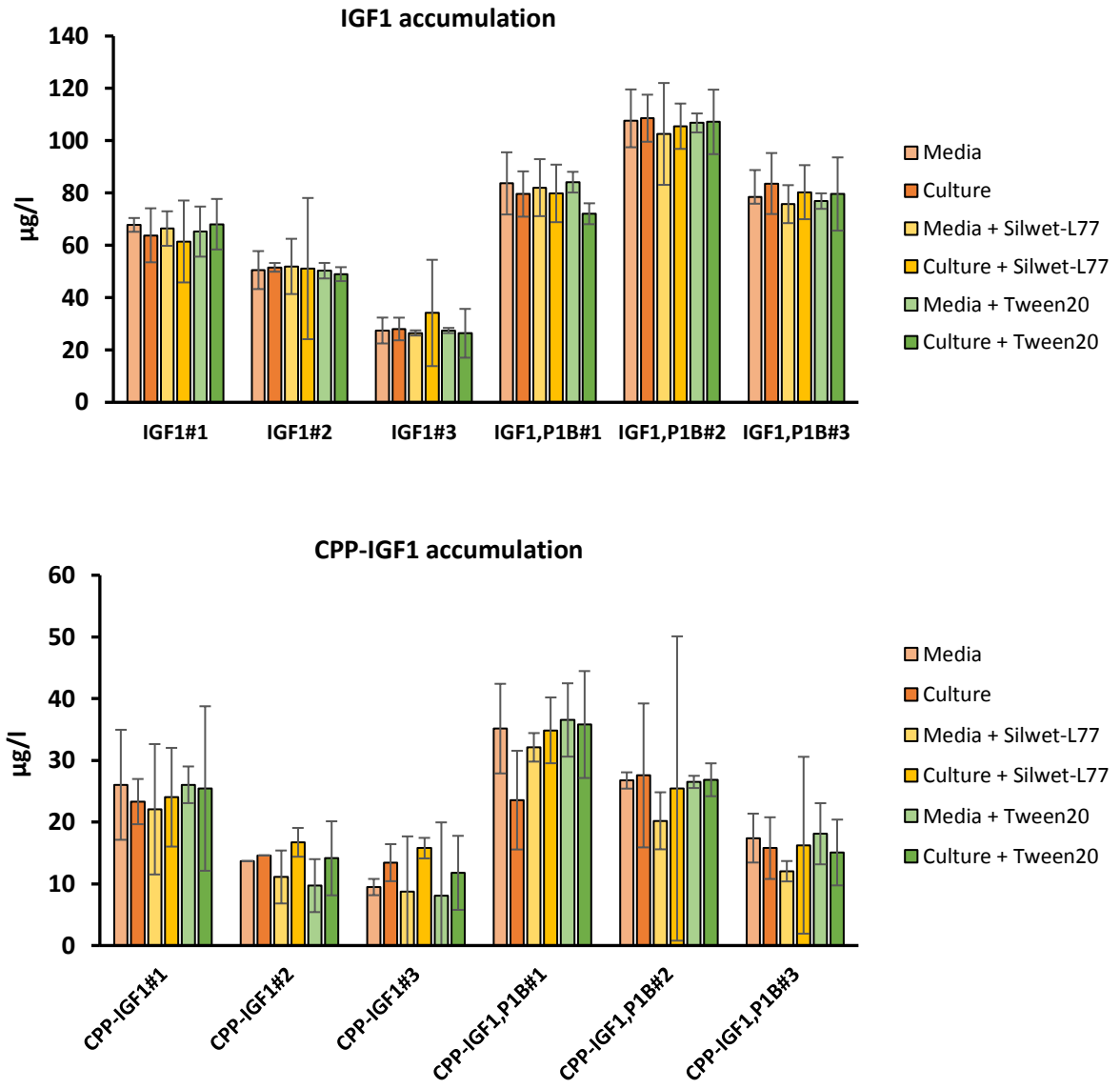


Figure R25. Accumulation of IGF1 and CPP-IGF1 peptides in culture media of the transformed Dc cell suspensions. **(a)** Western blot analysis of lyophilized culture media from transgenic callus lines carrying the indicated transgene, using anti-IGF1 antibodies. As a control, the synthetic IGF1 peptide at two different amounts (10 and 50 ng) was run in parallel. Lower panels correspond to the Ponceau stained transferred membranes. Molecular weight markers are indicated on the left in kDa. **(b)** ELISA analysis of lyophilized culture media (media) or suspension (culture) from transgenic callus lines. Media and cultures were treated with the tensoactives Silwet-L77 and Tween20, as indicated. Anti-IGF1 antibodies were used as primary antibody and Goat Anti-Rabbit IgG H&L conjugated with alkaline phosphatase as secondary antibody.

6.4. Evaluation of growth kinetics in transgenic cell cultures

Next, we evaluated the growing kinetics of the developed *D. carota* cell lines. Three independent transgenic suspensions, and three replicates per suspension, were generated for each transformation

event (IGF1; CPP-IGF1; IGF1,P1b; CPP-IGF1, P1b;P1b). Cell cultures were sampled and evaluated at different time points, including 0, 3, 7, 9, 11 and 15 days after subculture.

6.4.1 Cell suspensions expressing *IGF1* or *CPP-IGF1*

First, we evaluated the plant cell lines transformed with *IGF1* or *CPP-IGF1* constructs in combination with EV and compared to the wild-type (WT) cell line. Several parameters were scored to have a complete kinetics of growth, including the reduction of brix degree (ºB), the pH, the fresh weight (FW) and the cell packing volume (VEC). Results are shown in Fig R26. The ºB reduction in transgenic suspensions expressing *IGF1* showed a less pronounced slope than WT cell line (Fig R26a). WT suspensions had a sharp reduction in nutrients already at 3 days, whereas IGF1 lines showed a constant nutrient consumption until 7 days. The nutrient depletion appeared at 10 days in WT and at 15 dpi in IGF1 cell lines. The CPP-IGF1 suspensions showed a constant reduction of brix degrees along all the culture period, with the exception of those derived from CPP-IGF1#3 that showed shorter latency period and at 7 days lower nutrient levels (Fig R26a). When compared the six transgenic cell lines we could observe that the difference between independent replicates did not allow establishing clear differences between transgenic types. However, the graphs suggested that CPP-IGF1 cell lines would consume nutrients slower than IGF1 and P1b. The evolution of pH showed minor differences between most of cell lines. The pH of IGF1 cultures were approximately 0.1 units lower than the WT, while 0.05 for CPP-IGF1 (Fig R26b). The CPP-IGF1#2 showed larger differences with the rest of the cultures, having a different profile and reads up to 0.2 points lower than the WT. VEC analysis showed that the transgenic cell lines reached lower biomass levels than the WT. The maximum biomass levels were: 57 ± 0.7 %, 53 ± 4% and 66 ± 3.5 %for IGF1, CPP-IGF1 and WT respectively (Fig R26c). The graphs also showed that IGF1 cell lines had higher biomass than CPP-IGF1 along the culture. FW results corroborated VEC values. The maximum FW levels for IGF1 were 531 ± 3, while 518 ± 8 and 577 ± 3 for CPP-IGF1 and WT respectively (Fig R26d).

In the case of the plant cell suspensions transformed with EV in combination with P1b, we could observe that the transgenic cell cultures showed a slightly higher nutrient level than the WT cultures, but similar behavior (Fig R27a). The pH levels in the EV,P1b cell cultures were reduced in ~0.15 units (Fig R27b). The suspensions exhibited almost identical VEC and FW than the WT at the end of the propagation cycle (Fig R27c-d).

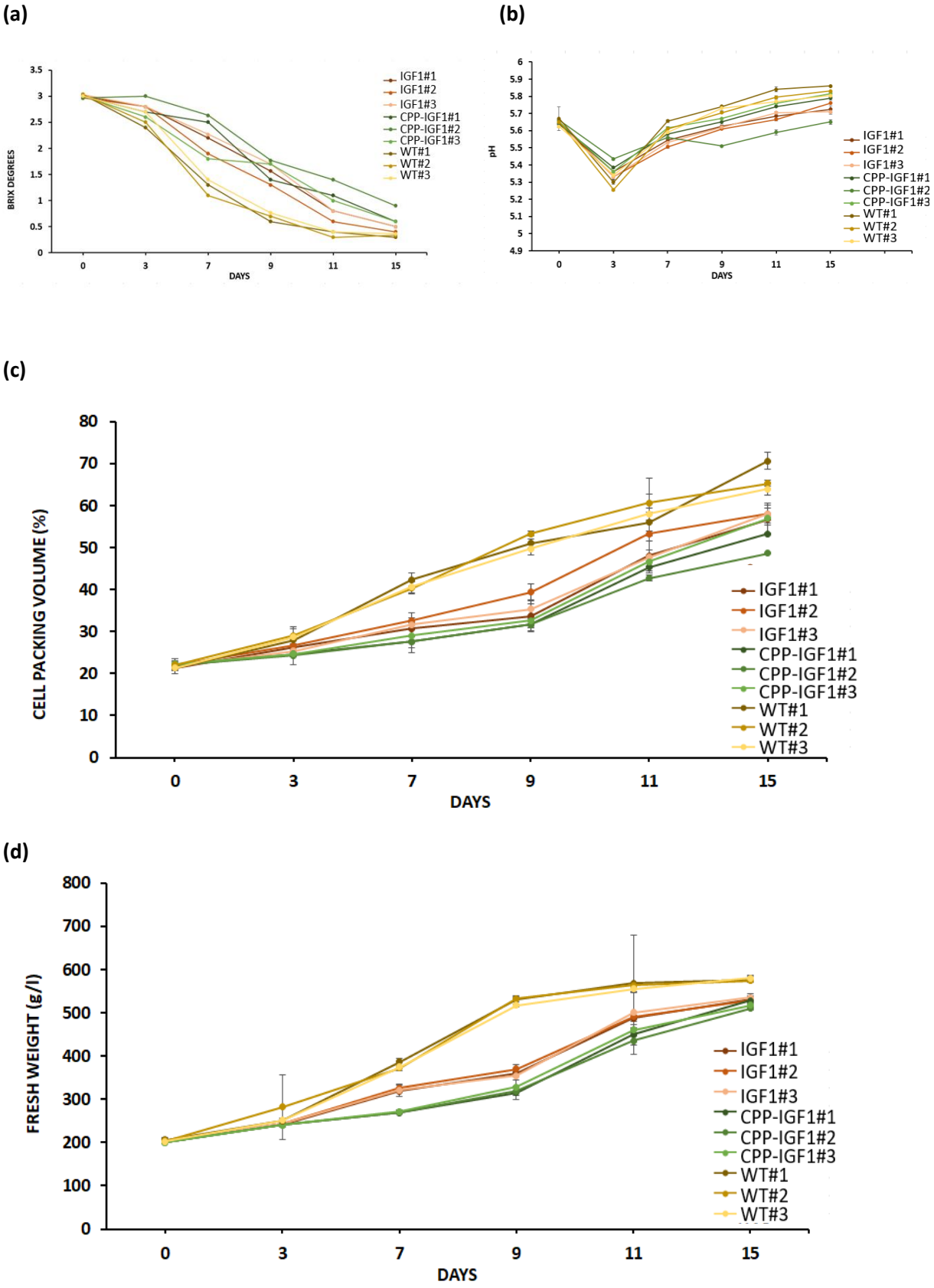


Figure R26. Growth curves of IGF1 and CPP-IGF1 cell suspensions. Graphs represent the mean values ± SD of brix degrees **(a)**, pH **(b),** cell packing volume expressed (VEC) in % **(c)** and fresh weight (FW) expressed in g/l, **(d)** of IGF1, CPP-IGF1 and WT cell suspensions measured at indicated time points. Values are the mean of three biological and technical replicates (n=6).

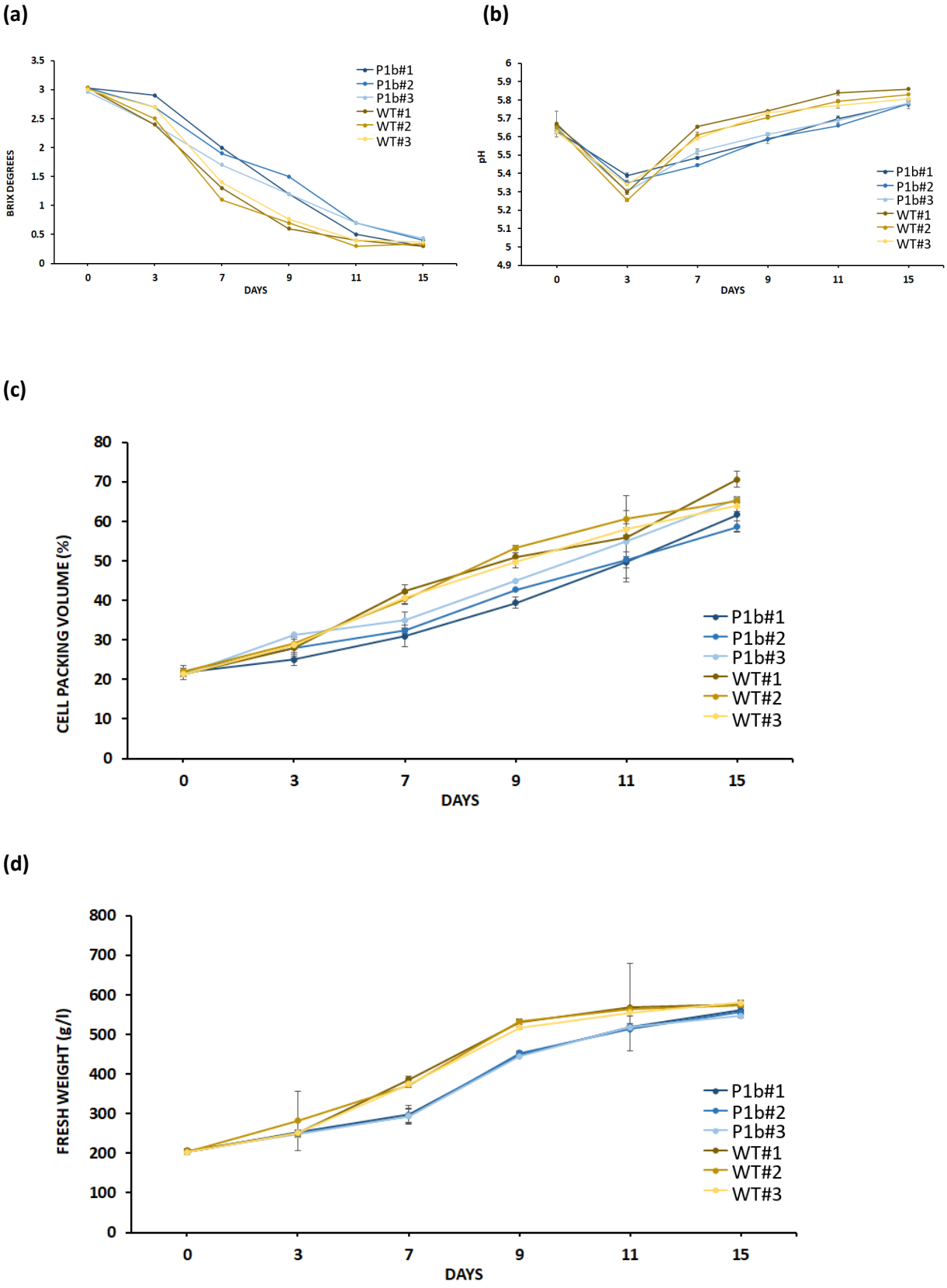


Figure R27. Growth curves of EV,P1b cell suspensions. Legend as in Fig R26, but expressing EV,P1b and WT cell suspension results.

Later, we evaluated the growth kinetics of the plant cell lines transformed with *IGF1* or *CPP-IGF1* constructs in combination with *P1b*. The brix degree reduction in IGF1,P1b cells showed a less pronounced slope than WT during all the culture. Nutrient levels in IGF1,P1b were ~0.2 to 0.7 ºB higher than the WT depending on the time analyzed. The three independent cell lines had a similar profile, with the exception of callus IGF1,P1b#3 that has increased nutritient consumption. Culture media was not depleted in IGF1,P1b#1 and IGF1,P1b#2 cultures (Fig R28a). The transgenic lines expressing *CPP-IGF1* and *P1b*, showed a constant nutrient consumption until 7 dps. The behavior of the three cell lines was similar, differing to the wt lectures in a range between ~ 0.3 to 1.1ºB. The cultures haven't depleted the media nutrients at 15 dps. The pH showed minor differences between the transgenic cell lines, and always, was lower than WT (Fig R28b). The VEC was similar for IGF1,P1b and CPP-IGF1,P1b cell lines (Fig R28c), being less acute than WT. The VEC kinetics of the cultures carrying the same transgene were practically identical. FW profiles showed the same patterns than VEC profiles (Fig R28d). The transgenic lines carrying the same transgene had similar FW increase being in all cases lower than the WT FW.

Then, we decided to compare the transgenic lines co-transformed with P1b in comparison to EV. In the case of the nutrient consumption, we could observe that the cell lines expressing peptides and the silencing suppressor did not depleted the media nutrients at the end of the propagation (Fig R29a). The pH profiles followed a similar behavior in all the suspensions, however the IGF1,P1b suspensions showed the lowest pH levels along the assay (Fig R29b). The differences between transgenic cell lines were minor. VEC and FW determinations showed similar results for all the transgenic cell cultures. In all the cases the maximum biomass accumulation was lower for the transgenic suspensions than for the WT (Fig R29c-d).

When the suspensions expressing *CPP-IGF1* or *CPP-IGF1* and *P1b* were compared, only minor differences for the different growth parameters evaluated were detectable (Fig R30). Only in the VEC evaluation we could observe some differences between the transgenic cultures, showing the CPP-IGF1,P1b cell suspensions lower VEC levels along the assay (Fig R30c).

Altoghether the results showed that althought the transgenic callus reached biomass levels lower than the WT at 15 dps, they can be propagated in liquid cultures and conserved similar kinetics compared to the WT. The results showed minor differences between biological replicates, indicating a high reproducibility between transgenic cultures. Also we could observe that the suspensions expressing *P1b*, showed minor affectations into the growth dinamics when compared with the WT. The suspensions expressing the peptides alone or in combination with *P1b*, reported similar results.

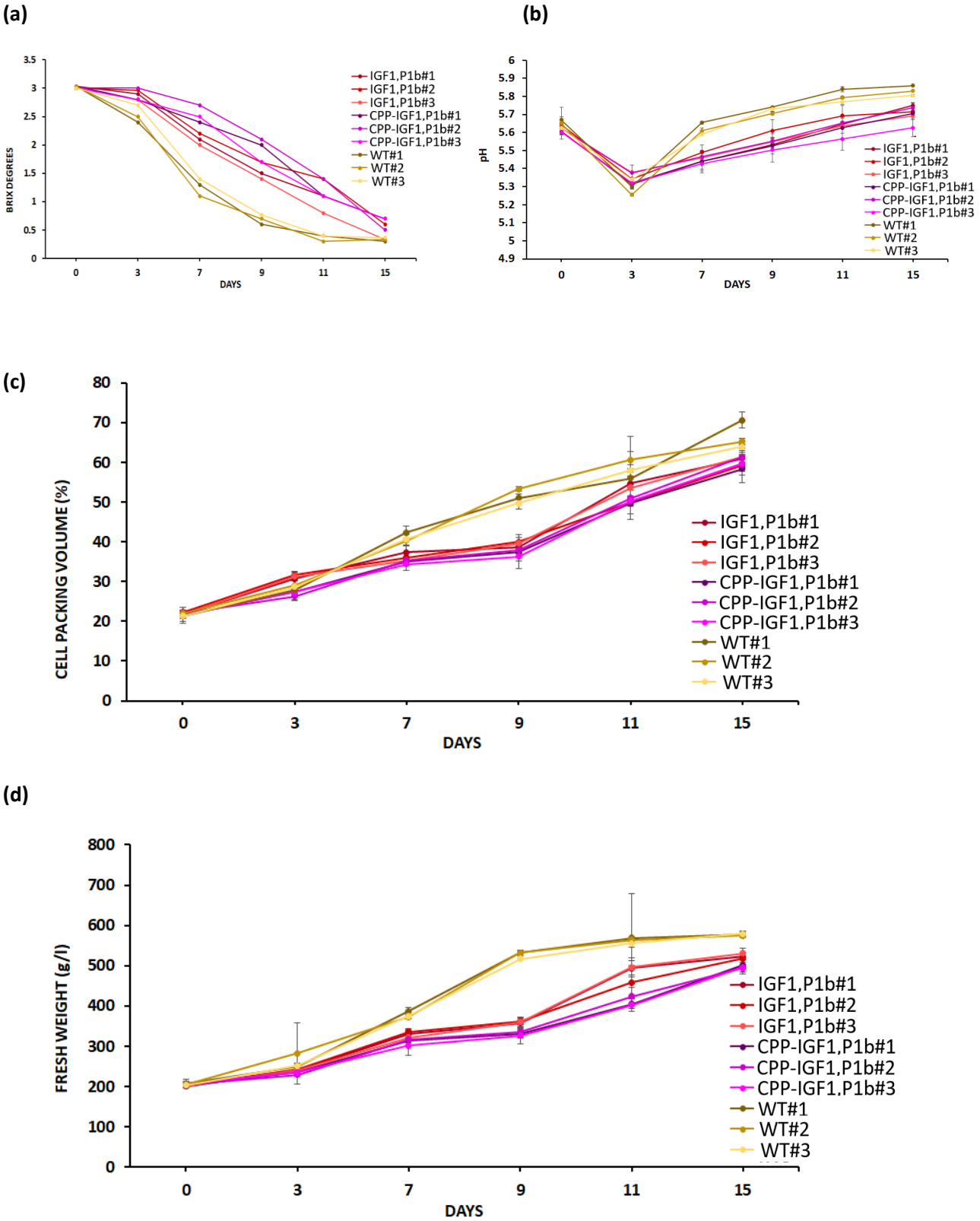


Figure R28. Growth curves of IGF1,P1b and CPP-IGF1,P1b cell suspensions. Legend as in Fig R26, but expressing IGF1,P1b; CPP-IGF1,P1b and WT cell suspension results.

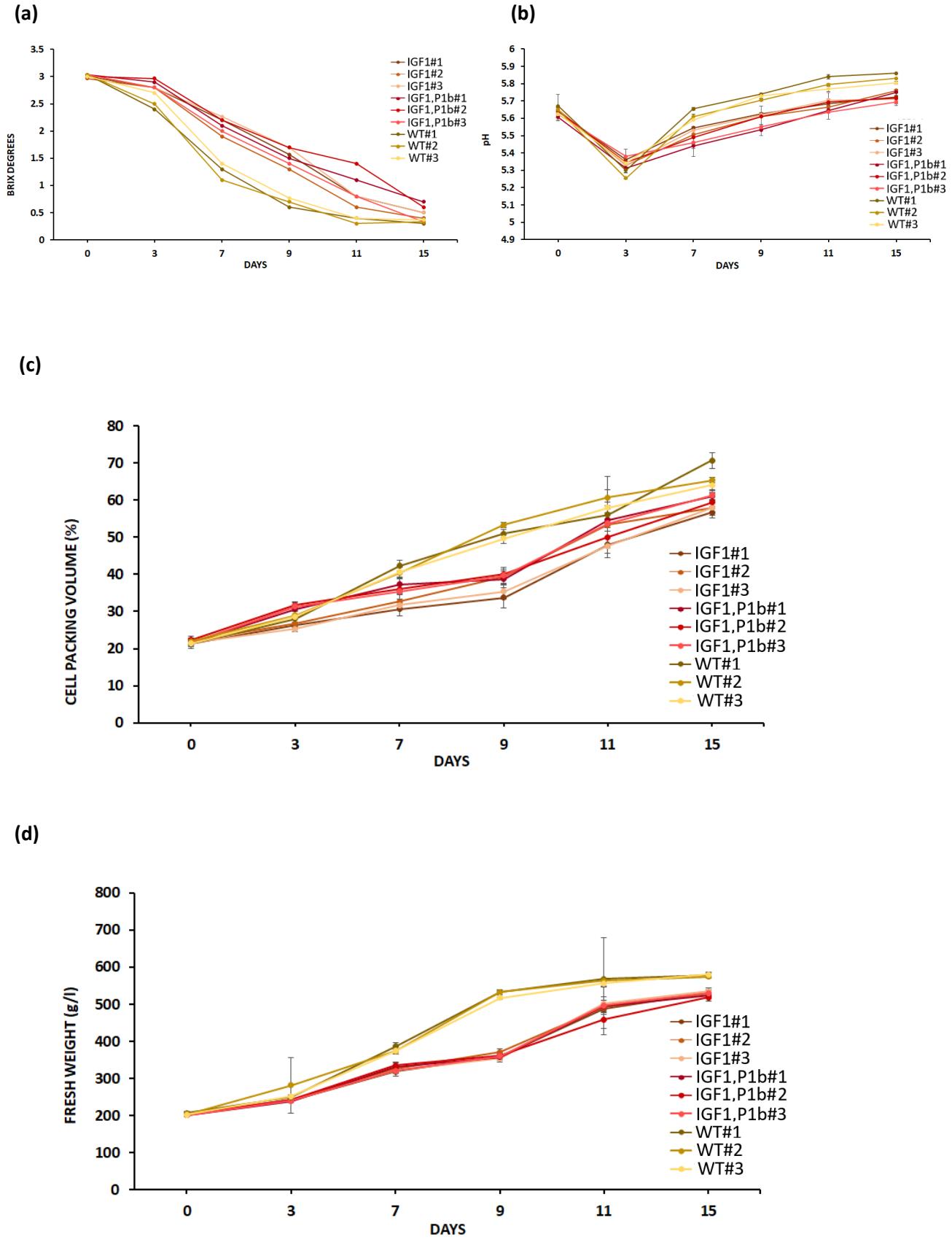
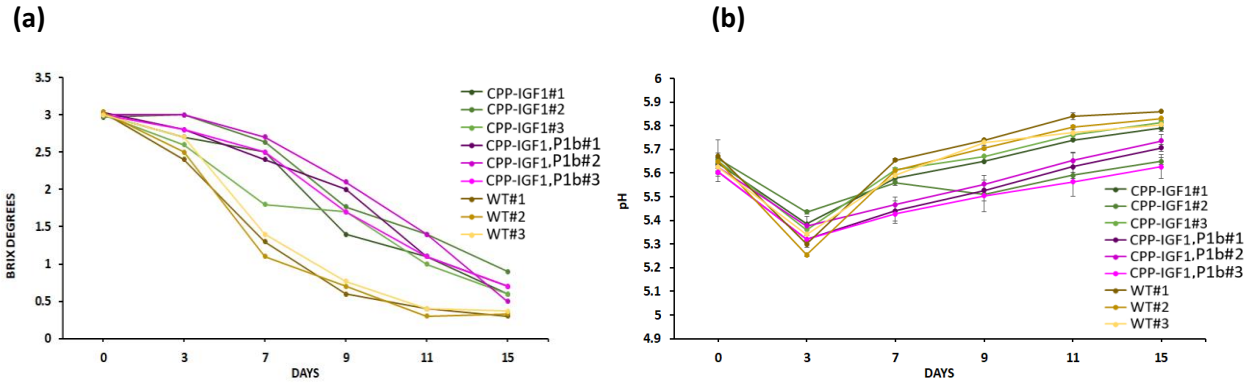
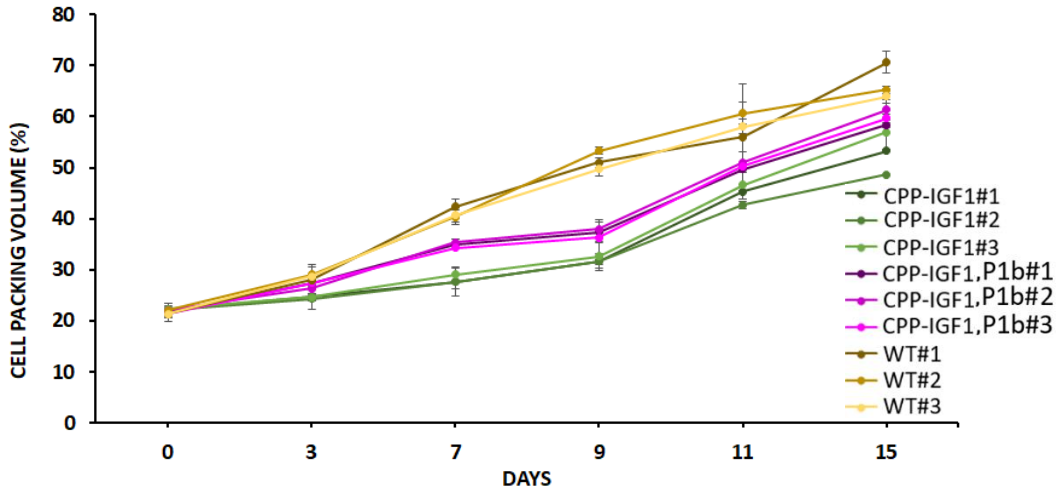


Figure R29. Growth curves of IGF1 and IGF1,P1b cell suspensions. Legend as in Fig R26, but expressing IGF1; IGF1,P1b and WT cell suspension results.





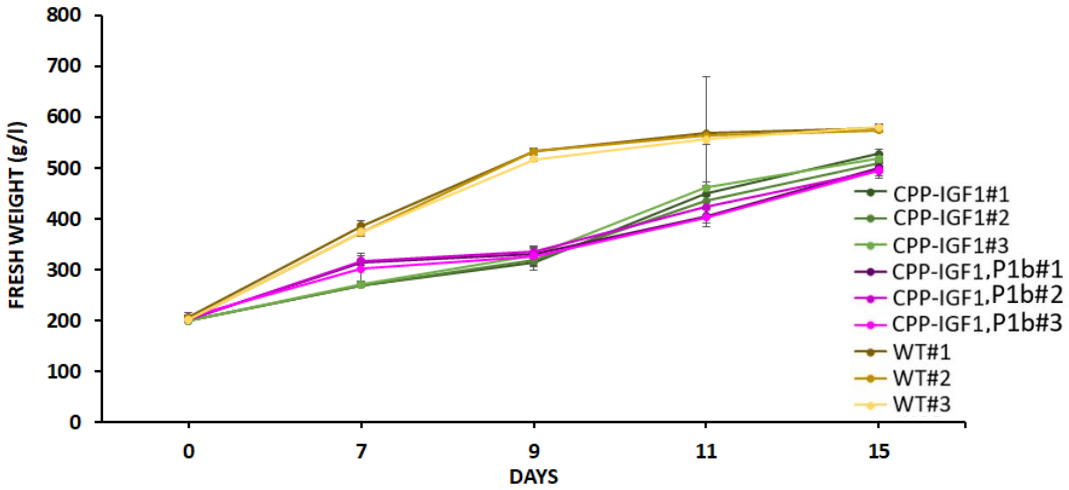


Figure R30. Growth curves of CPP-IGF1 and CPP-IGF1,P1b cell suspensions. Legend as in Fig R26, but expressing CPP-IGF1; CPP-IGF1,P1b and WT cell suspension results.

6.5 Peptides accumulation dynamics

Once determined growth kinetics, we evaluated the stability on the production of IGF1 and CPP-GF1 peptides for the different transgenic cell lines. Peptide accumulation was monitored by ELISA assays using specific anti-IGF1 antibodies. These lines were then subjected to time course experiments during 10 propagation cycles (5 months). The culture media of 15 days old plant cell suspension was analyzed by ELISA and the accumulation of the IGF1 and CPP-IGF1 was monitored in Fig R31 and R32. Along the 5 months, IGF1 cultures accumulated around 75.17 ± 11.19 µg/l, 32.04 ± 18.2 µg/l and 57.63 ± 8.09 µg/l of the peptide with statistically significant differences on yield among the three independent lines (Tukey test, $p \leq 0.05$) (Fig R31a). When comparing the peptide levels at different time samples for the same cell line no significant differences were observed (Tukey test, $p \leq 0.05$) (Fig R31b). In the case of the CPP-IGF1 cultures peptide yields were 15.22 ± 9.63 and 21.32 ± 6.38 µg/l (Fig R31a). The differences were statistically non-significant between them and between different propagation. In the case of CPP-IGF1#2, the CPP-IGF1 levels were under the detection threshold in most of the analyzed samples. The results showed that the IGF1 was accumulated in higher levels than CPP-IGF1 in all the samples, and during the entire time course. IGF1 accumulation was between 1.5- and 5-fold times higher than the accumulation of the CPP-IGF1.

In the case of the transgenic lines co-expressing *IGF1* or *CPP-IGF1* in combination with *P1b* silencing suppressor, the culture media contained amounts of 114.97 ± 14.28 µg/l, 134.98 ± 13.06 µg/l and 97.91 ± 16.43 µg/l in the three independent lines; whereas the amount of CPP-IGF1 were 29.53 ± 13.06 µg/l, 31.15 ± 10.32 µg/l and 16.42 ± 8.59 µg/l. The differences between production yields in the different cell lines were statistically significant (Tukey test, $p \leq 0.05$) (Fig R32a), whereas non-significant differences were detected among the yields of the same sample at different propagation cycles (Tukey test, $p \leq 0.05$) (Fig R32b).

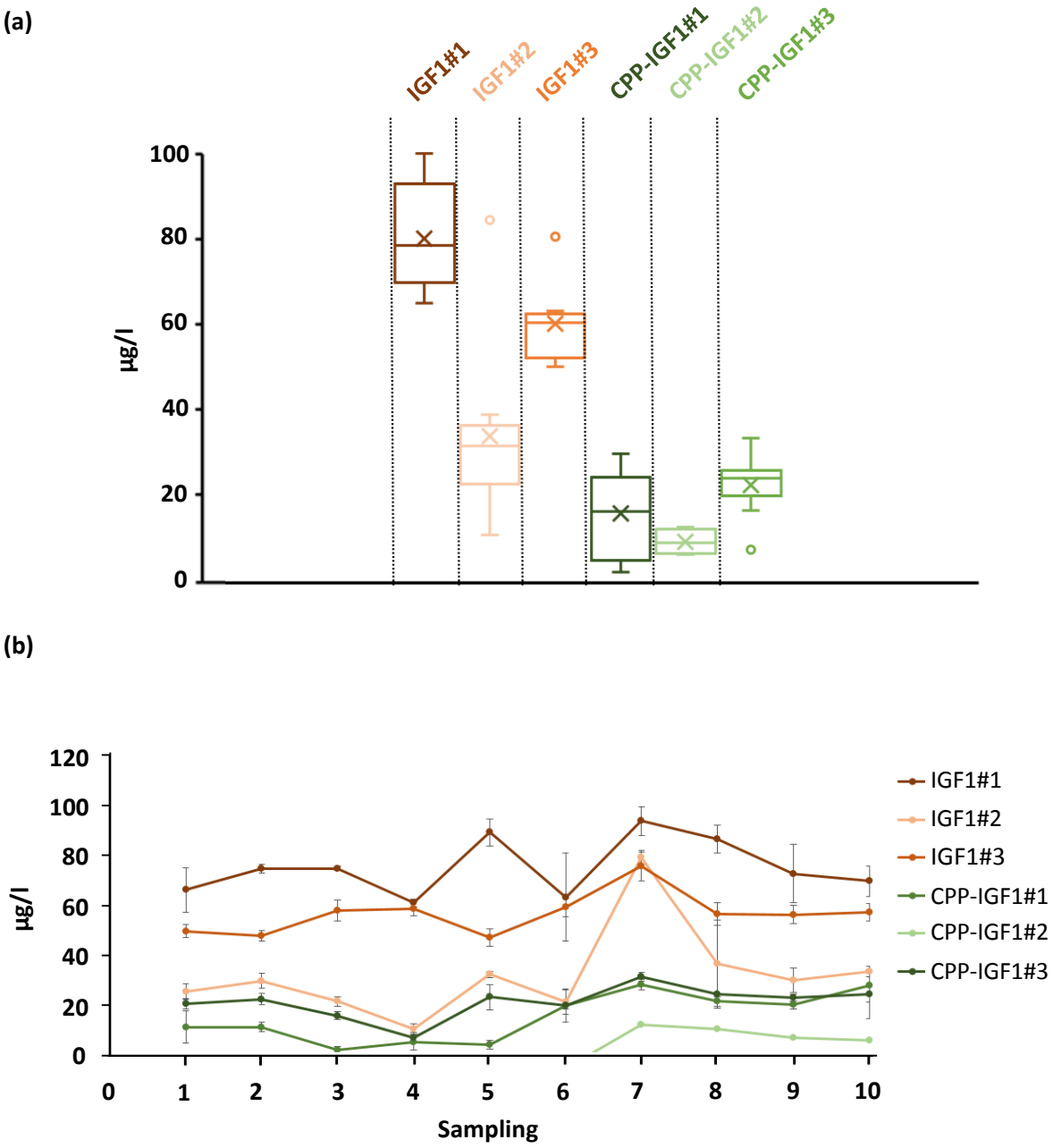


Figure R31. Accumulation of IGF1 and CPP-IGF1 in transgenic cell lines. **(a)** Boxplot of peptide accumulation along 5 months monitored every 15 days (n=10) in culture media from three independent cell lines transformed with IGF1 or CPP-IGF1 plasmids in combination with EV. Line represent mean, crosses represent the median and spots outliners. **(b)** Graph representing the IGF1 or CPP-IGF1 accumulation dynamics during 5 months of growth recorded every 15 days. Graph represents the mean value ± SD of 3 replicates.

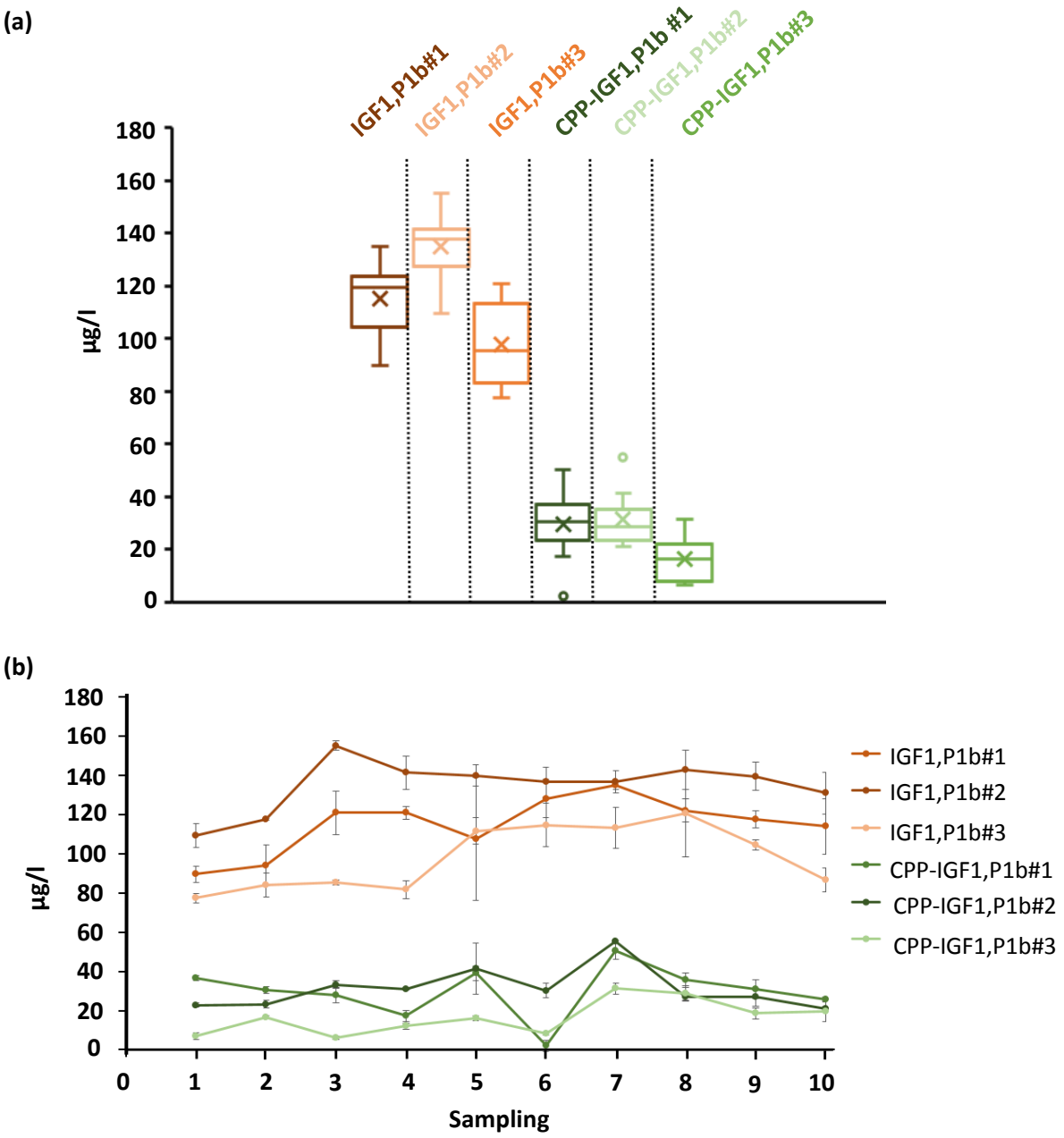


Figure R32. Accumulation of IGF1 and CPP-IGF1 in transgenic cell lines co-expressing *P1b* silencing suppressor. **(a)** Boxplot represents averages of peptide accumulation in culture media of indicated transgenic cell lines along 5 months sampling every 15 days. Line represent mean, crosses represent the median and spots outliners. **(b)** Graph represents mean values ± SD of 3 replicates of peptide accumulation in culture media of indicated transgenic lines.

Then, we compared the peptide accumulation between plant cell lines co-expressing *IGF1* or *CPP-IGF*1 in combination with the *P1b* or the EV. As shown in Fig R33, the cell lines co-transformed with the RSS accumulated higher levels than the control ones carrying the empty vector. In the case of IGF1,P1b cultures, peptide accumulation was between 1.3 to 4.2-fold times higher than the reached by the IGF1 cell lines. The peptide yields for CPP-IGF1,P1b cultures was between 0.77 to 2.4-fold times higher than CPP-IGF1 ones.

(a)

	IGF1#1	IGF1#2	IGF1#3
IGF1,P1b#1	1.529481	3.587656	2.030214
IGF1,P1b#2	1.795771	**4.212285**	2.383684
IGF1,P1b#3	**1.302591**	3.055448	1.729044

(b)

	CPP-IGF1#1	CPP-IGF1#2	CPP-IGF1#3
CPP-IGF1,P1b#1	1.939293	/	1.385164
CPP-IGF1,P1b#2	**2.045361**	/	1.460924
CPP-IGF1,P1b#3	1.078208	/	**0.770123**

Figure R33.Comparison of peptide accumulation averages between cell lines. Tables representing the average comparison between cell lines expressing *P1b* + *IGF1* with *IGF1* + EV **(a)**, and *P1b* + *CPP-IGF1* with *CPP-IGF1*+ EV **(b)**. Maximum and minimum differences are market in bold.

Altogether, the results showed that transgenic cultures accumulating IGF1 and CPP-IGF1 could propagate in liquid media and reach biomass levels that allow their maintenance and correct culture. Also, it was noticed that the cultures co-transformed with *P1b* in combination with *IGF1* or *CPP-IGF1*, reached higher peptide accumulation yields than the cell lines transformed with the EV and *IGF1* or *CPP-IGF1*, with minimum impact on growth when compared with the other transgenic cultures. Average peptide accumulation of the different cell lines is summarized in Fig R34.

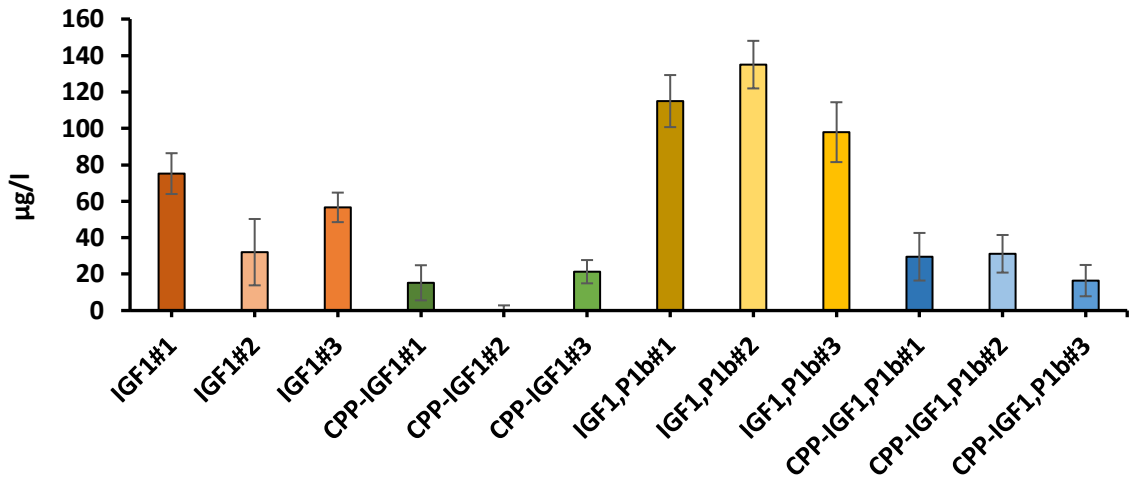


Figure R34. Accumulation of IGF1 and CPP-IGF1 in the different transgenic cell cultures. Graph represents the peptide accumulation average of the different transgenic cell lines expressing *IGF1* or *CPP-IGF1*. The error bars show the ± SD between 3 replicates.

6.6 Cell line cryopreservation and regeneration

A consistent performance of cell lines producing a compound of interest is crucial for an industrial application. The most common method to guarantee the stability of plant cell lines is the cryopreservation. Using a cryopreservation protocol established by Vytrus Biotech and described in Materials and Methods, we proceeded to preserve the transgenic cell lineIGF1,P1b#2. This cell line was selected because it was the line accumulating the highest levels of IGF1. The cryopreservation process was carried out after the first kinetics of growth. The objective of the assay was to evaluate the regeneration capacity of the transgenic callus after the cryopreservation process, as well as to compare the growing and peptide accumulation dynamics between cryopreserved cells and those maintained growing in liquid culture or maintained as callus.

The cell line was cryopreserved during 4 months after which we proceeded with the regeneration of the cell culture. The cryopreserved material was platted on sterile filter paper placed on culture media. The filter paper intends to retain the cryopreservation substances that can potentially affect callus growth (Mustafa *et al.*, 2011). The regenerated callus was then grown in sterility and darkness at 24ºC and propagated to new plates every 4 weeks. After 2 propagation cycles, we initiated cell suspension cultures, as previously done, by inoculating callus in 20 ml of the MSDc liquid media in 50 ml flasks and incubating with agitation in a rotary shaker at 24 °C, 110 rpm. In parallel, we initiated cell suspension cultures from the original callus of the IGF1,P1b#2 cell suspension that was maintained all the time as callus. Cell suspensions were subcultured every fourteen days, and a total of4 cycles were done before assessing growth and peptide accumulation dynamics. The Table R5 and Fig R35 show a summary of the different cell suspensions to be monitored.

Table R5. Summary of the cell lines used to assess stability over the time.

Cell line	Time in liquid culture (months)	Codification
IGF1,P1b#2	10.5	IPLiq
IGF1,P1b#2	2	IPCall
IGF1,P1b#2	2	IPCryo
EV,P1b	10.5	PLiq
EV,P1b	2	PCall
WT	10.5	WLiq
WT	2	WCall

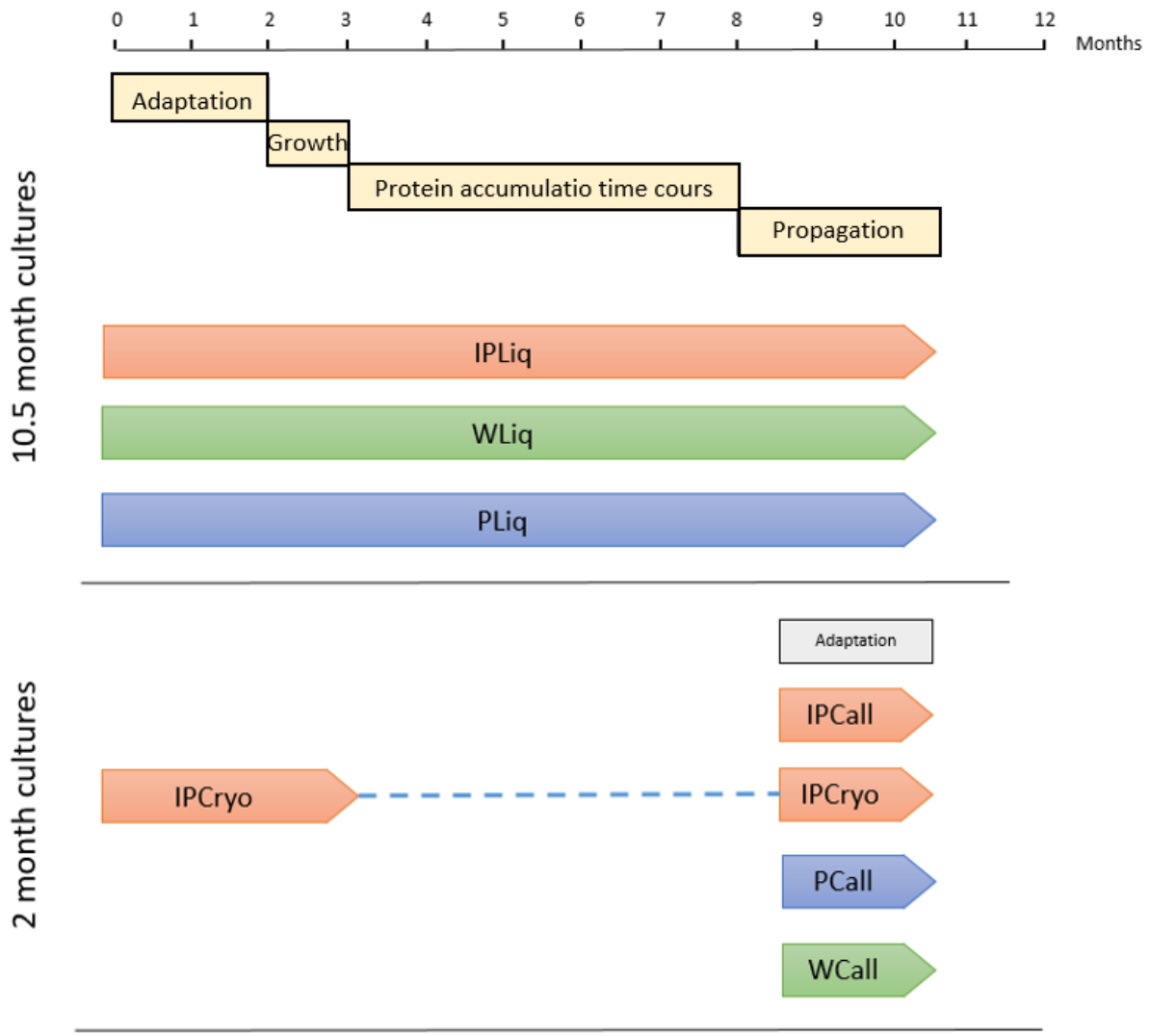


Figure R35. Experimental design to assess stability of the plant cell cultures. The graph illustrates the growth timeline of the different cell lines according to the codification in Table 6. Blue dashed line represents the cryopreserved period.

Cell suspension lines were grown at large scale using 0.5L flasks, 200 ml liquid media and 40 g of cell material. Several parameters were scored to have a complete kinetics of growth, including the reduction of brix degree, the pH, the fresh weight (FW) and the cell packing volume (VEC). The kinetic parameters are summarized in Fig R36.

We can observe that nutrient consumption dynamics followed a similar behavior for the suspensions established from the same cell lines, regardless when they had been initiated (Fig R36). The curves of the IGF1P1b#2 suspensions (IPCryo, IPCall, IPliq) were practically identical, as the P1b (PCall and PLiq) suspensions or WT (WCall, WLiq). We also can observe that the curve slopes of IGF1 cells were slightly less pronounced than the ones from P1b or WT suspensions. The pH showed identical curves for all cultures, except for the WT being slightly higher than the others (Fig R36b). The biomass measures as

VEC and FW showed a clear reduction for IGF1P1b#2 lines in comparison to WT and P1b (Fig R36c-d). These data showed that cryopreservation can guarantee cell line stability over time, as well as callus maintenance.

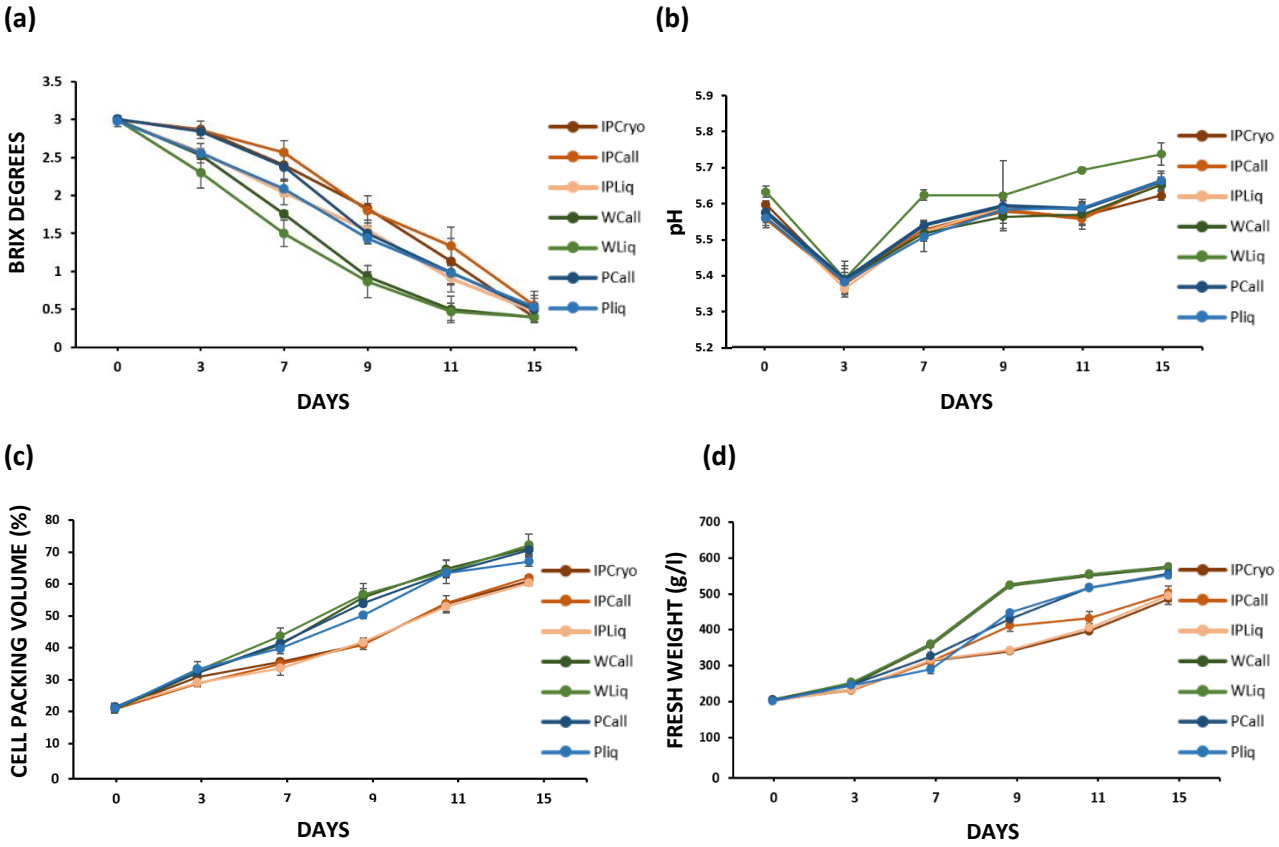


Figure R36. Growth curves of *D. carota* cell suspensions. Graphs represent the mean values ± SD of measures of brix degrees **(a)**, pH **(b),** cell packing volume expressed (VEC)**(c)** and fresh weight (FW) **(d)** at indicated times from indicated cell lines according to the codification in table 6.1. Values are the mean of three biological and technical replicates (n=9).

We also monitored the peptide production in the plant cell suspensions over time and after different conservation methods. For that, cultures were propagated in liquid media for 5 additional months (10 propagation cycles). Then, we determined the peptide accumulation in the culture by ELISA analysis. Fig R37 shows the amount of peptide in the extracellular media from the different cell lines over the time course. We can observe that the cryopreserved cultures always showed the highest peptide accumulation, being ~15 to 53% higher than values of cultures maintained in suspension, or ~21 to 50% higher than values of cultures maintained as calli. Peptide yields at the end of the time course were 136.39 ± 36.08 µg/l, 106.15 ± 18.28 µg/l and 88.72 ± 15.76 µg/l for IPCryo, IPCall and IPLiq respectively. These results suggested that the propagation of the transgenic cell line as suspension or as callus triggers a progressive reduction on peptide yields. In the case of the culture maintained in

suspension (IPLiq), the recombinant peptide yield was reduced at ~33% in 7 months (from 130.95 ± 10.96 µg/lat the end of the first propagation cycle to 88.72 ± 15.76 µg/l at the end of the second time course it was).

Altogether our data demonstrate that the maintenance of the transgenic plant cell lines is feasible using cryopreservation and it is the best strategy to guarantee the stable production of the recombinant peptide.

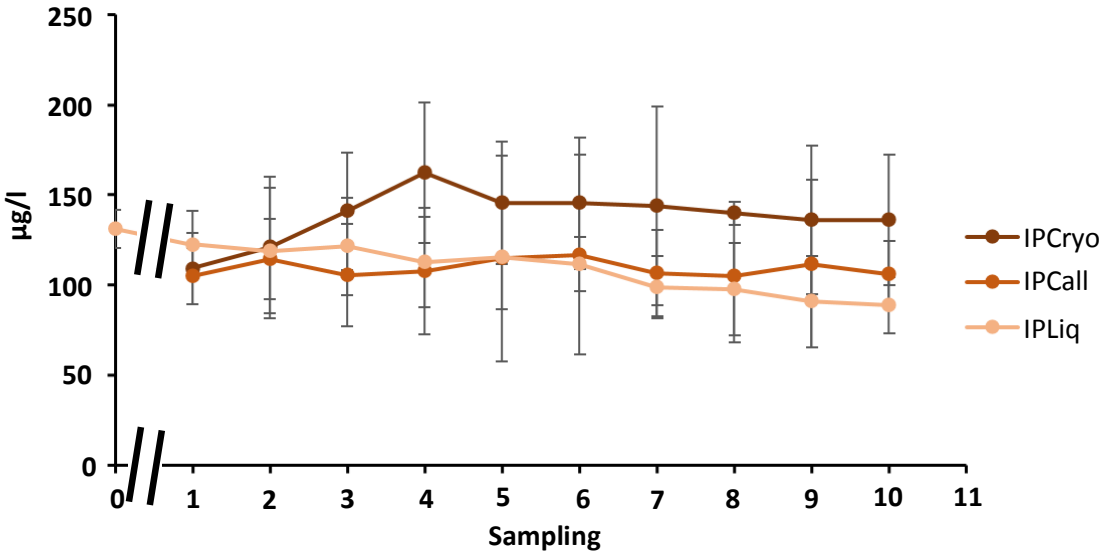


Figure R37.Time course of IGF1 accumulation in the IGF1P1b#2 transgenic line upon different maintenance methods. Data correspond to mean values obtained by ELISA analysis of culture media from three replicates, over 5 months and sampling at subculture time (day 15) according to the codification in Table R5. The error bars show the ± SD between the 3 replicates.

7. TRANSIENT PRODUCTION OF RECOMBINANT PROTEINS IN PLANT CELL CULTURES

7.1 Transient expression using Plant Cell Pack technology

Transient expression strategies allow the fast and efficient production of recombinant proteins in plants. To explore transient expression strategies for the production of recombinant proteins in plant cell cultures, we assessed the recently developed technology of Plant Cell Packs (PCPs) in our *D. carota* cell line. This technology involves the infusion of *A. tumefaciens* into plant cell aggregates deprived of cultivation medium Rademacher et al. 2019.To apply and optimize the technology with our cell line, different plant expression plasmids and Agrobacterium strains were used, and several parameters

were evaluated including: the composition of the infiltration buffer, the concentration of Agrobacterium and the time of incubation. The different assays were performed during a stage in Fraunhofer-Institut für Molekularbiologie und Angewandte Oekologie IME (Aachen, Germany), where PCPs technology had been previously developed.

First, we compared two different induction media (IM) (CRAG IM and IME IM), whose composition is described in Materials and Methods section, as well as different *A. tumefaciens* densities from 0.01 to 1.5OD_{600}. The *A. tumefaciens* strain used for introducing the genes into plant cells was GV3101. The assays were carried out with an expression cassette for the reporter fluorescent protein DsRed for easy monitoring gene expression. The expression cassette consisted in the gene encoding the DsRed controlled by the 35S promoter and the Nos-t, which was inserted into the plant expression plasmid pTRAK (Sack *et al.*, 2007). The Fig R38 shows a schematic representation of the plasmid used in the PCP transient gene expression assays.

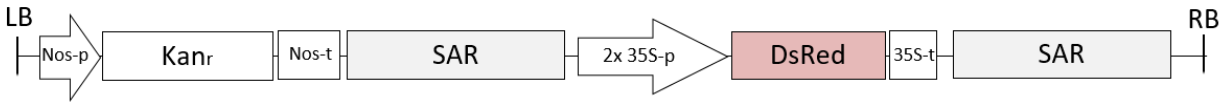


Figure R38. Schematic representation of pTRAK-RFP plasmid. LB and RB, left and right T-DNA borders; Nos-p, Nopaline synthase promoter; Nos-t, Nopaline synthase terminador; SAR, Scaffold Attachment Regions (gray boxes); 35S-p, CaMV 35S promoter; 35S-t, CaMV 35S terminator; DsRed, cDNA encoding the red fluorescent protein (red box) and Kanr, kanamycin resistance gene.

D. carota and BY2 cell cultures were used in the assay at 7 days after propagation at exponential growth phase. The PCPs were prepared using 1ml of these cell cultures per well in a 96-well microtiter plate as described in the Materials and Methods. PCPs of both cell lines were infused with 1ml of the Agrobacterium suspensions and subsequently incubated for 5 days. After the incubation, PCPs were taken out from the 96-well plates (Fig R39), total proteins extracted in PBS buffer and DsRed content determined based on fluorescence levels measured using SpectraMax® M3 system. The Fig R40 shows the results obtained for *D. carota* and BY2 cell suspensions.

The PCPs of *D. carota* and BY2 infiltrated with *A. tumefaciens* suspensions at 0.1 OD_{600} or higher showed a red coloration due to DsRed production. Agroinfusions with suspensions in the IME IM showed higher fluorescent signal than those with CRAG buffer at the different cell densities. The highest fluorescent signal was achieved by the *D. carota* and BY2 PCPs agroinfused at1 OD_{600} in IME IM. Both plant cell cultures showed similar fluorescence levels and followed similar OD-fluorescence patterns. Therefore, our selected conditions for further assays were 1 $OD_{600}$of *A. tumefaciens* suspension in IME IM.

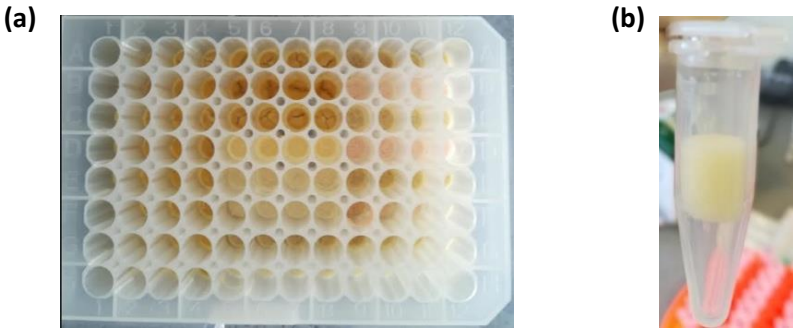


Figure R39. Plant cell pack visualization. **(a)** PCPs prepared in 96-well plates and infused with Agrobacterium suspensions carrying the pTRAK-RFP expression plasmid. The image was taken at 5 dpi. **(b)** Isolated *D. carota* cell pack in a 1.5 ml microtube.

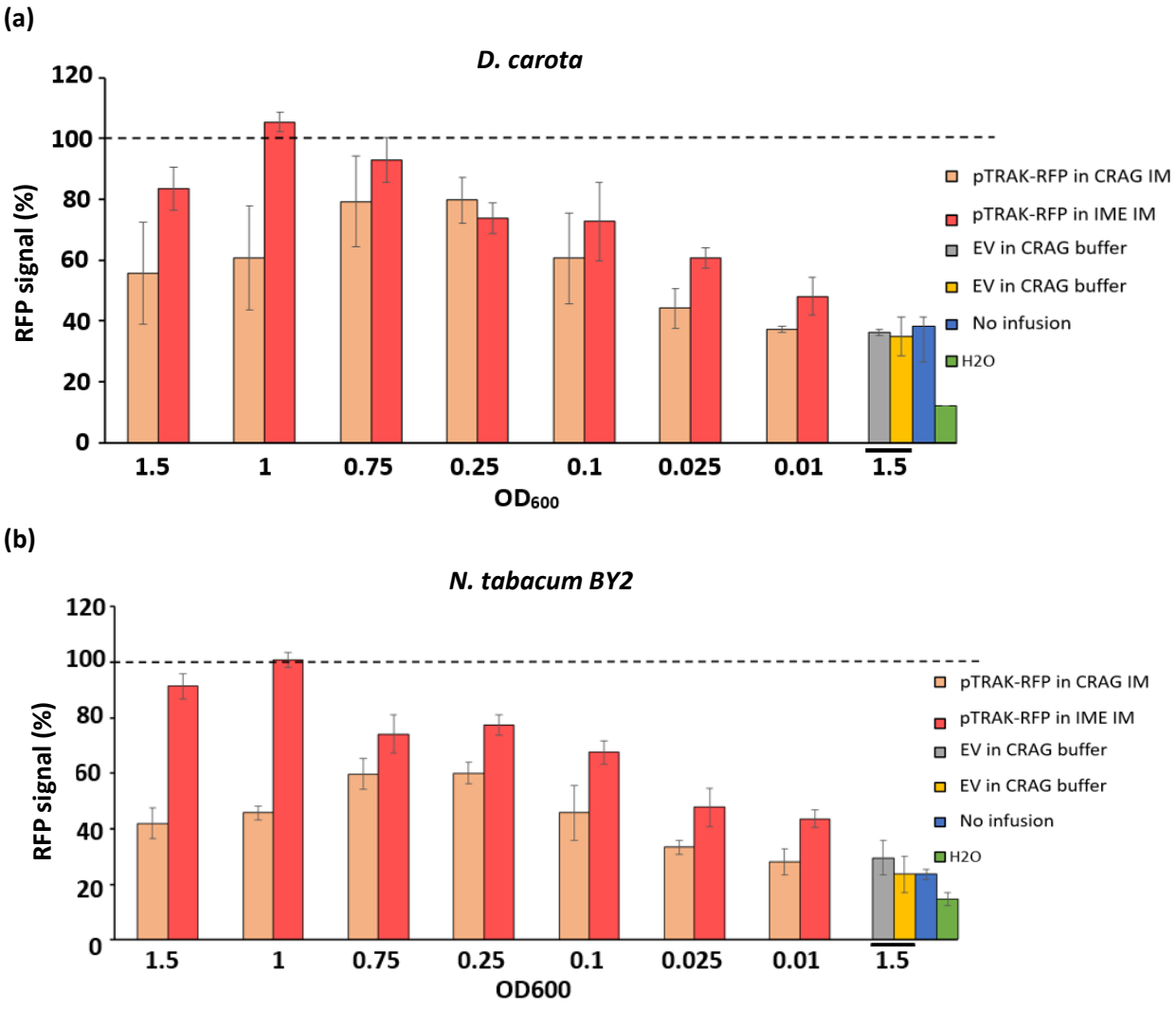


Figure R40. Optimization of the infusion buffer and Agrobacterium densities for *DsRed* expression in *D. carota* and BY2 PCPs. Graphs represent the DsRed relative fluorescence of *D. carota* **(a)** and BY2 **(b)** PCPs at 5 dpi with Agrobacterium suspensions at indicated densities and carrying the pTRAK-RFP or EV plasmids. As a control PCPs were not infused or infused with sterile water. Values are the mean ± SD of three biological replicates. The maximum average value was considered 100% signal.

Next, we also evaluated different *A. tumefaciens* strains and incubation times before infusion. Two different strains, namely GV301 and EHA105, and 7 different times were assessed. Results are shown in the Fig R41. The results indicated that for both plant cell lines the GV3101 strain triggered higher fluorescent signals at all incubation times. Also, the data revealed that an incubation of the PCPs with Agrobacterium suspensions for 50 or 60 min, before removal the excess of liquid, is optimal for DsRed production. At optimal incubation times, values were~20% higher than at the rest of the conditions.

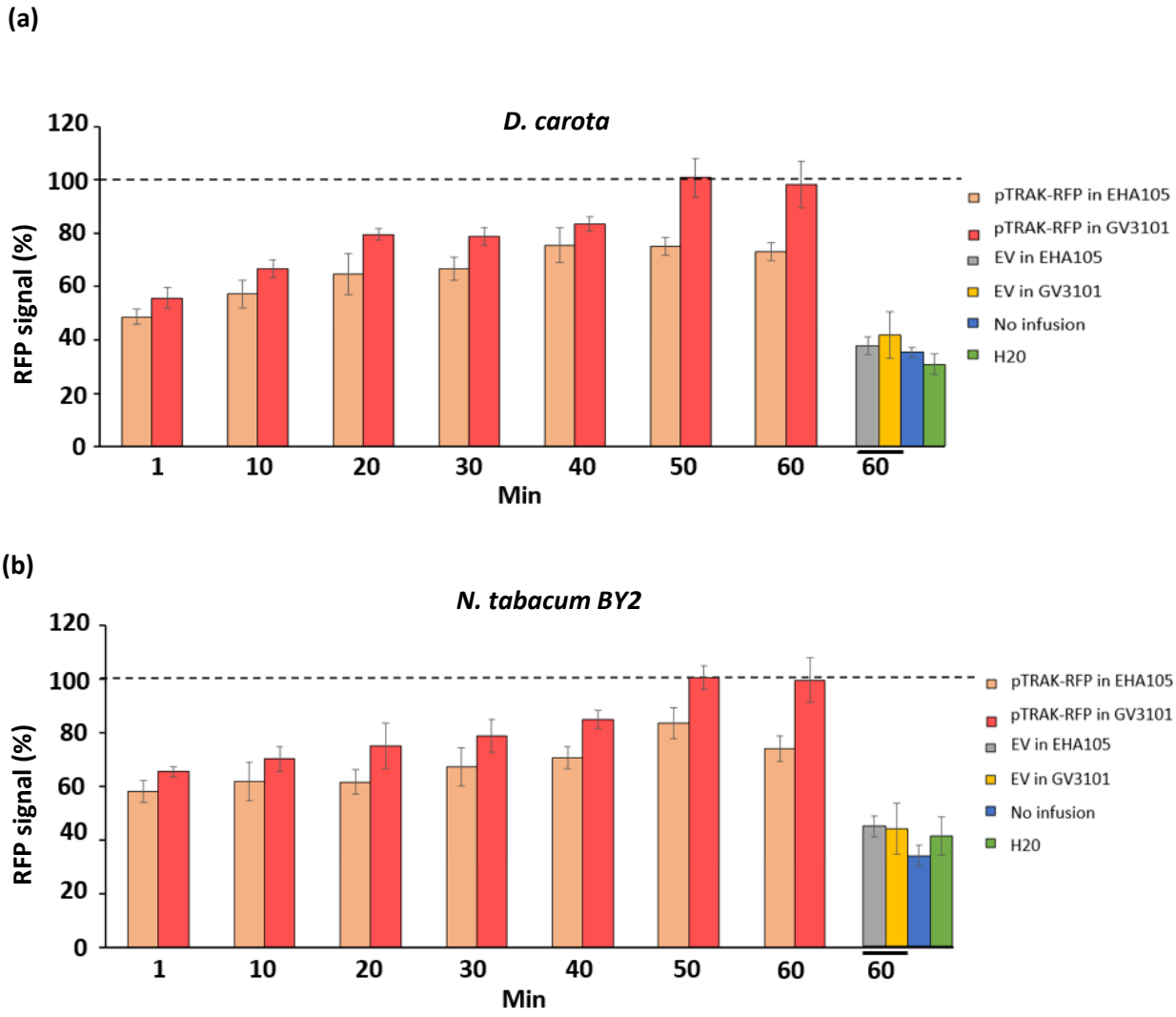


Figure R41. Selection of Agrobacterium strain and incubation time for *DsRed* expression in *D. carota* and BY2 PCPs. Graphs correspond to DsRed relative fluorescence of *D. carota* **(a)** and BY2 **(b)** PCPs at 5 dpi, infused with the indicated *A. tumefaciens* strains and incubation times. *A. tumefaciens* carried pTRAK-RFP or empty vector (EV). As a control PCPs were not infused or infused with sterile water. Values are the mean ± SD of three biological replicates. The maximum average value was considered 100% signal.

Subsequently, we assessed the effect of a surfactant on PCP transformation efficiency. The use of the surfactant SilwetL-77 in the co-cultivation solution is reported to boost transient transformation in

plants (Li *et al.*, 2009). Therefore, we evaluated the DsRed fluorescence in PCPs infused with Agrobacterium suspension in the presence of SilwetL-77 at concentrations ranging from 0.005% to 0.03%. Results shown in Fig R42 indicate that addition of the surfactant at concentrations ranging from 0.005% to 0.01% of SilwetL-77 increased the fluorescence signals in both cell lines. However higher concentrations of 0.01% resulted in a reduction on fluorescence levels.

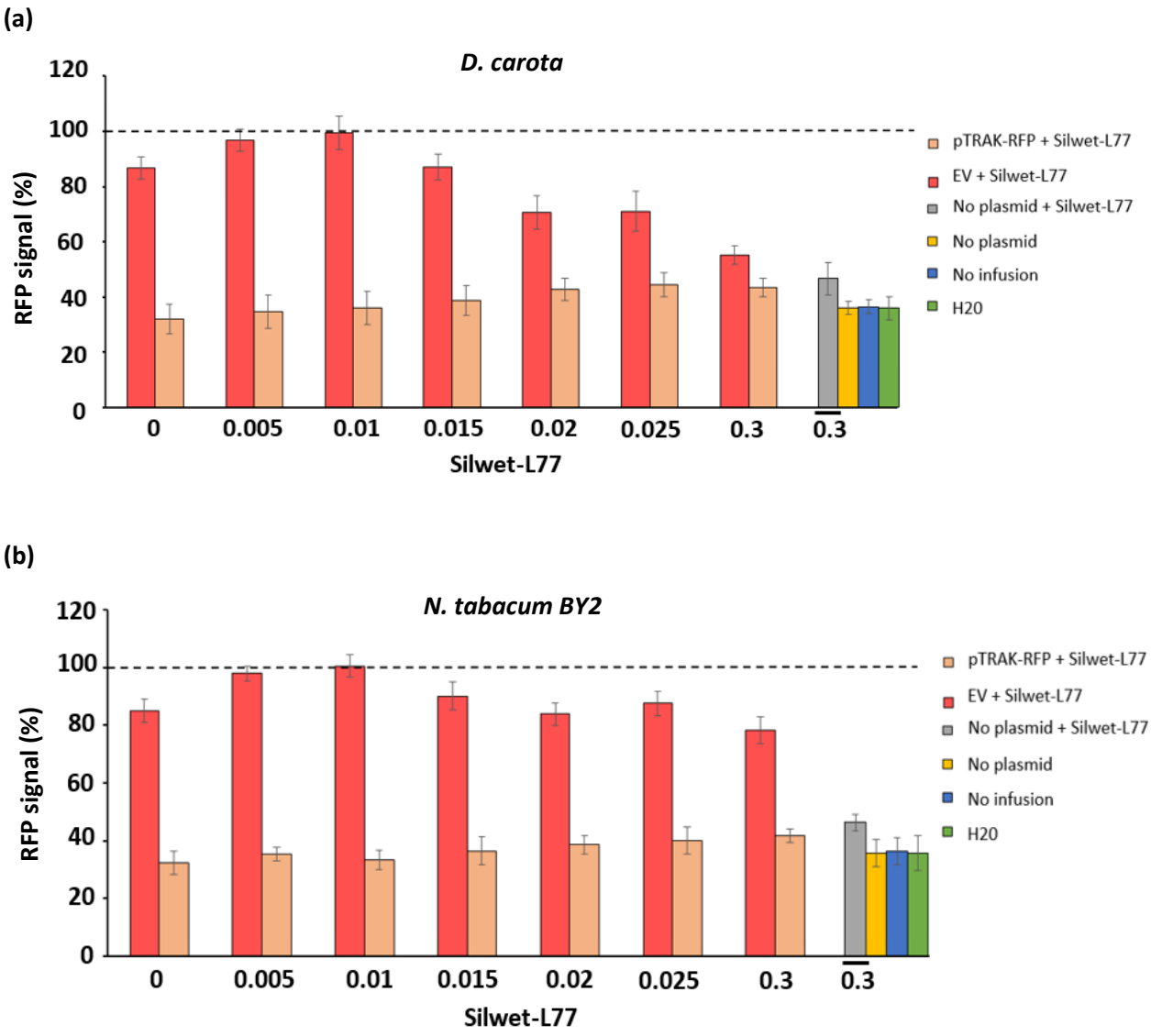


Figure R42. Evaluation of Silwet-L77 effect on transient *DsRed* expression in *D. carota* and BY2 PCPs. Graphs represent DsRed relative fluorescence of *D. carota* **(a)** and BY2 **(b)** PCPs at 5 dpi, infused with *A. tumefaciens* GV3101 carrying the pTRAK-RFP (pTRAK) or empty vector (EV) in the presence of indicated Silwet-L77 concentrations (%). As a control PCPs were not infused or infused with sterile water. Values are the mean ± SD of three biological replicates. The maximum average value was considered 100% signal.

Altogether, our results show that *D. carota* and BY2 PCPs are able to express recombinant proteins in transient expression assays. Also, we can conclude that among the tested conditions, the infusion with a suspension at 1 OD_{600} of *A. tumefaciens* strain GV3101 in IME IM containing 0.005-0.01% Silwet-L77 of surfactant, and the incubation for 50-60 min with the PCPs are the optimal conditions to achieve highest recombinant protein levels in both *D. carota* and BY2 cells. Under these conditions, we can see an intense fluorescent signal in the cell packs and protein extracts when UV-lighted, as shown in Fig R43.

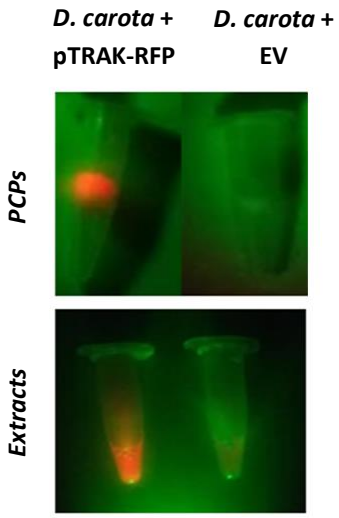


Figure R43. *D. carota* PCPs and protein extracts under UV light. Representative images of the appearance of *D. carota* PCPs infused with *A. tumefaciens* GV3101 carrying pTRAK-RFP or EV plasmids after 5 days and the derived protein extracts.

7.2 Evaluation of plasmid features on transient expression using PCPs

Once the appropriated conditions for the expression of fluorescent proteins in PCPs were established, we decided to evaluate plant expression plasmid features and how they could affect to the production of recombinant proteins.

7.2.1 Plasmid design and preparation

A major caveat with currently available binary expression plasmids used in Agrobacterium-mediated transformation is the random number and location of transgene insertions that they promoted in the host genome, and that clearly affects the expression of the transgenes (Kirchhoff *et al.*, 2012; Magy *et al.*, 2014). Strategies to avoid this limitation have been explored, and good results have been shown when Scaffold Attachment Regions (SAR) are introduced in the T-DNA regions (Breyne *et al.*, 1992).

Flanking a chimeric transgene with SAR sequences reduces significantly the variation in transgene expression among independent transformants, which leads to normalize transgene expression in stable transformation assays (Breyne *et al.*, 1992). Binary plasmids carrying SAR have successfully been used for the expression of recombinant proteins in plant cell cultures (Bortesi *et al.*, 2009). Using as backbone the pTRAK-RFP plasmid (Fig R44), we developed a new series of plasmids containing a different number and position of SAR sequences in order to evaluate their effect in PCP transient expression assays (Fig R44). We also prepared a plasmid with a shorter T-DNA to evaluate the potential effect of T-DNA length in transformation efficiency by removing the kanamycin resistance marker. All the plasmids express the *DsRed* gene under the control of 35S promoter and Nos terminator regulatory sequences. Plasmids were prepared using Gibson Assembly™ technology and transformed in *A. tumefaciens* strain GV3101 for PCP transient expression assays.

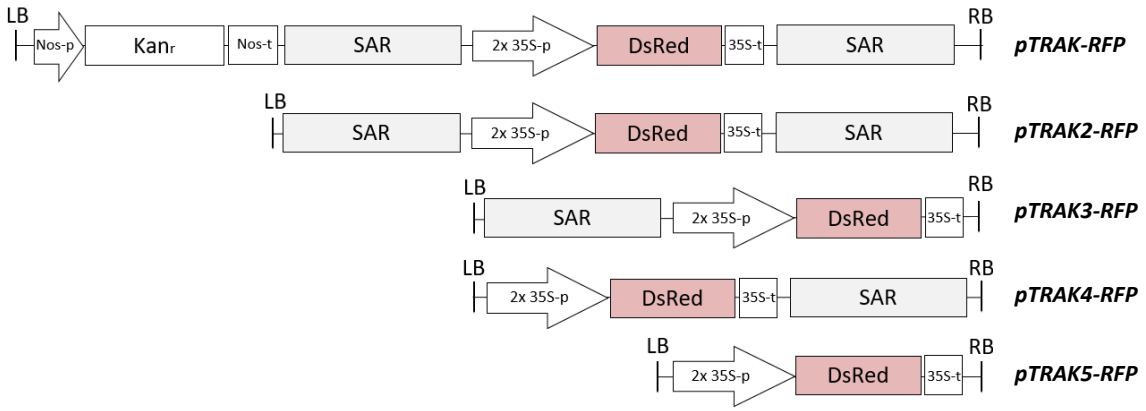


Figure R44. Schematic representation of pTRAK-RFP derived vectors. Schemes of the plasmid T-DNA regions used for the evaluation of SAR regions impact on recombinant protein production. LB and RB, left and right T-DNA borders; Nos-p, Nopaline synthase promoter; Nos-t, Nopaline synthase terminador; SAR, Scaffold Attachment Regions (gray boxes); 35S-p, CaMV 35S promoter; 35S-T, CaMV 35S terminator; DsRed, cDNA encoding the red fluorescent protein (red box) and Kan_r, kanamycin resistance gene.

7.2.2 Evaluation of pTRAK plasmids in PCP transient expression assays

We evaluated the fluorescence signal of PCPs extracts infused with *A. tumefaciens* carrying the different pTRAK-RFP derived plasmids (Fig R45). Results showed high fluorescence values when agroinfused with the plasmids carrying only one SAR (pTRAK3-RFP and pTRAK4-RFP). These values were no statistically different among them, indicating that the position of SAR region in relation to the transgene is not relevant for expression. However, pTRAK3-RFP and pTRAK4-RFP values were significantly higher than values obtained with the other plasmids. Fluorescence signal was the lowest for the plasmid pTRAK5-RFP without SAR sequences.

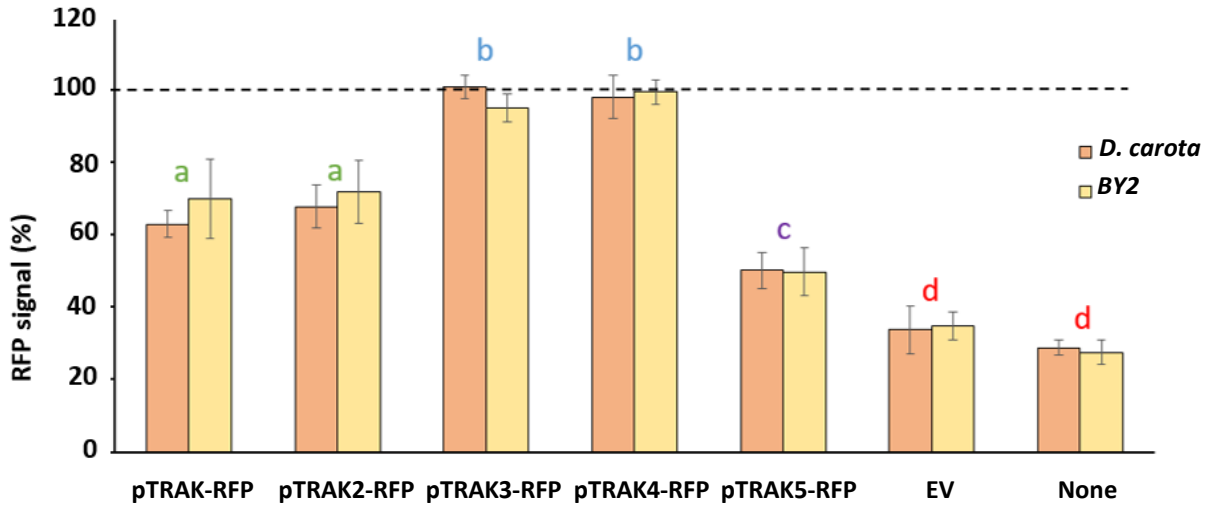


Figure R45. Evaluation of pTRAK-RFP derived plasmids in *D. carota* and BY2 PCPs. Graphs represent the DsRed relative fluorescence of *D. carota* and BY2 PCPs 5dpi of agroinfusion with the indicated plasmids. PCPs were not agroinfused as controls. Values are the mean ± SD of three biological replicates. The maximum average value was considered 100% signal for *D. carota* and BY2 PCPs. Same letters indicate no statistically significant differences according to Tukey test ($p \leq 0.05$).

Altogether, the results show that the presence of SAR in the expression plasmids boost the recombinant protein production. Also, the results indicate that the presence of one single SAR fragment triggers higher protein production in PCPs than two SAR fragments flanking the expression cassette.

7.2.3 *In planta* evaluation of pTRAK plasmids for fluorescent proteins production

The suitability of pTRAK plasmids for gene expression in plant cells was also assessed in whole plants. For that, *N. benthamiana* leaves were agroinfiltrated with the different pTRAK-RFP plasmids, alone or in combination with *P1b* silencing suppressor in transient expression assays. Then, the DsRed fluorescent signal from total protein extracts in saline buffer was quantified using a SpectraMax® M3 system. Protein extracts were prepared from infiltrated leaf sectors at 3, 5 and 7 dpi and analyzed in reader compatible plates. Fluorescence values at 3 dpi were very low in all the conditions (data not shown). The values at 5 dpi are shown in Fig R46, where we observe that the patches infiltrated with the plasmids containing only one SAR region (pTRAK4-RFP and pTRAK5-RFP) in combination with *P1b* reported the highest signal, without significant differences among them and significant differences in comparison to other plasmids. Similar results were obtained at 7 dpi (data not shown).

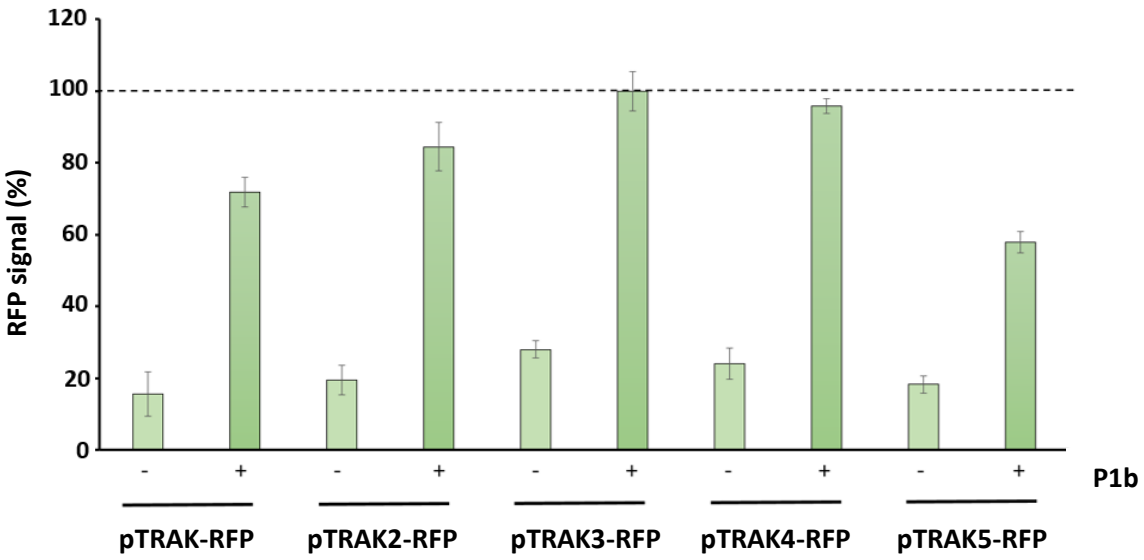


Figure R46. *DsRed* expression in *N. benthamiana* leaves using pTRAK-RFP plasmids. Graph represents the fluorescence signal of protein extracts from agroinfiltrated leaves with indicated plasmids at 5 dpi. Bars correspond to the average of three relative values in comparison to mock agroinfiltrated leaf signals and as percentage to the highest value (100% signal). Error bars show standard deviation three pools of leaves.

Our results show that plasmids carrying SAR are suitable for transient gene expression not only cell suspensions but also in whole plants. We also show that those plasmids carrying one single SAR sequence, namely pTRAK3-RFP and pTRAK4-RFP, mediate the higher protein levels, as previously observed in plant cell suspension assays. Moreover, the agroinfiltration of pTRAK-RFP plasmids in combination with silencing suppressors increases by a factor of three the protein production.

7.*1.4* Plasmid design for rapid cloning of transgenes

Although pTRAK plasmids show suitability for good and homogenous gene expression in plant cells, a major drawback is that they rely on traditional restriction digestion and ligation for cloning target genes, hampering rapid and easy gene cloning. To overcome this limitation, we prepared a modified version of pTRAK plasmid compatible with Gateway™ technology to enable rapid assembly of target genes for the production of recombinant protein in plants. The new plasmid was named pTAREK_GW and its features are shown in Fig R47a. A further improved version was also prepared (pTAREK-P1b, Fig R47b), this plasmid contains the expression cassette of the RNA silencing suppressor *P1b* into the T-DNA region for co-expression with the target gene, to enhance its expression as demonstrated early in this work.

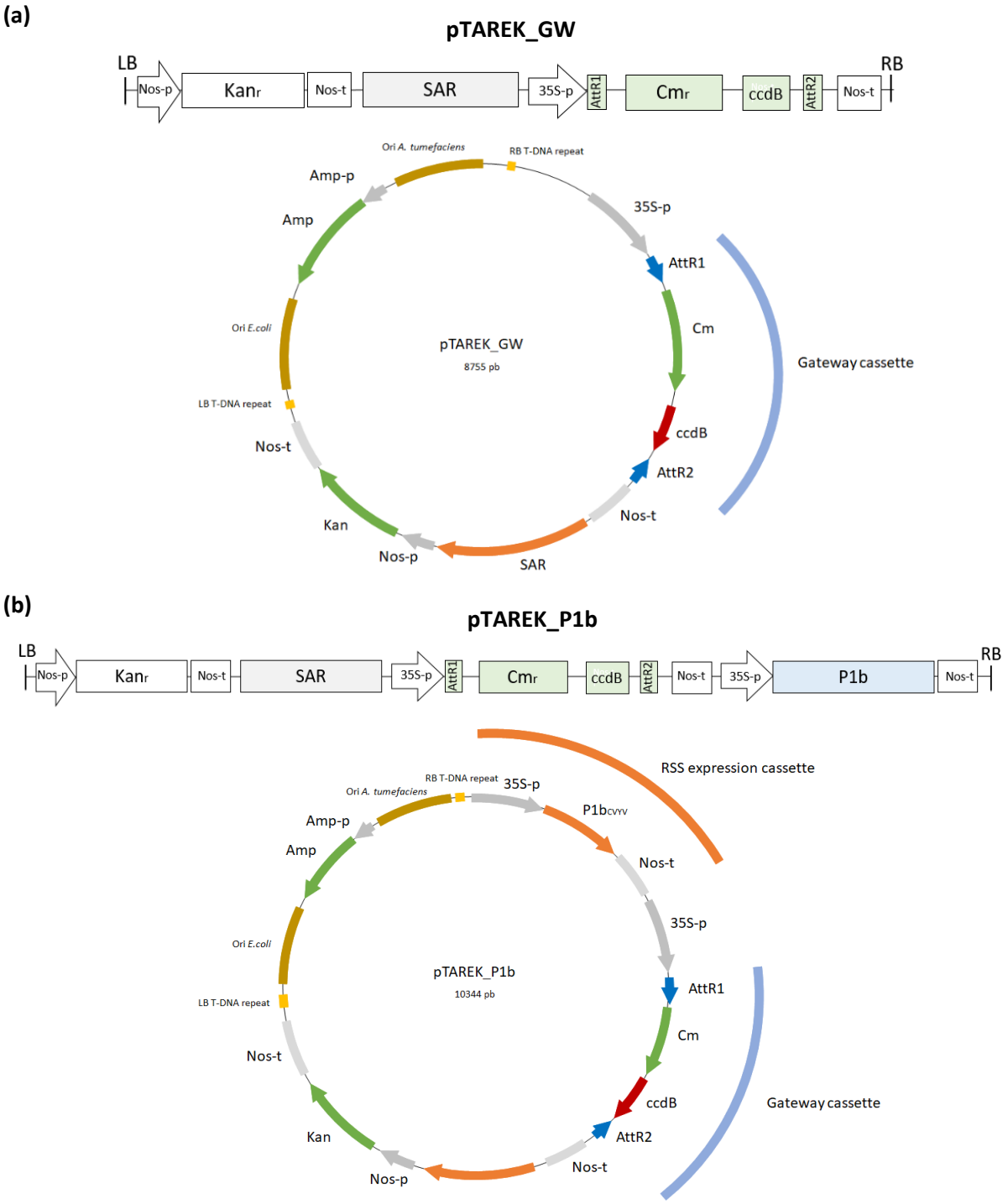


Figure R47. Schematic representation of pTAREK plasmids compatible with Gateway cloning system. Gateway cloning compatible plasmids for the expression of recombinant proteins based on pTRAK plasmid, namely pTAREK-GW **(a)** and pTAREK-P1b **(b)**. Features include:LB and RB, left and right T-DNA borders; Nos-p, Nopaline synthase promoter; Nos-t, Nopaline synthase terminador; SAR, Scaffold Attachment Regions (gray boxes); 35S-p, CaMV 35S promoter; 35S-t, CaMV 35S terminator; Kan, kanamycin resistance gene; Gateway recombination sites (attR1 and attR2 recombination regions; Cm, chloramphenicol resistance marker gene; and ccdB, *E. coli* DNA gyrase inhibitor gene); and P1b expression cassette.

Both plasmids pTAREK-GW and pTAREK-P1b contains the ampicillin resistance marker forselection in bacteria and kanamycin resistance marker in the T-DNA region for selection in plants. Also, they incorporate the AttR1 and AttR2 Gateway recombination sites for the cloning of target genes.

7.1.5 Evaluation of Gateway plasmids for gene expression in plants

To test the transient expression capability of Gateway compatible pTRAK plasmids, we prepared plasmids containing the *GFP* gene for easy monitoring expression by fluorescence. *GFP* expression mediated by pTAREK-GFP or pTAREK-P1b-GFP was compared tothe pGWB6 binary plant expression plasmid, which carries also the *GFP* sequence driven by the 35S promoter and Nos terminator regulatory sequences. Cloning the *GFP* sequence in pTAREK was simply done by recombination from an entry plasmid containing the GFP ORFs. A scheme of T-DNAs of the three plasmids are shown in Fig R48. These three plasmids were agroinfiltrated into *N. benthamiana* leaves following the distribution indicated in Fig R49 and GFP production was visually monitored under UV-light over the time. GFP was first detected at 3 dpi and accumulated gradually during the next 2 days, reaching maximum levels at 5 dpi (Fig R50). We detected a brighter fluorescence in the patches agroinfiltrated with the *P1b*, no matter is co-expressed in trans or cis; whereas those patches in combinations lacking the *P1b* showed a faint fluorescence at 3 dpi that disappeared with days.

Visually, GFP signal looked stronger in patches agroinfiltrated with pTAREK plasmids than those agroinfiltrated with pGWB6. For a more precise estimation, we quantified the fluorescent signal from protein extracts from different replicas. Leaf patches were slices, protein extracted in saline buffers and then fluorescence measured in a SpectraMax® M3 system. Results are shown in Fig R50, where we can observe more than two-fold increase in fluorescence values with pTAREK plasmids than with pGWB6. The differences in fluorescence were statistically significant when compared with pGWB6 but not among them. This indicates that co-expression of *P1b* significantly increase the production of GFP independently that co-expressed in cis or trans. We also determined that the maximum protein accumulation levels were reached at 5 dpi, with 300 ± 8.2, 681 ± 51.6 and 643 ± 56.9 µg protein/g fresh leaf for pGWB6, pTAREK_GW and pTAREK_P1b respectively (Fig R50).

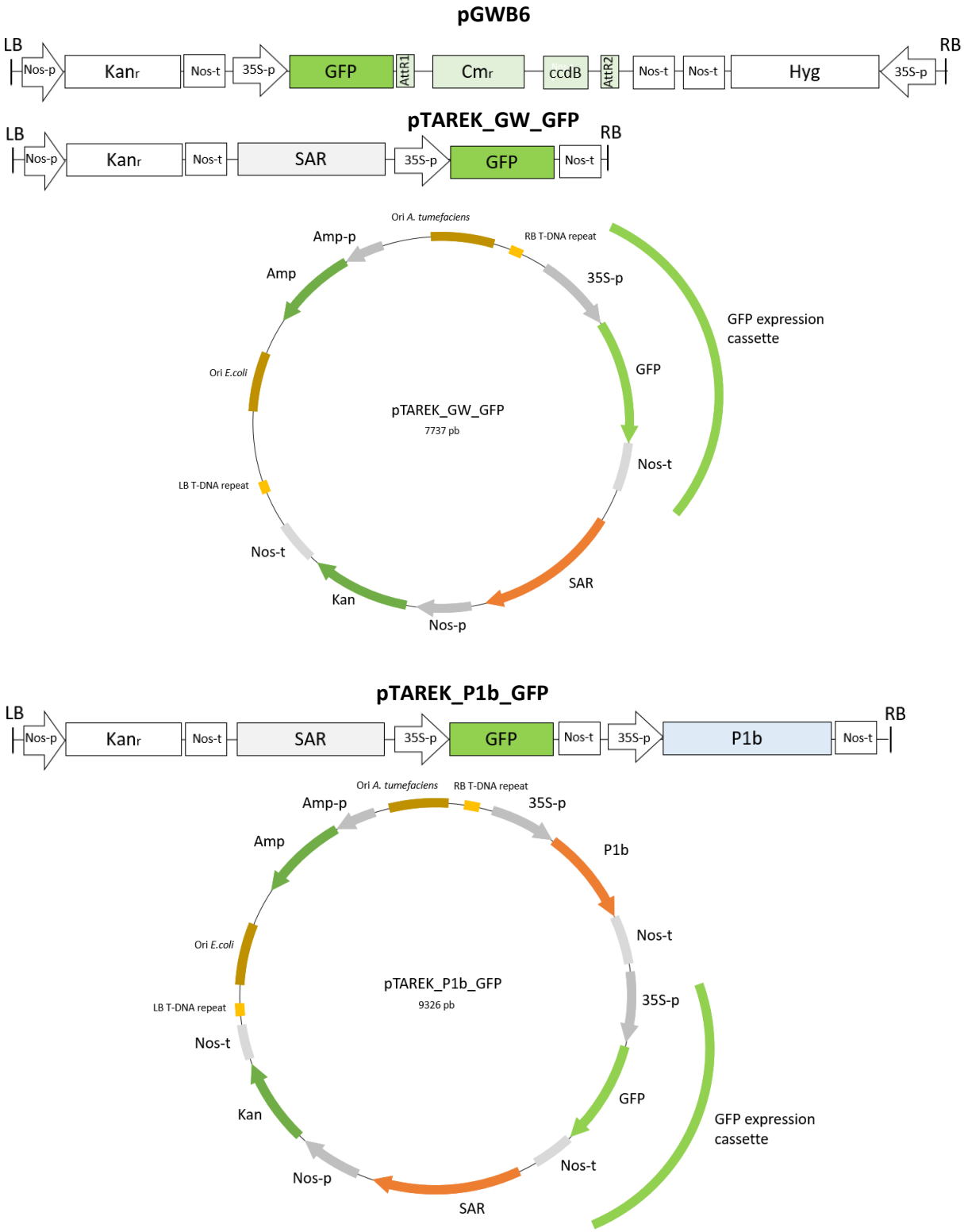


Figure R48. Schematic representation of pTAREK and pGWB plasmids used for GFP production evaluation. Gateway cloning compatible plasmids for the expression of recombinant GFP based on pTAREK and pGWB plasmids. Features include: LB and RB, left and right T-DNA borders; Nos-p, Nopaline synthase promoter; Nos-t, Nopaline synthase terminador; SAR, Scaffold Attachment Regions (gray boxes); 35S-p, CaMV 35S promoter; 35S-t, CaMV 35S terminator; Kan, kanamycin resistance gene; Gateway recombination sites (attR1 and attR2 recombination regions; Cm, chloramphenicol resistance marker gene; and ccdB, *E. coli* DNA gyrase inhibitor gene); GFP, green fluorent protein; Hyg, hygromycin resistance marker gene; and P1b; P1b RSS.

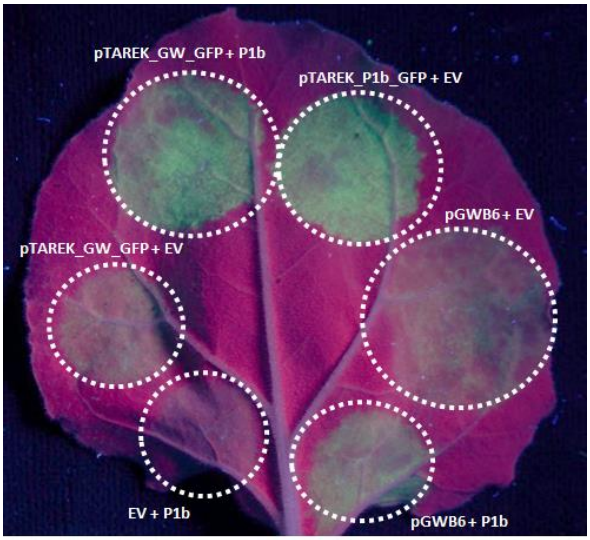


Figure R49.Transient expression of GFP in *N. benthamiana* leaves mediated by pTAREK plasmids. Image of a representative *N. benthamiana* leaf under UV light illumination from agroinfiltration assays using pTAREK_GW_GFP, pTAREK-P1b-GFP or pGWB6 plasmids in combination with EV or P1b vector at 5 dpi.

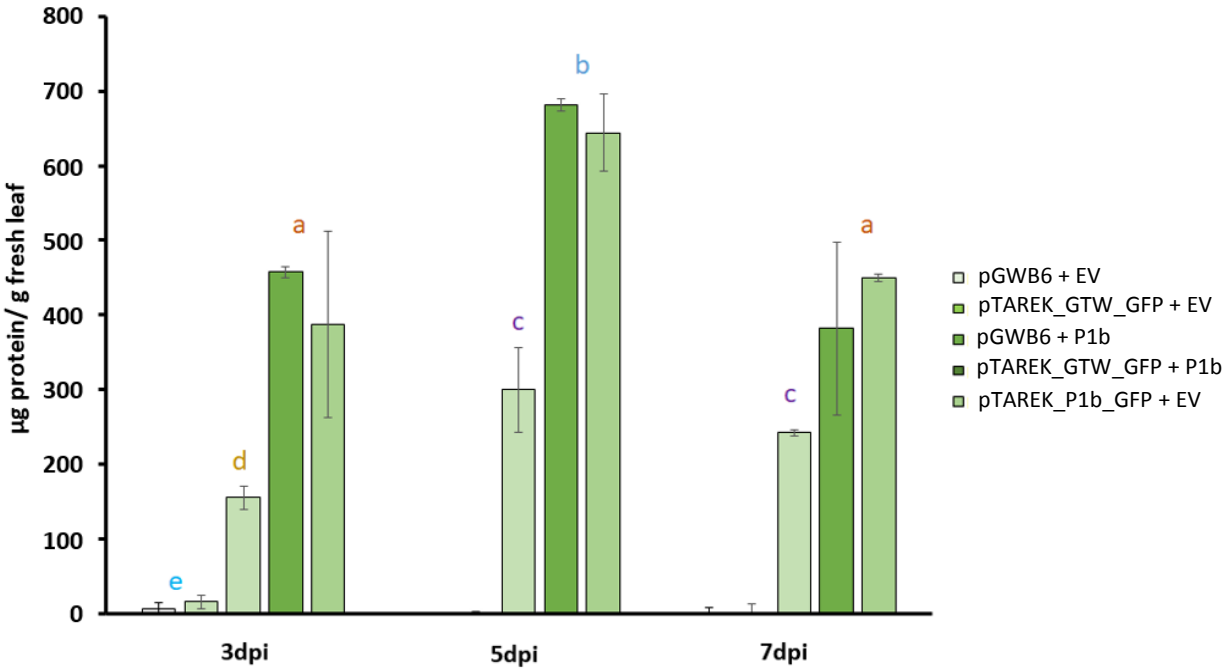


Figure R50. GFP expression in *N. benthamiana* leaves using pTAREK plasmids. GFP fluorescence was measured in soluble protein extracts from agroinfiltrated leaves with indicated plasmids at indicated times. Protein amounts was determined by comparison to fluorescence values of known amounts of GFP proteins in µg protein/g fresh leaf. Bars correspond to mean values ±SD of three independent plant pools (each pool comes from 3 patches from 3 leaves from same plant). Same letters indicate no statistically significant differences according to Tukey test ($p \leq 0.05$).

We can conclude that the use of pTAREK plasmids boost more than twice the *GFP* expression in comparison to classical binary plasmids. Also the results showed that the expression of CVYV silencing

suppressor *P1b* in combination with the engineered plasmids triggers similar protein accumulation than when co-expressed in *cis*.

7.1.6 Overexpression of the reporter *GFP* or the *P1b* silencing suppressor in combination with pTAREK plasmids

Once demonstrated that pTAREK_GW_GFP and pTAREK_P1b_GFP plasmids trigger similar accumulation levels of GFP levels in transient expression assays in *N. benthamiana* leaves, additional assays were done to explore whether the co-expression of the reporter GFP or the P1b silencing suppressor together with pTAREK_P1b_GFP would resulted in enhanced recombinant protein production in trans. To evaluate these effects pTAREK_P1b_GFP was co-infiltrated in combination with pGWB6 and pTAREK_GW_GFP plasmids. Results are shown in Fig R51, where we can visualize the GFP fluorescence in all the patches agroinfiltrated with the different combinations of expression cassettes for GFP and P1b.

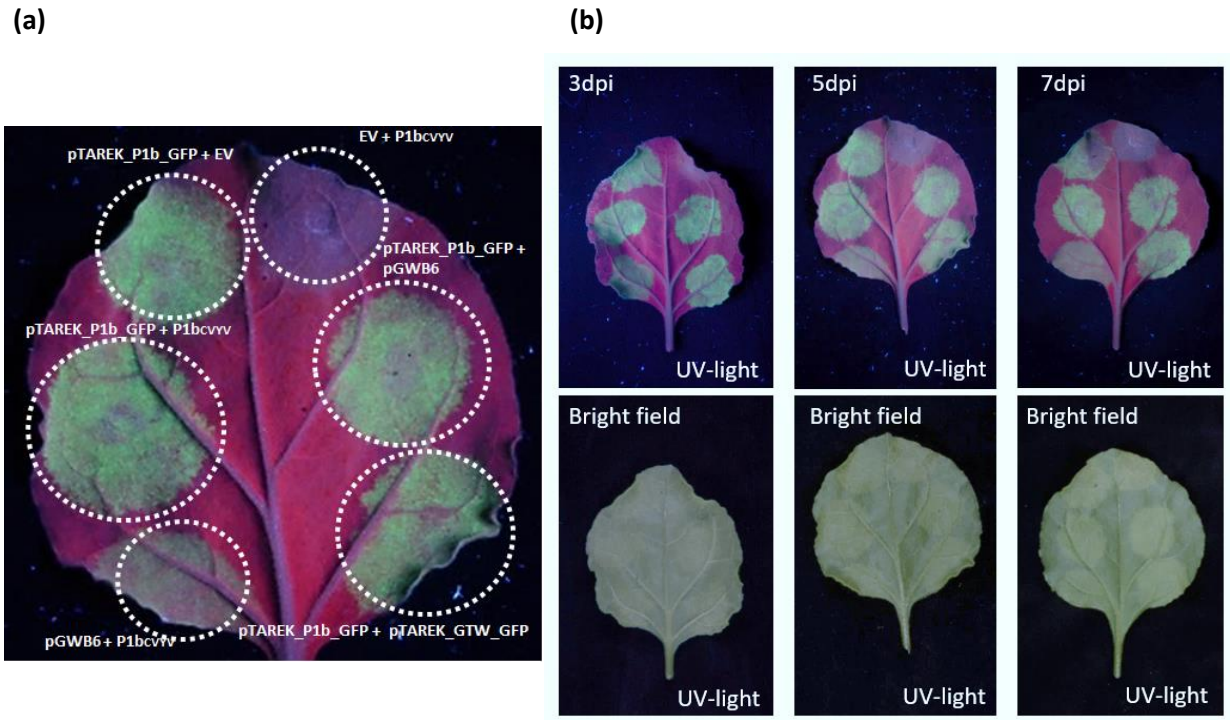


Figure R51. Overexpression of *P1b* and *GFP* in combination with pTAREK_P1b_GFP. (a) Photograph of a representative *N. benthamiana* leaf agroinfiltrated with indicated plasmids under UV-light at 5 dpi. The infiltrated patches are surrounded by a white circle. **(b)** Representative images of *N. bentamiana* leaves agroinfiltrated as in panel (a) at different times. Top images show the leaves under UV light and lower images under white light.

The fluorescence was quantified from protein extracts and values are presented in Fig R52. The control with pGW86 co-agroinfiltrated with P1b was producing significantly less GFP than the rest of samples. The comparison showed similar results for GFP levels in all patches with plasmid combinations that include pTAREK_P1b_GFP, with differences not being statistically significant (Tukey test, $p \leq 0.05$). These results suggested that the high GFP expression achieved with pTAREK_P1b_GFP was not increased by co-expression of additional constructs providing additional *GFP* or *P1b*.

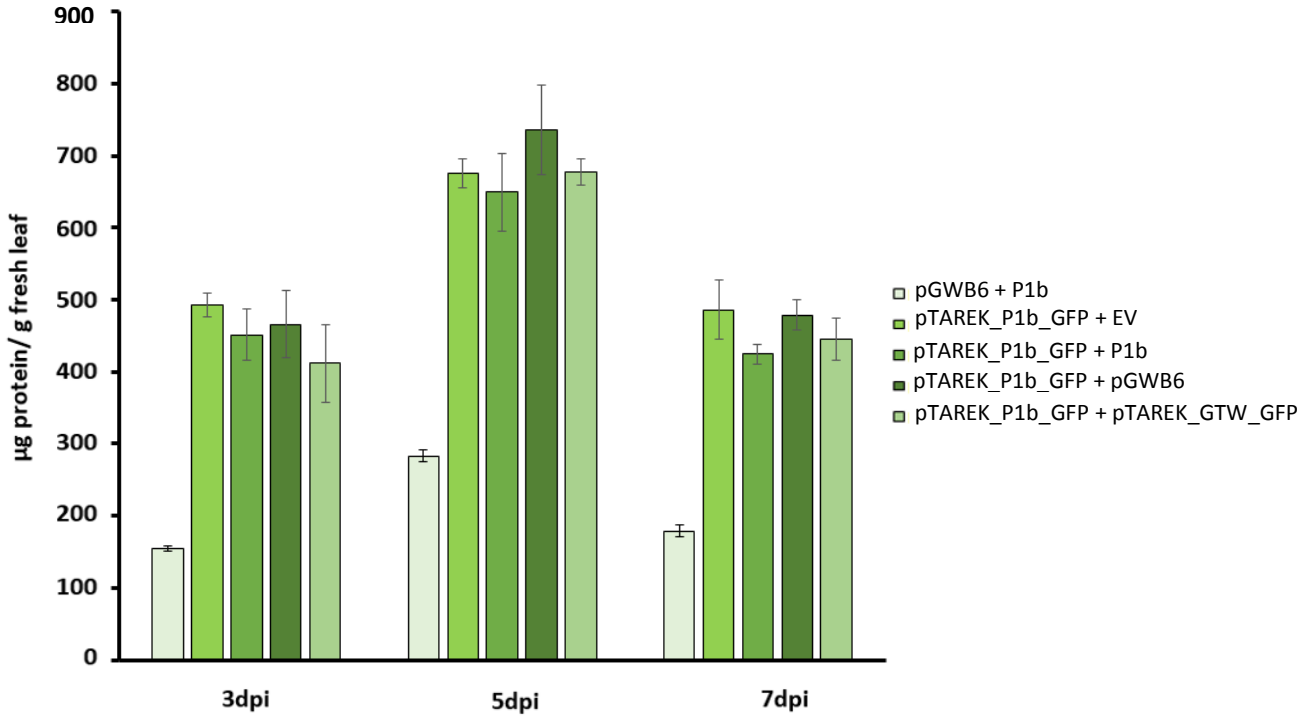


Figure R52. Co-expression of *P1b* with pTAREK_P1b_GFP. GFP expression levels in leaf patches co-infiltrated with pTAREK_P1b_GFP and different binary plasmids. GFP fluorescence was measured in soluble protein extract obtained from agroinfiltrated leaves, and expression is shown µg protein/g fresh leaf. The protein concentration was determined using a concentration curve using purified GFP. Error bars represent standard deviation between independent leaf pools (n=3).

8. PRODUCTION OF PEPTIDES IN *N. BENTHAMIANA* USING VIRAL VECTORS

8.1 Viral vector design

Transient expression in plants using viral vectors is a simple and fast method for the production of recombinant proteins. Usually, the protein yields achieved with these systems are higher compared to other strategies (Pasin *et al.*, 2019). Considering the aim of this project and the bibliographic information, we decided to explore the use of viral vectors for the production of our target

recombinant proteins in *N. benthamiana*. We selected a viral vector derived from *Tobacco mosaic virus* (TMV) that our group recently reported to be particularly adequate for this production strategy (Shi *et al.*, 2019). The viral vector consists in a disarmed cDNA version of the genome of TMV in which most of the coat protein (CP) coding sequence was replaced by the gene of interest. The CP of TMV plays important roles in RNA encapsidation and systemic infection, being necessary for virion formation and spreading of the infection throughout the plant. A small part of the *CP* sequence remained in the vector because it contains functional elements for the subgenomic promoter of the *CP* mRNA. This subgenomic promoter is essential for the high expression of the *CP* (in the wild type virus for each viral RNA, 2130 CP proteins are needed to encapside a progeny virion), and consequently it allows the production of very high yields of recombinant protein in this viral vector. The viral vector also retains the polymerase required for viral replication, with RNA-dependent RNA polymerase (RdRp) activity, and the movement protein (MP) needed for cell-to-cell spread through plasmodesmata.

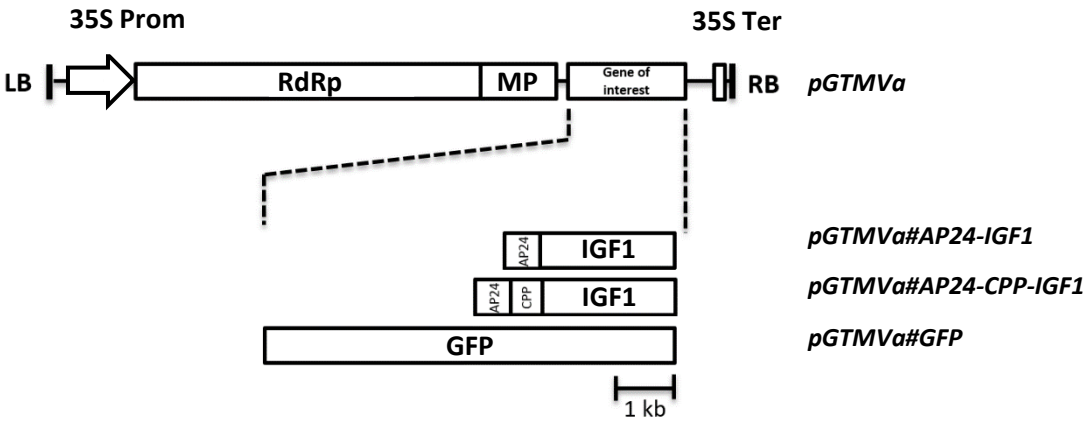


Figure R53. Schematic representation of TMV-derived constructs for the production of the indicated proteins. Two constructs were prepared in the pGTMVa plasmid for their expression under the control of the 35S promoter and terminator. They carry the sequences encoding IGF1 and CPP-IGF1. Both peptides contained the signal peptide of tobacco AP24 protein at the N-terminus for targeting to cell apoplasts. The GFP construct was used as a control. LB and RB, left and right borders of the Agrobacterium transfer DNA; RdRp, replicases; MP, movement protein.

Using pGTMVa as backbone, two different constructs carrying the codon optimized cDNA encoding the IGF1 peptide were prepared (Fig R53). One of them encoded the single IGF1 peptide alone, and the other was preceded by the cell penetrating peptide (CPP-IGF1) fused at the N-terminus. Both of them were designed to target the accumulation of the respective peptides in the cell apoplast by incorporating the signal peptide of the tobacco AP24 protein at the N-terminus. This signal peptide allows the protein to enter the secretory pathway toward the extracellular space. The accumulation of the peptides in the apoplasts may facilitate its subsequent purification from plant leaves. Constructs were verified by DNA sequencing and used for Agrobacterium transformation.

8.2 Production of IGF1 and CPP-IGF1 in *N. benthamiana* plants

The viral vectors pGTMVa#AP24-IGF1 and pGTMVa#AP24-CPP-IGF1 were transferred to *N. benthamiana* leaves via agroinfiltration. The pGTMVa#GFP control construct was used to easily monitorize the viral expression. Agroinfiltrated leaves were visually inspected every day during the following week. We observed that leaves agroinoculated with the vectors pGTMVa#AP24-IGF1 and pGTMVa#AP24-CPP-IGF1 showed a wrinkled or yellowish appearance, clearly differentiated from leaves agroinoculated with pGTMVa#GFP or Mock inoculated (Fig R54), which always showed a healthy green appearance, with only little lesions caused by syringe pressure during the infiltration process. These results suggested a possible toxicity associated to *IGF1* expression.

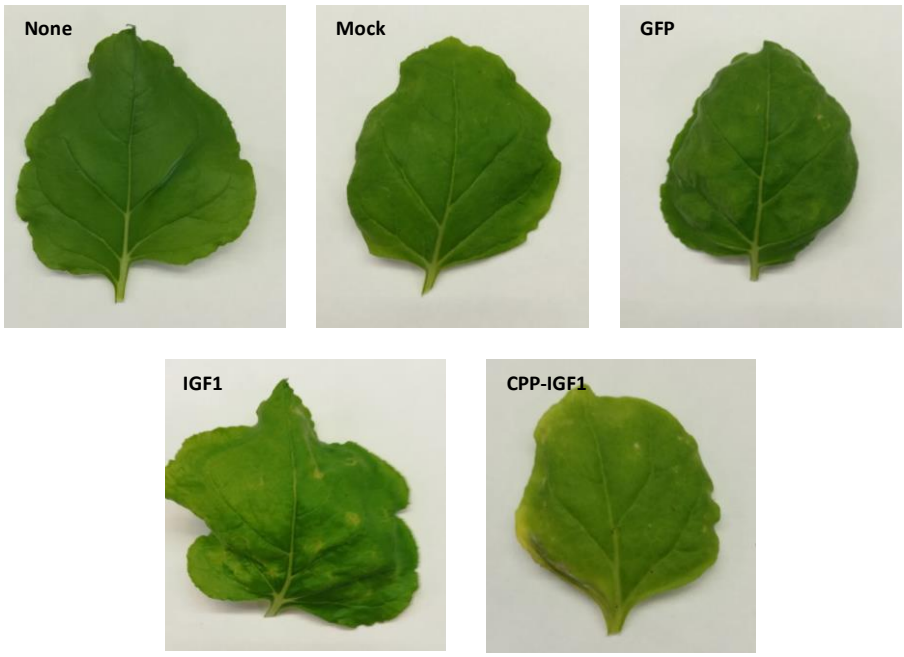


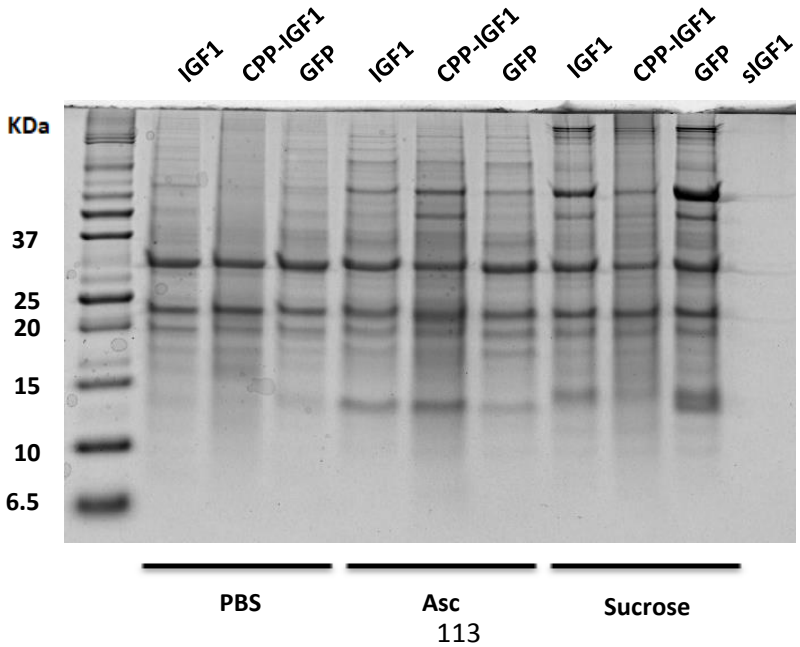
Figure R54. Phenotype of leaves expressing *IGF1* and *CPP-IGF1* genes. Appearance of *N. benthamiana* leaves 7 dpi with the indicated constructs or with the *A. tumefaciens* (mock).

To determine if IGF1 and CPP-IGF1 were produced in *N. benthamiana* leaves after the agroinfiltration, apoplasts were extracted and evaluated by SDS-PAGE at 7 dpi. Extraction of apoplastic fluids is a simple procedure based on vacuum infiltration of the leaves with a buffer solution followed by centrifugation. Three different extraction solutions were tested, namely PBS (phosphate-buffered saline), ascorbate, and sucrose buffers (see Materials and Methods section for composition). Proteins from extracellular fluids (ECFs) obtained from leaves agroinfiltrated with the viral vectors were separated by SDS-PAGE and stained with Coomassie blue. Protein profiles are shown in Fig R55a, where differences among extracts obtained with PBS, ascorbate or sucrose buffers were observed, but where no differences

among the extracts from samples IGF1, CPP-IGF1 or GFP were detected. A sample of the synthetic peptide run in parallel for comparative purposes was not detected by Coomassie blue staining. Then, the samples were analyzed by immunoblot detection using specific anti-IGF1. As shown in Fig R55b, several immunoreactive bands were detected in the extracts from IGF1 and CPP-IGF1 that were absent in GFP or mock extracts. The sizes of the bands in IGF1 extracts were similar to those observed in the synthetic IGF1 peptide sample, where bands around 8, 16, and 22 KDa were present. The IGF1 peptide was detected in the extracts obtained with the three different buffers. These results indicate that IGF1 is produced in *N. benthamiana* leaves and targeted to the apoplastic space. Our results also showed that IGF1 displays a tendency for multimerization during the analysis, both in the extracts and in the control sample of synthetic peptide.

In the case of CPP-IGF1 extracts, a band of higher mobility compared to the lower one of IGF1 (aproximately 8 KDa) was detected, with an apparent size of 11 KDa corresponding to the expected size for CPP-IGF1 peptide. This band was only detected in the extracts obtained with ascorbate and sucrose buffers (in this case less concentrated), being absent in the PBS buffer extracts. These results suggested that the CPP sequence could alter the physicochemical properties of IGF1 hampering its extraction in PBS buffer. Additional bands with higher sizes were also detected, suggesting that CPP-IGF1 has also a tendency to multimerize. Remarkably, the intensity of CPP-IGF1 bands was always less than the observed with IGF1 bands. This suggested that CPP-IGF1 accumulated to lower levels than IGF1, although alternatively we cannot discard that the differences were caused by a better detection of IGF1 than CPP-IGF1by the antibodies used. In any case, our results indicated that both peptides, IGF1 and CPP-IGF1 were produced and accumulated in the apoplast of *N. benthamiana* leaves using the TMV-derived vector expression system.

(a)



(b)

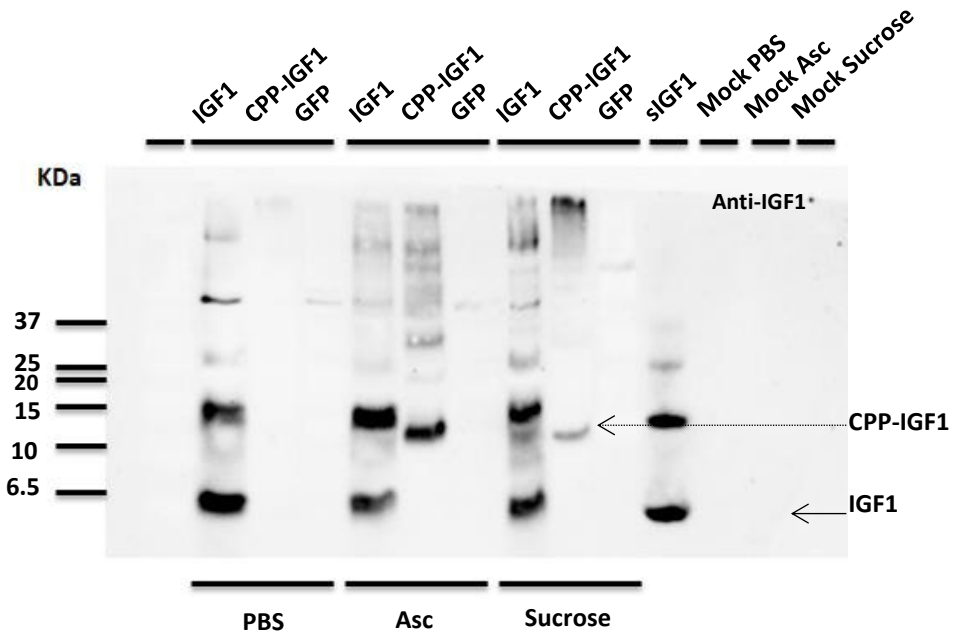


Figure R55. IGF1 and CPP-IGF1 accumulation in *N. benthamiana* leaves. **(a)** Protein profile of apoplasts from leaves at 7 dpi with pGTMVa#AP24-IGF1, pGTMVa#AP24-CPP-IGF1 or pGTMVa#GFP extracted using PBS, ascorbate (Asc) or sucrose buffers, as indicated. Proteins were separated by tricine-SDS-PAGE and Coomassie-blue stained. **(b)** Western-blot analysis of IGF1 and CPP-IGF1 using specific antibodies anti-IGF1. Synthetic IGF1 (50ng) was run in parallel as control. Molecular markers are shown in kDa on the left.

Once the production of IGF1 and CPP-IGF1 was demonstrated by Western Blot analysis, accumulation levels were determined using the ELISA assay. An indirect ELISA was used for the quantification of the amount of IGF1 and CPP-IGF1 peptides in the plant extracts. The ECFs and total protein extracts were analyzed in ELISA assays using specific anti-IGF1 antibodies.

Leaf samples were collected at 3, 5, 7, 9 and 12 dpi. At least 3 leaves per time point were pooled, weighted and ECFs or total protein extracted. The amount of peptide in each sample was estimated using ELISA assays and values are represented in the Fig R56 and R57. IGF1 was first detected at 3 dpi, and accumulation levels increased gradually over the next days reaching a maximum at 7 dpi, and then, decreasing progressively until 12 dpi. The maximum accumulation levels correspond to 7.92 ± 1.06 µg/g fresh weight of leaf for IGF1, and 1.52 ± 0.37 µg/g for CPP-IGF1 in ECFs. When total extracts were analyzed, yields were slightly higher, reaching 8.62 ± 0.37 µg for ICF1 and 1.8 ± 0.18 µg of CPP-IGF1 per gram *N. benthamiana* leaf (Fig R58). As expected, these differences in peptide concentration between ECFs or total protein extracts were not statistically significant (Tukey test, $p \leq 0.05$), which indicates that the recombinant proteins were correctly exported to the ECFs.

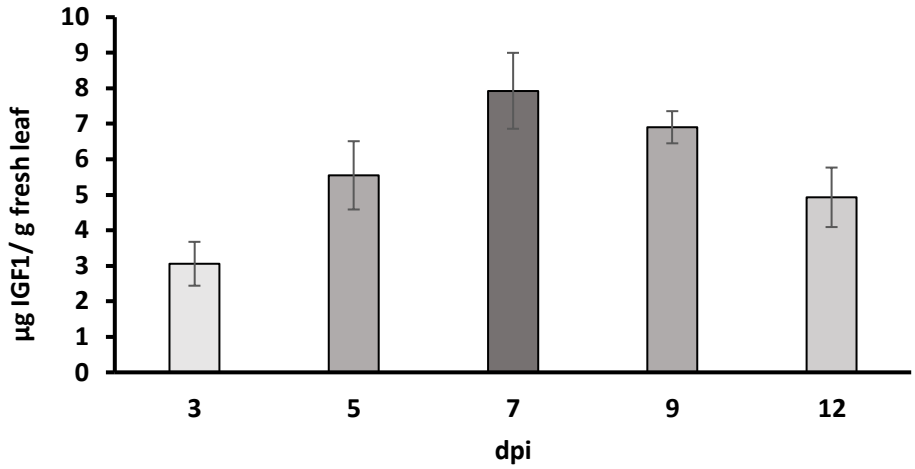


Figure R56. IGF1 accumulation in *N. benthamiana* leaves. Peptide accumulation in ECFs extracted with PBS from leaves as estimated by indirect ELISA analysis in comparison with known amount of synthetic IGF1. Values correspond to the mean of 3 independent estimations from pools of three leaves, and bars to SD.

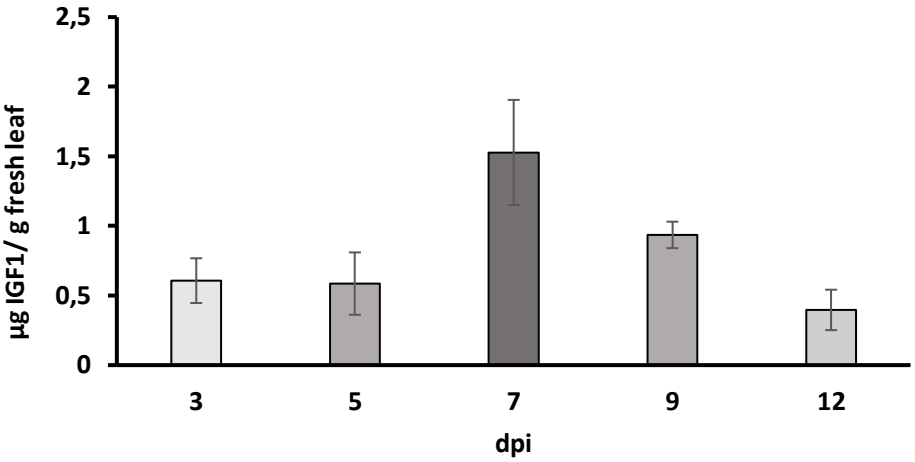


Figure R57. CPP-IGF1 accumulation in *N. benthamiana* leaves. Peptide accumulation in ECFs extracted with ascorbate buffer from leaves, as estimated by indirect ELISA analysis in comparison with known amount of synthetic IGF1. Values correspond to the mean of 3 independent estimations from pools of three leaves, and bars to SD.

Altogether, our results show that IGF1 and CPP-IGF1 are produced in *N. benthamiana* leaves using viral vectors. Accumulation of both peptides was targeted to leaf apoplasts facilitating extraction and downstream purification. Maximum accumulation was detected at 7 days after agroinfiltration of the leaves. And apparently, IGF1 reached higher levels than CPP-IGF1. The obtained yields are around 8 µg/g for IGF1 and 1.5 µg/g for CPP-IGF1.

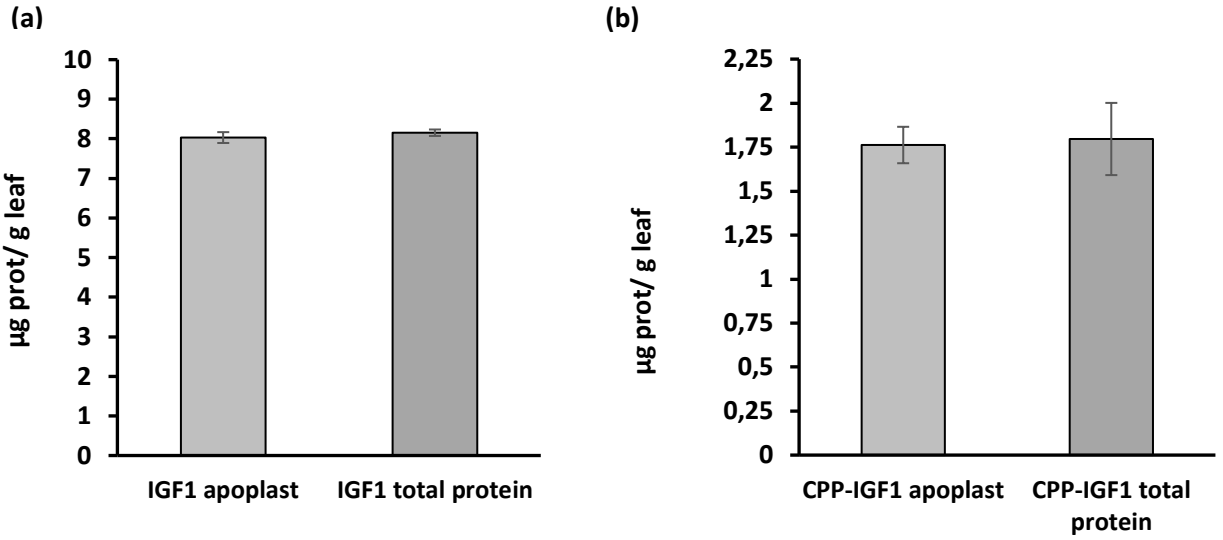


Figure R58. Comparative of IGF1 and CPP-IGF1 production levels in *N. benthamiana* leaves. Quantification of the amount of IGF1 and CPP-IGF1 accumulated in agroinfiltrated leaves as estimated by indirect ELISA analysis in comparison to known amounts of synthetic IGF1. Peptide concentration was determined in total protein extracts **(a)** or ECFs **(b)** from pools of three leaves at 7 dpi**.** Values correspond to the mean of 3 independent leaf pools, and bars to SD.

9. PRODUCTION OF RECOMBINANT PROTEINS IN LIPID DROPLETS OF *N. BENTHAMIANA* PLANTS

Given that the accumulation of oleosin fusion proteins in plant cell cultures seems to impact negatively on cell growth, we addressed the development of biotechnological tools and products based on oleosin technology in *N. benthamiana* plants as an alternative approach.

9.1 Phenotypical characterization of LD-enriched *N. benthamiana* plants

The production of oleosin fusion proteins in *N. benthamiana* leaves is not highly efficient mainly due to the low LD content of leaves. In order to evaluate the full potential of the oleosin technology in leaves, we decided to test transgenic *N. benthamiana* plants engineered for high LD content in leaves (LD+) (El Tahchy *et al.*, 2017). These plants were kindly provided by Dr Petrie, CSIRO, Australia, and they were genetically modified to express the *A. thaliana* genes corresponding to the diacylglycerol O-acyltransferase 1 (*DGAT1*) and WRINKLED1 *(WRI1)*, two key regulators of the TAG accumulation in plants. The constitutive expression of these two genes resulted in an increased accumulation of TAG in the leaves (Vanhercke *et al.*, 2013) and potentially the amount of leaf LDs. The transgenic line was denominated LD+ to indicate that it accumulated LDs.

First of all, we germinated seeds from the *N. benthamiana* LD+ line to amplified them. The plants LD+ grown in parallel to wild type (WT) showed growth reduction phenotype, being shorter than WT (Fig

R59). Growth reduction was statistically significant at 4 weeks, the height was ~21 ± 1.9 cm for WT plants, whereas only ~7 ± 0.6 cm for LD+. These values correspond to a 66% reduction. Then, the LD content of these plant leaves was analyzed under confocal laser microscopy after Nile red staining. As expected, LDs were not abundant on WT leaves but LD+ leaves accumulated a quite high number of LDs, homogenously distributed on the cellular cytoplasm (Fig R60). LDs were visualized sometimes as aggregates or as discrete vesicles in other cells. As a curiosity, a high number of LDs were always visualized on stomata cells.

(a)

(b)

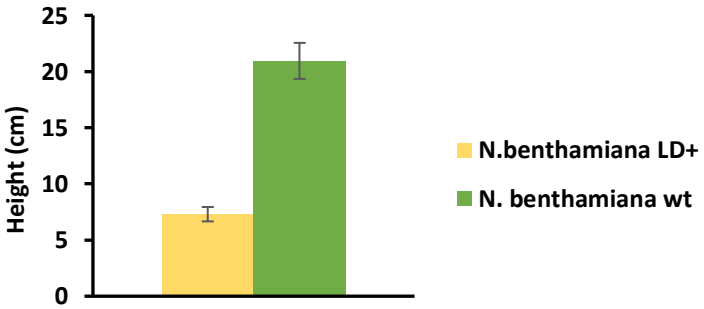


Figure R59. Phenotypical characterization of *N. benthamiana* LD+ plants. **(a)** Image showing representative appearance of 4 weeks *N. benthamiana* plants LD+ or WT. **(b)** Graph presents height mean values (cm) ± SD of WT and LD+ plants after 4 weeks of growth. Values are the mean of 9 plans in 3 independent assays.

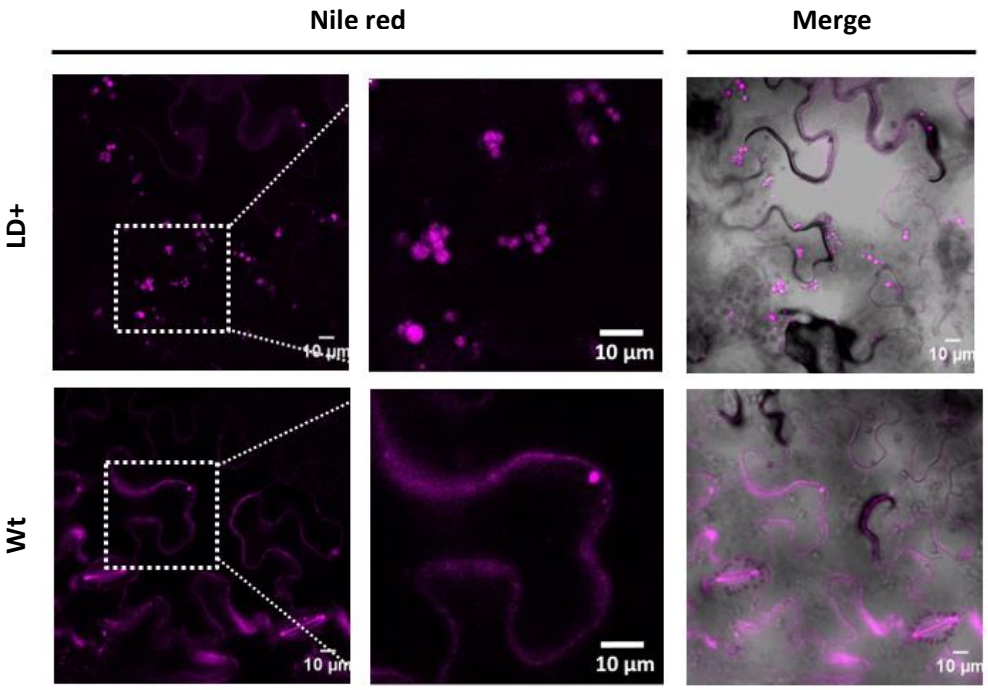


Figure R60. Visualization of LDs in leaves from *N. benthamiana* LD+ plants. Confocal laser microscopy images of epidermal cells of *N. benthamiana* WT and LD+ after Nile Red staining of LDs. White square marked area is amplified in the right images. Single slide confocal images are shown also merged with Bright field images. Scale bars correspond to 10 µm.

9.2 Production of oleosin fusion proteins in *N. benthamiana* LD+

We assessed the production of the chimeric oleosin fusion genes in *N. benthamiana* LD+ plants using transient expression assays. For that, we agroinfiltrated LD+ leaves with the plasmids designed for oleosin fusions to GFP previously described (*pGWB5#p35S:AtOle2-TEVs-GFP* and *pGWB6#p35S:GFP-TEVs-AtOle2* (Fig R15)), in combination with the plasmid for *P1b* expression. We used 9-week old *N. benthamiana* LD+ plants for the assay, the age at with they had reached a similar height and leave surface to the younger *N. benthamiana* WT used in agroinfiltration assays. Leaves were visualized under confocal microscopy at 3, 5 and 7 dpi. We detected the accumulation of both fusion proteins, AtOle2-GFP as well as GFP-AtOle2, in large cytoplasmic aggregates (up to 24 µm of diameter) that were stained with Nile red (Fig R61). These structures looked as they were aggregates of small LDs containing the recombinant proteins since they were surrounded by GFP signal, and of large LDs without GFP signal. These aggregates were visualized in different agroinfiltrated leaves from different plants and different experiments. LD aggregates were also visualized in leaves infiltrated with the mock solution, however as expected, they did not show GFP signal.

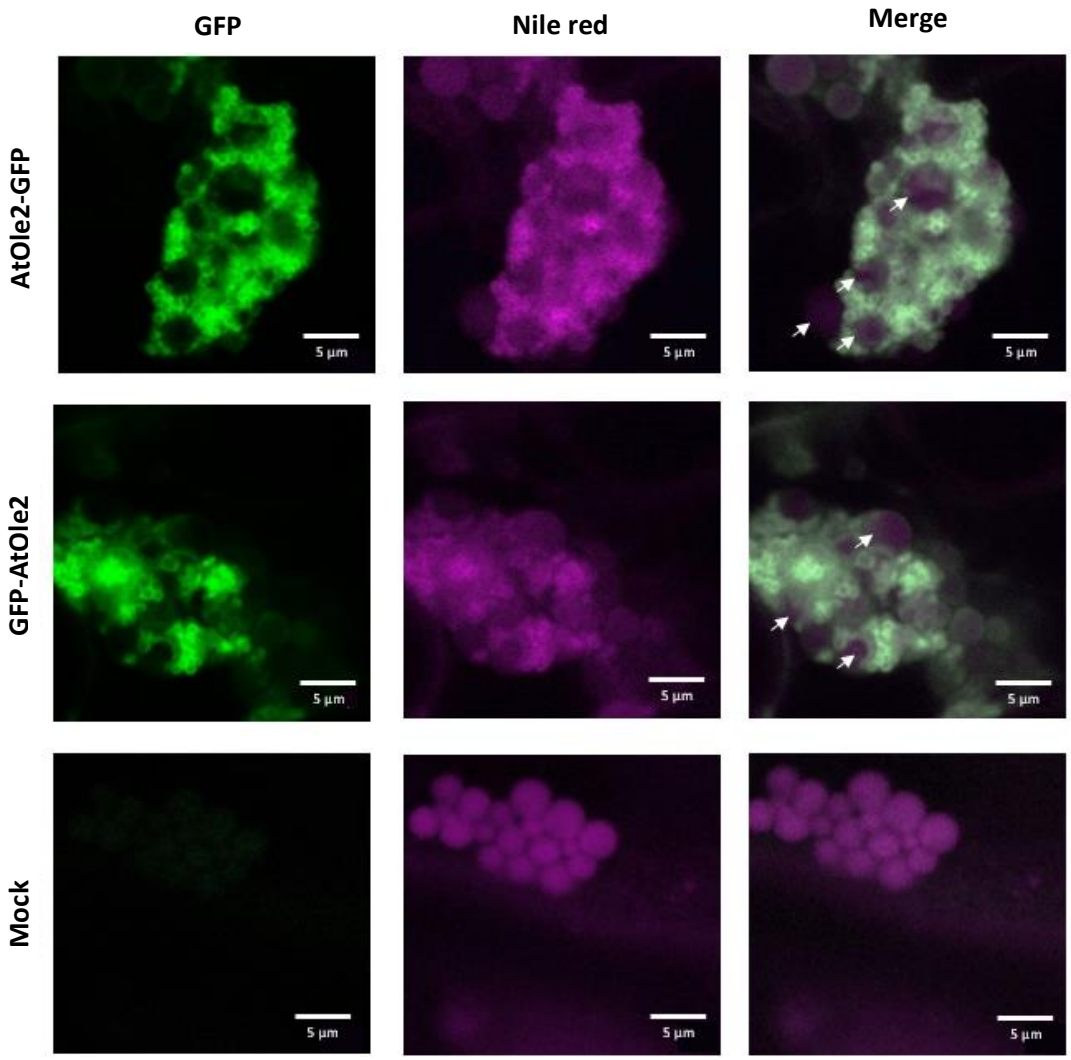


Figure R61. AtOle2-GFP and GFP-AtOle2 proteins are targeted to LDs in *N. benthamiana* LD+ leaves. Confocal fluorescence microscopy of *N. benthamiana* LD+ epidermal cells 3 days after agroinfiltration with either *AtOle2-GFP* or *GFP-AtOle2* gene, in combination with the *P1b* expression cassette, and Nile red stained. Arrows point large LDs. Scale bars correspond to 5 µm.

LD sizes were determined by image analysis (Fig R62), and we found that the GFP labelled LDs showed a similar size, independently whether they contained AtOle2-GFP or GFP-AtOle2 recombinant protein (average diameter of 0.99 ± 0.1 µm for AtOle2-GFP and of 1 ± 0.1 µm for GFP-AtOle2), whereas the non-fluorescent LDs showed similar size in all the analyzed leaves (average diameter of 3.34 ± 0.6 µm, 3.37 ± 0.6 µm and 3.19 ± 0.6 µm for AtOle2-GFP, GFP-AtOle2 and mock leaves, respectively).

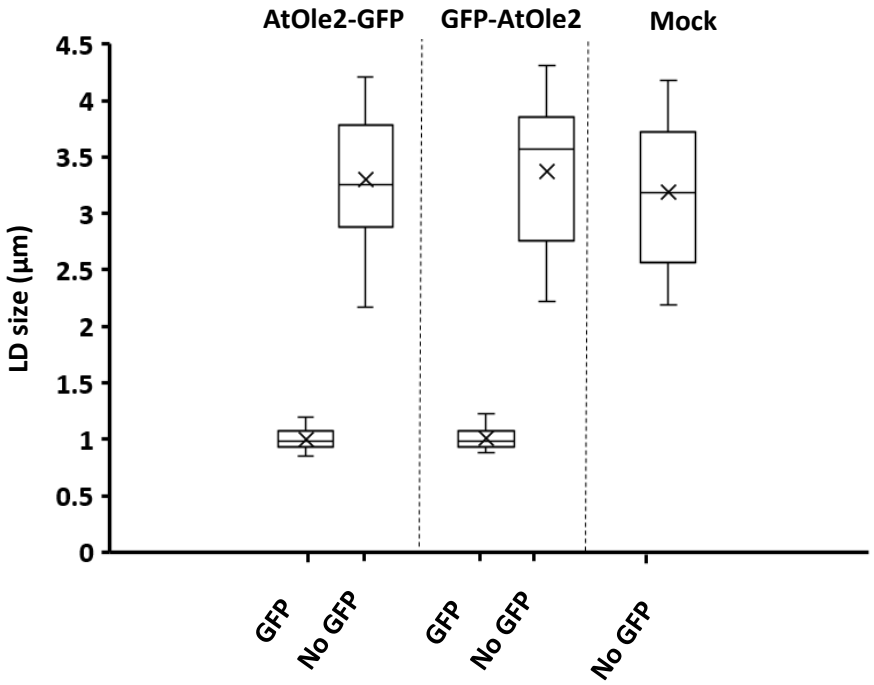


Figure R62. LD size distribution in *N. benthamiana* LD+ plants expressing *AtOle2-GFP* or *GFP-AtOle2*. Boxplot of LD sizes measured in LD+ leaves agroinfiltrated with the indicated gene constructs. As a control, leaves were also infiltrated with mock solution (Agrobacterium induction solution). Values correspond to the average diameters of 10 LDs from three different leaves and three experiments (n=90). LDs were measured individually from leaf aggregates.

Then, LDs were isolated from the infiltrated patches by flotation on a density gradient and visualized under confocal microscopy. The LDs were visualized as spherical vesicles forming large aggregates stained with Nile red. The aggregates were mainly formed by non-fluorescent LDs, while the fusion proteins were visualized only on the periphery of a small portion of the isolated LDs (Fig R63). As expected, no fluorescence was visualized in the isolated LDs from mock infiltrated patches.

Altogether, our results showed that oleosin fusions to the GFP were produced *in N. benthamiana* LD+ leaves and targeted to LDs. The recombinant LDs carrying the N-terminal and C-terminal tagged fusions showed a tendency to aggregate. The aggregates were formed mainly by small LDs (around 1 µm) showing fluorescence on their periphery (carrying the GFP fusions) together with non-fluorescent large LDs (> 1 µm). When isolated, most of LDs were non-fluorescent whereas only a small proportion of LDs showed fluorescence. This rate was significantly different to the *in vivo* observed leaves, suggesting that an enrichment of large LDs was favored during LD fractionation.

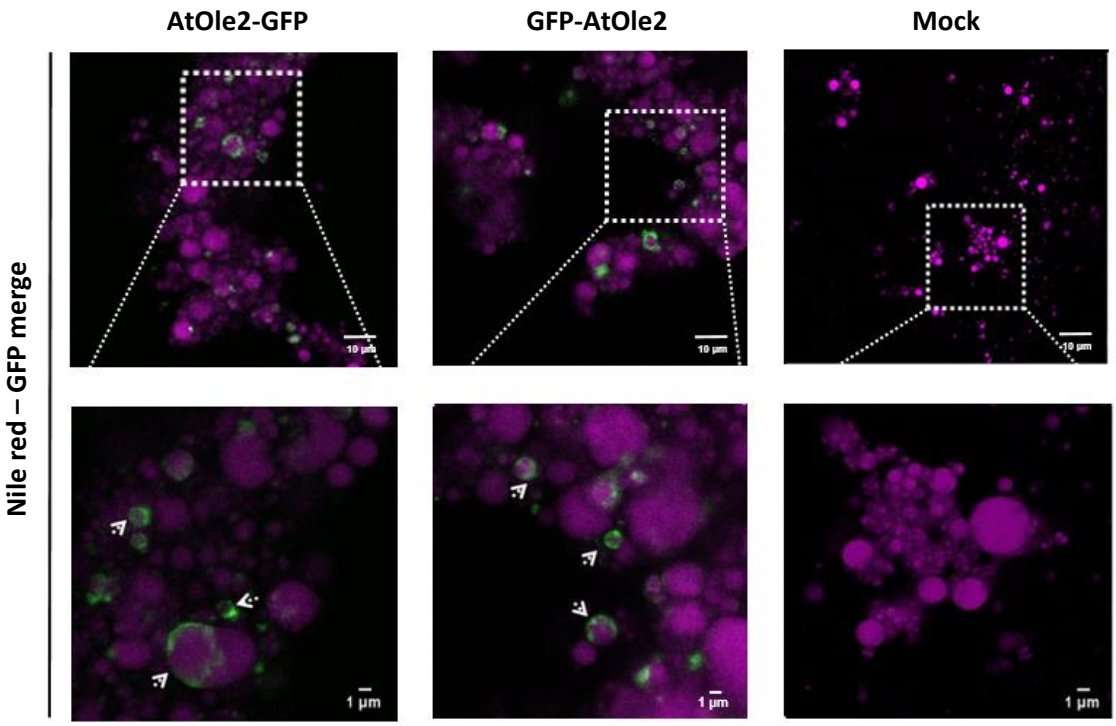


Figure R65. Isolated LDs from *N. benthamiana* LD+ leaves agroinfiltrated with *AtOle2-GFP* or *GFP-AtOle2* expression constructs. Images are merged from single scanned confocal microscopy slides for GFP and Nile Red signals from isolated Nile-red stained LDs. LDs were isolated from LD+ leaves at 5 dpi. White square area is amplified in the below images. Arrows point fluorescent LDs. Scale bars correspond to 10 µm and 1 µm as indicated.

9.4 Transient expression of lipid metabolism modulators in *N. benthamiana*

According to our results with LD+ plants, the concomitant formation of LDs with the accumulation of oleosin proteins might improve the accumulation of the recombinant oleosin fusion proteins. For that, we explored the transient expression of the *AtDGAT1* and *AtWRI1* genes, the same genes constitutively expressed in *N. benthamiana* LD+ plants, in combination with the gene constructs for oleosin fusion proteins. Two new gene constructs for the expression of the *DGAT1* and *WRI1* TAGs inductors were designed (Fig R64). We also prepared an additional construct to produce the single AtOle2 protein as a control. To prepare the constructs, we obtained the cDNAs 6530639541 and 3510696031 from "The Arabidopsis Biological Resource center (ABRC)" library in the pEntry plasmid (D-TOPO) and we moved them to the destination plasmid pGWB2 using the Gateway™ cloning system. The pGWB2 plasmid contains the sequences to drive the expression of the gene of interest by the 35S promoter and the Nos terminator.

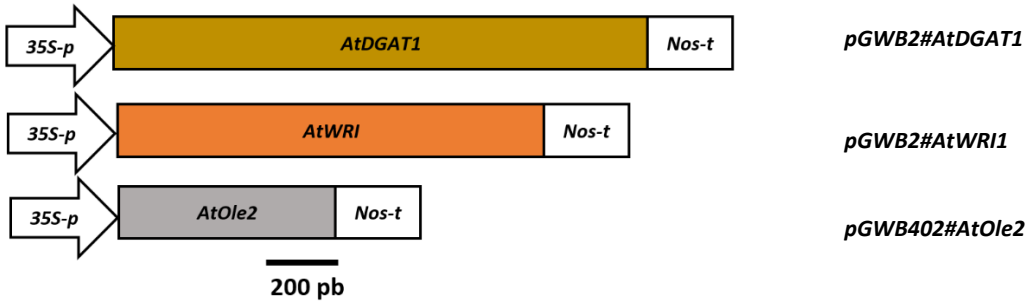


Figure R64. *A. thaliana DGAT1, WRI1* and *Ole2* gene constructs. Schematic diagram of the constructs for the expression of the *AtDGAT1*, *AtWRI1* and *AtOle2* genes under the control of the 35S promoter and the Nos terminator

Using these plasmids, the corresponding genes were transiently expressed in *N. benthamiana* leaves in combination with the P1b RNA silencing suppressor. Agroinfiltrated leaves were then visualized under confocal microscopy (Fig R65). We observed an increase in the number of LDs in the leaves expressing individually the *DGAT1*, *WRI1* and *AtOle2* genes, in comparison to the control leaves infiltrated with a mock solution. The number of LDs was further increased when the three *DGAT1*, *WRI1* and *AtOle2* genes were co-expressed. LDs were visualized as discrete spherical vesicles.

Then, the size of LDs was determined, showing a statistically significant increase on the size when WRI1 was expressed (Tukey test, $p \leq 0.05$) (Fig R66). The average diameter of the LDs was 2.44 ± 0.69 µm, 0.19 ± 0.07 µm, 0.32 ± 0.07 µm, 1.90 ± 0.71 µm and 1.62 ± 0.56 µm for WRI1, DGAT1, AtOle2, WRI1+DGAT1, and WRI1+DGAT1+AtOle2, respectively.

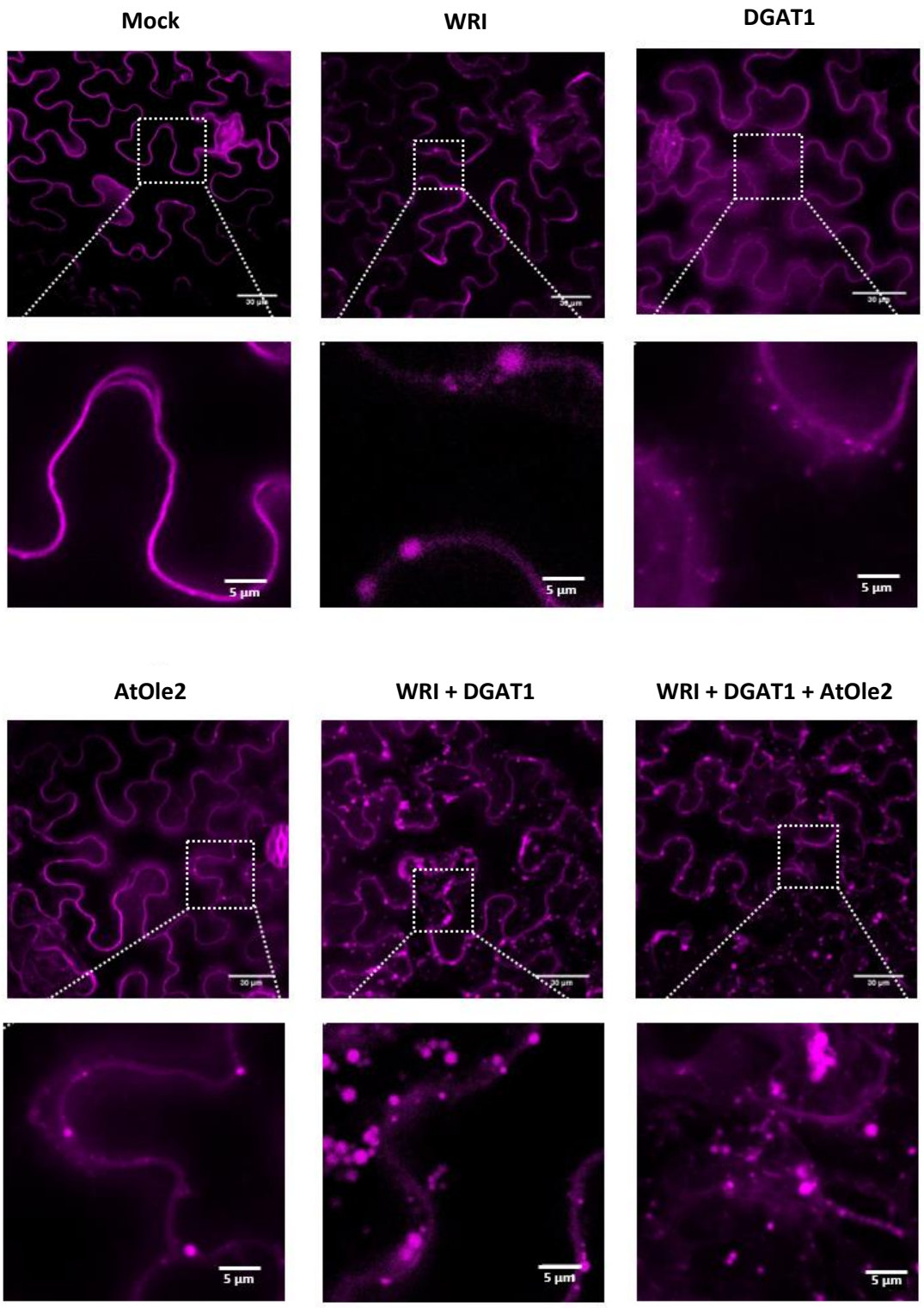


Figure R65. LD accumulation in *N. benthamiana* leaves transiently expressing *AtDGAT1*, *AtWRI1* and *AtOle2* genes. Representative confocal laser microscopy images of epidermal cells Nile red stained and transiently expressing the indicated genes in combination with the P1b silencing suppressor. Scale bars correspond to 30 and 5 µm as indicated.

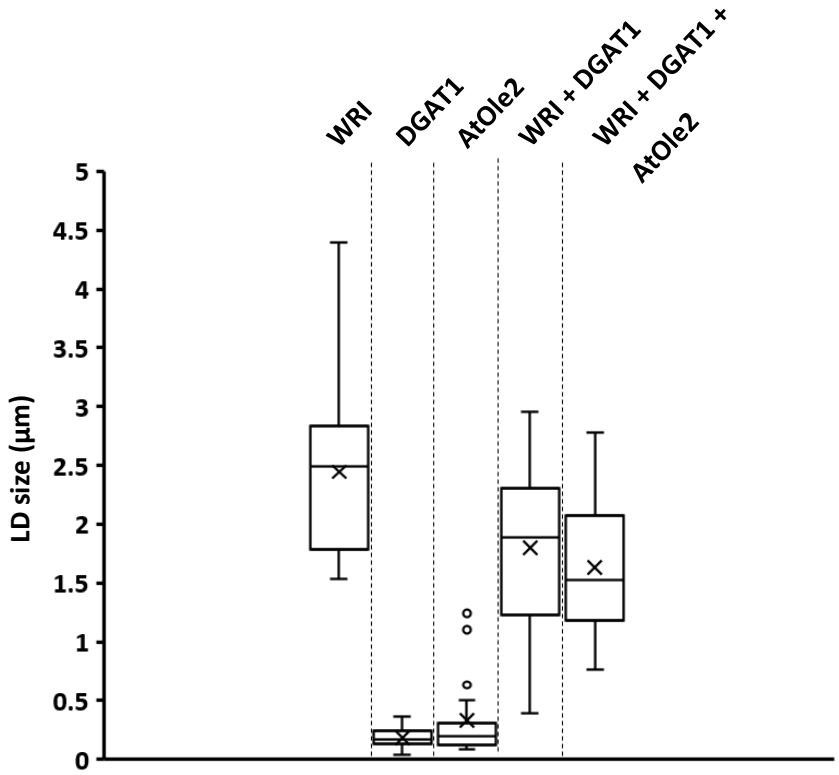


Figure R66. Increased size of LDs of *N. benthamiana* leaves infiltrated with *AtWRI1* gene. Boxplot of size diameter of LDs in expressing the indicated genes in combination with P1b. Values were determined for 10 LDs from three independent leaves and three different experiments (n=90).

Next, *AtOle-GFP* and *GFP-AtOle* constructs were transiently expressed in *N. benthamiana* leaves in combination with both TAG inductors. Based on our previous results, we hypothesized that the expression of the three components simultaneously would trigger an accumulation of recombinant LDs containing the oleosin fusion proteins, and therefore a high accumulation of the recombinant proteins. Given that the TAG accumulation will occur concomitant with the oleosin fusion production, most likely all the novo LDs should carry the oleosin recombinant proteins. Our observations of leaves co-expressing the *DGAT1*, *WRI1* and the oleosin fusions indeed showed a clear induction on LD formation (Fig R67). The LDs were visualized isolated or in small aggregates, being the isolated LDs more abundant in leaves accumulating the GFP-AtOle2 protein and the LD aggregates in the leaves with AtOle2-GFP. Most of the LDs were fluorescent, indicating that most of them carried the GFP protein and were recombinant LDs.

Then, LDs were purified from agroinfiltrated leaves by flotation-centrifugation on density gradient and visualized under confocal microcopy. Isolated LDs from leaves co-agroinfiltrated with *DGAT1*, *WRI1* and *AtOle2*, were visualized as discrete vesicles stained with Nile red without GFP fluorescence. Whereas LDs from leaves co-agroinfiltrated with *DGAT1*, *WRI1* and oleosin fusions, were visualized

also as Nile Red stained spherical vesicles exhibiting GFP fluorescence on their surface. Only few small LDs did not show fluorescence and aggregated with the GFP LDs (Fig R68).

Our results confirmed that the transient expression of TAG inductors (*WRI1* and *DGAT1*) in combination with oleosin fusions triggered the accumulation of recombinant LDs, containing the recombinant oleosin fusion proteins that can be easily purified by flotation on density gradients.

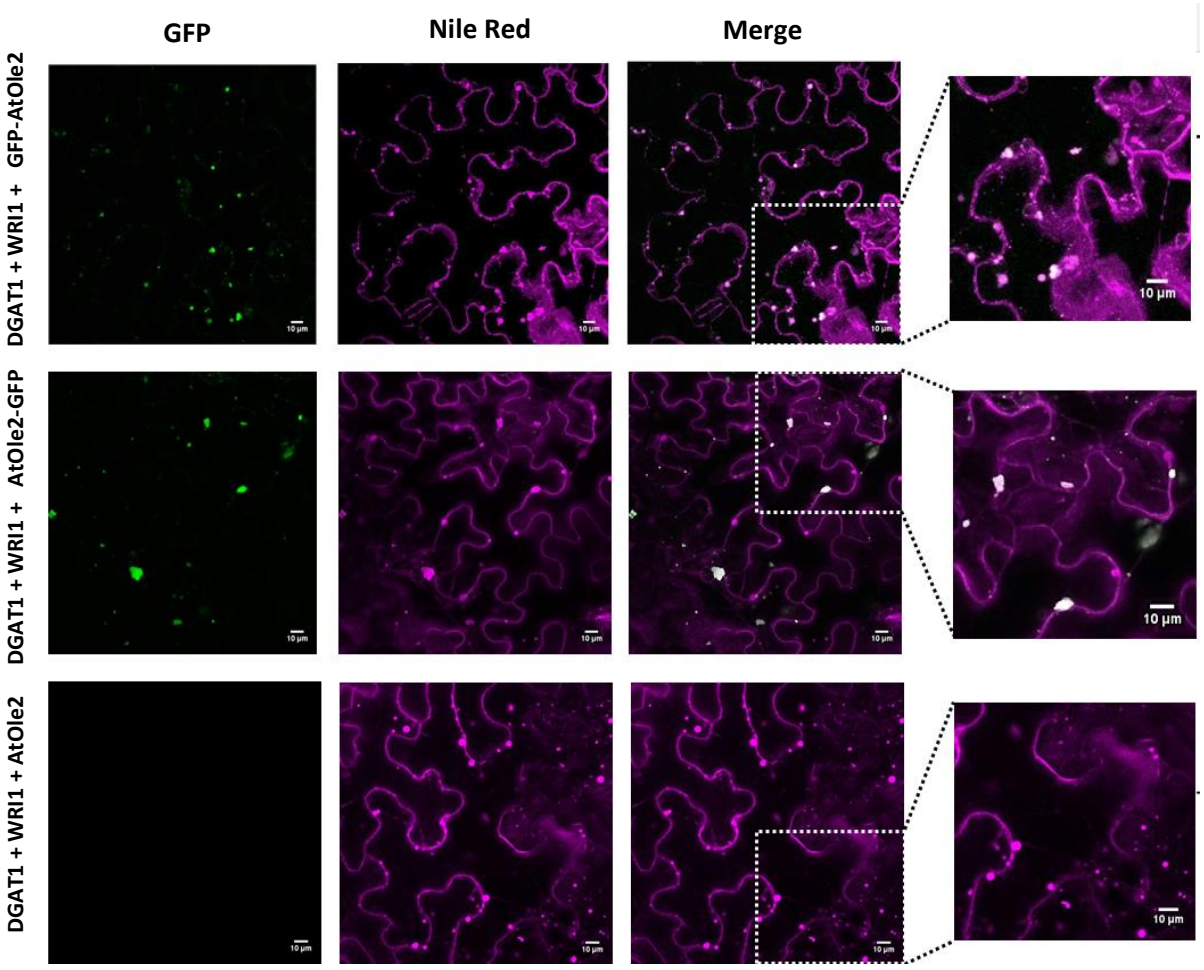


Figure R67. Transient induction of LDs and expression of oleosin fusion proteins. Confocal laser microscopy images of epidermal cells accumulating LDs at 5 dpi when transiently expressing the TAG inductors (*AtDGAT1 and AtWRI1*) in combination with *GFP-AtOle2*, *AtOle2-GFP* or *AtOle2*, as indicated. The cassette for expression of *P1b* was co-agroinfiltrated in all the cases. Leaves were stained with the LD marker, Nile red. Images shown GFP signal, Nile Red signal and the merged images. White squares are amplified in the contiguous images. Scale bars correspond to 10 µm.

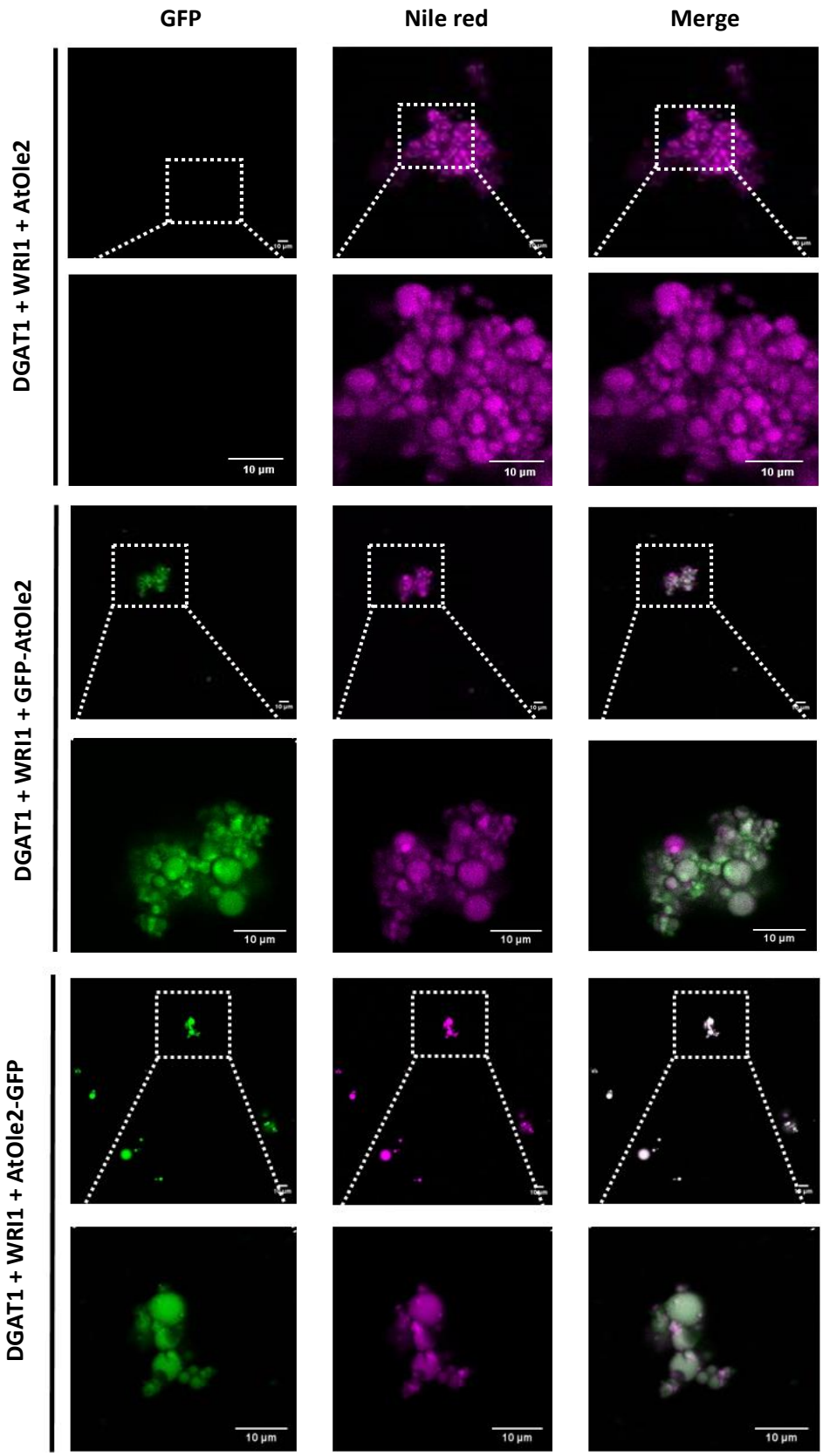
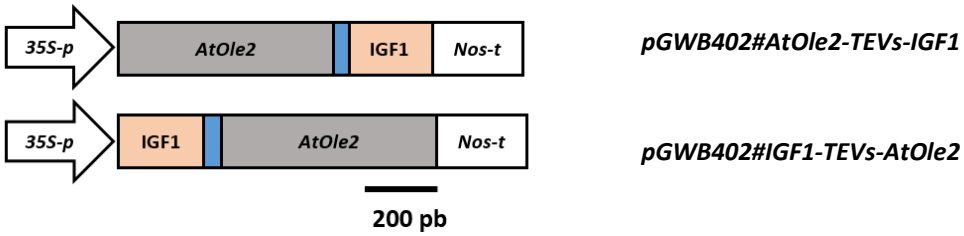


Figure R68. Visualization of fractionated LDs from leaves expressing TAG inductors and oleosin fusion proteins. Images correspond to single scanned confocal microscopy slides of fractionated and Nile red stained LDs expressing the indicated proteins. Images show the GFP signal, Nile Red signal and merged signals. White squares are amplified in the contiguous images. Scale bars correspond to 10 µm.

9.5 Production of IGF1 as an oleosin fusion protein in *N. benthamiana* LDs

In order to produce the IGF1 as an oleosin fusion protein targeted to LDs, we designed two different gene constructs for the production of a N- or C- terminal fusion of the AtOle2 protein (Fig R69). In both cases the peptide and the protein were linked through a TEV protease cleavage site. The expression of oleosin fusion genes was under the control of 35S promoter and the Nos terminator. The Gateway™ cloning system was then used for preparing the constructs using the pGWB402 plasmid was used as destination vector.

(a)



(b) ***AtOle2-IGF1***

ATGGCGGATACACACCGTGTCGACCGTACTGATAGACACTTTCAATTTCAGTCGCCCTATGAAGGCGGCCGAGGTCAAGGTCAGTATGAAGGT
GACCGTGGTTACGGTGGTGGCGGTTACAAGAGCATGATGCCTGAAAGTGGCCCATCTAGTACCCAAGTATTGTCCCTGTTGATTGGAGTCCCT
GTCGTCGGTTCGCTACTTGCCTTGGCTGGATTACTTCTAGCTGGTTCGGTGATCGGCTTAATGGTTGCTTTACCACTATTTCTCCTCTTCAGCCC
GGTTATAGTCCCAGCGGCTCTAACTATCGGGCTTGCAATGACAGGCTTTTTAGCCTCGGGGATGTTCGGTCTAACCGGGCTTAGCTCAATCTCA
TGGGTCATGAACTATCTTCGTGGGACAAGGAGAACTGTGCCTGAGCAATTGGAGTATGCTAAGAGGAGAATGGCTGATGCGGTTGGCTACGC
AGGACAAAAGGGCAAAGAAATGGGCCAGCATGTGCAGAACAAGGCCCAAGATGTTAAACAATATGATATTTCTAAGCCACATGACACTACCA
CTAAGGGTCATGAGACTCAGGGGAGGACGACGGCTGCAGAGAATCTTTATTTTCAGAGTGGGCCTGAGACTCTCTGTGGCGCCGAATTAGT
GGACGCATTGCAATTCGTTTGTGGCGATAGGGGGTTCTATTTCAACAAGCCTACTGGGTACGGTTCTTCTTCAAGAAGAGCTCCTCAAACTGG
CATTGTGGATGAATGTTGTTTCAGATCTTGTGATCTCAGGAGACTCGAAATGTACTGTGCTCCCTTAAAACCTGCCAAGTCTGCATGA

AtOle-IGF1 fusion protein

MADTHRVDRTDRHFQFQSPYEGGRGQGQYEGDRGYGGGGYKSMMPESGPSSTQVLSLLIGVPVVGSLLALAGLLLAGSVIGLMVALPLFLLFSP
VIVPAALTIGLAMTGFLASGMFGLTGLSSISWVMNYLRGTRRTVPEQLEYAKRRMADAVGYAGQKGKEMGQHVQNKAQDVKQYDISKPHDTTT
KGHETQGRTTAAENLYFQSGPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPAKSA*

IGF1-AtOle2

ATGGGGCCTGAGACTCTCTGTGGCGCCGAATTAGTGGACGCATTGCAATTCGTTTGTGGCGATAGGGGGTTCTATTTCAACAAGCCTACTGGG
TACGGTTCTTCTTCAAGAAGAGCTCCTCAAACTGGCATTGTGGATGAATGTTGTTTCAGATCTTGTGATCTCAGGAGACTCGAAATGTACTGTG
CTCCCTTAAAACCTGCCAAGTCTGCATGAGAGAATCTTTATTTTCAGAGTGCGGATACACACCGTGTCGACCGTACTGATAGACACTTTCAATTT
CAGTCGCCCTATGAAGGCGGCCGAGGTCAAGGTCAGTATGAAGGTGACCGTGGTTACGGTGGTGGCGGTTACAAGAGCATGATGCCTGAAA
GTGGCCCATCTAGTACCCAAGTATTGTCCCTGTTGATTGGAGTCCCTGTCGTCGGTTCGCTACTTGCCTTGGCTGGATTACTTCTAGCTGGTTCG
GTGATCGGCTTAATGGTTGCTTTACCACTATTTCTCCTCTTCAGCCCGGTTATAGTCCCAGCGGCTCTAACTATCGGGCTTGCAATGACAGGCTT
TTTAGCCTCGGGGATGTTCGGTCTAACCGGGCTTAGCTCAATCTCATGGGTCATGAACTATCTTCGTGGGACAAGGAGAACTGTGCCTGAGCA
ATTGGAGTATGCTAAGAGGAGAATGGCTGATGCGGTTGGCTACGCAGGACAAAAGGGCAAAGAAATGGGCCAGCATGTGCAGAACAAGGC
CCAAGATGTTAAACAATATGATATTTCTAAGCCACATGACACTACCACTAAGGGTCATGAGACTCAGGGGAGGACGACGGCTGCATGA

IGF1-AtOle2 fusion protein

MGPETLCGAELVDALQFVCGDRGFYFNKPTGYGSSSRRAPQTGIVDECCFRSCDLRRLEMYCAPLKPAKSAENLYFQSMADTHRVDRTDRHFQF
QSPYEGGRGQGQYEGDRGYGGGGYKSMMPESGPSSTQVLSLLIGVPVVGSLLALAGLLLAGSVIGLMVALPLFLLFSPVIVPAALTIGLAMTGFLAS
GMFGLTGLSSISWVMNYLRGTRRTVPEQLEYAKRRMADAVGYAGQKGKEMGQHVQNKAQDVKQYDISKPHDTTTKGHETQGRTTAA*

Figure R69. Gene constructs for the production of the AtOle2-IGF1 and IGF1-AtOle2 fusion proteins. **(a)** Schematic diagram of the constructs for the expression of the oleosin fusion genes under the control of the 35S promoter and the Nos terminator. **(b)** DNA and amino acid sequence of the fusion proteins AtOle2-IGF1 and IGF1-AtOle2 where IGF1 peptide is shown in orange, the AtOle2 protein in gray and the TEV protease recognition site in blue. The triangle indicates the TEV protease cleavage site.

The constructs were co-agroinfiltrated in combination with the TAG inductors in *N. benthamiana* leaves. Leaves were visualized under confocal microscopy at 5 dpi and representative images are shown in Fig R70. In all cases, the *N. benthamiana* cells showed a high content of Nile red stained LDs. These LDs were viewable along all the days of the sampling period (3, 5, 7 and 9 dpi). The presence of the oleosin fusion protein in the LDs cannot be monitored by confocal fluorescence microscopy because no fluorescent traceable signal of GFP was present. The size of LDs was measured by image analysis. No significant statistically differences in size were found among the LDs from leaves expressing *AtOle2*, *IGF1-AtOle2* and *AtOle2-IGF1* (Tukey test, $p \leq 0.05$) (Fig R71). The average LD diameters were 2.5± 0.9 µm for AtOle2, 2.01 ± 0.77 µm for AtOle2-IGF1, and 2.49 ± 1.0 µm for IGF1-AtOle2 leaves.

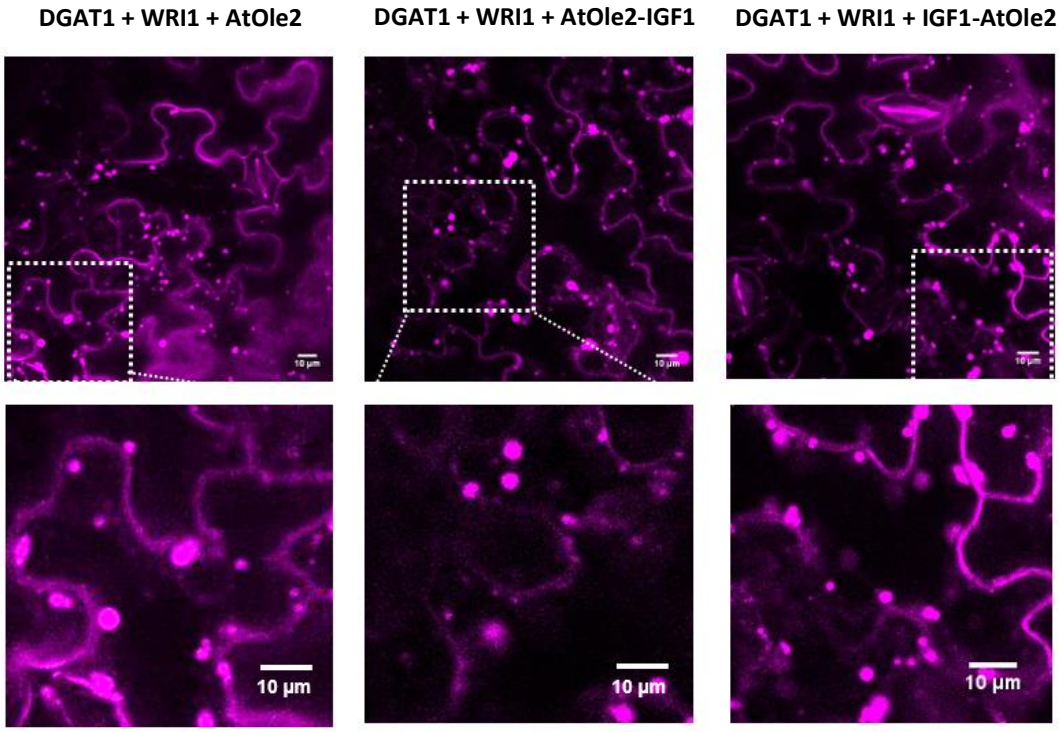


Figure R70. Transient expression of *AtOle2*, *AtOle2-IGF1* and *IGF1-AtOle2* genes in *N. benthamiana* plants. Confocal laser microscopy images of Nile-red stained epidermal cells expressing the TAG inductors, *AtDGAT1* and *AtWRI1*, in combination with *AtOle2*, *AtOle2-IGF1* or *IGF1-AtOle2*. White squares are amplified in the below images. Scale bars correspond to 10 µm as indicated.

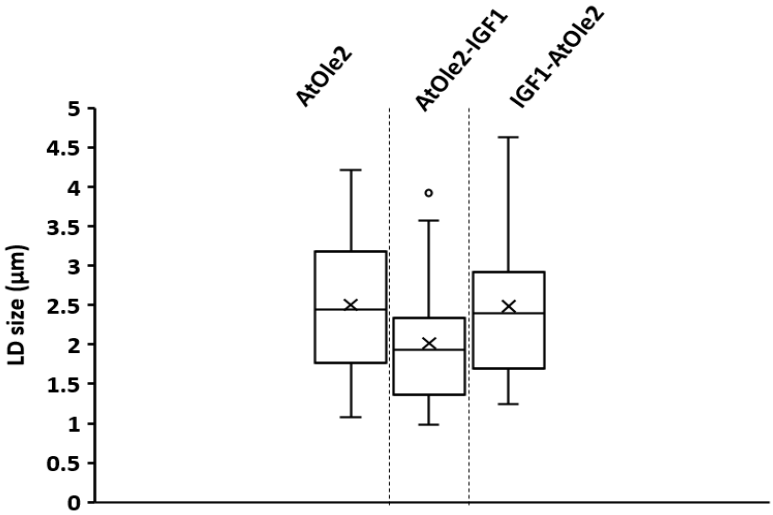


Figure R71. Size of LDs from *N. benthamiana* leaves co-expressing TAG inductors and *IGF1* oleosin fusion proteins. Boxplot of LD diameters from leaves co-agroinfiltrated with indicated constructs. Values correspond to 10 LDs measured from three independent leaves and three experiments (n=90). No significant differences were found (Tukey test, $p \leq 0.05$) in all the comparisons.

The accumulation of the recombinant proteins was assessed by Western blot analysis of LDs proteins. For that LDs from leaves agroinfiltrated with the different constructs were isolated by flotation-centrifugation in sucrose density gradients, their proteins separated by SDS-PAGE, transferred to nitrocellulose membrane and immunodetected using anti-IGF1 antibodies (Fig R72). We detected specific bands with the same expected apparent molecular weight in the samples expressing the *IGF1* as fusion to *AtOle2* (*AtOle2-IGF1* and *IGF1-AtOle2*), which were absent in the other samples. The estimated size for this band was 30 kDa in agreement with the expected size for the fusion proteins (7.7 kDa IGF1 + 0.9 kDa TEVs + 22 kDa AtOle2). These results indicated that both fusion proteins were produced and targeted to LDs. Moreover, the intensity of the band in IGF1-AtOle2 was stronger.

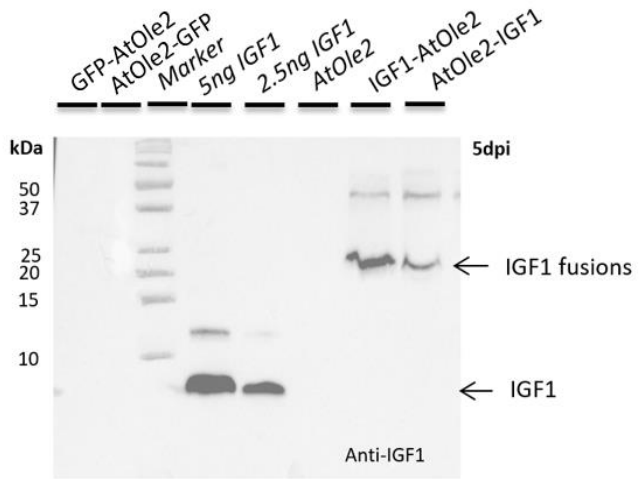


Figure R72. Western blot analysis of AtOle2 fusions to IGF1. LD proteins from *N. benthamiana* leaves infiltrated with indicated constructs in combination with *AtWRI*, *AtDGAT1* and *P1b* were separated by SDS-PAGE, transferred to nitrocellulose membrane and immunodetected with anti-IGF1.LD fractions were obtained from

leaves at 5 dpi. Synthetic IGF1 (5 or 2.5 ng) was run in parallel as positive control. Molecular weight markers are shown in kDa on the left. Arrows point synthetic IGF1 and oleosin-IGF1 fusions.

To determine the amount of IGF1 produced as a fusion protein, we used an indirect ELISA assays with LD fractions. ELISA results showed that IGF1 was first detected at 3 dpi, and accumulation levels increased gradually over the next days reaching a maximum at 5 dpi, decreasing then progressively until 9 dpi. The maximum accumulation levels correspond to ~0.52 ± 0.09 µg/g fresh weight of leaf for AtOle2-IGF1, and ~1.07 ± 0.26 µg/g fresh weight of leaf for IGF1-AtOle2 in extracts (Fig R73).

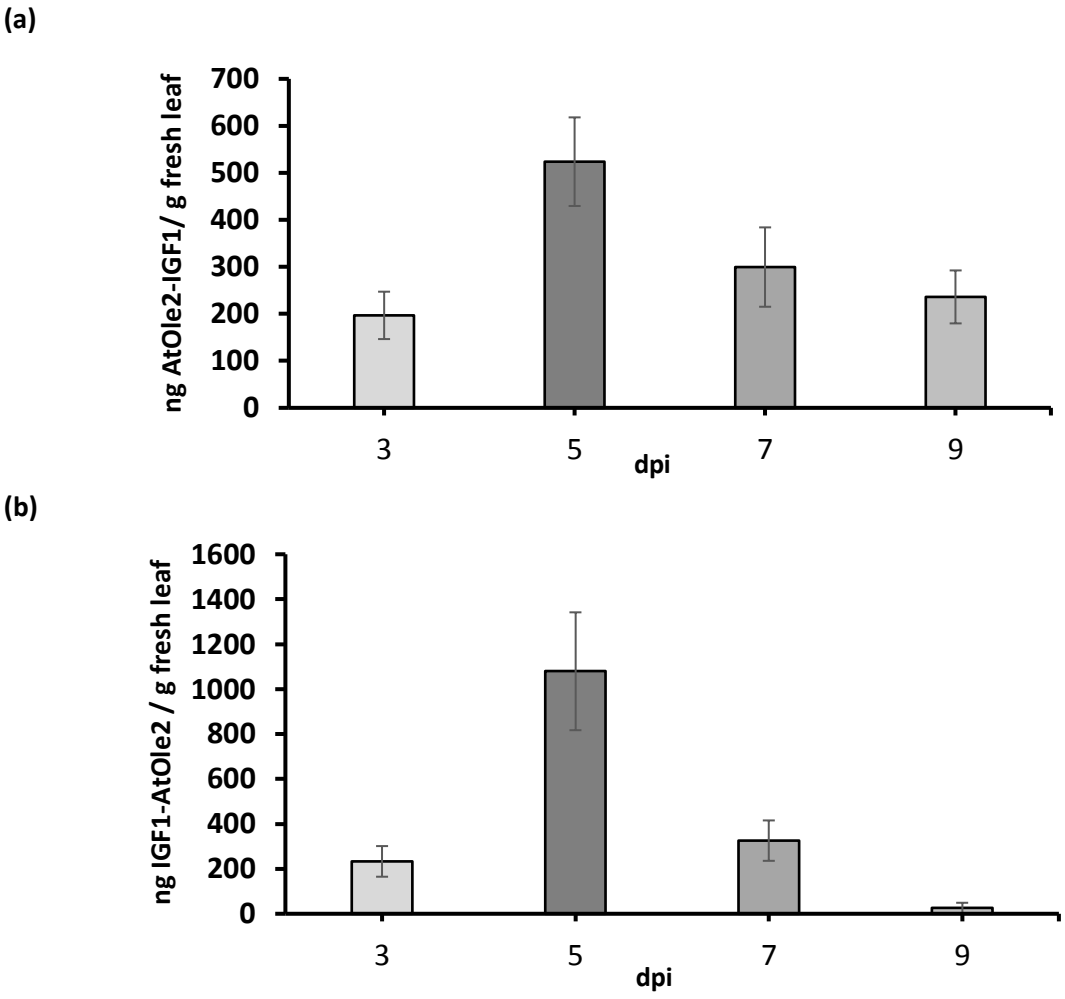


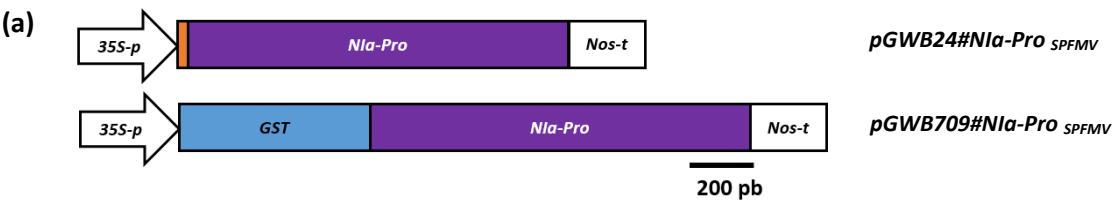
Figure R73. Accumulation of IGF1 as oleosin fusion protein in transient expression assays in *N. benthamiana* plants. Accumulation was estimated by indirect ELISA analysis of LDs containing AtOle2-IGF1 **(a)** and IGF1-AtOle2 **(b),** in comparison with known amount of synthetic IGF1 peptide. Graphs correspond to the mean ± SD of 3 independent estimations from pools of three leaves.

Altogether, our results showed that the transient expression of *WRI1* and *DGAT1* genes with oleosin fusion proteins leads to the formation of LDs carrying the recombinant fusion proteins. This strategy

will allow the use of the oleosin technology in leaves. The maximum oleosin fusions accumulation was detected at 5 days after agroinfiltration of the leaves and, apparently, the fusion of IGF1 in the N-terminal of oleosin reached higher levels than the fusion in the C-terminal. The obtained yields were around 1 µg/g for IGF1-AtOle2 and 0.5 µg/g for AtOle2-IGF1.

9.6 Development of an oleosin fusion based biotechnological tool for in vivo evaluation of protease activity

Although the *N. benthamiana* LD+ plants were not used for the production of recombinant LDs containing oleosin fusions to IGF1, we took advantage of them in combination with the oleosin technology, to develop a biotechnological tool for evaluation of protease activity in transient expression assays. Two different viral proteases were assessed with our system, namely the well-known protease from the potyvirus *Tobacco etch virus* (TEV), and the Nla-Pro from another potyvirus *Sweet Potato Feathery Mottle Virus* (*SPFMV*) as a proof of concept. The system was designed to allow a simple screening of the activity of proteases by *in vivo* visualization and molecular analysis of the transient expression of reporter constructs. These reporter constructs consisted on an expression cassette with a fusion between oleosin and GFP protein, linked through the specific protease cleavage sites (heptapeptide) for each viral protease. The selection of the cleavage consensus sequence was based in the most frequent heptapeptides listed for each particular virus in a dedicated database (http://www.dpvweb.net/potycleavage/species.html). Several gene constructs, including the reporter constructs, were designed and prepared in plasmids for transient expression in *N. benthamiana* (Fig R74). The SPFMV NIa-pro viral product was cloned in two different expression cassettes; one to be produced as a glutathione S-transferase (GST) fusion protein, and the other as a His-tagged protein, both of them containing the 35S promoter and Nos terminator regulatory sequences. Also, the reported construct for the expression of *AtOle2* protein linked to the GFP through the NIa-Pro cleavage sequence (SPFMVs). As a negative control we used the plasmid *pGWB5#AtOle2-TEVs-GFP* described in section 4.



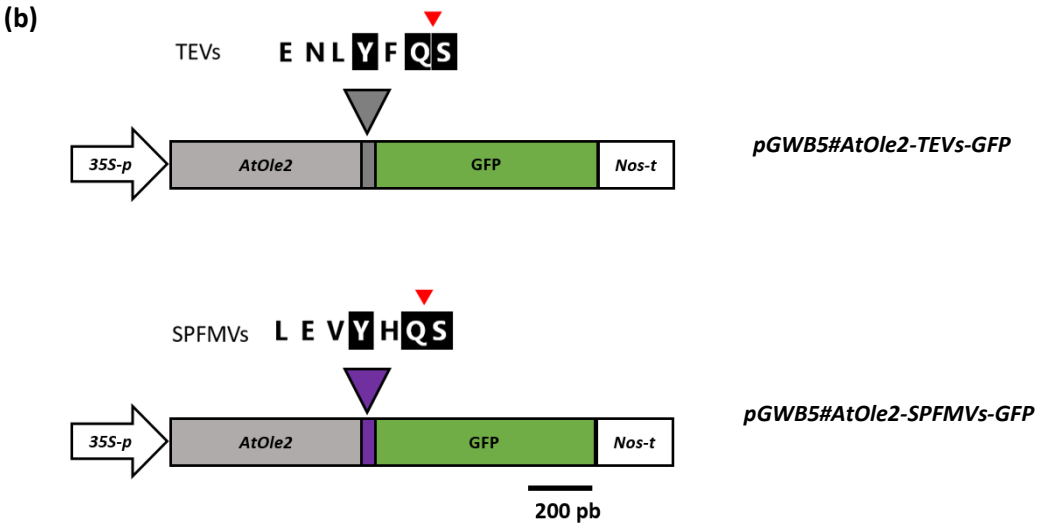
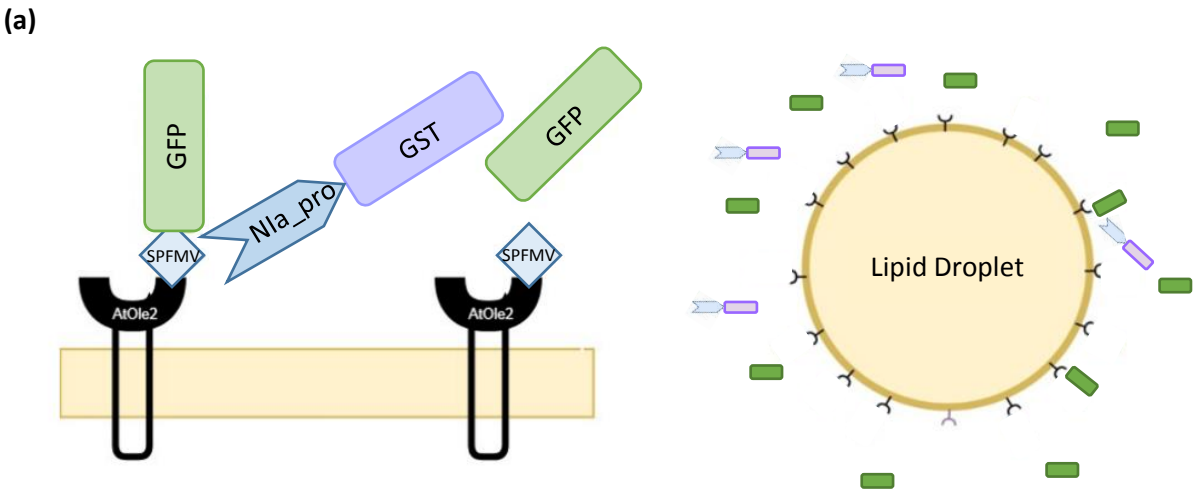


Figure R74. Constructs for the evaluation of the protease activity of Nla-Pro from SPFMV. Map of the gene constructs for the expression of SPFMV *Nla-Pro protease* (purple boxes) as a His- (orange box) or a GST- (blue box) tagged protein, and the AtOle2-GFP fusion proteins linked through the TEV (TEVs, orange box) or SPFMV (SPFMVs, purple box) protease cleavage sites. All of them were driven by the 35S promoter (35S-p) and Nos terminator (Nost-t) sequences.

Based on previous results, our reporter constructs were expected to produce fluorescently labelled LDs when transiently expressed in *N. benthamiana* alone. We hypothesized that the co-agroinfiltration of the reporter construct with the appropriated protease would result in the cytoplasmic re-localization of the GFP fluorescence, a change that can be easily monitored by confocal microscopy (Fig R75a). However, when the reporter construct is co-agroinfiltrated with a non-active protease, or in absence of the appropriate and specific cleavage site, the GFP signal would remain associated to LDs (Fig R75b).



(b)

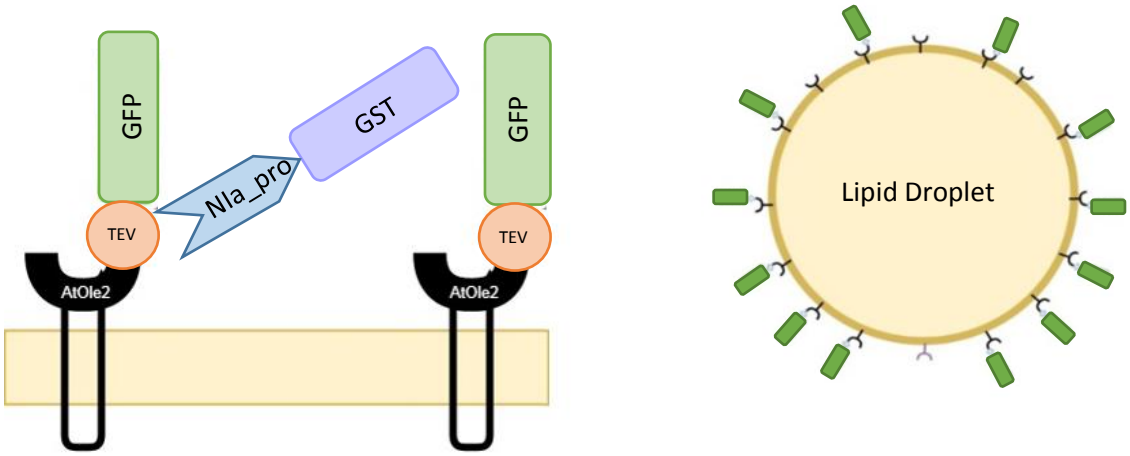


Figure R75. Model of the oleosin-based biotechnological tool for *in vivo* evaluation of protease activity. Schematic representation of the AtOle2-TEVs-GFP and AtOle2-SPFMVs-GFP fusion proteins localized in the periphery of LDs. The GFP release from LDs under the presence of the SPFMV Nla_Pro when the fusion protein contained the target cleavage site **(a),** or the LD associated GFP when the fusion protein did not contain the target cleavage site **(b)** are illustrated.

Transient expression assays using the prepared gene constructs validated our working hypothesis, as shown in Fig R76-R77. We could observe that both fusion proteins (AtOle2-TEVs-GFP and AtOle2-SPFMVs-GFP) associated to LDs in *N. benthamiana* LD+ leaves when transiently produced (Fig 9.9-10). However, in co-infiltrations, the co-expression of *AtOle2-SPFMVs-GFP* with the *Nla-Pro* protease, independently if tagged with His or GST, showed a cytoplasmic localization of GFP signal (Fig R76). These observations suggested that the protease was acting on the appropriately cleavage site releasing the GFP from the fusion protein that was re-distributed homogenously in the cellular cytoplasm. On the other hand, the co-expression of *AtOle2-TEVs-GFP* with the *Nla-Pro* protease of the heterologous virus SPFMV, independently if tagged with His or GST, showed LD-associated GFP signals (Fig R77). These observations confirmed that our system allow the evaluation of the protease activity, as well as the specificity of cleavage sites, in *in vivo* assays by simple monitoring GFP localization under confocal microscopy.

Those leaf patches co-agroinfiltrated with the reporter constructs together with proteases were collected for western-blot analysis to molecularly confirm the visual results. Patches from different leaves and plants were pooled, grinded and total protein extracted. The SPFMV NIa-pro was immunodetected at 3 dpi using specific antibodies against the GST tag, either in single agroinfiltrated patches (GST-NIa-pro), or in co-agroinfiltrations with oleosin fusions (AtOle2-SPFMVs-GFP + GST-NIa-pro) (Fig R78a). Surprisingly, the accumulation of the GST-Nla-pro was only detected at short times, disappearing apparently nearly at 5 dpi and 7dpi. When revealing the membrane with anti-GFP antibodies, bands with an apparent molecular size of the recombinant oleosin fusion protein (around

50 kDa for AtOle2-SPFMWs-GFP) were detected in both single and co-agroinfiltrations (Fig R78b). In those patches in which the protease was expressed in combination with the fusion proteins, a band corresponding to the mobility of the free GFP was also detected (Fig R78b). These results suggested that the fusion protein was digested by the protease and the GFP released, although still some partially digested fusion protein could be also detected. However, in those samples from the fusions carrying the TEV protease cleavage site in combination with the SPFMV NIa-pro, only the bands corresponding to fusion proteins were detected (Fig R78c). Therefore, the molecular analyses confirmed the results of the confocal microscopy visualization, demonstrating the effectiveness of our system to evaluate protease activities *in vivo* by simple visualization under confocal microscopy.

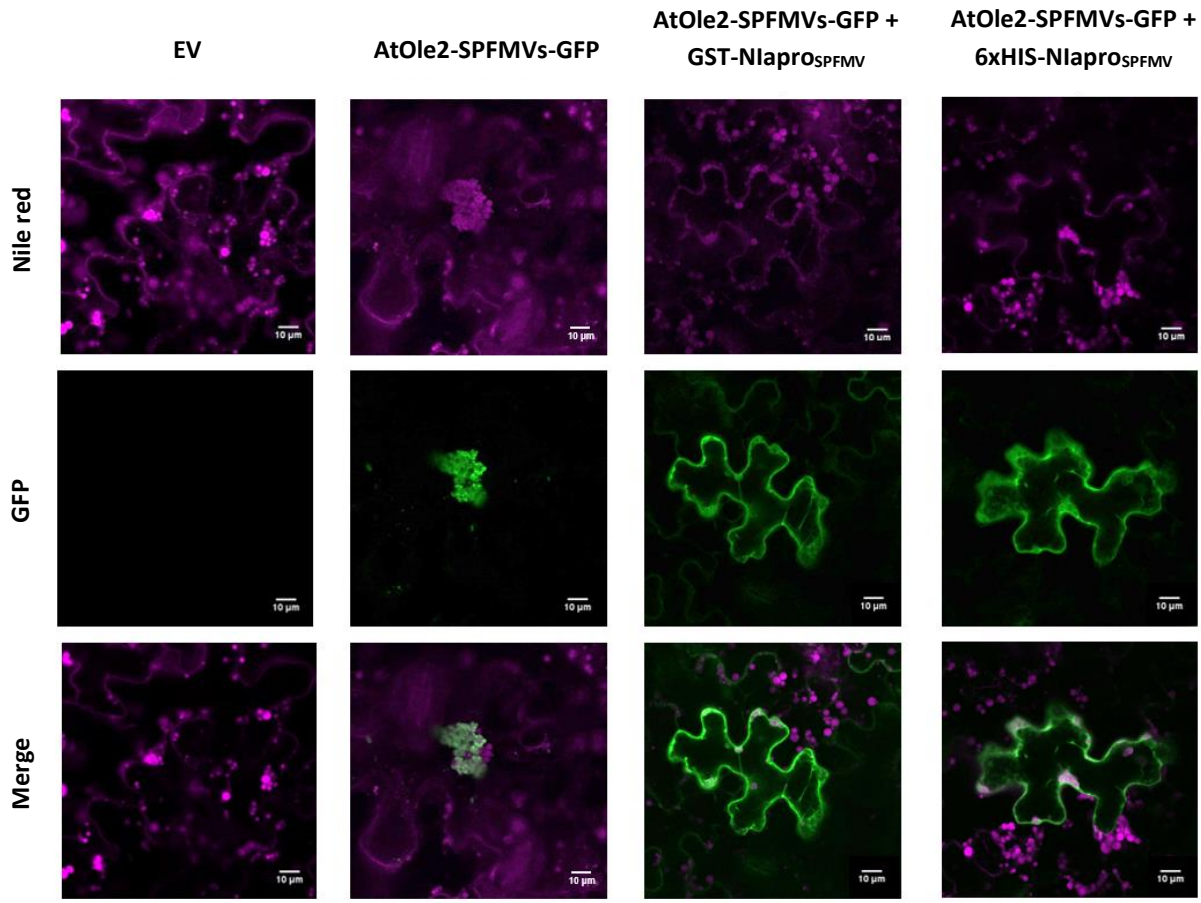


Figure R76. Visualization of GFP relocalization in *N. benthamiana* LD+ leaves transiently expressing the reporter cleavage construct in combination with the tagged-NIa-pro protease. Confocal laser microscopy images of epidermal cells expressing the *AtOle2-SPFMVs-GFP* alone or in combination with *NIa-Pro* protease, His- or GST-tagged, as indicated. Images of Nile red stained leaves are also shown, as well as the merged images with GFP signal images. Images were taken 5 dpi. Scale bars correspond to 10 µm as indicated.

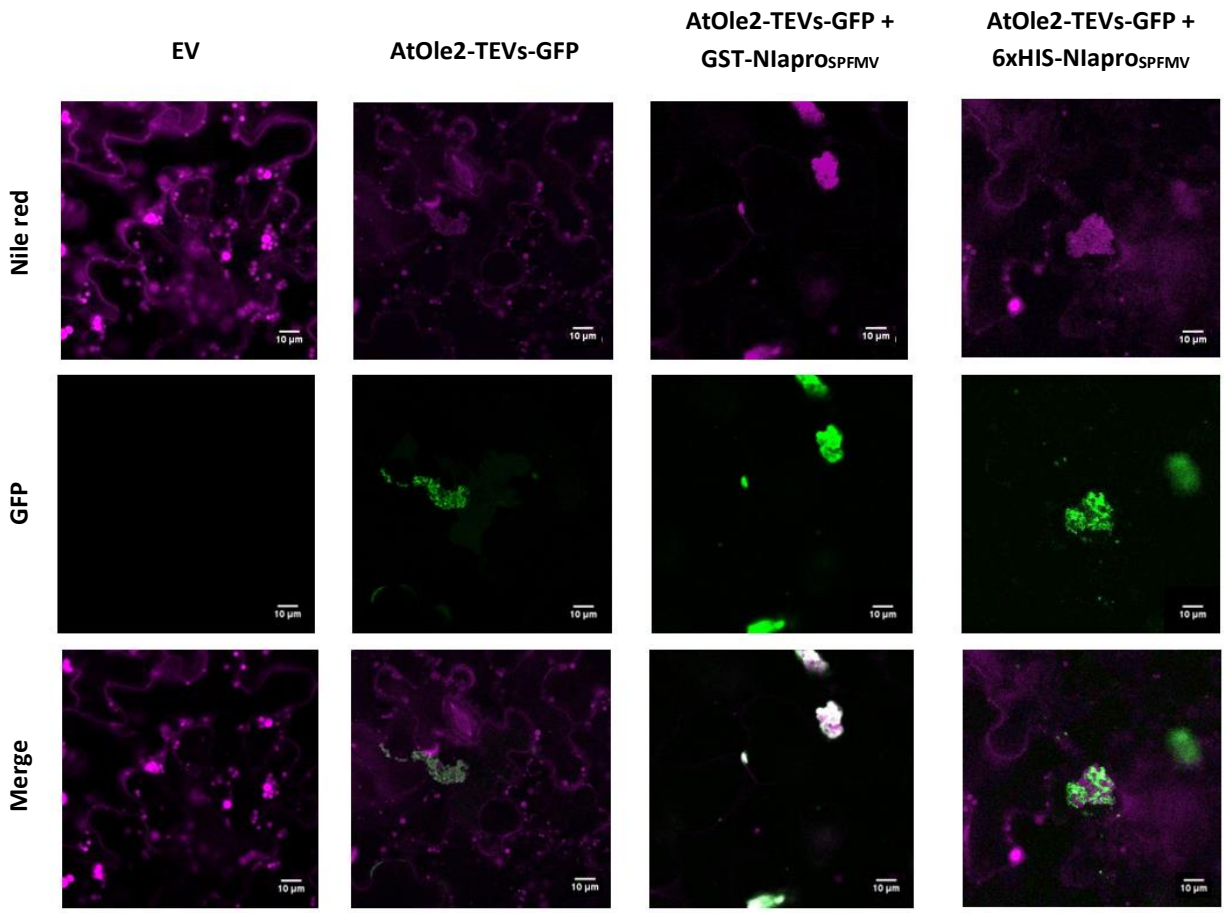
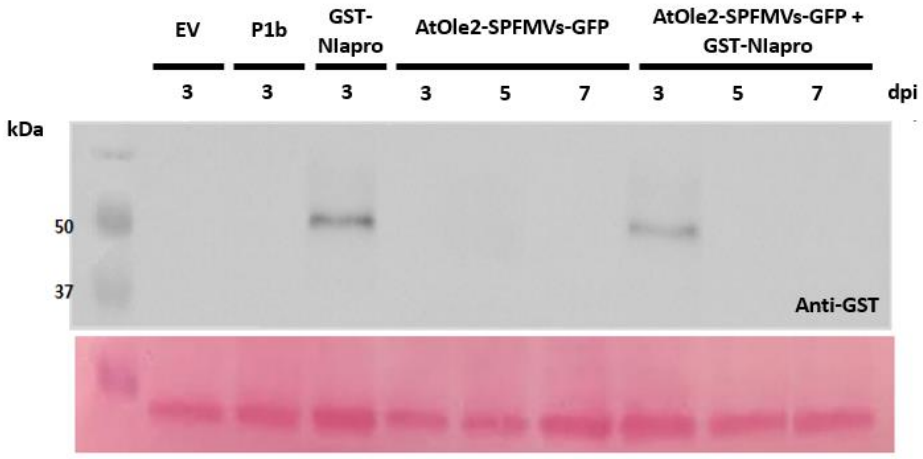
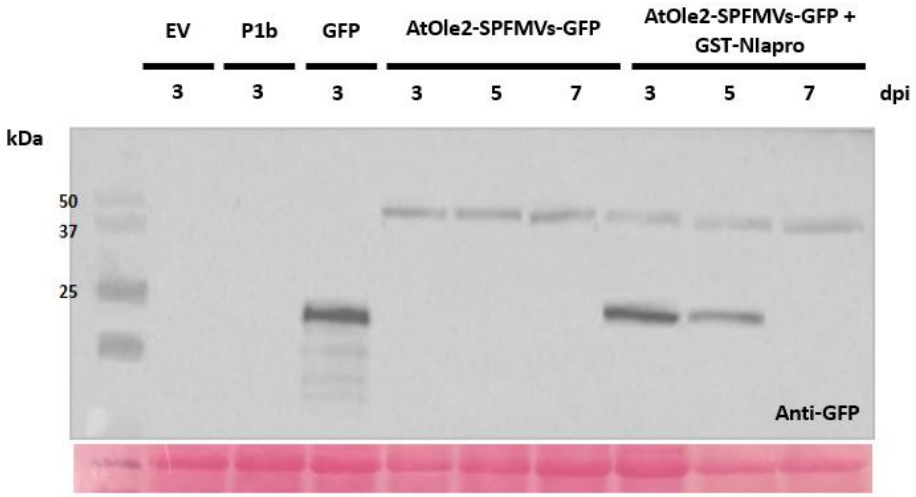


Figure R77. Visualization of GFP localization in *N. benthamiana* LD+ leaves transiently accumulating the AtOle2-TEVs-GFP protein together with the tagged-NIa-pro protease. Confocal laser microscopy images of epidermal cells expressing the *AtOle2-TEVs-GFP* alone or in combination with *Nla-Pro* protease, His- or GST-tagged, as indicated. Images of Nile red stained leaves are also shown, as well as the merged images with GFP signal images. Images were taken 5 dpi. Scale bars correspond to 10 µm as indicated

(a)



(b)



(c)

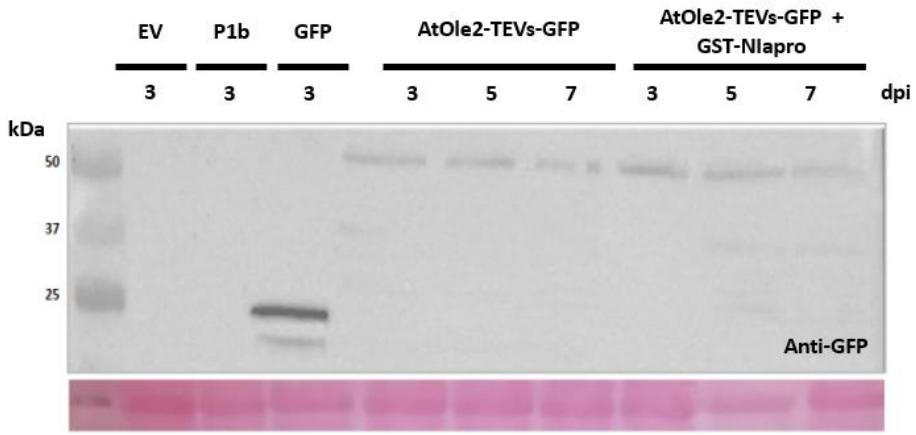


Figure R78. Western blot evaluation of SPFMV NIa-pro production and proteolytic activity. SDS-Total protein extracts from leaves agroinfiltrated with empty vector (EV), or expression constructs of the silencing suppressor *P1b*, SPFMV *GST-NIa-pro*, *GFP*, and the reporter constructs *AtOle2-SPFMVs-GFP* or *AtOle2-TEVs-GFP*, alone or in combination as indicated, at 3, 5 or 7 dpi. Proteins were immunodetected with anti-GST **(a)** or anti-GFP antibodies **(b, c)**. Molecular weight markers are shown on the left in kDa.

10. ACTIVITY EVALUATION OF THE PLANT PRODUCED PEPTIDES IN HUMAN FIBROBLAST CELL LINES

The biological activity of the produced peptides was tested using the different fractions containing the IGF1, CPP-IGF1 (*D. carota* cell culture media and *N. benthamiana* ECFs) and oleosin fusions to IGF1 (isolated LDs derived from *N. benthamiana* leaves). A human fibroblast cell line was used for the assays (MRC-5, ATCC® CCL-171™). Fibroblasts respond to IGF1 increasing their mitogenesis (Zheng *et al.*, 1997) and collagen synthesis (Chetty *et al.*, 2006). Hence, we decided to evaluate the activity of peptides to enhance fibroblast viability. A MTT-microculture tetrazolium assay was selected as an easy screening method (Mosmann, 1983), which allows the evaluation in parallel of the cytotoxic or mitotic

activity of peptides present in different matrixes. Considering the activity described for the synthetic IGF1, three different concentrations were tested (5, 20 and 80ng/ml). Table R6 shows the different peptides and buffers used in the assays.

Table R6. Samples containing peptides and controls evaluated in MTT assay.

Name	Identity	Production matrix	Peptide or fusion	Plant fractions *
dIGF1	Sample	*D. carota*	IGF1	IGF1,P1b#2 depleted medium
dCPP-IGF1	Sample	*D. carota*	CPP-IGF1	CPP-IGF1,P1b#1 medium
dEV,P1b	Control	*D. carota*	-	EV,P1b medium
dsIGF1	Control	Synthetic	IGF1	EV,P1b medium
pIGF1	Sample	*N. benthamiana*	IGF1	ECF in PBS buffer
pCPP-IGF1	Sample	*N. benthamiana*	CPP-IGF1	ECF in PBS buffer
Mock	Control	*N. benthamiana*	-	ECF in PBS buffer
psIGF1	Control	Synthetic	IGF1	Mock ECF in PBS buffer
AtOle2-IGF1	Sample	*N. benthamiana*	AtOle2-IGF1	*N. benthamiana* LDs
IGF1-AtOle2	Sample	*N. benthamiana*	IGF1-AtOle2	*N. benthamiana* LDs
AtOle2	Control	*N. benthamiana*	-	*N. benthamiana* LDs
AtOle2 + sIGF1	Control	Synthetic	IGF1	*N. benthamiana* LDs

*Different fractions containing the peptide or fusion protein from indicated plant matrixes.

The activities of peptides or fusion proteins were assayed as described in Materials and Methods, evaluating the fibroblast viability at 24 and 48h after exposure (ae) to the different fractions. Also, images of the fibroblast cultures were taken to determine possible cell abnormalities induced by the treatments. The results were plotted as variations in cell viability, as determined by comparison to the corresponding controls considered 100%.

The assays with the fractions from *D. carota* cell cultures did not present differences with the control vehicle (dEV,P1b) at 24h. At 48h, the two peptides (IGF1 and the variant CPP-IGF1) at the three tested concentrations increased fibroblast viability, with values above the control. The CPP-IGF1 behaved similar to the synthetic peptide control (no statistical differences at 5 and 80 ng/ml, while at 20 ng/ml it appeared as less active), whereas the IGF1 resulted in increases of viability, as shown in Fig R79.

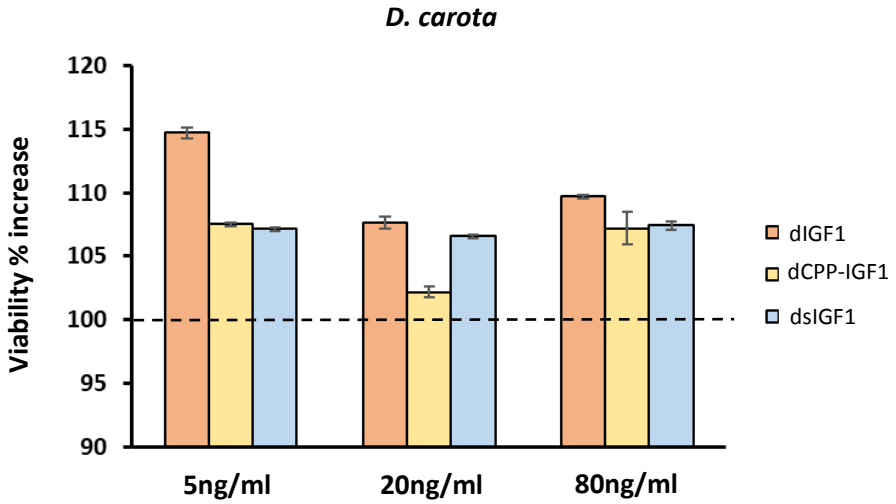


Figure R79. Evaluation of *D. carota* produced peptides by MTT analysis. Bars represent the increase in fibroblast viability, expressed as %, at 48h ae to dIGF1 and dCPP-IGF1. Synthetic IGF1 resuspended in EV,P1b depleted media (dsIGF1) was used as control. The peptide concentrations evaluated are shown below the graph. The results are the mean of five biological replicates. Dashed line indicates 100% viability corresponding to the one of the vehicle control (dEV,P1b) without addition of IGF1.

The increase in fibroblast viability was around an average of 7% for the dIGF1, dCPP-IGF1 and dsIGF1 at the concentrations tested. Only at 5ng/ml the dIGF1 showed remarkable enhanced activity stimulation (~14%). No apparent dose-dependent responses were observed between the peptide concentrations and the cell viability. No abnormalities were observed in the fibroblasts exposed to the different fractions (Fig R80).

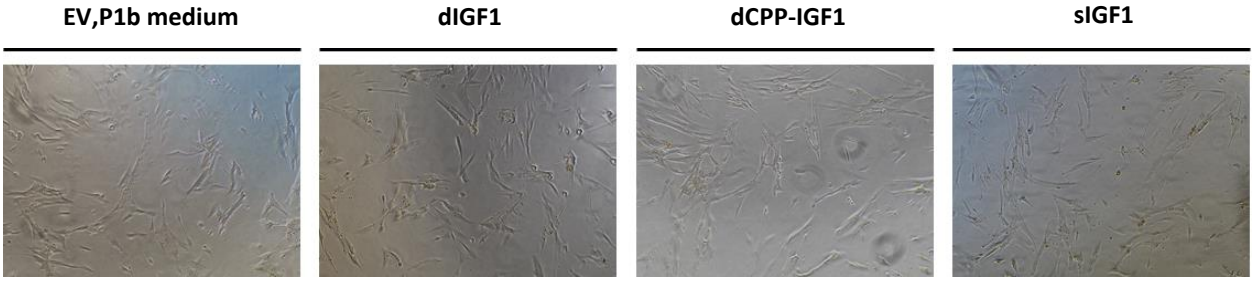


Figure R80. Fibroblast appearance after treatment with *D.carota* cell culture fractions. Images showing the appearance of the fibroblast cultures exposed for 48 h to the different *D. carota* suspensions media containing the indicated peptides.

When the peptides produced in *N. benthamiana* were evaluated, again no differences between the pIGF1 and pCPP-IGF1 and the psIGF1 at 24h ae were observed (not shown). At 48h ae, all the samples containing peptides at the different concentrations increased fibroblast viability (Fig R81). Both pIGF1 and pCPP-IGF1 showed similar results to the reported by the psIGF1. No significant differences (Tukey

test, $p \leq 0.05$) were observed between peptides or concentrations. The average increase in viability ranged between approximately 17-20%. We did not observed abnormalities in fibroblasts exposed to the different solutions (Fig R82).

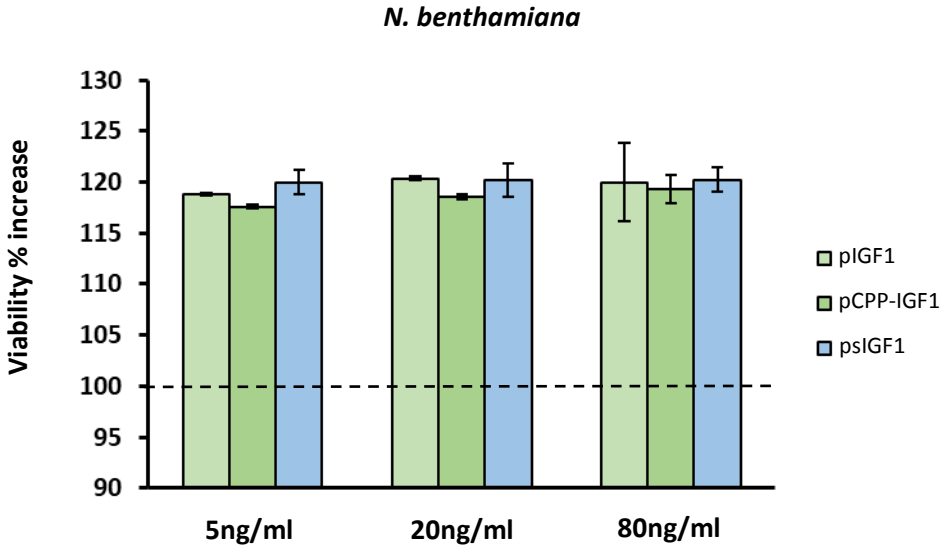


Figure R81. Evaluation of *N. benthamiana* produced peptides by MTT analysis. Bars represent the increase in fibroblast viability, expressed as %, at 48h ae to ECF containing the pIGF1, pCPP-IGF1 and psIGF1 peptides. The peptide concentrations evaluated are shown below the graph. The results are the mean of five biological replicates. Dashed line indicates 100% viability corresponding to the one of the vehicle control (Mock) without addition of IGF1.

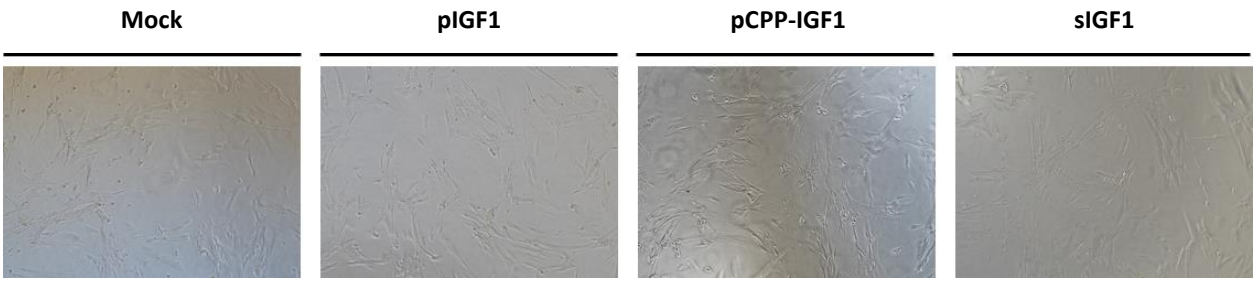


Figure R82. Fibroblast appearance after treatment with *N. benthamiana* ECFs. Images showing the appearance of the fibroblast cultures exposed to the different *N. benthamiana* ECFs 48h ae.

Finally, the AtOle2-IGF1 and IGF1-AtOle2 samples were evaluated for activity. The control LDs (AtOle2) did not cause cytotoxicity into fibroblast (not shown). When the synthetic IGF1 was added to this control, an increase of viability was observed at the lower and higher concentrations tested, while for the intermediate concentration no differences were found (Fig R83). The LDs with fusions of IGF1 at the N- or C-terminus of AtOle2 resulted less active that the free synthetic IGF1, but still a significant

increase in viability was found at the higher concentration tested. Again, it should be noticed that there was no clear doses-response for the three concentrations.

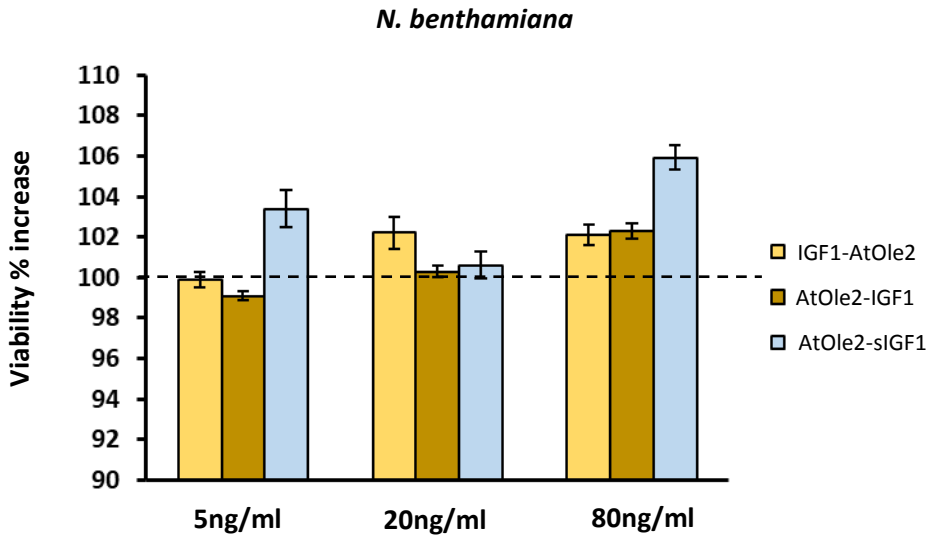


Figure R83. Evaluation of recombinant LDs by MTT analysis. Bars represent the increase in fibroblast viability, expressed as % at 48h ae to LDs carrying the fusion proteins AtOle2-IGF1 or IGF1-AtOle2. Synthetic IGF1 added to LDs carrying the AtOle2 protein (AtOle2 + IGF1) was used as control. The peptide concentrations evaluated are shown below the graph. The results are the mean of five biological replicates. Dashed line indicates 100% viability corresponding to the one of the vehicle control (AtOle2) without addition of IGF1.

When fibroblast cultures exposed to the LDs were visualized, we did not see abnormalities in fibroblast morphology (Fig R84).

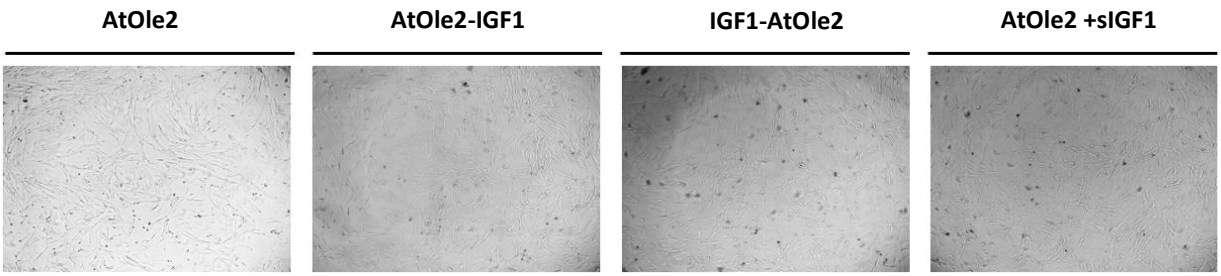


Figure R84. Fibroblast appearance after exposure to recombinant LDs carrying the indicated proteins. Images were taken after 48h of exposure to LD fractions *N. benthamiana*.

Even though it seems that the plant produced peptides were more active than those produced in plant cell cultures, it should be taken in consideration that the different vehicles by themselves might have effects on fibroblast viability. Fig R85 shows the reduction of viability triggered by the different vehicles

(Mock and dEV,P1b). As it could be observed, the reductions in viability were -16.3% for the Mock and -6.71% for the dEV,P1b. These differences should be considered as they set different baselines for each independent assay.

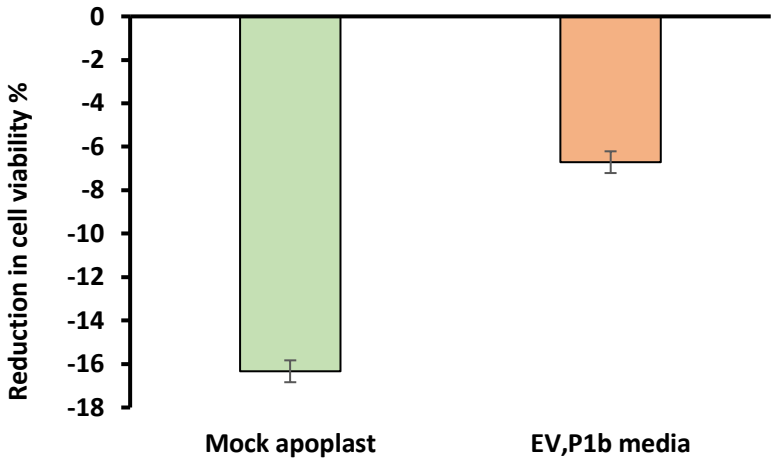


Figure R85. Comparison of vehicles used in the MTT analysis. Graph representing the reduction in fibroblast viability, expressed in % at 48 h ae to the different vehicles used in the *N. benthamiana* (Mock) and *D. carota* (EV,P1b medium). The results are the mean of five biological replicates.

Altogether, the results showed that the different IGF1 and derived peptides produced in *D. carota* and *N. benthamiana* maintained their activities when confronted to human cell lines.

DISCUSSION

DISCUSSION

This thesis describes the production of diverse peptides in plant cell cultures and also in whole plants and focuses on how to increase the yields while maintaining the desired biological activities of the products. Each of the two production matrices tested (plant cell cultures or tissues from whole plants) has properties that make it advantageous for certain applications, making them optimal candidates to be used for different marketable products or applications. On one hand, the use of whole plants allows the production of recombinant proteins at low cost and in a very short and controlled timeframe, with simple procedures to scale up the production. However, it is difficult for plant products to reach the market due to limitations for industrial certification of the processes and/or contamination of products with toxic secondary metabolites or other undesired molecules associated with the production (Fischer *et al.*, 2012). In contrast, the production of recombinant proteins using biotechnological approaches with plant cell cultures would be more straight forward to certify, and all production parameters could be accurately controlled, but there could be cost limitations and other problems deriving from the intrinsic difficulty of plant cell culture maintenance (James and Lee, 2006). In this thesis we have explored several different strategies to overcome some of these problems to produce high yields and increase the activity of the products. The main strategies tested include the use of viral RNA silencing suppressors (RSSs) to enhance gene expression, the adaption of lipid body technology to provide a system to stabilize products and facilitate their purification, and also the use of viral vectors for efficient expression. The application of all these procedures for the production of selected compounds will be discussed.

DEVELOPMENT OF BIOTECHNOLOGICAL TOOLS AND MATERIALS

The first chapter of the result section describes the development and evaluation of biotechnological tools and materials for use in the different production strategies. Screening was done to select an active RSS for the expression assays, and the plant cell lines with the highest potential for use in oleosin technology based on their LD content. Additionally, a set of gene constructs and new plant cell cultures were developed and evaluated as part of this research.

Regarding the use of RNA silencing suppressors of viral origin, first we evaluated two RSSs derived from two plant viruses of the family *Potyviridae* already being extensively used in our laboratory: HCPro of WMV (Domingo-Calap *et al.*, 2020; Lakatos *et al.*, 2006) and P1b of CVYV (Valli et al., 2006). The criteria for selecting them included previous characterization: the HCPro belongs in the category of the first viral RSSs to be characterized experimentally (Kasschau and Carrington, 1998; Souza *et al.*, 2013), while

P1b presents analogies and differences in its mode of action. The use of RSSs is expected to prevent the transgene-induced gene silencing, and enhance expression levels of transgenes in plant leaves by Agroinfiltration (Sainsbury and Lomonossoff, 2014). The experiments were performed with *N. benthamina* plants through Agroinfiltration for transient expression, being a fast and easy way to examine intracellular localization and accumulation of fluorescently-tagged reporter proteins. The comparison of the two RSSs showed a greater efficacy of co-expressing plasmids with the fluorescent reporter GFP and P1b from CVYV compared to HCPro from WMV, indicating that P1b enhanced recombinant protein accumulation (Fig R1-R3). The assay also allowed us to obtain preliminary data regarding the levels of protein accumulation using the two evaluated RSSs. As a result, P1b was selected as the optimal candidate for our assays in *N. benthamiana* plants and also in plant cell cultures, assuming that the silencing suppression strategy of P1b (preventing siRNA loading into the RNA-induced silencing complex) would act in the same way in both production matrices regardless of possible differences in cell compartmentalization. The decision to select P1b was reinforced by the fact that HCPro is a multitasking protein (Valli *et al.*, 2018), which could interfere in metabolic pathways necessary for our project (i.e. lipid accumulation and others) or even in cell callus viability. In the case of P1b, apart from interfering RNA silencing activity and the serine protease activity in its C terminus (Valli *et al.*, 2008), few other activities have been currently described. In addition, the smaller size of P1b compared to HCPro facilitates its expression.

The LD content in plant cell cultures was screened to evaluate the natural differences between species and allowed us to divide the species in three groups based on the characteristic of their lipid droplet. Group one, with large but few LDs, included *C. sativa* and *C. longa*, while in group 2, with small but numerous LDs, were *A. montana, G. herbaceum, D. carota and S. crassifolia*. Finally, the third group of species, with almost no detectable LDs, was limited to two model cell lines, BY2 derived from *N. tabacum* and *A. thaliana* (Fig R8). Previous reports stated that LDs were not abundant in BY2 cell cultures (Zhang *et al.*, 2016), which might have precluded further investigation. Our results suggested that the BY2 case might be the exception rather than the rule, and it opens up the possibility of using a technology rather unexplored until now. The species with large LDs, with a diameter above 2 µm, had less than 70 LDs/cell, while those with smaller LDs, diameters below 2µm, had more than 70 LDs/cell, 200 LDs/cell in some cases. Our data do not provide reasons to explain these differences in LD content and characteristics, and no correlations were evident when looking at the plant material used for callus induction (Table R4), or the evolutionary distances between the species (Fig R4). The importance of the plant material used for callus induction derives from cell diversity in the different plant cell cultures (different plant tissues do contain different cell types after differentiation), but which of the original cell types would rise to the callus culture could not be predicted using this

procedure. We cannot rule out that the different LD accumulation patterns might be associated with the cell type that originated the callus and not with the plant tissue used in the induction.

After screening the different plant cell lines, *D. carota* was selected as a good candidate for our future assays. Given that the cell line used initially has some restrictions of use, we generated a new *D. carota* cell line using traced carrots as starting material (Fig R9). When the new cell line was compared in terms of growth and LD content with the previously available cell line, we found that their growth rates were practically identical after an adaptation process to the same culture media (Fig R11-R12). The remarkably good match in growth dynamics between both cell lines suggested that *D. carota* could provide robust cell lines with little or no variability in terms of callus induction reproducibility and growth behavior. This robustness is a key factor for selecting a species to be used as source material for plant cell cultures, because it would allow the periodical induction of new plant cell lines whenever needed, or the induction of new replacement lines in case of accidental loss. Simply starting from easy to obtain plant material, minimizes the risk of problems such as unexpected changes in the responses to stimuli (transformation, elicitation, etc), or variability between different cell lines. The high reproducibility of callus induction and suspension starting from plant material is important to avoid the need for continuous propagation of plant cell callus. A plant species that allows rapidly new cell lines to be established when needed is a valuable asset, even when transgenic modifications are considered.

Regarding LD contents, there were differences in the sizes of the LDs in the two carrot-derived cell lines, with the new cell line having larger LDs than the previously tested line. This could be due to the clear differences in their times under culture conditions, from years to only a few months. Continuous propagation as callus or cell suspension could result in changes in LD profiles, but we must wait to see if this is the case. At this point we cannot rule out other possible explanations for the differences in LDs, such as the type of cells de-differentiated for growing as callus: although similar tissues were used as initial materials for establishing both cell lines, the specific cell type that originated the two calli remained unknown.

Altogether, we found that most of the plant cell cultures inspected accumulated detectable LDs in their cytoplasm. The observed differences in their number and size might depend on the species and the cell line. Also, our results show that a *de novo* induced *D. carota* plant cell culture can achieve growth rates equivalent to that of cell lines used in industrial productions, after a simple adaptation process.

We then went on to focus on oleosin technology for the production and accumulation of recombinant proteins in plants. *A. thaliana* oleosin 2 (AtOle2) was selected as an adequate scaffold for the fusion of proteins of interest. This oleosin was available in the laboratory and chosen because of its size: AtOle2

is the largest Arabidopsis oleosin (molecular weight of 21 kDa), making it easier to identify the fusion proteins. To test the potential uses of the technology, we decided to create a chimeric fusion between AtOle2 and GFP for easy monitoring and evaluation under fluorescent microscopy. Although in most previous reports oleosin fusions have been at the C-terminal (Boothe *et al.*, 2010; Montesinos *et al.*, 2016; Nykiforuk *et al.*, 2006, 2011; Parmenter *et al.*, 1995), we decided to prepare and compare two different constructs tagging the AtOle2 in the N- or C-terminal with the GFP reporter. The N- and C-terminal domains of oleosins are considered to have properties of an amphipathic α-helix and proposed to be located on the surface of the oil body (Tzen *et al.*, 1992; Lacey *et al.*, 1998). Oleosin ends vary in length and physicochemical characteristics (Huang and Huang, 2015), consequently our comparison was relevant to determine how elongations affect the accumulation and localization of oleosins in the LDs, as well as the stability of LDs.

Our results showed that both types of fusions accumulated in *N. benthamiana* leaves shortly after infiltration (<2 dpi), but fluorescence quickly disappeared (Fig R16). It is unclear why the production of oleosin fusion proteins was so transient. Even if there was a rapid turnover of LDs, the absence of *de novo* synthesized oleosin fusions is hard to explain. One possibility is that the presence of Agrobacterium might alter the dynamics of LDs. The lack of detection of GFP signals after 2 dpi contrasted with the results obtained during the RSS evaluation, where free GFP could be visualized at least until 7 dpi. This huge difference in the period of accumulation could be due to the different behavior of free and oleosin-fused GFP, needing to be incorporated into LDs practically inexistent in leaves to correctly accumulate the oleosin-fused forms. Our data is in agreement with Wahlroos et al. 2003, who found that the production of oleosin-GFP fusion protein occurred early, but was not accumulated due to the lack of sufficient LDs in *N. benthamiana* leaf cells for anchoring. We can hypothesize that oleosin fusions only accumulate to a specific level, and a strong repression of the transgene and/or a degradation mechanism once this threshold is reached could virtually eliminate the produced fusions. Further experiments are required to elucidate why the oleosin fusions were only detected at 2 dpi.

Our results also show that both AtOle2 fusions were incorporated in the LDs, with different LD accumulation patterns depending on whether the fusion was in the N or C terminal. The AtOle2-GFP accumulation triggered the aggregation of LDs, whereas they remained as discrete vesicles with the GFP-AtOle2. These results suggest that the incorporation of fusion of AtOle2 with GFP in the C-terminal can alter the properties of the protein, producing observable effects on the LDs. Also, the observed aggregations could be associated with some kind of steric interaction between the GFP and the LD surface elements. Previous reports have described the aggregation of oleosin fusions in the C-terminal when transiently expressed in Nicotiana plants from six hours after bombardment, and the authors hypothesized that the aggregation was due to difficulties of the cells being able to deal with extensive

oil body storage (Wahlroos *et al.*, 2003b). However, considering our results of the N-terminal tagging of the AtOle2, where the LDs incorporating the oleosin fusions remained as discrete vesicles, we hypothesize that GFP tagging might affect the protein-protein or protein-lipid interactions necessary for correct LD dynamics. This issue could be a reflection of the different characteristics of the two terminal ends of the oleosins as described by Huang and Huang 2015.

Our data suggest that oleosin technology could be used in cell cultures with high LD content for proper anchoring an d stabilization of oleosin fusion proteins, favoring their accumulation. To test this hypothesis, we concentrated our efforts in *D. carota* cells. Future studies are needed to verify our predictions, and to further explore the limits of the technology in other species.

PRODUCTION OF RECOMBINANT PEPTIDES IN *D. CAROTA* CELL CULTURES

In the next part of our research, we explored different strategies to produce recombinant peptides in *D. carota* plant cell cultures, including transient and stable production of peptides in the culture media and production using the oleosin technology.

First we evaluated the feasibility of genetically modifying the newly developed *D. carota* cell line. We tested the resistance of calli to hygromycin and kanamycin by supplementing the solid media with different concentrations of the antibiotics (Fig R17). Considering the results of these assays we decided to use a concentration of 50 mg/l for both antibiotics. Then we transformed *D. carota* cultures using *A. tumefaciens* carrying plasmids for GFP expression, obtaining different calli accumulating GFP levels sufficient for visual inspection under UV light (Fig R19). Subsequently, we transformed the same cell line with constructs for oleosin fusion production. Although in the first transformation experiments we did not obtain antibiotic resistant calli, probably due to a low transformation efficiency, we finally obtained transformant calli producing the oleosin fusions correctly incorporated in the LDs (Fig R20). No differences in LD accumulation patterns were observed between the N- and C-terminal fusions. Transgenic calli producing the oleosin fusions were in general smaller than the control calli producing free cytoplasmic GFP, and their growth stopped completely after one propagation cycle (Fig R21). These results suggest that the accumulation of oleosin fusions in LDs could dramatically affect callus growth. It is important to consider that LDs are highly dynamic organelles and are actively involved in many physiological processes including membrane biogenesis, cell growth, diurnal regulation processes, hormone signaling, and plant growth and development (Shao *et al.*, 2019). Therefore, it is likely that any modification in their dynamics could indirectly affect other essential biological processes, so reducing the viability of cells.

Next, using the stable transformation protocol previously tested, we transformed the *D. carota* cell cultures with genetic constructs for the production of peptides of industrial interest. Among the different candidates, the human insulin like growth factor 1 (IGF1) was selected for our purposes because of its potential uses in cosmetics (Trüeb, 2018) and therapeutics (Lui *et al.*, 2019; Ranke *et al.*, 2009). A peptide described as a cell penetrating peptide (CPP) to increase the potential of the recombinant proteins to penetrate the mammalian cells was fused to the IGF1. The hPP3 peptide was selected due to its proven capability for transporting protein cargoes and penetration of human fibroblasts (Wang *et al.*, 2016). For the cell culture transformations, we used the prepared constructs for the production of IGF1 and CPP-IGF1 in combination with a construct for the expression of the P1b silencing suppressor. Our initial hypothesis was that cell lines co-expressing IGF1 or CPP-IGF1 with P1b would accumulate higher recombinant protein levels than the ones co-transformed with the EV. In contrast to transient expression assays in *N. benthamiana* leaves, there are only few studies reporting the use of RSSs to enhance recombinant protein accumulation in plant cell cultures (Boivin *et al.*, 2010; Huang *et al.*, 2018), and the activity of P1b has not been evaluated in plant cell cultures in stable transformation. Different cell lines were obtained in which the presence of the transgenes was confirmed for all the transformed lines (Fig R24). These lines produced the peptides, as showed by western blot and ELISA analysis using specific anti-IGF1 antibodies (Fig R25), and the growth and dynamics of peptide accumulation in the different cell suspensions was also evaluated (Fig R26 to R32). All of the transgenic suspensions reached biomass levels lower than the WT, but they could be propagated and conserve similar kinetics to the WT in liquid cultures. The suspensions co-transformed with P1b and EV (control transformations), were the most similar to the untransformed WT in terms of growth, suggesting that the presence of an active RSS-like P1b was not critical for callus growth. This contrasted with the results observed in plants after stable transformations to express RSSs, where strong negative effects have been reported both on growth and on development (Siddiqui *et al.*, 2008). When the peptide accumulation was evaluated, we found that cell lines co-expressing IGF1 or CPP-IGF1 with P1b gave the highest peptide yield, validating our initial hypothesis that the stable expression of P1b serves to enhance peptide accumulation in transgenic cell cultures. When the IGF1 and CPP-IGF1 accumulation levels were compared, we observed that IGF1 levels where higher in all cases. This difference could be associated with the size and other physicochemical characteristics of the peptides, the CPP-IGF1 being longer and having an isoelectric point lower than the IGF1, as well as an improved cell penetrating capacity. Alternatively, it is possible that differences were only associated to quantification limitations: as indicated previously, ELISA assays was used for quantification using anti-IGF1 polyclonal antibodies compared to known amounts of synthetic IGF1, and we cannot rule out that the presence of the CPP sequence could alter the recognition in CPP-IGF1, affecting its detection and quantification. Regarding the yields of peptide accumulation, they were similar to the values achieved

with other growth hormones or interferons in plant cell cultures (Hong *et al.*, 2002; Honga *et al.*, 2002; Magnuson *et al.*, 1998). Also, we should consider that previous reports have described that yields might be affected by degradation of proteins when secreted into the culture media (Kwon *et al.*, 2003; LaCount *et al.*, 1997; Rainer *et al.*, 1999). If this is the case, we might be underestimating the amount of peptide produced. Future research is needed to study the stability of IGF1 in *D. carota* culture media, and to design new strategies to improve peptide yields.

We also explored the possibility of cryopreserving the transgenic cell lines. This is relevant to avoid reduction in peptide accumulation over successive propagation cycles, as the initial materials would always be available when needed. Previous reports have described that transgenic cell lines could suffer a gradual reduction in the yield of recombinant proteins due to their growth as clusters (James and Lee, 2006). Given this growth as clusters, difficult to separate during propagation, the suspensions can be heterogeneous, containing a mosaic of cells producing different amounts of recombinant proteins and having different growth rates, so resulting in drifts in performance. Those cells with high growth rates but low production levels can gradually dominate the population. For our assays we selected a transgenic cell line co-expressing P1b and IGF1. An aliquot of the cell culture suspension was cryopreserved during three months, while the rest of the material was maintained in liquid or solid media. After the three months, all the cells were recovered (those cryopreserved and those maintained in solid and liquid culture) then propagated as suspensions and their growth and peptide accumulation profiles evaluated. The results show that cryopreservation did not negatively affect the growth dynamics after adaptation of the callus to liquid media, the tolerance to the cryopreservation indicating the robustness of the developed *D. carota* new cell line for practical applications. When the peptide accumulation yields were evaluated, we confirmed that the cryopreserved cell line reached higher protein levels than those maintained in liquid or solid media, in which some reductions were observed, consistent with the expected dynamics of production. Interestingly, the peptide yields in the cryopreserved line were similar to those from before the cryopreservation process, and remained constant for several months. These results demonstrated the feasibility of cryopreservation of transgenic *D. carota* cell lines as a means to secure master cell banks. The results also corroborate the data in James and Lee, 2006, who found reductions of peptide yields in cultures maintained either as liquid or solid cultures for some months, which, in the case of our cell cultures maintained as suspension, were highly significant with ~33% reduction in a timeframe of seven months.

Additionally, we explored the transient expression of recombinant peptides in *D. carota* cell cultures. Transient expression strategies could potentially allow a fast and efficient production of recombinant proteins in plant cell suspensions, avoiding the problems associated with transgenic cell lines. The

expression was assayed during a stage in the Fraunhofer-Institut IME (Aachen, Germany), using the PCP technology previously developed there (Rademacher et al. 2019). Initially, we designed some assays to optimize peptide expression in BY2 and carrot cell lines, including evaluation of the agroinfusion buffer (Fig R40), concentration of agrobacteria (Fig R40), agrobacterium strains (Fig R41), incubation time (Fig R41) and the effects of surfactants (Fig R42). Different plasmids were also evaluated for their efficiency in transient gene expression. Our starting plasmid was the pTRAK-RFP, previously described (Rademacher et al., 2019). This plasmid carries two SAR sequences into the T-DNA region, sequences which facilitate normalized transgene expression in stable transformation assays (Breyne *et al.*, 1992), and binary plasmids carrying SARs have been successfully used for the expression of recombinant proteins in plant cell cultures (Bortesi *et al.*, 2009). Using pTRAK-RFP as the backbone, we constructed a series of plasmid variants containing different numbers of SAR sequences in different positions with respect to the sequence to be expressed (Fig R44). The performance of these variants was evaluated by comparing the accumulation of fluorescent marker proteins in BY2 and carrot cell lines (Fig R45). In these experiments we found that the plasmids containing only one SAR triggered higher accumulation of the reporter protein than the plasmids having two SARs or none. Although the presence of SARs is considered important for the correct attachment of the T-DNA to discrete regions in the nuclear matrix favoring T-DNA gene expression (Breyne *et al.*, 1992), more than one SAR seems to be less efficient. The reason might be related to competing for the matrix attachment or a structural modification when two SARs are present. Using the plasmid containing only one SAR in the 3'-terminal position with respect to the expression cassette, we engineered two new plasmids compatible with Gateway™ technology, one with and the other without the expression cassette of the RSS P1b from CVYV *in cis*. These new plasmids, denominated pTAREK_GW and pTAREK_P1b, enabled the rapid and easy assembly of target genes for the production of recombinant protein in plants (Fig R47). The expression cassette of P1b was inserted in the plasmid pTAREK_P1b to enhance target gene expression without the need for co-infiltration with a plasmid containing an RSS expression cassette. Both plasmids were evaluated with reporter fluorescent proteins allowing high level of expression, particularly the pTAREK_P1b that led to high, homogeneous expression of target genes in transient assays in plant cells (Fig R50-R51). In future research we could use the pTAREK_P1b plasmid as backbone for designing improved versions, for instance adapting the use of 5' and 3' untranslated regions from plant viruses that might serve as enhancers (Sainsbury *et al.*, 2009) or even synthetic sequences with similar functional characteristics (Peyret *et al.*, 2019).

RECOMBINANT PEPTIDE PRODUCTION IN *N. BENTHAMIANA* PLANTS

In the third part of this thesis, we explored different strategies to express recombinant proteins in *N. benthamiana* plants by transient expression assays. First we explored the production of recombinant proteins in *N. benthamiana* using a viral vector derived from TMV developed in our research group (Shi *et al.*, 2019). Different constructs were used for the IGF1 or CPP-IGF1 peptides and target accumulation in leaf apoplasts. Extracellular accumulation of IGF1 and CPP-IGF1 could greatly facilitate their downstream purification, and this would greatly reduce production costs and times (Fig D1). Extracellular fluids (ECFs) enriched in IGF1 or CPP-IGF1 were obtained from leaves and evaluated by western blotting and ELISA using specific antibodies against IGF1.

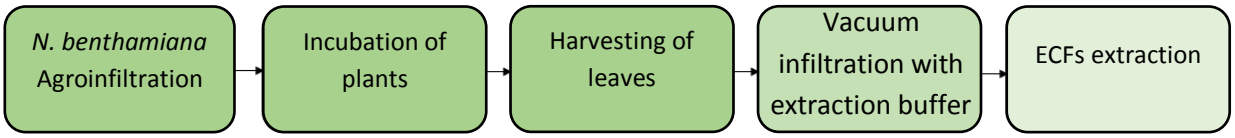


Figure D1. Scheme of ECFs extraction from *N. benthamiana leaves*.

As with plant cell cultures, the accumulation of IGF1 was higher than the accumulation of CPP-IGF1 in plant leaves (Fig R55 to R58). Equally, these differences could correspond to differences in the amount of peptide due to the physicochemical characteristics of the peptides and might be associated with inaccurate quantification using the anti-IGF1 antibodies to detect both peptides. The peptide-enriched ECFs could be processed for further isolation, or used directly as active end-product, consequently reducing the manufacturing costs.

Another approach for the production of our candidate peptides was the use of oleosin technology in plant leaves. Our results showed that the accumulation of recombinant LDs (LDs containing oleosin fusions) was only detectable for short periods post infiltration, and we hypothesized that perhaps the leaves could not hold the accumulation of oleosin fusions because they do not contain enough LDs to anchored them. Therefore, we explored the oleosin fusion technology in plants engineered for high LD content in leaves. These transgenic *N. benthamiana* plants were generated at the CSIRO, Australia and kindly provided to us by Dr. Petrie (Vanhercke *et al.*, 2014). They were constitutive overexpressors of the limiting enzyme in the TAG biosynthesis and a master regulator of TAG biosynthesis, the DGAT1 and WRI1 proteins, respectively (Figure D2).

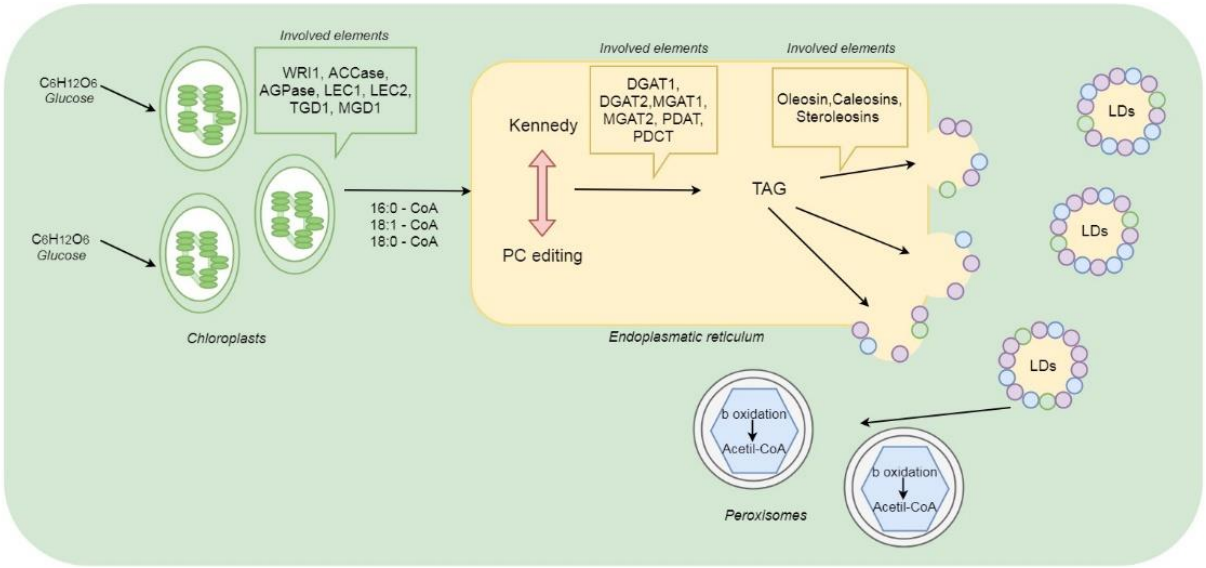


Figure D2. Schematic representation of LD production in plant cells. Gene products including the regulatory WRI1 and other elements of the pathway that trigger the transformation of glucose to acetyl-CoA and derivatives in plant chloroplasts are indicated. These molecules are used as the carbon source for TAG production by DGAT1 and other ER enzymes. The ER-produced TAGs are stored in LDs by oleosins and other structural proteins (caleosins, steroleosins, and others) represented with purple, green and blue circles, respectively. Produced LDs bud from ER as individual vesicles and accumulate in the cytoplasm. LD TAGs can be later β-oxidated in the peroxisomes.

Using these transgenic plants, we transiently expressed oleosin fusions to GFP to monitor the accumulation of the recombinant proteins. The accumulation of oleosin fusions was clearly enhanced in the high LD content plants (*N. benthamiana* LD+), producing large recombinant LDs aggregates (up to 24 µm of diameter) and until 7 dpi (Fig R61). These results supported our hypothesis that the capacity of leaves to hold oleosin fusion proteins requires sufficient LDs for anchoring. However, we observed that these aggregates contained two different types of LDs. The first were small with fluorescence on their surface containing the recombinant fusion proteins, whereas the others were larger and with no associated fluorescence (Fig R62). A possible explanation is that the large LDs were present before the agroinfiltration of our gene constructs, hence they did not carry the oleosin recombinant proteins. In the case of the small, fluorescent LDs, they should be *de novo* formed during the transient production of the recombinant oleosins and therefore incorporate newly synthesized oleosins. The overexpression of the oleosin fusions in leaf cells could induce the formation of small LDs. Correlations between LD size and oleosin levels has been already described, suggesting that oleosins modulate the size of LDs *in vivo* (Montesinos *et al.*, 2016; Popa *et al.*, 2019; Ting *et al.*, 1996; Tzen *et al.*, 1993). Even though more oleosin fusions were accumulated in leaves of the LD+ plants, and for longer periods of time, a large number of LDs were not recombinants, interfering with the isolation of the recombinant proteins. When LDs were isolated from the infiltrated leaves, most of the

LDs were large and non-fluorescent (Fig R63). Most of the small fluorescent LDs were not isolated by flotation centrifugation in sucrose gradients, as they remained in the pellet. The reason might be a higher density of LDs with a high content of oleosin proteins. In addition, the oleosins forming recombinant LDs are GFP fusion proteins, larger than regular oleosins. Moreover, the recombinant LDs showed a tendency for aggregation. It is clear that the LD isolation methods need to be optimized for recombinant LDs, and adapted to the specific recombinant LDs carrying particular peptides.

The results with the LD+ plants strongly suggested that the formation of new LDs was a result of the concomitant increase of TAG production with the overexpression of oleosin proteins, causing accumulation of the recombinant oleosin fusion proteins as newly formed LDs. To further explore this possibility, we tested the transient expression of the *AtDGAT1* and *AtWRI1* genes, the same genes constitutively expressed in *N. benthamiana* LD+ plants, in combination with the gene constructs for oleosin fusion proteins, considering that the simultaneous expression of the three components would trigger an accumulation of LDs, and all *de novo* produced LDs should carry oleosin fusions (Fig D3). Furthermore, the co-expression of the *AtDGAT1* and *AtWRI1* genes in combination with the fluorescently tagged oleosin fusions AtOle2-GFP and GFP-AtOle2 allowed us to trace the recombinant LDs (Fig R67). We visualized all the LDs isolated or in small aggregates, mostly fluorescent, indicating that most of them carried the GFP fusion proteins and were recombinant. The LDs were purified from agroinfiltrated leaves by flotation-centrifugation on density gradients and visualized with confocal microcopy, to confirm the *in vivo* observations (Fig R68). The results showed that the oleosin fusion proteins can be correctly produced and accumulated in *N. benthamiana* WT leaves and targeted to LDs. Furthermore, by co-expressing the Arabidopsis *WRI1* and *DGAT1* genes, the plant cells accumulated mainly recombinant LDs. These results demonstrate that our new system could represent a considerable improvement in terms of yield and timing for oleosin technology, which is currently based on LD-enriched tissues such as seeds requiring completion of the plant cycle that could take several months. In contrast, the transient production in *N. benthamiana* leaves only required several days. In order to test the potential use of the new technology for industrial applications, oleosin fusions to IGF1 were co-expressed with the *WRI1* and *DGAT1* genes allowing the production and isolation of LDs containing the oleosin-IGF1 (Fig R70 to R74). Our results show that this co-expression allowed the accumulation of oleosin fusions to proteins of industrial interest. The isolated LDs containing the oleosin-IGF1 protein could be directly used for different applications, as previously reported (Cai *et al.*, 2018). Our system would allow the production of therapeutically-ready treatments in a transient system, without the need for transgenic plans, that sometimes have a negative perception and could be subjected to legal limitations (Fischer *et al.*, 2013; Wunderlich and Gatto, 2015). As an additional benefit, the system only required a straightforward purification scheme. This is an important asset because the downstream purification processes are estimated to account for, on average, up to 80%

of the costs for recombinant protein manufacturing in other production systems (Menkhaus *et al.*, 2004; Raven *et al.*, 2015).

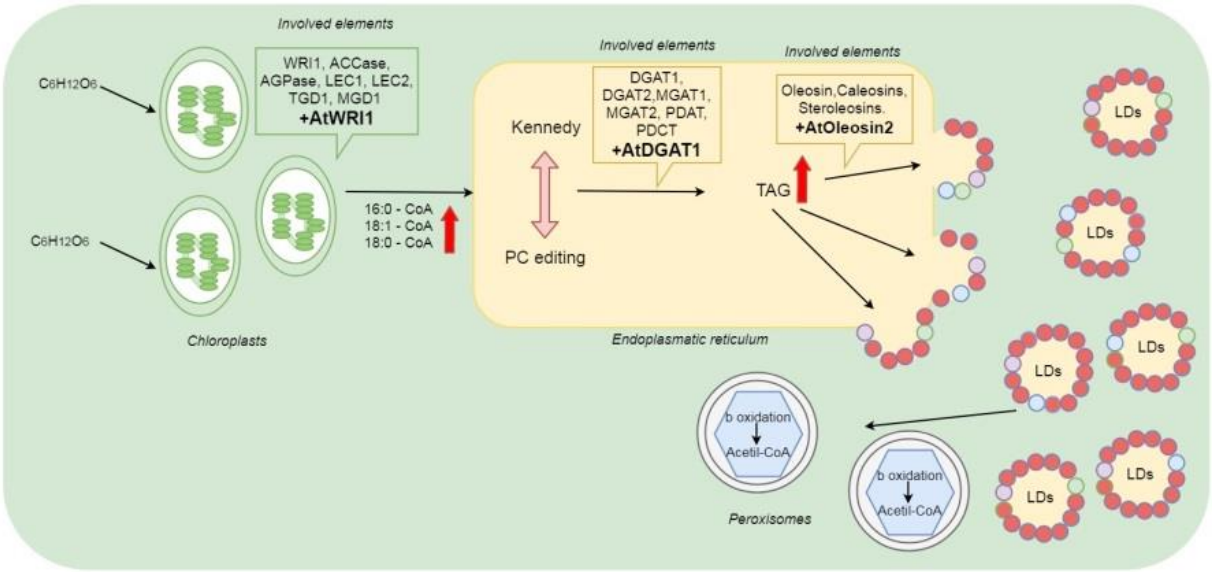


Figure D3. Co-expression of TAG inductors and oleosin fusions. Schematic representation of TAG accumulation and LD production in plant cells after co-agroinfiltration of DGAT1, WRI1 and oleosin fusion expression constructs. Blue, purple and green circles represent the structural native proteins forming LDs (oleosins, caleosins, steroleosins, among others) and red circles represent the oleosin fusions. Red arrows indicate the induced increase of the metabolic pathway elements.

Taking advantage of the fact that the LDs containing the oleosin-GFP fusions were easily monitored under fluorescent microscopy, we decided to use the recombinant LDs as biotechnological tools to test protease activities. There is a constant high demand, globally, for biotechnological tools with applications in platforms for production of recombinant proteins, in particular purification methods and proteases for specific release of products (Li et al. 2013). Our experience with potyviruses was an opportunity to generate new specific proteases. Potyviruses belong to the largest family of RNA plant viruses (Wylie et al. 2017) and they use a strategy of polyprotein processing by specific viral proteases for their genome expression. The large number of different potyviruses and the high specificity of their proteases (listed in the dedicated database http://www.dpvweb.net/potycleavage/species.html) was tempting to provide new protease candidates, and we decided to engineer a method to screen proteases activity using the NIa-Pro protease of the potyvirus Sweet potato feathery mottle virus (SPFMV) as proof of concept. Through co-agroinfiltration of oleosin fusions to GFP separated with the predicted NIa-Pro cleavage sites of SPFMV or Tobacco etch virus (TEV), together with the protease, we could test activities in vivo. We found that the SFPMV NIa-pro was specific for its own cleavage site,

since it did not recognize that of TEV, and the activity was easily visualized as a change in distribution of the GFP signal from being concentrated in LDs to present in the whole cytoplasm (Fig R74 to R78). The results of the test show that oleosin technology could be used for evaluating and characterizing plant virus-derived proteases.

EVALUATION OF THE ACTIVITY OF THE PLANT PRODUCED PEPTIDES

Finally, we evaluated the activity of the plant produced peptides at different levels of purification. Fibroblasts were selected as the target of IGF1 peptides because they are known to respond efficiently to IGF1 increasing their mitogenesis (Zheng *et al.*, 1997) and their collagen synthesis (Chetty *et al.*, 2006). The peptides produced in *D. carota* (IGF1 or CPP-IGF1) were present in cell-depleted media, while those produced in *N. benthamiana* were contained in ECFs (IGF1 or CPP-IGF1) or isolated LDs (AtOle-IGF1 or IGF1-AtOle). All the tested solutions enhanced fibroblast division in peptide concentrations up to 5ng/ml, showing their potential for cosmetic or clinical applications, with very high activity in the LDs containing oleosin fusions to IGF1.

To summarize, here we demonstrate the potential of plant matrices for the production of human peptides for industrial purposes. We constructed different binary plasmids and viral vectors that allowed us to produce IGF1 and CPP-IGF1 in two different plant matrices, plant cell cultures and plant leaves, with different market potential. In the case of plant cell cultures, we enhanced the production of the different peptides with co-expression of P1b, a plant virus-derived RNA silencing suppressor, and evaluated the growth parameters of the cultures and peptide accumulation dynamics over a number of propagation cycles. We evaluated transgenic cell lines after stable transformation, testing a cryopreservation strategy with successful results. Also, a strategy for extracting recombinant protein from plant leaves was engineered and tested. To do so, we developed plasmids for the expression of oleosin fusions, combined with a metabolic engineering strategy to accumulate recombinant LDs in plant leaves after transient expression of key regulatory and functional elements of the lipid-related pathways, resulting in a rapid, time-efficient and low cost alternative to the currently available plant seed-based strategies. Furthermore, we demonstrated that the plant produced IGF1, CPP-IGF1 and oleosin-IGF1 were biologically active, and therefore constituted a starting point for development of promising products, applicable to many potential fields such as skin care treatments.

CONCLUSIONS

CONCLUSIONS

1. P1b is a strong and durable RNA-silencing suppressor able to enhance transient gene expression in *N. benthamiana* leaves.

2. Plant cultured cells contain lipid droplets (LDs) in different amount and sizes, without any clear correlation with the plant tissue used to generate the cell line or the phylogeny of the plant species.

3. *D. carota* cell cultures can be easily stablished from root material, adapted to compatible industrial growth conditions, and transformed with *A. tumefaciens*.

4. Constitutive expression of the *Arabidopsis oleosin 2* (*AtOle2*) fusions to *GFP* in *D. carota* cells severely affected their growth.

5. *D. carota* cell lines can produce the human insulin growth factor 1 (IGF1) and tailored derivate peptides (CPP-IGF1) by stable transformation. Peptides can be secreted and accumulated into the culture media, facilitating their purification. When co-expressed with *P1b*, productions were boosted. Cryopreservation procedures could prevent the reduction in peptide accumulation through propagation cycles.

6. The *D. carota* cell suspensions can be transiently transformed using Plant Cell Packs (PCPs) technology.

7. Only one SAR is enough to normalize high expression of recombinant protein in plant production systems.

8. A new plasmid with SAR and *P1b* expression cassette was engineered for the production of recombinant proteins in plants.

9. IGF1 and CPP-IGF1 are produced in *N. benthamiana* plants using TMV-based viral vectors and easily purified when targeted to ECFs.

10. Transient expression in *N. benthamiana* leaves of *AtOle2* fusions to *GFP* (*AtOle2-GFP* or *GFP-AtOle2*) results in the formation of recombinant fluorescent LDs that accumulated only for a limited time (2 dpi). The LDs containing AtOle2-GFP aggregated, whereas GFP-AtOle2 LDs remained isolated.

11. We adapted the oleosin technology for transient expression of recombinant proteins in LDs of *N. benthamiana* leaves. The LD content in leaves can be increased by co-expressing the *A. thaliana* transcription factor *WRI1* and the enzyme *DGAT1.* These two elements together with P1b and oleosin fusions to GFP or IGF1 trigger the accumulation of recombinant products in LDs and allow their isolation.

12. The transient expression of AtOle2 fused to GFP in transgenic *N. benthamiana* plants constitutively expressing *WRI1* and *DGAT1* resulted in localization of GFP signal in LDs, redistributed to the citoplasm when coexpressed with a protease that specifically cleaves in a site engineered to allow separation of the GFP. The system can be used for *in vivo* evaluation of the specificity of plant virus proteases.

13. Plant produced IGF1, CPP-IGF1 and IGF1 fused to AtOle2 are active towards human fibroblast cell lines (MRC-5).

BIBLIOGRAPHY

BIBLIOGRAPHY

A. Mansour A. Al-Musa (1993) *Cucumber Vein Yellowing Virus; Host Range and Virus Vector Relationships*. *J. Phytopathol.*, **137**, 73– 8.

Abell, B.M., Holbrook, L.A., Abenes, M., Murphy, D.J., Hills, M.J., and Moloney, M.M. (1997) *Role of the proline knot motif in oleosin endoplasmic reticulum topology and oil body targeting*. *Plant Cell*, **9**, 1481–1493.

Adams, M.J., Antoniw, J.F., and Beaudoin, F. (2005) *Overview and analysis of the polyprotein cleavage sites in the family Potyviridae*. *Mol. Plant Pathol.*, **6**, 471–487.

Aguirre, G.A., Ita, J.R., Garza, R.G., and Castilla-Cortazar, I. (2016) *Insulin-like growth factor-1 deficiency and metabolic syndrome*. *J. Transl. Med.*, **14**, 1–23.

Allen, G.C., Spiker, S., and Thompson, W.F. (2000) *Use of matrix attachment regions (MARs) to minimize transgene silencing*. *Plant Mol Biol*, **43**, 361–376.

Andersen, M., Nørgaard-Pedersen, D., Brandt, J., Pettersson, I., and Slaaby, R. (2017) *IGF1 and IGF2 specificities to the two insulin receptor isoforms are determined by insulin receptor amino acid 718*. *PLoS One*, **12**, 1–14.

Arlotta, K.J. and Owen, S.C. (2019) *Antibody and antibody derivatives as cancer therapeutics*. *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology*, **11**, 1–16.

Arzola, L., Chen, J., Rattanaporn, K., Maclean, J.M., and 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*. *Int. J. Mol. Sci.*, **12**, 4975–4990.

Baghban, R., Farajnia, S., Rajabibazl, M., Ghasemi, Y., Mafi, A.A., Hoseinpoor, R., et al. (2019) *Yeast Expression Systems: Overview and Recent Advances*. *Mol. Biotechnol.*, **61**, 365–384.

Bahramnejad, B., Naji, M., Bose, R., and Jha, S. (2019) *A critical review on use of Agrobacterium rhizogenes and their associated binary vectors for plant transformation*. *Biotechnol. Adv.*, **37**, 107405.

Banilas, G., Daras, G., Rigas, S., Moloney, M.M., and Hatzopoulos, P. (2011) *Oleosin di-or tri-meric fusions with GFP undergo correct targeting and provide advantages for recombinant protein production*. *Plant Physiol. Biochem.*, **49**, 216–222.

Baud, S. and Lepiniec, L. (2010) *Physiological and developmental regulation of seed oil production*. *Prog. Lipid Res.*, **49**, 235–249.

Beatty, P.H. and Lewis, J.D. (2019) *Cowpea mosaic virus nanoparticles for cancer imaging and therapy*. *Adv. Drug Deliv. Rev.*, **145**, 130–144.

Beaudoin, F., Wilkinson, B.M., Stirling, C.J., and Napier, J.A. (2000) *In vivo targeting of a sunflower oil body protein in yeast secretory (sec) mutants*. *Plant J.*, **23**, 159–170.

Bethell, D.R. and Huang, J. (2004) *Recombinant human lactoferrin treatment for global health issues: Iron deficiency and acute diarrhea*. *BioMetals*, **17**, 337–342.

Boivin, E.B., Lepage, É.T., Matton, D.P., De Crescenzo, G., and Jolicoeur, M. (2010) *Transient expression of antibodies in suspension plant cell suspension cultures is enhanced when co-transformed with the tomato bushy stunt virus p19 viral suppressor of gene silencing*. *Biotechnol. Prog.*, **26**, 1534–

1543.

Boothe, J., Nykiforuk, C., Shen, Y., Zaplachinski, S., Szarka, S., Kuhlman, P., et al. (2010) *Seed-based expression systems for plant molecular farming*. *Plant Biotechnol. J.*, **8**, 588–606.

Bortesi, L., Rossato, M., Schuster, F., Raven, N., Stadlmann, J., Avesani, L., et al. (2009) *Viral and murine interleukin-10 are correctly processed and retain their biological activity when produced in tobacco*. *BMC Biotechnol.*, **9**.

Bosch, D., Castilho, A., Loos, A., Schots, A., and Steinkellner, H. (2013) *N-Glycosylation of Plant-produced Recombinant Proteins*. *Curr. Pharm. Des.*, **19**, 5503–5512.

Boucher, J., Tseng, Y.H., and Kahn, C.R. (2010) *Insulin and insulin-like growth factor-1 receptors act as ligand-specific amplitude modulators of a common pathway regulating gene transcription*. *J. Biol. Chem.*, **285**, 17235–17245.

Bourras, S., Rouxel, T., and Meyer, M. (2015) *Agrobacterium tumefaciens Gene Transfer: How a Plant Pathogen Hacks the Nuclei of Plant and Nonplant Organisms* . *Phytopathology*, **105**, 1288–1301.

Breyne, P., Van Montagu, M., Depicker, A., and Gheysen, G. (1992) *Characterization of a plant scaffold attachment region in a DNA fragment that normalizes transgene expression in tobacco*. *Plant Cell*, **4**, 463–471.

Bundó, M., Shi, X., Vernet, M., Marcos, J.F., López-García, B., and Coca, M. (2019) *Rice seeds as biofactories of rationally designed and cell-penetrating antifungal PAF peptides*. *Front. Plant Sci.*, **10**.

Bunney, P.E., Zink, A.N., Holm, A.A., Billington, C.J., and Kotz, C.M. (2017) *Orexin activation counteracts decreases in nonexercise activity thermogenesis (NEAT) caused by high-fat diet*. *Physiol. Behav.*, **176**, 139–148.

Cagliari, A., Margis, R., Dos Santos Maraschin, F., Turchetto-Zolet, A.C., Loss, G., and Margis-Pinheiro, M. (2011) *Biosynthesis of triacylglycerols (TAGs) in plants and algae*. *Int. J. Plant Biol.*, **2**, 40–52.

Cai, J., Wen, R., Li, W., Wang, X., Tian, H., Yi, S., et al. (2018) *Oil body bound oleosin-rhFGF9 fusion protein expressed in safflower (Carthamus tinctorius L.) stimulates hair growth and wound healing in mice*. *BMC Biotechnol.*, **18**, 1–11.

Carbonell, A., Dujovny, G., García, J.A., and Valli, A. (2012) *The Cucumber vein yellowing virus Silencing Suppressor P1b Can Functionally Replace HCPro in Plum pox virus Infection in a Host-Specific Manner*. */ 151 MPMI*, **25**, 151–164.

Cernac, A. and Benning, C. (2004) *WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis*. *Plant J.*, **40**, 575–585.

Chambers, A.C.; Aksular, M.; Graves, L.P.; Irons, S.L.; Possee, R.D.; King, L.. (2018) *Overview of the baculovirus expression system. Curr. Protoc. Protein Sci.*, **91**, 5.4.1–5.4.6.

Charron, C., Nicolaï, M., Gallois, J.L., Robaglia, C., Moury, B., Palloix, A., and Caranta, C. (2008) *Natural variation and functional analyses provide evidence for co-evolution between plant eIF4E and potyviral VPg*. *Plant J.*, **54**, 56–68.

Chaudhary, S., Parmenter, D.L., and Moloney, M.M. (1998) *Transgenic Brassica carinata as a vehicle for the production of recombinant proteins in seeds*. *Plant Cell Rep.*, **17**, 195–200.

Cheng, Z.M., Schnurr, J.A., and Kapaun, J.A. (1998) *Timentin as an alternative antibiotic for suppression*

of Agrobacterium tumefaciens in genetic transformation. *Plant Cell Rep.*, **17**, 646–649.

Chetty, A., Cao, G.J., and Nielsen, H.C. (2006) *Insulin-like growth factor-I signaling mechanisms, type I collagen and alpha smooth muscle actin in human fetal lung fibroblasts*. *Pediatr. Res.*, **60**, 389–394.

Chiang, C.J., Chen, H.C., Chad, Y.P., and Tzen, J.T.C. (2005) *Efficient system of artificial oil bodies for functional expression and purification of recombinant nattokinase in Escherichia coli*. *J. Agric. Food Chem.*, **53**, 4799–4804.

Chiang, C.J., Chen, H.C., Chao, Y.P., and Tzen, J.T.C. (2007) *One-step purification of insoluble hydantoinase overproduced in Escherichia coli*. *Protein Expr. Purif.*, **52**, 14–18.

Chilton, M.D., Drummond, M.H., Merlo, D.J., Sciaky, D., Montoya, A.L., Gordon, M.P., and Nester, E.W. (1977) *Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis*. *Cell*, **11**, 263–271.

Chitnis, M.M., Yuen, J.S.P., Protheroe, A.S., Pollak, M., and Macaulay, V.M. (2008) *The type 1 insulin-like growth factor receptor pathway*. *Clin. Cancer Res.*, **14**, 6364–6370.

Choudhary, S., Thakur, S., and Bhardwaj, P. (2019) *Molecular basis of transitivity in plant RNA silencing*. *Mol. Biol. Rep.*, **46**, 4645–4660.

Chuang, R.L.C., Chen, J.C.F., Chu, J., and Tzen, J.T.C. (1996) *Characterization Isoforms from of Seed Oil Bodies Rice Embryos1 Surface Oleosin these two isoforms were obtained , and compared with the deduced sequences of two maize Key words : J. Biochem.*, **81**, 74–81.

Chung, B.Y.-W., Miller, W.A., Atkins, J.F., and Firth, A.E. (2008) *An overlapping essential gene in the Potyviridae*. *Proc. Natl. Acad. Sci.*, **105**, 5897–5902.

Clark, C.A., Davis, J.A., Abad, J.A., Cuellar, W.J., Fuentes, S., Kreuze, J.F., et al. (2012a) *Sweetpotato Viruses: 15 Years of Progress on Understanding and Managing Complex Diseases*. *Plant Dis.*, **96**, 168–185.

Coca, M., Bortolotti, C., Rufat, M., Peñas, G., Eritja, R., Tharreau, D., et al. (2004) *Transgenic Rice Plants Expressing the Antifungal AFP Protein from Aspergillus Giganteus Show Enhanced Resistance to the Rice Blast Fungus Magnaporthe Grisea*. *Plant Mol. Biol.*, **54**, 245–259.

Craik, D.J., Fairlie, D.P., Liras, S., and Price, D. (2013) *The Future of Peptide-based Drugs*. *Chem. Biol. Drug Des.*, **81**, 136–147.

Csorba, T., Kontra, L., and Burgyán, J. (2015) *Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence*. *Virology*, **479–480**, 85–103.

Cui, X., Lu, L., Wang, Y., Yuan, X., and Chen, X. (2018) *The interaction of soybean reticulon homology domain protein (GmRHP) with Soybean mosaic virus encoded P3 contributes to the viral infection*. *Biochem. Biophys. Res. Commun.*, **495**, 2105–2110.

Cummings, J.F., Guerrero, M.L., Moon, J.E., Waterman, P., Nielsen, R.K., Jefferson, S., et al. (2014) *Safety and immunogenicity of a plant-produced recombinant monomer hemagglutinin-based influenza vaccine derived from influenza A (H1N1)pdm09 virus: A Phase 1 dose-escalation study in healthy adults*. *Vaccine*, **32**, 2251–2259.

Díaz-Pendón, J.A. and Ding, S.-W. (2008) *Direct and Indirect Roles of Viral Suppressors of RNA Silencing in Pathogenesis*. *Annu. Rev. Phytopathol.*, **46**, 303–326.

Ding, S.W. and Voinnet, O. (2007) *Antiviral Immunity Directed by Small RNAs*. *Cell*, **130**, 413–426.

Domingo-Calap, M.L., Moreno, A.B., Pendón, J.A.D., Moreno, A., Fereres, A., and López-Moya, J.J. (2020) *Assessing the impact on virus transmission and insect vector behavior of a viral mixed infection in melon*. *Phytopathology*, **110**, 174–186.

Dyer, J.M., Stymne, S., Green, A.G., and Carlsson, A.S. (2008) *High-value oils from plants*. *Plant J.*, **54**, 640–655.

E. Kurstak ed. (1982) *Handbook of Plant Virus Infections: Comparative Diagnosis (Book). Plant. Cell Environ.*, **5**, 188–189.

van Eerde, A., Gottschamel, J., Bock, R., Hansen, K.E.A., Munang'andu, H.M., Daniell, H., and Liu Clarke, J. (2019) *Production of tetravalent dengue virus envelope protein domain III based antigens in lettuce chloroplasts and immunologic analysis for future oral vaccine development*. *Plant Biotechnol. J.*, **17**, 1408–1417.

Ellstrand, N.C. (2003) *Going to " Great Lengths " to Prevent the Escape of Genes That Produce Specialty Chemicals*. *Plant Physiol.*, **132**, 1770–1774.

Fehér, A. (2015) *Somatic embryogenesis - stress-induced remodeling of plant cell fate*. *Biochim. Biophys. Acta - Gene Regul. Mech.*, **1849**, 385–402.

Ferrari, S. (2010) *Biological Elicitors of Plant Secondary Metabolites: Mode of Action and Use in the Production of Nutraceutics*. In: *Advances in experimental medicine and biology (Book)*, pp. 152–166.

Fischer, R., Schillberg, S., Buyel, J., and Twyman, R. (2013) *Commercial Aspects of Pharmaceutical Protein Production in Plants*. *Curr. Pharm. Des.*, **19**, 5471–5477.

Fischer, R., Schillberg, S., Hellwig, S., Twyman, R.M., and Drossard, J. (2012) *GMP issues for recombinant plant-derived pharmaceutical proteins*. *Biotechnol. Adv.*, **30**, 434–439.

Gao, J. and Lee, J.M. (1992) *Effect of Oxygen Supply on the Suspension Culture of Genetically Modified Tobacco Cells*. *Biotechnol. Prog.*, **8**, 285–290.

Gidda, S.K., Park, S., Pyc, M., Yurchenko, O., Cai, Y., Wu, P., et al. (2016) *Lipid droplet-associated proteins (LDAPs) are required for the dynamic regulation of neutral lipid compartmentation in plant cells*. *Plant Physiol.*, **170**, 2052–2071.

Giner, A., Lakatos, L., García-Chapa, M., López-Moya, J.J., and Burgyán, J. (2010) *Viral protein inhibits RISC activity by argonaute binding through conserved WG/GW motifs*. *PLoS Pathog.*, **6**, 1–13.

Gleba, Y., Klimyuk, V., and Marillonnet, S. (2007) *Viral vectors for the expression of proteins in plants*. *Curr. Opin. Biotechnol.*, **18**, 134–141.

Gleba, Y., Marillonnet, S., and Klimyuk, V. (2004) *Engineering viral expression vectors for plants: The "full virus" and the "deconstructed virus" strategies*. *Curr. Opin. Plant Biol.*, **7**, 182–188.

Goeddel, D. V., Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Crea, R., et al. (1979) *Expression in Escherichia coli of chemically synthesized genes for human insulin. Proc. Natl. Acad. Sci.*, **76**, 106–110.

González-Gamboa, I., Manrique, P., Manrique, P., and Ponz, F. (2017) *Plant-made potyvirus-like particles used for log-increasing antibody sensing capacity*. *J. Biotechnol.*, **254**, 17–24.

Goodman, J.M. (2008) *The gregarious lipid droplet*. *J. Biol. Chem.*, **283**, 28005–28009.

Greenspan, P., Mayer, E.P., and Fowler, S.D. (1985) *Nile red: A selective fluorescent stain for intracellular lipid droplets*. *J. Cell Biol.*, **100**, 965–973.

Grennan, A.K. (2006) *Plant response to bacterial pathogens. Overlap between innate and gene-for-gene defense response*. *Plant Physiol.*, **142**, 809–811.

Guidotti, G., Brambilla, L., and Rossi, D. (2017) *Cell-Penetrating Peptides: From Basic Research to Clinics*. *Trends Pharmacol. Sci.*, **38**, 406–424.

Han, Z., Madzak, C., and Su, W.W. (2013) *Tunable nano-oleosomes derived from engineered Yarrowia lipolytica*. *Biotechnol. Bioeng.*, **110**, 702–710.

Hara, Y., Morimoto, T., and Fujita, Y. (2004) *Production of shikonin derivatives by cell suspension cultures of Lithospermum erythrorhizon*. *Plant Cell Rep.*, **6**, 8–11.

Hefferon, K. (2017) *Plant Virus Expression Vectors: A Powerhouse for Global Health*. *Biomedicines*, **5**, 44.

Heilmann, M., Iven, T., Ahmann, K., Hornung, E., Stymne, S., and Feussner, I. (2012) *Production of wax esters in plant seed oils by oleosomal cotargeting of biosynthetic enzymes*. *J. Lipid Res.*, **53**, 2153–2161.

Hiatt, A., Pauly, M., Whaley, K., Qiu, X., Kobinger, G., and Zeitlin, L. (2015) *The emergence of antibody therapies for Ebola*. *Hum. Antibodies*, **23**, 49–56.

Hibino, K. and Ushiyama, K. (1999) *Commercial Production of Ginseng by Plant Tissue Culture Technology*. In: *Plant Cell and Tissue Culture for the Production of Food Ingredients (Book)* , pp. 215–224. Boston, MA: Springer.

Himada, T.L.S. and Ishimura, I.H.A.R.A. (2010) *Intracellular Lipid Droplet-Associated Proteins : Unique Members and Their Biological Functions Oil-Body-Membrane Proteins and Their Physiological Functions in Plants*. *Biol. Pharm. Bull.*, **33**, 360–363.

Hong, S.Y., Kwon, T.H., Lee, J.H., Jang, Y.S., and Yang, M.S. (2002) *Production of biologically active hG-CSF by transgenic plant cell suspension culture*. *Enzyme Microb. Technol.*, **30**, 763–767.

Honga, S.Y., Kwona, T.H., Leeb, J.H., Jangc, Y.S., and Moon-Sik Y. (2002) *Production of biologically active hG-CSF by transgenic plant cell suspension culture*. *Enzyme Microb. Technol.*, **30**, 263/767.

Hsieh, S.K., Yu, Y.J., Tang, N.Y., Lin, J.R., and Jinn, T.-R. (2017) *Expression of Mastoparan B, a Venom Peptide, Via Escherichia coli C43 (DE3) Coupled with an Artificial Oil Body-Cyanogen Bromide Technology Platform*. *Protein Pept. Lett.*, **24**, 1021–1029.

Huang, A.H.C. (2018) *Plant lipid droplets and their associated proteins: Potential for rapid advances*. *Plant Physiol.*, **176**, 1894–1918.

Huang, M. Der and Huang, A.H.C. (2015) *Bioinformatics reveal five lineages of oleosins and the mechanism of lineage evolution related to structure/function from green algae to seed plants*. *Plant Physiol.*, **169**, 453–470.

Huang, T.K., Falk, B.W., Dandekar, A.M., and McDonald, K.A. (2018) *Enhancement of recombinant protein production in transgenic Nicotiana benthamiana plant cell suspension cultures with co-cultivation of Agrobacterium containing silencing suppressors*. *Int. J. Mol. Sci.*, **19**, 21–25.

James, E. and Lee, J.M. (2006) *Loss and recovery of protein productivity in genetically modified plant cell lines*. *Plant Cell Rep.*, **25**, 723–727.

Jauset, T. and Beaulieu, M.E. (2019) *Bioactive cell penetrating peptides and proteins in cancer: a bright future ahead*. *Curr. Opin. Pharmacol.*, **47**, 133–140.

Kalafatovic, D. and Giralt, E. (2017) *Cell-Penetrating Peptides: Design Strategies beyond Primary Structure and Amphipathicity*. *Molecules*, **22**, 1929.

Kasschau, K.D. and Carrington, J.C. (1998) *A counterdefensive strategy of plant viruses: Suppression of posttranscriptional gene silencing*. *Cell*, **95**, 461–470.

Kendall, A., McDonald, M., Bian, W., Bowles, T., Baumgarten, S.C., Shi, J., et al. (2008) *Structure of Flexible Filamentous Plant Viruses*. *J. Virol.*, **82**, 9546–9554.

Kim, M., O'Callaghan, P.M., Droms, K.A., and James, D.C. (2011) *A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies*. *Biotechnol. Bioeng.*, **108**, 2434–2446.

Kirchhoff, J., Raven, N., Boes, A., Roberts, J.L., Russell, S., Treffenfeldt, W., et al. (2012) *Monoclonal tobacco cell lines with enhanced recombinant protein yields can be generated from heterogeneous cell suspension cultures by flow sorting*. *Plant Biotechnol. J.*, **10**, 936–944.

Komori, T., Imayama, T., Kato, N., Ishida, Y., Ueki, J., and Komari, T. (2007) *Current Status of Binary Vectors and Superbinary Vectors*. *Plant Physiol.*, **145**, 1155–1160.

Kopertekh, L. and Schiemann, J. (2019) *Transient Production of Recombinant Pharmaceutical Proteins in Plants: Evolution and Perspectives*. *Curr. Med. Chem.*, **26**, 365–380.

Krenek, P., Samajova, O., Luptovciak, I., Doskocilova, A., Komis, G., and Samaj, J. (2015) *Transient plant transformation mediated by Agrobacterium tumefaciens: Principles, methods and applications*. *Biotechnol. Adv.*, **33**, 1024–1042.

Kwon, T.H., Kim, Y.S., Lee, J.H., and Yang, M.S. (2003) *Production and secretion of biologically active human granulocyte-macrophage colony stimulating factor in transgenic tomato suspension cultures*. *Biotechnol. Lett.*, **25**, 1571–1574.

Lacey, D.J., Wellner, N., Beaudoin, F., Napier, J. a, and Shewry, P.R. (1998) *safflower (Carthamus tinctorius L .) and sunflower (Helianthus annuus L .)*. *Proteins*, **477**, 469–477.

Lacomme, C., Molnar, A., Csorba, T., Lakatos, L., Varallyay, E., and Burgyan, J. (2005) *Plant Virus-Derived Small Interfering RNAs Originate Predominantly from Highly Structured Single-Stranded Viral RNAs*. *J. Virol.*, **79**, 7812–7818.

LaCount, W., An, G., and Lee, J.M. (1997) *The effect of polyvinylpyrrolidone (PVP) on the heavy chain monoclonal antibody production from plant suspension cultures*. *Biotechnol. Lett.*, **19**, 93–96.

LAEMMLI, U. (1970) *Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4.* Nature, **227**, 680–685.

Lakatos, L., Csorba, T., Pantaleo, V., Chapman, E.J., Carrington, J.C., Liu, Y.P., et al. (2006) *Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors*. *EMBO J.*, **25**, 2768–2780.

Lam, P., Gulati, N.M., Stewart, P.L., Keri, R.A., and Steinmetz, N.F. (2016) *Bioengineering of Tobacco Mosaic Virus to Create a Non-Infectious Positive Control for Ebola Diagnostic Assays*. *Sci. Rep.*, **6**, 1–8.

Landry, N., Ward, B.J., Trépanier, S., Montomoli, E., Le Dargis, M., Lapini, G., and Vézina, L.P. (2010)

Preclinical and clinical development of plant-made virus-like particle vaccine against avian H5N1 influenza. *PLoS One*, **5**.

Leader, B., Baca, Q.J., and Golan, D.E. (2008) *Protein therapeutics: A summary and pharmacological classification*. *Nat. Rev. Drug Discov.*, **7**, 21–39.

Lecoq H (1992) *Les virus des cultures de melon et de courgette de plein champ (I)*. *PHM Rev Hortic.*, **324**, 15–25.

Lee, T.T.T., Leu, W.M., Yang, H.H., Chen, B.C.M., and Tzen, J.T.C. (2006) *Sesame oleosin and prepro-2S albumin expressed as a fusion polypeptide in transgenic rice were split, processed and separately assembled into oil bodies and protein bodies*. *J. Cereal Sci.*, **44**, 333–341.

Lee, W.S., Tzen, J.T.C., Kridl, J.C., Radke, S.E., and Huang, A.H.C. (1991) *Maize oleosin is correctly targeted to seed oil bodies in Brassica napus transformed with the maize oleosin gene*. *Proc. Natl. Acad. Sci. U. S. A.*, **88**, 6181–6185.

Leonard, S., Plante, D., Wittmann, S., Daigneault, N., Fortin, M.G., and Laliberte, J.-F. (2000) *Complex Formation between Potyvirus VPg and Translation Eukaryotic Initiation Factor 4E Correlates with Virus Infectivity*. *J. Virol.*, **74**, 7730–7737.

Li, F. and Wang, A. (2019) *RNA-Targeted Antiviral Immunity: More Than Just RNA Silencing*. *Trends Microbiol.*, **27**, 792–805.

Li, F., Xu, D., Abad, J., and Li, R. (2012) *Phylogenetic relationships of closely related potyviruses infecting sweet potato determined by genomic characterization of Sweet potato virus G and Sweet potato virus 2*. *Virus Genes*, **45**, 118–125.

Li, Hongrui, Yang, J., Chen, Y., Guan, L., Du, L.N., Guo, Y.X., et al. (2014) *Expression of a functional recombinant oleosin-human hyaluronidase hPH-20 fusion in Arabidopsis thaliana*. *Protein Expr. Purif.*, **103**, 23–27.

Li, J.F., Park, E., Von Arnim, A.G., and Nebenführ, A. (2009) *The FAST technique: A simplified Agrobacterium-based transformation method for transient gene expression analysis in seedlings of Arabidopsis and other plant species*. *Plant Methods*, **5**, 1–15.

Li, M., Murphy, D.J., Lee, K.H.K., Wilson, R., Smith, L.J., Clark, D.C., and Sung, J.Y. (2002) *Purification and structural characterization of the central hydrophobic domain of oleosin*. *J. Biol. Chem.*, **277**, 37888–37895.

Li, W., Li, L.G., Sun, X.F., and Tang, K.X. (2012) *An oleosin-fusion protein driven by the CaMV35S promoter is accumulated in Arabidopsis (Brassicaceae) seeds and correctly targeted to oil bodies. Genet. Mol. Res.*, **11**, 2138–2146.

Li, W., Yang, J., Cai, J., Wang, H., Tian, H., Huang, J., et al. (2017) *Oil body-bound oleosin-rhFGF-10: A novel drug delivery system that improves skin penetration to acceleratewound healing and hair growth in mice*. *Int. J. Mol. Sci.*, **18**.

Lin, L.J., Tai, S.S.K., Peng, C.C., and Tzen, J.T.C. (2002) *Erratum: Steroleosin, a sterol-binding dehydrogenase in seed oil bodies (Plant Physiology (2002) 128 (1200-1211))*. *Plant Physiol.*, **129**, 1930.

Lindbo, J. A. (2007) *TRBO: A High-Efficiency Tobacco Mosaic Virus RNA-Based Overexpression Vector*. *Plant Physiol.*, **145**, 1232–1240.

Ling, Z., Qi-Qing, J., Yu, W., Fei, B., Shu-Jie, Q., Shu-Bo, W., et al. (2016) *Transgenic Expression and*

Identification of Recombinant Human Proinsulin in Peanut. *Brazilian Arch. Biol. Technol.*, **59**, 1–9.

Liu, J.H., Selinger, L.B., Cheng, K.J., Beauchemin, K.A., and Moloney, M.M. (1997) *Plant seed oil-bodies as an immobilization matrix for a recombinant xylanase from the rumen fungus Neocallimastix patriciarum*. *Mol. Breed.*, **3**, 463–470.

Lu, R., Martin-Hernandez, A.M., Peart, J.R., Malcuit, I., and Baulcombe, D.C. (2003) *Virus-induced gene silencing in plants*. *Methods*, **30**, 296–303.

Lui, J.C., Colbert, M., Cheung, C.S.F., Ad, M., Lee, A., Zhu, Z., et al. (2019) *Cartilage-Targeted IGF-1 Treatment to Promote Longitudinal Bone Growth*. *Mol. Ther.*, **27**, 673–680.

Magnuson, N.S., Linzmaier, P.M., Reeves, R., An, G., HayGlass, K., and Lee, J.M. (1998) *Secretion of biologically active human interleukin-2 and interleukin-4 from genetically modified tobacco cells in suspension culture*. *Protein Expr. Purif.*, **13**, 45–52.

Magy, B., Tollet, J., Laterre, R., Boutry, M., and Navarre, C. (2014) *Accumulation of secreted antibodies in plant cell cultures varies according to the isotype, host species and culture conditions*. *Plant Biotechnol. J.*, **12**, 457–467.

Mamat, U., Wilke, K., Bramhill, D., Schromm, A.B., Lindner, B., Kohl, T.A., et al. (2015) *Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins*. *Microb. Cell Fact.*, **14**, 1–15.

Mansour, A.A., Banik, S., Suresh, R. V., Kaur, H., Malik, M., McCormick, A.A., and Bakshi, C.S. (2018) *An improved tobacco mosaic virus (TMV)-conjugated multiantigen subunit vaccine against respiratory tularemia*. *Front. Microbiol.*, **9**, 1–21.

Mansour, M.P., Shrestha, P., Belide, S., Petrie, J.R., Nichols, P.D., and Singh, S.P. (2014) *Characterization of oilseed lipids from "DHA-producing Camelina sativa": A new transformed land plant containing long-chain omega-3 oils*. *Nutrients*, **6**, 776–789.

Mardanova, E.S., Blokhina, E.A., Tsybalova, L.M., Peyret, H., Lomonossoff, G.P., and Ravin, N. V. (2017) *Efficient transient expression of recombinant proteins in plants by the novel pEff vector based on the genome of potato virus X*. *Front. Plant Sci.*, **8**, 1–8.

Mardanova, E.S., Kotlyarov, R.Y., Kuprianov, V. V., Stepanova, L.A., Tsybalova, L.M., Lomonosoff, G.P., and Ravin, N. V. (2015) *Rapid high-yield expression of a candidate influenza vaccine based on the ectodomain of M2 protein linked to flagellin in plants using viral vectors*. *BMC Biotechnol.*, **15**, 1–10.

Marillonnet, S., Giritch, A., Gils, M., Kandzia, R., Klimyuk, V., and Gleba, Y. (2004) *In planta engineering of viral RNA replicons: Efficient assembly by recombination of DNA modules delivered by Agrobacterium*. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 6852–6857.

Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V., and Gleba, Y. (2005) *Systemic Agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants*. *Nat. Biotechnol.*, **23**, 718–723.

Martin, R.R., Zhou, J., and Tzanetakis, I.E. (2011) *Blueberry latent virus: An amalgam of the Partitiviridae and Totiviridae*. *Virus Res.*, **155**, 175–180.

Martin, S. and Parton, R.G. (2006) *Lipid droplets: a unified view of a dynamic organelle*. *Nat. Rev. Mol. Cell Biol.*, **7**, 373–378.

Mayo, M.A. (1995) *The Potyviridae*. *New Phytol.*, **131**, 289–290.

McDonald, M., Kendall, A., Bian, W., McCullough, I., Lio, E., Havens, W.M., et al. (2010) *Architecture of the potyviruses*. *Virology*, **405**, 309–313.

Mei, Y.-Z., Zhu, Y.-L., Huang, P.-W., Yang, Q., and Dai, C.-C. (2019) *Strategies for gene disruption and expression in filamentous fungi*. *Appl. Microbiol. Biotechnol.*, **103**, 6041–6059.

Menkhaus, T.J., Bai, Y., Zhang, C., Nikolov, Z.L., and Glatz, C.E. (2004) *Considerations for the recovery of recombinant proteins from plants*. *Biotechnol. Prog.*, **20**, 1001–1014.

Milletti, F. (2012) *Cell-penetrating peptides: Classes, origin, and current landscape*. *Drug Discov. Today*, **17**, 850–860.

Mingot, A., Valli, A., Rodamilans, B., San León, D., Baulcombe, D.C., García, J.A., and López-Moya, J.J. (2016) *The P1N-PISPO trans -Frame Gene of Sweet Potato Feathery Mottle Potyvirus Is Produced during Virus Infection and Functions as an RNA Silencing Suppressor*. *J. Virol.*, **90**, 3543–3557.

Mojzita, D., Rantasalo, A., and Jäntti, J. (2019) *Gene expression engineering in fungi*. *Curr. Opin. Biotechnol.*, **59**, 141–149.

Montague, N.P., Thuenemann, E.C., Saxena, P., Saunders, K., Lenzi, P., and Lomonossoff, G.P. (2011) *Recent advances of cowpea mosaic virus-based particle technology*. *Hum. Vaccin.*, **7**, 383–390.

Montesinos, L., Bundó, M., Izquierdo, E., Campo, S., Badosa, E., Rossignol, M., et al. (2016) *Production of Biologically Active Cecropin A Peptide in Rice Seed Oil Bodies*. *PLoS One*, **11**, e0146919.

Moremen, K.W., Tiemeyer, M., and Nairn, A. V. (2012) *Vertebrate protein glycosylation: diversity, synthesis and function*. *Nat. Rev. Mol. Cell Biol.*, **13**, 448–462.

Mosmann, T. (1983) *Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays*. *J. Immunol. Methods*, **65**, 55–63.

Murashige, T. and Skoog, F. (1962) *Murashige1962Revised.Pdf*. *Physiol. Plant.*, **15**, 474–497.

Murphy, J.F., Klein, P.G., Hunt, A.G., and Shaw, J.G. (1996) *Replacement of the tyrosine residue that links a potyviral VPg to the viral RNA is lethal*. *Virology*, **220**, 535–538.

Murphy, J.F., Rychlik, W., Rhoads, R.E., Hunt, A.G., and Shaw, J.G. (1991) *A tyrosine residue in the small nuclear inclusion protein of tobacco vein mottling virus links the VPg to the viral RNA. J. Virol.*, **65**, 511–3.

Mustafa, N.R., De Winter, W., Van Iren, F., and Verpoorte, R. (2011) *Initiation, growth and cryopreservation of plant cell suspension cultures*. *Nat. Protoc.*, **6**, 715–742.

Nakagawa, T., Ishiguro, S., and Kimura, T. (2009) *Gateway vectors for plant transformation*. *Plant Biotechnol.*, **26**, 275–284.

Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., et al. (2007) *Improved gateway binary vectors: High-performance vectors for creation of fusion constructs in transgenic analysis of plants*. *Biosci. Biotechnol. Biochem.*, **71**, 2095–2100.

Nakayashiki, H. (2005) *RNA silencing in fungi: Mechanisms and applications*. *FEBS Lett.*, **579**, 5950–5957.

Newman, C.N. (2011) *Production of human erythropoietin in transgenic canola employing the technology of oleosin fusion*. *J. Undergrad. Res. Ohio State*, **3**, 70–85.

Noordam, R., Gunn, D.A., Tomlin, C.C., Maier, A.B., Griffiths, T., Catt, S.D., et al. (2013) *Serum insulin-*

like growth factor 1 and facial ageing: High levels associate with reduced skin wrinkling in a cross-sectional study. *Br. J. Dermatol.*, **168**, 533–538.

Nürnberger, T., Brunner, F., Kemmerling, B., and Piater, L. (2004) *Innate immunity in plants and animals: Striking similarities and obvious differences*. *Immunol. Rev.*, **198**, 249–266.

Nykiforuk, C.L., Boothe, J.G., Murray, E.W., Keon, R.G., Goren, H.J., Markley, N.A., and Moloney, M.M. (2006) *Transgenic expression and recovery of biologically active recombinant human insulin from Arabidopsis thaliana seeds*. *Plant Biotechnol. J.*, **4**, 77–85.

Nykiforuk, C.L., Shen, Y., Murray, E.W., Boothe, J.G., Busseuil, D., Rhéaume, E., et al. (2011) *Expression and recovery of biologically active recombinant Apolipoprotein AIMilano from transgenic safflower (Carthamus tinctorius) seeds*. *Plant Biotechnol. J.*, **9**, 250–263.

Ogawa, Y., Sakurai, N., Oikawa, A., Kai, K., Morishita, Y., Mori, K., et al. (2012) *High-throughput cryopreservation of plant cell cultures for functional genomics*. *Plant Cell Physiol.*, **53**, 943–952.

Olspert, A., Chung, B.Y.-W., Atkins, J.F., Carr, J.P., and Firth, A.E. (2015) *Transcriptional slippage in the positive-sense RNA virus family Potyviridae*. *EMBO Rep.*, **16**, 995–1004.

Palmer, K. and Cooper, M.D. (2014) *Plant Viral Vectors. Current Topics in Microbiology and Immunology*, (Palmer,K. and Gleba,Y., eds). Berlin, Heidelberg: Springer.

Parmenter, D.L., Boothe, J.G., van Rooijen, G.J.H., Yeung, E.C., and Moloney, M.M. (1995) *Production of biologically active hirudin in plant seeds using oleosin partitioning*. *Plant Mol. Biol.*, **29**, 1167–1180.

Parthibane, V., Rajakumari, S., Venkateshwari, V., Iyappan, R., and Rajasekharan, R. (2012) *Oleosin is bifunctional enzyme that has both monoacylglycerol acyltransferase and phospholipase activities*. *J. Biol. Chem.*, **287**, 1946–1954.

Pasin, F., Menzel, W., and Daròs, J.A. (2019) *Harnessed viruses in the age of metagenomics and synthetic biology: an update on infectious clone assembly and biotechnologies of plant viruses*. *Plant Biotechnol. J.*, **17**, 1010–1026.

Peng, C.-C., Chen, J.C.., Shyu, D.J.., Chen, M.-J., and Tzen, J.T.. (2004) *A system for purification of recombinant proteins in Escherichia coli via artificial oil bodies constituted with their oleosin-fused polypeptides*. *J. Biotechnol.*, **111**, 51–57.

Peng, C.C., Shyu, D.J.H., Chou, W.M., Chen, M.J., and Tzen, J.T.C. (2004) *Method for Bacterial Expression and Purification of Sesame Cystatin via Artificial Oil Bodies*. *J. Agric. Food Chem.*, **52**, 3115–3119.

Permyakova, N. V., Zagorskaya, A.A., Belavin, P.A., Uvarova, E.A., Nosareva, O. V., Nesterov, A.E., et al. (2015) *Transgenic carrot expressing fusion protein comprising M. tuberculosis antigens induces immune response in mice*. *Biomed Res. Int.*, **2015**.

Peyret, H., Brown, J.K.M., and Lomonossoff, G.P. (2019) *Improving plant transient expression through the rational design of synthetic 5′ and 3′ untranslated regions*. *Plant Methods*, **15**, 1–13.

Peyret, H. and Lomonossoff, G.P. (2015) *When plant virology met Agrobacterium: The rise of the deconstructed clones*. *Plant Biotechnol. J.*, **13**, 1121–1135.

Popa, C., Shi, X., Ruiz, T., Ferrer, P., and Coca, M. (2019) *Biotechnological production of the cell penetrating antifungal PAF102 peptide in pichia pastoris*. *Front. Microbiol.*, **10**.

Prochiantz, A. (2000) *Messenger proteins: Homeoproteins, TAT and others*. *Curr. Opin. Cell Biol.*, **12**,

400–406.

Pumplin, N. and Voinnet, O. (2013) *RNA silencing suppression by plant pathogens: Defence, counter-defence and counter-counter-defence*. *Nat. Rev. Microbiol.*, **11**, 745–760.

Purcifull, D.E. (1979) *Serological Distinction of Watermelon Mosaic Virus Isolates*. *Phytopathology*, **69**, 112.

Qiu, X., Wong, G., Audet, J., Bello, A., Fernando, L., Alimonti, J.B., et al. (2014) *Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp ™ wrote the manuscript HHS Public Access*. *Nature*, **514**, 47–53.

Rademacher, T., Sack, M., Blessing, D., Fischer, R., Holland, T., and Buyel, J. (2019) *Plant cell packs: a scalable platform for recombinant protein production and metabolic engineering*. *Plant Biotechnol. J.*, **17**, 1560–1566.

Rainer, F., Schumann, D., Zimmermann, S., Drossard, J., Sack, M., and Schillberg, S. (1999) *Expression and characterization of bispecific single-chain Fv fragments produced in transgenic plants*. *Eur. J. Biochem.*, **262**, 810–816.

Rajakumari, S., Grillitsch, K., and Daum, G. (2008) *Synthesis and turnover of non-polar lipids in yeast*. *Prog. Lipid Res.*, **47**, 157–171.

Ranke, M.B., Wölfle, J., Schnabel, D., and Bettendorf, M. (2009) *Treatment of Dwarfism With Recombinant Human Insulin-Like Growth Factor-1*. *Dtsch. Arztebl.*, **106**, 703–709.

Raven, N., Rasche, S., Kuehn, C., Anderlei, T., Klöckner, W., Schuster, F., et al. (2015) *Scaled-up manufacturing of recombinant antibodies produced by plant cells in a 200-L orbitally-shaken disposable bioreactor*. *Biotechnol. Bioeng.*, **112**, 308–321.

Rinderknecht, E. and Humbel, R.E. (1978) *The amino acid sequence of human insulin-like growth factor I and its structural homology with proinsulin*. *J. Biol. Chem.*, **253**, 2769–2776.

Rodamilans, B., Valli, A., Mingot, A., San León, D., Baulcombe, D., López-Moya, J.J., and García, J.A. (2015) *RNA Polymerase Slippage as a Mechanism for the Production of Frameshift Gene Products in Plant Viruses of the Potyviridae Family*. *J. Virol.*, **89**, 6965–6967.

Van Rooijen, G.J.H. and Moloney, M.M. (1995) *Plant Seed oil-bodies as carriers for foreign proteins*. *Bio/Technology*, **13**, 72–77.

Rosano, G.L. and Ceccarelli, E.A. (2014) *Recombinant protein expression in Escherichia coli: Advances and challenges*. *Front. Microbiol.*, **5**, 1–17.

Rosenfeld, R.G., Hwa, V., Wilson, L., Lopez-bermejo, A., Buckway, C., Burren, C., et al. (1999) *The Insulin-like Growth Factor Binding Protein Superfamily: New Perspectives*. *Pediatrics*, **104**.

Routh, S. and Nandagopal, K. (2017) *Patent Survey of Resveratrol, Taxol, Podophyllotoxin, Withanolides and Their Derivatives Used in Anticancer Therapy*. *Recent Pat. Biotechnol.*, **11**, 85–100.

Rubio-Huertos M, L.-A.D. (1966) *Ultraestructura de células de pimiento infectadas con un virus y su localización en las mismas*. *Microbiol. Española*, **19**, 77–86.

Ruiz-Ferrer, V. and Voinnet, O. (2009) *Roles of Plant Small RNAs in Biotic Stress Responses*. *Annu. Rev. Plant Biol.*, **60**, 485–510.

Ruiz, V., Baztarrica, J., Rybicki, E.P., Meyers, A.E., and Wigdorovitz, A. (2018) *Minimally processed crude leaf extracts of Nicotiana benthamiana containing recombinant foot and mouth disease virus-like*

particles are immunogenic in mice. *Biotechnol. Reports*, **20**.

Sabanadzovic, S., Ghanem-Sabanadzovic, N.A., and Valverde, R.A. (2010) *A novel monopartite dsRNA virus from rhododendron*. *Arch. Virol.*, **155**, 1859–1863.

Sabanadzovic, S., Valverde, R.A., Brown, J.K., Martin, R.R., and Tzanetakis, I.E. (2009) *Southern tomato virus: The link between the families Totiviridae and Partitiviridae*. *Virus Res.*, **140**, 130–137.

Sack, M., Paetz, A., Kunert, R., Bomble, M., Hesse, F., Stiegler, G., et al. (2007) *Functional analysis of the broadly neutralizing human anti-HIV-1 antibody 2F5 produced in transgenic BY-2 suspension cultures*. *FASEB J.*, **21**, 1655–1664.

Sainsbury, F. and Lomonossoff, G.P. (2014) *Transient expressions of synthetic biology in plants*. *Curr. Opin. Plant Biol.*, **19**, 1–7.

Sainsbury, F., Thuenemann, E.C., and Lomonossoff, G.P. (2009) *PEAQ: Versatile expression vectors for easy and quick transient expression of heterologous proteins in plants*. *Plant Biotechnol. J.*, **7**, 682–693.

Sarmiento, C., Ross, J.H.E., Herman, E., and Murphy, D.J. (1997) *Expression and subcellular targeting of a soybean oleosin in transgenic rapeseed. Implications for the mechanism of oil-body formation in seeds*. *Plant J.*, **11**, 783–796.

Sato, F. and Yamada, Y. (1984) *High berberine-producing cultures of coptis japonica cells*. *Phytochemistry*, **23**, 281–285.

Saunders, K., Sainsbury, F., and Lomonossoff, G.P. (2009) *Efficient generation of cowpea mosaic virus empty virus-like particles by the proteolytic processing of precursors in insect cells and plants*. *Virology*, **393**, 329–337.

Schägger, H. and von Jagow, G. (1987) *Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa*. *Anal. Biochem.*, **166**, 368–379.

Scholthof, K.B.G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., et al. (2011) *Top 10 plant viruses in molecular plant pathology*. *Mol. Plant Pathol.*, **12**, 938–954.

Scholz, J. and Suppmann, S. (2017) *A new single-step protocol for rapid baculovirus-driven protein production in insect cells*. *BMC Biotechnol.*, **17**, 1–9.

van der Schoot, C., Paul, L.K., Paul, S.B., and Rinne, P.L.H. (2011) *Plant lipid bodies and cell-cell signaling a new role for an old organelle? Plant Signal. Behav.*, **6**, 1732–1738.

Schuster, S., Miesen, P., and van Rij, R.P. (2019) *Antiviral RNAi in insects and mammals: Parallels and differences*. *Viruses*, **11**.

Shao, Q., Liu, X., Su, T., Ma, C., and Wang, P. (2019) *New Insights Into the Role of Seed Oil Body Proteins in Metabolism and Plant Development*. *Front. Plant Sci.*, **10**, 1–14.

Shi, X., Cordero, T., Garrigues, S., Marcos, J.F., Daròs, J.A., and Coca, M. (2019) *Efficient production of antifungal proteins in plants using a new transient expression vector derived from tobacco mosaic virus*. *Plant Biotechnol. J.*, **17**, 1069–1080.

Shih, Y.P. (2005) *Self-cleavage of fusion protein in vivo using TEV protease to yield native protein*. *Protein Sci.*, **14**, 936–941.

Shimada, T.L. and Hara-Nishimura, I. (2015) *Leaf oil bodies are subcellular factories producing*

antifungal oxylipins. *Curr. Opin. Plant Biol.*, **25**, 145–150.

Shimada, T.L., Ogawa, Y., Shimada, T., and Hara-Nishimura, I. (2011) *A non-destructive screenable marker, OsFAST, for identifying transgenic rice seeds*. *Plant Signal. Behav.*, **6**, 1454–1456.

Shimada, T.L., Takano, Y., and Hara-Nishimura, I. (2015) *Oil body-mediated defense against fungi: From tissues to ecology*. *Plant Signal. Behav.*, **10**, 3–5.

Siddiqui, S.A., Sarmiento, C., Truve, E., Lehto, H., and Lehto, K. (2008) *Phenotypes and functional effects caused by various viral RNA silencing suppressors in transgenic Nicotiana benthamiana and N. tabacum*. *Mol. Plant-Microbe Interact.*, **21**, 178–187.

Siloto, R.M.P. (2006) *The Accumulation of Oleosins Determines the Size of Seed Oilbodies in Arabidopsis*. *Plant Cell Online*, **18**, 1961–1974.

Singh, P., Alex, J.M., and Bast, F. (2014) *Insulin receptor (IR) and insulin-like growth factor receptor 1 (IGF-1R) signaling systems: Novel treatment strategies for cancer*. *Med. Oncol.*, **31**.

Sorel, M., Garcia, J.A., and German-Retana, S. (2014) *The Potyviridae Cylindrical Inclusion Helicase: A Key Multipartner and Multifunctional Protein* . *Mol. Plant-Microbe Interact.*, **27**, 215–226.

Souza, A.C., Vasques, R.M., Inoue-Nagata, A.K., Lacorte, C., Maldaner, F.R., Noronha, E.F., and Nagata, T. (2013) *Expression and assembly of Norwalk virus-like particles in plants using a viral RNA silencing suppressor gene*. *Appl. Microbiol. Biotechnol.*, **97**, 9021–9027.

Strasser, R. (2016) *Plant protein glycosylation*. *Glycobiology*, **26**, 926–939.

El Tahchy, A., Reynolds, K.B., Petrie, J.R., Singh, S.P., and Vanhercke, T. (2017) *Thioesterase overexpression in Nicotiana benthamiana leaf increases the fatty acid flux into triacylgycerol*. *FEBS Lett.*, **591**, 448–456.

Takaiwa, F., Wakasa, Y., Takagi, H., and Hiroi, T. (2015) *Rice seed for delivery of vaccines to gut mucosal immune tissues*. *Plant Biotechnol. J.*, **13**, 1041–1055.

Takeyama, N., Kiyono, H., and Yuki, Y. (2015) *Plant-based vaccines for animals and humans: recent advances in technology and clinical trials*. *Ther. Adv. Vaccines*, **3**, 139–154.

Tanaka, Y., Nakamura, S., Kawamukai, M., Koizumi, N., and Nakagawa, T. (2011) *Development of a series of Gateway binary vectors possessing a tunicamycin resistance gene as a marker for the transformation of Arabidopsis thaliana*. *Biosci. Biotechnol. Biochem.*, **75**, 804–807.

Taylor, N.J. and Fauquet, C.M. (2003) *Microparticle Bombardment as a Tool in Plant Science and Agricultural Biotechnology*. *DNA Cell Biol.*, **21**, 963–977.

Tekoah, Y., Shulman, A., Kizhner, T., Ruderfer, I., Fux, L., Nataf, Y., et al. (2015) *Large-scale production of pharmaceutical proteins in plant cell culture-the protalix experience*. *Plant Biotechnol. J.*, **13**, 1199–1208.

Ting, J.T.L., Balsamo, R.A., Ratnayake, C., and Huang, A.H.C. (1997) *Oleosin of plant seed oil bodies is correctly targeted to the lipid bodies in transformed yeast*. *J. Biol. Chem.*, **272**, 3699–3706.

Ting, J.T.L., Lee, K., Ratnayake, C., Platt, K.A., Balsamo, R.A., and Huang, A.H.C. (1996) *Oleosin genes in maize kernels having diverse oil contents are constitutively expressed independent of oil contents: Size and shape of intracellular oil bodies are determined by the oleosins/oils ratio*. *Planta*, **199**, 158–165.

Trüeb, R.M. (2018) *Further clinical evidence for the effect of IGF-1 on hair growth and alopecia*. *Ski.*

Appendage Disord., **4**, 90–95.

Tseng, C.W., Liao, C.Y., Sun, Y., Peng, C.C., Tzen, J.T.C., Guo, R.T., and Liu, J.R. (2014) *Immobilization of clostridium cellulolyticum d -psicose 3-epimerase on artificial oil bodies*. *J. Agric. Food Chem.*, **62**, 6771–6776.

Tseng, J.M., Huang, J.R., Huang, H.C., Tzen, J.T.C., Chou, W.M., and Peng, C.C. (2011) *Facilitative production of an antimicrobial peptide royalisin and its antibody via an artificial oil-body system*. *Biotechnol. Prog.*, **27**, 153–161.

Tugume, A.K., Mukasa, S.B., and Valkonen, J.P.T. (2008) *Natural Wild Hosts of Sweet potato feathery mottle virus Show Spatial Differences in Virus Incidence and Virus-Like Diseases in Uganda* . *Phytopathology*, **98**, 640–652.

Tzen, J.T.C., Cao, Y.Z., Laurent, P., Ratnayake, C., and Huang, A.H.C. (1993) *Lipids, proteins, and structure of seed oil bodies from diverse species*. *Plant Physiol.*, **101**, 267–276.

Tzen, J.T.C., Lai, Y.K., Chan, K.L., and Huang, A.H.C. (1990) *Oleosin isoforms of high and low molecular weights are present in the oil bodies of diverse seed species*. *Plant Physiol.*, **94**, 1282–1289.

Tzen, J.T.C., Lie, G.C., and Huang, A.H.C. (1992) *Characterization of the charged components and their topology on the surface of plant seed oil bodies*. *J. Biol. Chem.*, **267**, 15626–15634.

Tzfira, T. and Citovsky, V. (2003) *State of the Field The Agrobacterium-Plant Cell Interaction . Taking Biology Lessons from a Bug 1 , 2*. *Plant Physiol.*, **133**, 943–947.

Uhde-Holzem, K., Schlösser, V., Viazov, S., Fischer, R., and Commandeur, U. (2010) *Immunogenic properties of chimeric potato virus X particles displaying the hepatitis C virus hypervariable region I peptide R9*. *J. Virol. Methods*, **166**, 12–20.

Umate, P. (2012) *Comparative Genomics of the Lipid-body-membrane Proteins Oleosin, Caleosin and Steroleosin in Magnoliophyte, Lycophyte and Bryophyte*. *Genomics, Proteomics Bioinforma.*, **10**, 345–353.

Untiveros, M., Olspert, A., Artola, K., Firth, A.E., Kreuze, J.F., and Valkonen, J.P.T. (2016) *A novel sweet potato potyvirus open reading frame (ORF) is expressed via polymerase slippage and suppresses RNA silencing*. *Mol. Plant Pathol.*, **17**, 1111–1123.

Usmani, S.S., Bedi, G., Samuel, J.S., Singh, S., Kalra, S., Kumar, P., et al. (2017) *THPdb: Database of FDA-approved peptide and protein therapeutics*. *PLoS One*, **12**, 1–12.

Valli, A., Dujovny, G., and Garcia, J.A. (2008) *Protease Activity, Self Interaction, and Small Interfering RNA Binding of the Silencing Suppressor P1b from Cucumber Vein Yellowing Ipomovirus*. *J. Virol.*, **82**, 974–986.

Valli, A., López-Moya, J.J., and García, J.A. (2007) *Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae*. *J. Gen. Virol.*, **88**, 1016–1028.

Valli, A., Martin-Hernandez, A.M., Lopez-Moya, J.J., and Garcia, J.A. (2006) *RNA Silencing Suppression by a Second Copy of the P1 Serine Protease of Cucumber Vein Yellowing Ipomovirus, a Member of the Family Potyviridae That Lacks the Cysteine Protease HCPro*. *J. Virol.*, **80**, 10055–10063.

Valli, A.A., Gallo, A., Rodamilans, B., López-Moya, J.J., and García, J.A. (2018) *The HCPro from the Potyviridae family: an enviable multitasking Helper Component that every virus would like to have*. *Mol. Plant Pathol.*, **19**, 744–763.

Vanhercke, T., Divi, U.K., El Tahchy, A., Liu, Q., Mitchell, M., Taylor, M.C., et al. (2017) *Step changes in leaf oil accumulation via iterative metabolic engineering*. *Metab. Eng.*, **39**, 237–246.

Vanhercke, T., El Tahchy, A., Liu, Q., Zhou, X.R., Shrestha, P., Divi, U.K., et al. (2014) *Metabolic engineering of biomass for high energy density: Oilseed-like triacylglycerol yields from plant leaves*. *Plant Biotechnol. J.*, **12**, 231–239.

Vanhercke, T., El Tahchy, A., Shrestha, P., Zhou, X.R., Singh, S.P., and Petrie, J.R. (2013) *Synergistic effect of WRI1 and DGAT1 coexpression on triacylglycerol biosynthesis in plants*. *FEBS Lett.*, **587**, 364–369.

Vijayapalani, P., Maeshima, M., Nagasaki-Takekuchi, N., and Miller, W.A. (2012) *Interaction of the trans-frame potyvirus protein P3N-PIPO with host protein PCaP1 facilitates potyvirus movement*. *PLoS Pathog.*, **8**.

Vindigni, J.D., Wien, F., Giuliani, A., Erpapazoglou, Z., Tache, R., Jagic, F., et al. (2013) *Fold of an oleosin targeted to cellular oil bodies*. *Biochim. Biophys. Acta - Biomembr.*, **1828**, 1881–1888.

Vogel, E., Santos, D., Mingels, L., Verdonckt, T.W., and Broeck, J. Vanden (2019) *RNA interference in insects: Protecting beneficials and controlling pests*. *Front. Physiol.*, **10**, 1–21.

Voinnet, O. (2008) *Use, tolerance and avoidance of amplified RNA silencing by plants*. *Trends Plant Sci.*, **13**, 317–328.

Wahlroos, T., Soukka, J., Denesyuk, A., Wahlroos, R., Korpela, T., and Kilby, N.J. (2003a) *Oleosin expression and trafficking during oil body biogenesis in tobacco leaf cells*. *Genesis*, **35**, 125–132.

Wang, H., Ma, J., Yang, Y., Zeng, F., and Liu, C. (2016) *Highly Efficient Delivery of Functional Cargoes by a Novel Cell-Penetrating Peptide Derived from SP140-Like Protein*. *Bioconjug. Chem.*, **27**, 1373–1381.

Wang, M. and Duan, B. (2019) *Materials and Their Biomedical Applications*. In: *Encyclopedia of Biomedical Engineering (Book)*, pp. 135–152. Elsevier.

Webb RE, S.H. (1965) *Isolation and identification of Watermelon mosaic virus 1 and 2*. *Phytopathology*, **55**, 895–900.

Weisman, R. (2009) *More contamination troubles for Genzyme (Newspaper article)*. *November. Bost. Blobe oline.*

White, K.A. (2015) *The polymerase slips and PIPO exists*. *EMBO Rep.*, **16**, 885–886.

Wijesundera, C., Boiteau, T., Xu, X., Shen, Z., Watkins, P., and Logan, A. (2013) *Stabilization of fish oil-in-water emulsions with oleosin extracted from canola meal*. *J. Food Sci.*, **78**, 1340–1347.

Wilson, S.A. and Roberts, S.C. (2012) *Recent advances towards development and commercialization of plant cell culture processes for the synthesis of biomolecules*. *Plant Biotechnol. J.*, **10**, 249–268.

Winichayakul, S., Pernthaner, A., Livingston, S., Cookson, R., Scott, R., and Roberts, N. (2012) *Production of active single-chain antibodies in seeds using trimeric polyoleosin fusion*. *J. Biotechnol.*, **161**, 407–413.

Wolfson, W. (2013) *Grow your own: Protalix biotherapeutics produces drugs in carrot cells*. *Chem. Biol.*, **20**, 969–970.

Wunderlich, S. and Gatto, K.A. (2015) *Consumer Perception of Genetically Modified Organisms and Sources of Information*. *Adv. Nutr.*, **6**, 842–851.

Wylie, S.J., Adams, M., Chalam, C., Kreuze, J., López-Moya, J.J., Ohshima, K., et al. (2017) *ICTV Virus Taxonomy Profile: Potyviridae*. *J. Gen. Virol.*, **98**, 352–354.

Xu, C. and Shanklin, J. (2016) *Triacylglycerol Metabolism, Function, and Accumulation in Plant Vegetative Tissues*. *Annu. Rev. Plant Biol.*, **67**, 179–206.

Xu, J., Dolan, M.C., Medrano, G., Cramer, C.L., and Weathers, P.J. (2012) *Green factory: Plants as bioproduction platforms for recombinant proteins*. *Biotechnol. Adv.*, **30**, 1171–1184.

Xu, J., Ge, X., and Dolan, M.C. (2011) *Towards high-yield production of pharmaceutical proteins with plant cell suspension cultures*. *Biotechnol. Adv.*, **29**, 278–299.

Yang, J., Guan, L., Guo, Y., Du, L., Wang, F., Wang, Y., et al. (2015) *Expression of biologically recombinant human acidic fibroblast growth factor in Arabidopsis thaliana seeds via oleosin fusion technology*. *Gene*, **566**, 89–94.

Yang, Z. and Li, Y. (2018) *Dissection of RNAi-based antiviral immunity in plants*. *Curr. Opin. Virol.*, **32**, 88–99.

Yen, C.L.E., Stone, S.J., Koliwad, S., Harris, C., and Farese, R. V. (2008) *DGAT enzymes and triacylglycerol biosynthesis*. *J. Lipid Res.*, **49**, 2283–2301.

Yi, S., Yang, J., Huang, J., Guan, L., Du, L., Guo, Y., et al. (2015) *Expression of bioactive recombinant human fibroblast growth factor 9 in oil bodies of Arabidopsis thaliana*. *Protein Expr. Purif.*, **116**, 127–132.

Zhang, M., Cao, X., Jia, Q., and Ohlrogge, J. (2016) *FUSCA3 activates triacylglycerol accumulation in Arabidopsis seedlings and tobacco BY2 cells*. *Plant J.*, **88**, 95–107.

Zheng, J.L., Helbig, C., and Gao, W.Q. (1997) *Induction of cell proliferation by fibroblast and insulin-like growth factors in pure rat inner ear epithelial cell cultures*. *J. Neurosci.*, **17**, 216–226.

Zhu, J. (2012) *Mammalian cell protein expression for biopharmaceutical production*. *Biotechnol. Adv.*, **30**, 1158–1170.