UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Design of Genetic Elements and Software Tools for Plant Synthetic Biology

Marta Vázquez Vilar

Thesis Advisor

Dr. Diego Orzáez Calatayud

Valencia, May 2016

El Dr. Diego Orzáez Calatayud, Científico Titular del Consejo Superior de Investigaciones Científicas, perteneciente al Instituto de Biología Molecular y Celular de Plantas (IBMCP, UPV-CSIC) de Valencia,

CERTIFICA que Marta Vázquez Vilar, ha realizado bajo su dirección en el Instituto de Biología Molecular y Celular de Plantas, el trabajo titulado "Design of Genetic Elements and Software Tools for Plant Synthetic Biology", y que autoriza su presentación para optar al grado de Doctor en Biotecnología. Y para que así conste, firma el presente certificado en Valencia a 5 de Mayo de 2016.

Dr. Diego Orzáez Calatayud

a Asun

Agradecimientos

Inocente yo cuando me fui a pasar un verano a Valencia en 2010, sin saber que un año después volvería para quedarme algo más de cuatro años en el 2.10. Durante estos años siempre pensé que el mejor momento de la tesis sería terminarla, pero ahora que lo veo con perspectiva me doy cuenta de lo apasionante que fue recorrer el camino para llegar a la meta…un camino con muchas piedras, pero con gente increíble a mi alrededor que me ayudó a levantarme en cada tropiezo. Así que ahora toca echar la vista atrás y acordarme de todos los que habéis estado a mi lado durante esta etapa de mi vida.

En primer lugar quiero darle las gracias a mi director de tesis. **Diego**, todavía recuerdo esa llamada en la que me preguntabas incrédulo si era yo la que iba a ir dos meses al labo...y meses después me diste la oportunidad de volver para hacer una tesis. Gracias por haber sido el director de tesis que necesitaba en cada momento. Supiste guiarme dejándome siempre libertad para tomar mis propias decisiones…y aunque me equivoqué en muchas de ellas eso me permitió crecer mucho científica y personalmente. Te agradezco también la confianza que siempre has mostrado en mí habiéndome dado la oportunidad de participar en tantos congresos y el apoyo que siempre me has dado, especialmente el de estos últimos meses como director de tesis en la distancia.

Además de Diego sois muchos los del 2.10 a los que os tengo muchísimo que agradecer. Y empezaré por ti **Asun** porque sin duda has sido la persona más importante en esta etapa de mi vida. Gracias por haberme escuchado durante tantas horas, dándome siempre buenos consejos y solución a cualquier duda que te plantease. Gracias por haber sido mi compañera de comidas los findes, por haber ido a las cuatro de la mañana al lab a hacer protoplastos, por haber aprendido css para hacer gbcloning más bonita y por infinitas cosas más. Cualquier agradecimiento que te escribiese se quedaría corto porque siento que esta tesis es tan tuya como mía, gracias por haberte convertido en mi hermana mayor durante estos cuatro años. **Toni Granell**, gracias por haber tenido siempre palabras de ánimo y buenos consejos para mí. **Ale**, tú fuiste la primera persona del lab a la que conocí y tengo que darte las gracias por todo lo que me enseñaste en los primeros meses. Me alegro de haber sido la heredera de GoldenBraid y espero que mantengamos el contacto nos lleven a donde nos lleven nuestras vidas. **Paloma**, muchísimas gracias por el apoyo que me diste, especialmente en los malos momentos. **Estefania**, recuerdo que a tu llegada al labo pensé que nunca podría llevarme bien con una persona como tú, tan distinta a mí...pero parece que eso de que la primera impresión es lo que cuenta no siempre se cumple. De ti aprendí que uno puede hacer tantas cosas como se proponga en un mismo día...cuestión de organización. Y menos mal que tuve una buena profe, porque sino mi vida en Holanda sería un desastre. Estefania, gracias por haber sido mi confidente durante estos años y gracias por todo el apoyo que me has dado siempre. Mucho ánimo, que tú ya estás también en la recta final! **Flor**, gracias por haberme arrastrado al gym tantas veces...haciendo que mi vida fuese un poco más sana física y mentalmente. De ti aprendí a ver la vida con optimismo y a darle a los problemas la importancia que se merecen, ni un poquito más. Son muchas las partidas de palas, las cenas en tu casa y los viajes que ya hemos compartido (me encantaron los días que pasamos en Toulouse y no olvidaré nuestra aventura en Amsterdam)...pero sé que todavía nos quedan muchos más viajes por hacer

juntas. **Silvia**, todavía recuerdo los años sentada a tu lado y lo bien que te portaste siempre conmigo (no sé si porque yo hacía las tareas de vez en cuando o porque me tenías cariño, jeje). Muchas gracias por todo lo que me has enseñado. **Julve**, te preguntase lo que te preguntase siempre me sorprendían tus respuestas, gracias por ponerle emoción a nuestra vida en el labo. **Joan**, gracias por tu ayuda con los experimentos de mi etapa final de la tesis. **Gianfranco** (te incluyo en el 2.10 porque ya eres un poco parte de él), gracias por tu ayuda con los análisis metabolómicos. También quiero daros las gracias a todos los estudiantes que habéis pasado por el labo en este tiempo, especialmente a **Jose** y **Choni** por haber puesto vuestro pequeño granito de arena en esta tesis. Muchas gracias también al equipo **iGEM 2014** (especialmente a Alfredo, Alba, Lucía y Jose otra vez) por haberme devuelto la ilusión por mi trabajo en un momento en el que la había perdido y por esos fantásticos días en Boston.

Aparte del 2.10, sois muchos los miembros de '4labos mas VIP' los que habéis estado ahí. **Marta**, una 'very important person for me' que venía cada tarde a despedirse y decirme que no trabajase mucho. Gracias por haber intentado sacarme del IBMCP cada finde y haberme permitido formar parte de tus planes. Contigo descubrí que vivir en Valencia no estaba tan mal como yo creía (hay playa, buenos bares, sitios para salir de fiesta y es primavera de octubre a mayo...en verano es mejor irse de estancia, no?). Compartimos muchos viajes de los que tengo bonitos recuerdos, especialmente de mi primera aventura holandesa...y cuando no me convertía en tu compañera de viaje siempre había una postal en mi buzón que me demostraba que te habías acordado de mí, ¿próximo destino? **Nora**, compañera de comidas durante tres años y de piso durante un mes. Después de haber vivido tu mudanza, ya tenía experiencia cuando me tocó a mí, jeje. Gracias por ser la mejor organizando viajes y por esa tranquilidad que transmites. **Eugenio**, gracias por tantos buenos ratos en las comidas con sus black stories incluidas. **Amelia**, aunque nos conocimos tarde me alegro de haber compartido contigo gym y cenitas varias. **Noel**, **Marco, David**, **Jorge**…gracias a todos por tantos momentos surrealistas en las comidas. **Mari Ángeles**, gracias por tus palabras de cariño y por tu gran ayuda para conseguir cualquier cosa que necesitase para mi trabajo. **Miguel A. Pérez** y **MD**, gracias por haberme cedido un hueco en vuestro labo con la tranquilidad necesaria para escribir esta tesis.

Además tengo que daros las gracias a mis bioinformáticos favoritos. **Javier Forment**, contigo empecé el camino de gbcloning y, aunque ha cambiado mucho desde entonces no me olvido de los orígenes. **Peio**, muchísimas gracias por todo lo que has aportado a esta tesis y por todos esos ratos que hemos pasado juntos, tú programando y yo tratando de seguir lo que hacías. **Jose Blanca**, gracias por todas las buenas ideas aportadas a gbcloning.

Gracias a todos los de seguridad del IBMCP por haberme abierto la puerta a cualquier hora, cualquier día del año, especialmente gracias a **Jorge**, compañero de comidas algún que otro finde. Y gracias también a los del **Bar Tranvía** por haberme alimentado gran parte de estos años.

Thanks to all the people I met in Norwich. I have to thank you **Cathie Martin** for gave me the opportunity to join your lab for three months. There, I shared great barbecues and 'salsa' nights with **Gabriel**, **Alessandro**, **Dani**, **Katharina** and **Sabrina**.

I want to thank also the people I already met on this new stage of my scientific career as a pre-postdoc. **Christian**, thanks a lot for giving me the opportunity of joining your group while I was finishing this thesis. **Nong** and **Rita**, thanks for became my new family in Wageningen.

Y fuera del ámbito laboral y no por ello menos importantes en esta etapa de mi vida hay otras muchas personas a las que tengo que darles las gracias. **Papá** e **mamá**, gracias pola educación que me déchedes cando era unha nena, sen o voso exemplo de vida nunca tivera chegado a ser quen son. Gracias por permitirme sempre tomar as miñas propias decisións e apoiarme nelas. Gracias polos vosos bos consellos desde a experiencia e o voso apoio incondicional. Gracias polas infinitas horas de Skype (gracias tamén a Skype) que me facían sentirme coma na casa desde a distancia e polas innumerables viaxes que fixéchedes a Valencia para pasar tempo conmigo. **Zalo**, o meu irmán, gracias por todos esos consellos de irmán maior e por todos eses reencontros en Valencia, Barcelona ou Leganés... admírote moito e eres o meu referente na carreira científica. **Alba**, gracias por ter sempre palabras de ánimo para min. En xeral quero darlle as gracias a toda a familia por estar sempre ahí, pero especialmente a **Sofía**, unha prima e unha amiga coa que sigo compartindo cada nova experiencia desde a distancia. **María** e **Graciela**, aínda que as nosas vidas se distanciasen fai xa moitos anos, cada vez que nos vemos sigo tendo a sensación de que o tempo non pasou todo sigue sendo coma sempre.

Y por último tengo que darles las gracias a esos biotecs que siguen ahí. Gracias a **Amparo** y **Mercedes** por haberme acogido en vuestras casas en mis viajes congresiles a Londres y Berlin. Y aunque haya dejado para el final a mis pepis, ellos ya saben lo mucho que les quiero. Manu y Emi, vosotros me acompañasteis en mi primera estancia en el IBMCP y durante estos años compartimos unos cuantos reencuentros y muchas conversaciones...escribiros sin emoticonos se me hace muy raro (gatiti mironsoso). **Manu**, gracias por sacarme una sonrisa con cada uno de tus mensajes, incluidos los que me llegan desde el pasado. **Emi**, gracias por ser como eres y por todo lo que has intentado enseñarme durante estos años, aunque ya sabes que tengo memoria de pez y ya no me acuerdo ni de la mitad de los emperadores, ¿cuándo volvemos a Roma?

En resumen…llegué al 2.10 siendo una niña y gracias a todos vosotros me fui de él siendo una niña grande con una maleta cargada de muchos buenos recuerdos y con la fuerza y los conocimientos necesarios para solventar los problemas que se me puedan presentar en la vida. Aprendí mucho con todos vosotros y lo más importante es que sé que algunos de vosotros habéis llegado a mi vida para quedaros, ¿para cuándo el próximo reencuentro?

Summary

Synthetic Biology is an emerging interdisciplinary field that aims to apply the engineering principles of modularity, abstraction and standardization to genetic engineering. The nascent branch of Synthetic Biology devoted to plants, Plant Synthetic Biology (PSB), offers new breeding possibilities for crops, potentially leading to enhanced resistance, higher yield, or increased nutritional quality. To this end, the molecular tools in the PSB toolbox need to be adapted accordingly, to become modular, standardized and more precise. Thus, the overall objective of this Thesis was to adapt, expand and refine DNA assembly tools for PSB to enable the incorporation of functional specifications to the description of standard genetic elements (phytobricks) and to facilitate the construction of increasingly complex and precise multigenic devices, including genome editing tools.

The starting point of this Thesis was the modular DNA assembly method known as GoldenBraid (GB), based on type IIS restriction enzymes. To further optimize the GB construct-making process and to better catalog the phytobricks collection, a database and a set of software-tools were developed as described in Chapter 1. The final web-based software package, released as GB2.0, was made publicly available at www.gbcloning.upv.es. A detailed description of the functioning of GB2.0, exemplified with the building of a multigene construct for anthocyanin overproduction was also provided in Chapter 1. As the number and complexity of GB constructs increased, the next step forward consisted in the refinement of the standards with the incorporation of experimental information associated to each genetic element (described in Chapter 2). To this end, the GB package was reshaped into an improved version (GB3.0), which is a self-contained, fully traceable assembly system where the experimental data describing the functionality of each DNA element is displayed in the form of a standard datasheet. The utility of the technical specifications to anticipate the behavior of composite devices was exemplified with the combination of a chemical switch with a prototype of an anthocyanin overproduction module equivalent to the one described in Chapter 1, resulting in a dexamethasone-responsive anthocyanin device. Furthermore, Chapter 3 describes the adaptation and functional characterization of CRISPR/Cas9 genome engineering tools to the GB technology. The performance of the adapted tools for gene editing, transcriptional activation and repression was successfully validated by transient expression in *N. benthamiana*. Finally, Chapter 4 presents a practical implementation of GB technology for precision plant breeding. An intragenic construct comprising an intragenic selectable marker and a master regulator of the flavonoid biosynthesis was stably transformed in tomato resulting in fruits enhanced in flavonol content.

All together, this Thesis shows the implementation of increasingly complex and precise genetic designs in plants using standard elements and modular tools following the principles of Synthetic Biology.

Resumen

La Biología Sintética es un campo emergente de carácter interdisciplinar que se fundamenta en la aplicación de los principios ingenieriles de modularidad, abstracción y estandarización a la ingeniería genética. Una nueva vertiente de la Biología Sintética aplicada a las plantas, la Biología Sintética Vegetal (BSV), ofrece nuevas posibilidades de mejora de cultivos que podrían llevar a una mejora de la resistencia, a una mayor productividad, o a un aumento de la calidad nutricional. Sin embargo, para alcanzar este fin las herramientas moleculares disponibles en estos momentos para BSV deben ser adaptadas para convertirse en modulares, estándares y más precisas. Por ello se planteó como objetivo general de esta Tesis adaptar, expandir y refinar las herramientas de ensamblaje de DNA de la BSV para permitir la incorporación de especificaciones funcionales en la descripción de elementos genéticos estándar (fitobricks) y facilitar la construcción de estructuras multigénicas cada vez más complejas y precisas, incluyendo herramientas de editado genético.

El punto de partida de esta Tesis fue el método de ensamblaje modular de ADN GoldenBraid (GB) basado en enzimas de restricción tipo IIS. Para optimizar el proceso de ensamblaje y catalogar la colección de fitobricks generados se desarrollaron una base de datos y un conjunto de herramientas software, tal y como se describe en el Capítulo 1. El paquete final de software se presentó en formato web como GB2.0, haciéndolo accesible al público a través de www.gbcloning.upv.es. El Capítulo 1 también proporciona una descripción detallada del funcionamiento de GB2.0 ejemplificando su uso con el ensamblaje de una construcción multigénica para la producción de antocianinas. Con el aumento en número y complejidad de las construcciones GB, el siguiente paso necesario fue el refinamiento de los estándar con la incorporación de la información experimental asociada a cada elemento genético (se describe en el Capítulo 2). Para este fin, el paquete de software de GB se reformuló en una nueva versión (GB3.0), un sistema de ensamblaje auto-contenido y completamente trazable en el que los datos experimentales que describen la funcionalidad de cada elemento genético se muestran en forma de una hoja de datos estándar. La utilidad de las especificaciones técnicas para anticipar el comportamiento de dispositivos biológicos compuestos se ejemplificó con la combinación de un interruptor químico y un prototipo de un módulo de sobreproducción de antocianinas equivalente al descrito en el Capítulo 1, resultando en un dispositivo de producción de antocianinas con respuesta a dexametasona. Además, en el Capítulo 3 se describe la adaptación a la tecnología GB de las herramientas de ingeniería genética CRISPR/Cas9, así como su caracterización funcional. La funcionalidad de estas herramientas para editado génico y activación y represión transcripcional se validó con el sistema de expresión transitoria en *N.benthamiana*. Finalmente, el Capítulo 4 presenta una implementación práctica del uso de la tecnología GB para hacer mejora vegetal de manera precisa. La transformación estable en tomate de una construcción intragénica que comprendía un marcador de selección intragénico y un regulador de la biosíntesis de flavonoides resultó en frutos con un mayor contenido de flavonoles.

En conjunto, esta Tesis muestra la implementación de diseños genéticos cada vez más complejos y precisos en plantas utilizando elementos estándar y herramientas modulares siguiendo los principios de la Biología Sintética.

Resum

La Biologia Sintètica és un camp emergent de caràcter interdisciplinar que es fonamenta amb l'aplicació a la enginyeria genètica dels principis de modularitat, abstracció i estandardització. Una nova vessant de la Biologia Sintètica aplicada a les plantes, la Biologia Sintètica Vegetal (BSV), ofereix noves possibilitats de millora de cultius que podrien portar a una millora de la resistència, a una major productivitat, o a un augment de la qualitat nutricional. Tanmateix, per poder arribar a este fi les eines moleculars disponibles en estos moments per a la BSV han d'adaptar-se per convertir-se en modulars, estàndards i més precises. Per això es plantejà com objectiu general d'aquesta Tesi adaptar, expandir i refinar les eines d'ensamblatge d'ADN de la BSV per permetre la incorporació d'especificacions funcionals en la descripció d'elements genètics estàndards (fitobricks) i facilitar la construcció d'estructures multigèniques cada vegada més complexes i precises, incloent eines d'edidat genètic.

El punt de partida d'aquesta Tesi fou el mètode d'ensamblatge d'ADN modular GoldenBraid (GB) basat en enzims de restricció tipo IIS. Per optimitzar el proces d'ensamblatge i catalogar la col.lecció de fitobricks generats es desenvolupà una base de dades i un conjunt d'eines software, tal i com es descriu al Capítol 1. El paquet final de software es presentà en format web com GB2.0, fent-se accessible al públic mitjançant la pàgina web www.gbcloning.upv.es. El Capítol 1 també proporciona una descripció detallada del funcionament de GB2.0, exemplificant el seu ús amb l'ensamblatge d'una construcció multigènica per a la producció d'antocians. Amb l'augment en nombre i complexitat de les construccions GB, el següent pas fou el refinament dels estàndards amb la incorporació de la informació experimental associada a cada element genètic (es descriu en el Capítol 2). Per a aquest fi, el paquet de software de GB es reformulà amb una nova versió anomenada GB3.0. Aquesta versió consisteix en un sistema d'ensamblatge auto-contingut i complemtament traçable on les dades experimentals que descriuen la funcionalitat de cada element genètic es mostren en forma de fulla de dades estàndard. La utilitat de les especificacions tècniques per anticipar el comportament de dispositius biològics compostos s'exemplificà amb la combinació de un interruptor químic i un prototip d'un mòdul de sobreproducció d'antocians equivalent al descrit al Capítol 1. Aquesta combinació va tindre com a resultat un dispositiu de producció d'antocians que respón a dexametasona. A més a més, al Capítol 3 es descriu l'adaptació a la tecnologia GB de les eines d'enginyeria genètica CRISPR/Cas9, així com la seua caracterització funcional. La funcionalitat d'aquestes eines per a l'editat gènic i activació i repressió transcripcional es validà amb el sistema d'expressió transitòria en *N. benthamiana*. Finalment, al Capítol 4 es presenta una implementació pràctica de l'ús de la tecnologia GB per fer millora vegetal de mode precís. La transformació estable en tomaca d'una construcció intragènica que comprén un marcador de selecció intragènic i un regulador de la biosíntesi de flavonoïdes resultà en plantes de tomaca amb un major contingut de flavonols en llur fruits.

En conjunt, esta Tesi mostra la implementació de dissenys genètics cada vegada més complexos i precisos en plantes utilitzant elements estàndards i eines modulars seguint els principis de la Biologia Sintètica.

Table of Contents

1 | Introduction

1. Brief introduction to Synthetic Biology

Synthetic Biology (SynBio) is an emerging interdisciplinary field that aims to apply the basic principles of engineering to molecular biology and genetic engineering (What's in a name?, 2009). The main objective of SynBio is not to depict the fundamental principles underlying biology but to make biology useful for the society by making it practical and predictable (Cook et al., 2014). To achieve this goal SynBio proposes to rewire biological systems by introducing new biological components or pathways created following the engineering principles of standardization, modularity and abstraction of function (Silver et al., 2014, Khalil & Collins, 2010).

The idea of applying engineering principles to biology in order to facilitate and speed up the design and development of new biological devices is inspired in the development of the computing sciences (Gardner & Hawkins, 2013). The great progress of computing during the 20th century was based on the aforementioned engineering principles. After cracking how to write and how to read instructions, several standard electronic components were created and easily combined to build hardware components. In parallel, computer languages emerged to provide standard rules for the creation of machine-readable instructions, which can be referred as software. Software and hardware are complementary and none of them can be used without the other.

With the milestone of computer science in mind, synthetic biology started to consider molecular biology and its rules as the software to write new instructions with DNA. The standardization and modularity of the composition rules in molecular biology are powering SynBio. Besides software, as in computer science, SynBio also needs hardware. A universal biological machine does not exist yet, although many chassis are being used and improved to come up to this objective (Leonard et al., 2008). Engineered bacteria and yeast are being employed for fermentation (Altaras & Cameron, 1999, Jarboe et al., 2007) and drug synthesis (Chang & Keasling, 2006) among other biotechnological applications (Andrianantoandro et al., 2006). Production of an antimalarial drug precursor in *Saccharomyces cerevisiae* following a SynBio approach that involved the engineering of the farnesyl pyrophosphate (FPP) pathway and the introduction of two enzymes catalyzing the conversion of FPP first to amorphadiene and next to artemisinic acid constitutes one of the SynBio hallmarks (Ro et al., 2006).

2. SynBio meets Plant Biology

SynBio founded its beginning in microbial systems, but most of the issues that this discipline aims to solve are beyond the ability of microbes to deal with them. That is the reason why SynBio has advanced to multicellular systems, notably to plants (Liu & Stewart, 2015). Plants are primary sources of biomass, an excellent source of valuable secondary metabolites and an advantageous platform for recombinant protein production. The nascent branch of SynBio devoted to plants (Plant Synthetic Biology, Plant SynBio) aims to accomplish new engineering goals, more complex and precise than current single-gene approaches, as building entire biosynthetic or signaling pathways within the plant 'chassis', reshufle endogenous (intragenic) genetic elements to produce new crop traits, or create new sources of genetic diversity through targeted gene editing at genome escale. Examples of challenging SynBio-inspired projects are the engineering of C4 metabolism in rice (Taniguchi et al., 2008), the engineering of the CAM machinery into bioenergy crops for biofuel production in marginal lands (DePaoli et al., 2014), or transforming plant leaves in oil-accumulating organs (Slocombe et al., 2009) to mention only some of the most relevant examples.

The molecular tools required to apply SynBio to plants are insufficiently developed, dispersed, scattered and not yet standardized. Plant synthetic biologists are just starting to develop molecular biology tools adapted to plants particularities. Complex engineering in plants requires, in first place, to improve the ability to transfer large sets of transgenes to the plant genome. Remarkable examples of multigene engineering using biolistics (Agrawal et al., 2005) have been reported in the literature, including combinatorial approaches leading to nutritional enhancement (Naqvi et al., 2010). However, plant transformation is mainly performed using *Agrobacterium tumefaciens* as intermediate, which requires plasmids with specific features such as an Agrobacterium replication origin and conserved borders flanking the DNA that is being transferred to the plant (Patron, 2014). Since DNA integration into plant genomes is random, some tools for targeted integration based on synthetic nucleases are also being developed by the plant community (Puchta & Fauser, 2013). In addition to the development of new genetic tools, the Plant SynBio community is also joining forces to move towards common standards, to develop computer resources for data sharing and to search new funding opportunities. During its years of existence Plant SynBio has grown very quickly and the construction of SynBio-inspired biological devices has been reported using the plant cell as chassis. A pioneer example by June Medford group was the transfer to plants of a bacterial-based signal transduction pathway for TNT detection (Antunes et al., 2006, Antunes et al., 2011, Antunes et al., 2009). More recently, a red-light

controlled synthetic switch was built based on the red-light mediated reversible dimerization of the *Arabidopsis thaliana* phytochrome B and its interacting factor PIF6 (Muller et al., 2014). Outside the purely academic world, a start-up company gave the first steps towards the creation of a 'glowing plant' emitting light showing promising results in a crowdfunding initiative (Callaway, 2013). Shared efforts are making Plant SynBio the heart of the second wave of SynBio, as it has the potential to face the world challenges of providing food, energy and other materials and compounds from limited resources. Plant SynBio can help towards these goals by facilitating the development of crops with enhanced resistance to biotic and abiotic stresses and with a higher yield. Plant SynBio also enables the production in plants of novel compounds for therapeutics and cosmetics. However, public concerns about the use of transgenesis are slowing its applications (Rastogi Verma, 2013). To ease public opposition to transgenesis, alternative technologies such as CRISPR/Cas9 are being used for genome editing.

3. Foundational Technologies and Tools for Plant Synthetic Biology

Further advances in Plant SynBio require the parallel development of Tools and Technologies specially adapted to the plant chassis, most notably DNA assembly methods and associated computational tools. First, it is necessary to increase the capacity of "writing" new genetic information in the form of DNA, primarily by *de novo* chemical synthesis of building blocks and later combining those building blocks into complex devices through increasingly efficient DNA assembly methods. Second, it is required to develop new software tools that complement DNA assembly with efficient data storage and computer-assisted design, thus enabling faster cycles of building and testing for the creation of new biological devices. This introduction will review the state of the art of DNA assembly methods and computational design tools for Plant SynBio.

3.1. Assembly of Genetic Building Blocks

The ability to write and transfer complex DNA instructions involving two or more genes is a key factor to implement the vision of Plant SynBio (Naqvi et al., 2010). Until very recently, the introduction of multiple genes into plants (gene stacking or multigene engineering) was mostly achieved in a sequential way by cross-breeding or by re-transforming plants. The principal problem of these strategies is that transgenes are integrated randomly at different genomic positions, which can result in independent segregation in the progeny. An alternative for the introduction of multiple transgenes is the co-transformation of linked genes in a single DNA insert with the subsequent co-

integration of the transgenes at the same chromosomal position, which allows for co-segregation in subsequent generations. The co-transformation strategy requires the creation of large DNA constructs comprising multiple artificial genetic parts. Classical, non-standard restriction-ligation cloning methods using Type II restriction enzymes resulted inefficient to build multigenic constructs. In principle, precise multigenic constructs could be created by chemical DNA synthesis, which makes tailored DNA sequences without the need of a precursor DNA template. Making use of this technology, synthetic genes can be redesign for codon optimization and for minimizing secondary RNA structures (Villalobos et al., 2006). Although DNA synthesis is becoming increasingly affordable, and certainly has become more economical than classical cloning and mutagenesis procedures, it is still unaffordable for multigene-scale DNA synthesis. Moreover, tailored DNA synthesis lacks the advantages of modularity and standardization proposed by SynBio.

The practical problems associated with the construction of large multigene cassettes for plant transformation are being addressed by the development of new DNA assembly strategies (Chao et al., 2014). Early from the foundation of the discipline, Synthetic Biologists introduced the concept of standard DNA building blocks or Biobricks (name adopted by the iGEM Foundation) (Lewens, 2013), to refer to genetic parts with a basic biological function (promoters, coding regions, terminators…) that can be assembled together to create modules with a defined function (Zhang & Jiang, 2010). These basic parts can be either synthesized or PCR amplified from a suitable template. Multigenic constructs can be subsequently created from those basic parts making use of DNA assembly methods. The standardization of the genetic elements results in modularization of the assembly strategies. Modular design speeds up engineering and decreases the effort required for the assembly of multigenic structures, easing the exchange and reuse of assembled parts and devices (Muller & Arndt, 2012).

Here I present an overview of the main assembly strategies currently in use in Plant SynBio, which can be classified in two groups, those that use restriction enzymes (RE) in restriction-ligation cloning, and those using PCR-overlapping strategies.

- **Restriction-ligation based assembly strategies**

Restriction-ligation methods have in common that all use digestion with REs and subsequent ligation reactions to release and assemble genetic elements. RE-based methods are easy to understand and to adopt by users that are already using classical restriction-ligation in their cloning procedures. Their major drawback is the usual requirement of removing internal restriction sites for the REs used in each specific system.

The first standard cloning method proposed in SynBio was the BioBrick standard. The BioBrick standard proposes a sequential cloning scheme for the binary assembly of standard parts to create multigene assemblies. The standard makes use of four type II restriction enzymes: EcoRI, PstI, SpeI and XbaI (Lewens, 2013). Digestion of the destination vector and the donor insert with the appropriate enzymes generates complementary sticky ends that recreate the original sites at the beginning of the component and creates a six nucleotides scar at the junction (Figure 1a). Although BioBrick is widely used in bacterial SynBio, this assembly standard has some drawbacks that are especially disadvantageous for working with plants. It leaves scars between parts, which is not useful for applications such as protein engineering where scar benign joints are required. Moreover, the number of parts available on the BioBricks registry for plant biotechnology applications is limited and, in addition, the assembly of complex constructs requires the use of a sequential cloning schema, which is time consuming and generates many intermediate useless constructs.

An alternative solution came with the Golden Gate cloning system (Engler et al., 2009). Golden Gate overcomes BioBrick limitations making use of a Type IIS restriction enzyme (BsaI). Type IIS enzymes have separate cutting sites and recognition sites, therefore transferring the decision of specific sequence of sticky ends to the user, and allowing the combination of several parts, up to nine, in a single restriction-ligation reaction (Figure 1b). These characteristics enable scar-benign assemblies minimizing the time required to assemble new transcriptional units in comparison with BioBricks.

In its initial description, Golden Gate was limited to a combination of no more than nine elements, and was not suitable for multigene engineering. It was efficient enough to assemble one or two transcriptional units (TUs), but not to combine several TUs in a single construct. Multigene assembly based in Golden Gate was made possible by introducing a second Type IIS enzyme in the cloning schema, which allowed performing sequential assembly levels by alternating the two enzymes. This general strategy, which was developed independently by two Plant groups, gave rise to two alternative cloning methods known as MoClo and GoldenBraid (GB) respectively. MoClo proposed a large collection of destination vectors that allow the one-step assembly of different numbers of transcriptional units (TUs) (Weber et al., 2011a). GoldenBraid was an alternative modular system that allowed the indefinite growth of the multigenic structures making use of just four destination vectors by proposing a binary and iterative loop strategy (Figure 1c) (Sarrion-Perdigones et al., 2011). Each cloning system has its pros and cons (Table 1). Whereas the time required for creating complex assemblies is larger with GoldenBraid than with MoClo, any intermediate element generated with GoldenBraid is reusable, favoring the exchange of parts and modules. An extra benefit of GoldenBraid is

that it incorporates an integrated web-platform offering software solution to the design and assembly of multigenic structures. This platform also includes a collection of genetic parts and modules to be used in Plant SynBio and Plant Metabolic Engineering boosting the users exchange of created modules. The so-called GoldenBraid 2.0 web platform was developed during this Thesis and it is described in Chapter 1 and in (Vazquez-Vilar et al., 2015).

Modular Cloning methods such as MoClo and GoldenBraid benefit from the assignment of arbitrary overhangs to defined positions within a prototypical Transcriptional Unit, therefore creating a positional notation also known as "Assembly Syntax". Such categorization allows the exchange and reuse of genetic elements. Initially, MoClo and GB (version one) syntaxes had little in common, as they were developed independently (see Weber et al. (2011a) and Sarrion-Perdigones et al. (2011)). During this Thesis several confluency efforts were done to increase compatibility among the two main Plant modular cloning systems. The GB2.0 grammar described in Chapter 1 represents an initial effort towards this goal. Finally, a concerted effort to bring together a wider community consensus crystalized in a position paper describing the so-called Standard Plant Syntax (Patron et al., 2015). The final adaptation of the GoldenBraid system to this common Syntax is described in Chapter 2 as a part of the development of the third version of the GoldenBraid system (GB3.0).

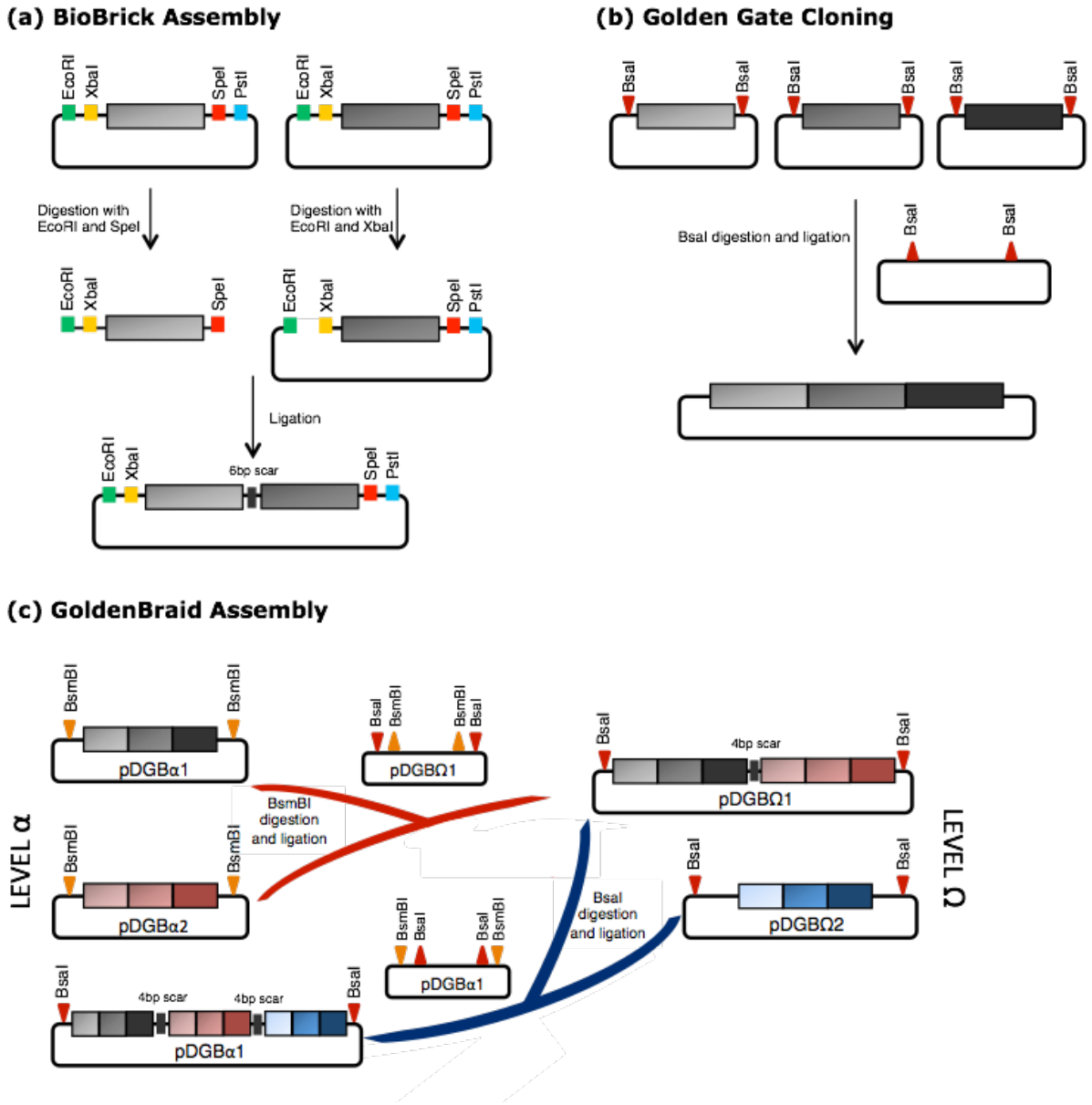


Figure 1. Restriction-ligation DNA assembly systems.

(a) Digestion of two BioBrick parts with pairs of enzymes leaves both plasmids with 4 base pair compatible overhangs. Upon ligation of compatible sites, a 6 base pair scar is generated between the original parts that form the new BioBrick. (b) Multiple DNA parts flanked by Type IIS restriction sites generating 4 base pair compatible overhangs are simultaneously assembled in a destination vector with the Type IIS sites in inverted orientation. (c) Vectors obtained from Golden Gate reactions can be assembled between them in GoldenBraid reactions where genetic elements assembled in compatible vectors (1 and 2) of the same level (α or Ω) can be combined between them in any destination vector of the opposite level.

- **Overlap assembly methods**

Alternatively to restriction-ligation methods, several methods based on homologous recombination have been applied to create large multigenic constructs. They overcome the issue of removing internal restriction sites required on the restriction-ligation dependent strategies. However, these methods require amplification of fragments with primers specifically designed to each fragment, which is costly and time-consuming (Table 1). Parts and destination vector are amplified with flanking overlap regions that are typically 40 nucleotides in length and define their position in the final assembly. Parts and vector are mixed together and joined by enzymatic reactions *in vitro* or using recombination systems *in vivo*.

The sequence ligase independent cloning (SLIC) method requires the treatment of the parts and vector to be assembled with the T4 polymerase that, in the absence of dNTPs, have a 3'-5' exonuclease activity, leaving 5' overhangs (Li & Elledge, 2012). Careful design of the intended overhangs, lacking cytosines, allows the T4 exonuclease activity be counterbalanced by the addition of dCTP. After letting the T4 act for a controlled time, all the PCR products and the destination vector having appropriately long 5' overhangs are assembled together creating a stable circular molecule which does not require ligase activity to be efficiently transformed into *E. coli* cells.

A very popular cloning alternative is the Gibson assembly, where all the parts and the linearized destination vector are treated with a three-enzymes cocktail comprising a T5 exonuclease with 5'-3' activity, a polymerase to fill the gaps after 3' overhangs annealing, and a ligase for nicks reparation (Gibson et al., 2009). In CPEC, or circular polymerase extension cloning, all the parts and the vector prime each other in a low-cycle PCR (Quan & Tian, 2009). SLICE is based on the same principles as CPEC but the enzyme mixture is replaced by cell extracts (Zhang et al., 2012). In general, although overlap-based methods are very efficient, they are not widespread among plant biotechnologists, which have mainly adopted Restriction-Ligation Methods for Plant SynBio approaches.

Assembly method	Mechanism	Scar	Fragments assembled in parallel	Idempotency of the generated construct	Reference
BioBricks	Type II RE	6bp	Two	Yes	(Lewens, 2013)
Golden Gate	Type IIS RE	No (or 4bp with standard parts)	Multiple	No	(Engler et al., 2009)
Moclo	Type IIS RE	4bp	Multiple	No	(Weber et al., 2011a)
GoldenBraid	Type IIS RE	4bp	Two	Yes	(Sarrion-Perdigones et al., 2011)
SLIC	Overlap	No	Multiple	Yes	(Li & Elledge, 2012)
Gibson	Overlap	No	Multiple	Yes	(Gibson et al., 2009)
CPEC	Overlap	No	Multiple	Yes	(Quan & Tian, 2009)
SLICE	Overlap	No	Multiple	Yes	(Zhang et al., 2012)

Table 1. Comparison of standardized DNA assembly methods.

3.2. Software tools for assisted DNA assembly, modeling of synthetic gene circuits and data management

The advantages of the standardization along with the growing ability of researchers to manipulate DNA are generating a large quantity of genetic parts and modules. Their utility relies on a robust characterization and on the ability to incorporate them in higher order devices with a predictable response. Standardization is enabling the definition of common rules for the composition and characterization of those genetic parts, facilitating the development of registries and associated software tools for *in silico* DNA assembly and modeling.

3.2.1. Registries of plant specific genetic elements

In recent years, many information-sharing platforms for data management emerged to handle synthetic biology knowledge and favor the communication among researchers. Databases are the way to keep the information generated in a SynBio laboratory

organized, favoring the exchange of standard genetic elements. The Registry of Standard Biological Parts is probably the best-known collection of genetic parts (Peccoud et al., 2008). It includes a catalog of more than 20.000 parts organized by part type, chassis and function. However, the set of parts specific for plants in this Registry is limited (Boyle et al., 2012). After the establishment of the Registry as a reference platform, other collections rose up. The International Open Facility Advancing Biotechnology (BioFAB) (Mutalik et al., 2013b) offers a characterized collection of bacterial promoters and termination regions and the Joint BioEnergy's Inventory of Composable Elements (JBEI-ICE) (Ham et al., 2012) is a biological database to store strains and seeds, as well as DNA parts and devices. At the beginning of this thesis, there were, to our knowledge, no collections or registries of DNA elements devoted to Plant SynBio. Even today there are few examples, notably the UK-based OpenPlant Initiative (http://openplant.org/) and the GB2.0/GB3.0 public database (https://gbcloning.upv.es/search/) described in this Thesis (see Chapter 1, Chapter 2 and Sarrion-Perdigones et al. (2013)), although only GB seems to have currently an online database in place.

3.2.2. In silico DNA assembly

Besides easing the laboratory procedures, it is important to automate all the assembly process so as to minimize the user's input, avoiding human errors, and leading to the expected result. The coexistence of many assembly standards is a challenge for the computational community. To create specific software tools for each assembly standard could be the solution to manage the huge quantity of available information. Modular assembly methods facilitate the development of computational tools to design the assemblies, therefore facilitating automation (Hillson et al., 2012). Recently, many new computational tools for multigenic assemblies in SynBio have been created. ClothoCAD assembles gene circuits from parts with a formal grammar offering data retrieval from the Registry (Xia et al., 2011). j5 makes possible to perform *in silico* assemblies with any of the following cloning strategies: SLIC, CPEC, Gibson and Golden Gate (Hillson et al., 2012). While all these strategies result in scar-benign assemblies, the solution provided by the j5 software is non-intuitive, requiring as input an user-created datasheet. Again, at the starting of this thesis there were no assembly software tools devoted to Plant SynBio. To fill this gap, we created the integrated gbcloning web-platform offering solution to data management, design and assembly of multigenic structures with the GoldenBraid strategy (see Chapter 1 and Sarrion-Perdigones et al. (2013)).

As mentioned before, during the elaboration of this Thesis, Golden Gate and GoldenBraid developers agreed the use of a common syntax for Plant Biotechnology.

This Standard Plant Syntax establishes the physical composition rules that govern the way in which individual genetic components (hereafter referred to as phytobricks) are to be connected together to create higher order modules and devices. Consequently, the GB2.0 software tools were adjusted to the new standard syntax (see Chapter 2) and now also assist Golden Gate/MoClo multipartite assemblies.

3.2.3. Functional data management and modeling of synthetic gene circuits

SynBio ultimately aims to enable the implementation of Computer Assisted Design (CAD) in biological systems. To reach this goal, functional composition rules need to be established so that biological components can be reliably and predictably assembled into functional devices, a process that implies the hierarchical abstraction of biological functions. Standard parts and devices need an exhaustive characterization by measuring their outputs under diverse conditions to obtain reliable data for modeling and CAD (Pasotti & Zucca, 2014). Standard genetic elements conforming to a certain category should go through the same characterization procedure in order to establish reliable comparisons. Furthermore it is important to store all the quantitative data in the form of standard datasheets. The creation of datasheets associated to each genetic element makes the information accessible to researchers and computers, which can use it to predict the behavior of combinations of those standard elements. The quantitative characterization of a standard genetic element in bacteria and the display of the collected information together with a summary of the assembly characteristics on a datasheet have been reported (Arkin, 2008, Lee et al., 2011).

As difficult this goal is for single-celled microbes, it is far more challenging for complex, multi-celled organisms as plants. Despite this, Plant Biotechnologists are in some way attempting to elaborate and implement functional compositional rules in a non-standard manner. For instance, every attempt to use a developmental-regulated promoter to drive the expression of a transcription factor is ultimately a non-systematic exercise of implementing a functional composition rule. Therefore, although probably not with the level of detail that could be implemented in microbial cells, it is important to make the effort for creating standardized descriptions of plant components which includes a quantitative description, as standard and as exhaustive as possible, of the behavior of the device, so we can learn from the premises of these descriptions which kind of predictive models can be build. Chapter 2 in this Thesis describes efforts towards developing standard datasheets in the GoldenBraid format.

4. Plant chassis for Synthetic Biology

The ability to read, write and understand DNA as a programming language is not enough to assemble a biological device. This process also requires an organism that provides the infrastructure and energy to execute the written code generating the desired function. Theoretically, an organism depleted from superfluous functions would be the perfect chassis into which Synthetic Biologists could incorporate genetic circuits (Unbottling the genes, 2009). In bacteria, some chassis have been developed as engineered *E. coli* (Posfai et al., 2006) or *Pseudomonas* (Nikel et al., 2014). However, plants have large genomes, complex signal-pathways, many organs and most of them are difficult to transform. For this reason, all the chassis proposed for plants still require to be studied and perhaps engineered to ensure a predictable behavior of the designed circuit. I discuss here some of the most used Plant chassis in the context of Plant SynBio.

4.1. Plant Protoplasts

Plant protoplasts offer a cell-based experimental system that maintains the tissue identity and have positive biochemical, genetic and physiological characteristics (Yoo et al., 2007). Protoplast isolation and transformation protocols are well-established for different plant species including *A. thaliana* and tobacco (Davey et al., 2005). They have been used to dissect several signaling pathways (Wang et al., 2005, Yanagisawa et al., 2003, Hwang & Sheen, 2001) and for quickly testing genetic devices (Muller et al., 2014). To overcome the experimental variability observed when testing parts in protoplasts, Schaumberg et al. (2015) created a mathematical model to normalize data enabling the extraction of quantitative parameters describing the behavior of the tested devices.

4.2. Arabidopsis thaliana

Arabidopsis is a model plant widely used in Plant Biology and Genetics. The advantageous characteristics that make it suitable as a model plant make it appropriate also for SynBio applications. An example of the use of this chassis in SynBio is the implementation of an entire synthetic pathway for TNT detection (Antunes et al., 2011). The "Glowing plant" project also makes use of Arabidopsis to obtain a bioluminescent plant (Callaway, 2013).

4.3. Marchantia polymorpha

Marchantia is a primitive and easy-to-engineer liverwort with simple genomic structures that contain less genetic redundancy than the Arabidopsis genome. The simple development, the ease of cultivation in suspension, agar or on soil and the transformation facility place Marchantia as an ideal chassis (Cook et al., 2014, Sharma et al., 2014). It was recently proposed as a model plant for SynBio, although its use is not yet widely extended.

4.4. Nicotiana benthamiana

Nicotiana benthamiana is becoming the chassis of choice for many synthetic biology and molecular farming applications, especially those associated with Agrobacterium-mediated transient expression, a method that facilitates the transfer and expression of multiple genes to the plant cells in a few days (Sainsbury & Lomonossoff, 2014). Briefly, this method, also known as "agroinfiltration", consists in the use of a syringe or released-vacuum to infiltrate by overpressure a diluted Agrobacterium suspension into the intercellular space of leaves. In this form Agrobacterium reaches virtually all cells in the leaf, transiently transferring its T-DNA in a few hours and resulting in widespread expression of T-DNA-encoded genes during a few days. Despite the advantages of *N. benthamiana* as a chassis, there are many factors affecting the behavior of genetic devices during transient expression that need to be studied, one of the most important being the number of copies of T-DNA entering each cell. The method allows assaying gene expression in a fast and reliable manner, and it has been proposed as an experimental standard for the quantification of transcriptional activities in plants (see Chapter 2). An important advantage of agroinfiltration is that allows the transient modification of the chassis, allowing the adjustment of possible orthogonal effects as for instance its post-translational modifications. Some examples are the expression of proteases inhibitors (Goulet et al., 2012) or glycosil-transferases for human-like antibody glycosylation (Castilho et al., 2012). Ideally, however, these modifications should be stably incorporated into the *N. benthamiana* genome to be fully integrated in the chassis. Some efforts have already been made moving on this direction such as the engineering of *N. benthamiana* plants with humanized glycosilation patterns (Strasser et al., 2008).

4.5. Crop Plants

Although the chassis described above are best suited for characterizing new genetic circuits, it is important to remember that the aim of SynBio is to deal with the problems affecting our society. With these objectives in mind, the most widely distributed and consumed crops, such as rice, wheat or tomato, are the ultimate chassis for engineering, not only aiming at traditional breeding traits such as yield improvement and biotic and abiotic stress resistance but also as platforms for biofortification and bioproduction of novel compounds.

5. Genome engineering in plants

Precise manipulation of the chassis genome and its endogenous gene expression patterns is expected to play an important role in the advance of Plant SynBio. Transgenesis was developed around 30 years ago to insert genes into plants and, during this time, successful examples of plants with novel functions have been reported (Rao et al., 2009). However, with this technology, DNA is mainly randomly integrated into the host genome causing sometimes unwanted effects in an unpredictable way (Alonso et al., 2003). Recently, new synthetic nucleases have enabled tailored sequence modifications in plants, including gene targeted mutations, gene insertions and gene replacements. This unprecedented exhaustive control of the genetic material has a great potential for the creation of crop plants with improved tolerance to stress (Zurbriggen et al., 2008), rice with more efficient photosynthesis (von Caemmerer et al., 2012) or cereals with nitrogen-fixing ability (Oldroyd & Dixon, 2014). In addition to speed up the creation of improved crops, public perception of genome engineering is more favorable than with transgenesis.

Synthetic endonucleases induce gene modification by causing double stranded DNA breaks in specific sites, which triggers DNA repair mechanisms, mainly non-homologous end joining (NHEJ) and at low frequencies homologous recombination (HR) (Gorbunova & Levy, 1999). With the supply of a donor DNA, HR mediates gene insertions and gene replacements in eukaryotic cells. Engineered nucleases include zinc-finger nucleases, TAL effectors or CRISPR/Cas9 technology (Gaj et al., 2013). In addition to site-directed genome modifications, engineered nucleases can also be used for gene regulation enabling the activation or repression of specific genes.

5.1. Zinc-Finger and TALE nucleases

A first generation of synthetic tools for genome engineering involved zinc-finger and TALE engineered nucleases. Synthetic zinc-finger nucleases (ZFNs) are relatively small proteins that include an array of zinc-fingers each of which recognizes a triplet of nucleotides (Figure 2a). Conveniently fused to a nuclease domain, ZFNs could selectively produce double strand breaks in specific genomic targets. However, these proteins are not totally modular, being the triplet recognized by each zinc-finger influenced by the adjacent motifs (Fu et al., 2009). This issue was partially overcome by searching the zinc-finger arrays in public libraries, a strategy that led to higher targeted mutation rates in Arabidopsis (Zhang et al., 2010). Transcription activator-like effectors (TALEs) are effector proteins secreted by Xanthomonas during the infection of different plant species (Joung & Sander, 2013). As ZFNs, TALE nucleases (TALENs) also need to be redesigned to target each nucleotidic sequence. The major benefit is that their DNA binding domain, in contrast to ZFNs, is completely modular, with two aminoacidic residues required to recognize each targeted nucleotide (Figure 2b). Tailored-construction of custom TALENs was facilitated by Golden Gate cloning following the design of a library of TALE modules whose combinatorial arrangement allowed the assembly of any TALE DNA binding domain (Weber et al., 2011b).

5.2. CRISPR/Cas9

While ZFNs and TALENs require recoding a protein to each gene target, CRISPR/Cas9 only involves the redesign of a 20 base pairs guide RNA that directs the Cas9 nuclease to its target site (Figure 2c) (Jinek et al., 2012). This great advantage makes CRISPR/Cas9 probably the most powerful technology for genome engineering. CRISPR/Cas9 can be used for modulating gene expression by using a nuclease defective Cas9 (Bikard et al., 2013). Modulation of gene expression, especially repression, opens the possibility of creating new genetic circuits with great potential for Plant SynBio (Nielsen & Voigt, 2014, Brophy & Voigt, 2014, Liu et al., 2014). Since its use for gene editing at genomic level was proposed in 2012, the number of publications using this technology has been rising in an unprecedented way. Starting in mammal cell cultures, its use rapidly moved to model and crop plants demonstrating the possibility of easily inducing targeted mutations at high rates (Belhaj et al., 2013, Li et al., 2013, Nekrasov et al., 2013). It is expectable that the standardization of the CRISPR/Cas9 toolkit for plants by its integration in a Modular Cloning schema such as GoldenBraid, will increase the reliability of experimental results, will facilitate the exchange of new tools, and will contribute to

more efficient design and cloning of (multiplexing) constructs. The integration of CRISPR/Cas9 toolkit in the general GB schema is described in Chapter 3.

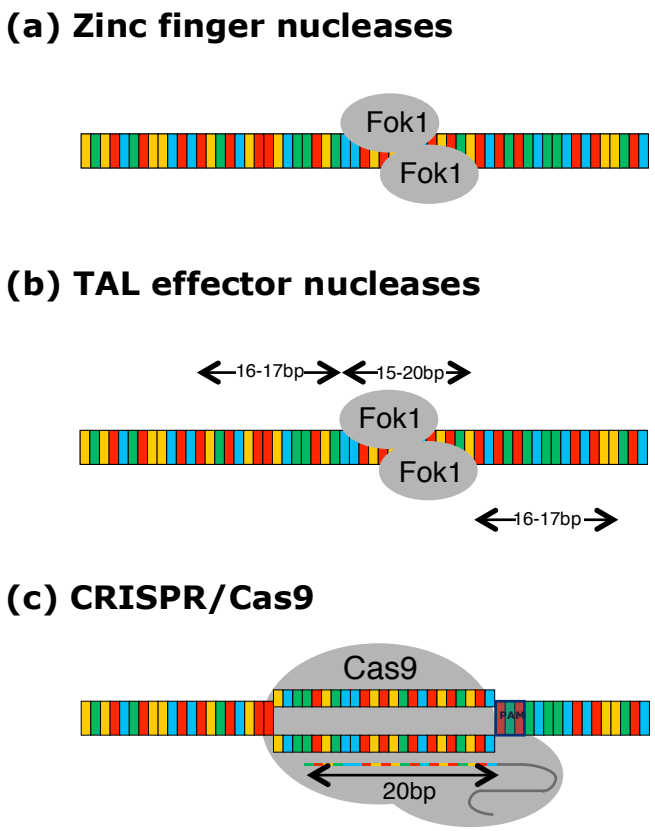


Figure 2. Schematic representation of three genome engineering strategies.

(a) Zinc finger nucleases (ZFNs) are fusions of three or four zinc-fingers (each binding three nucleotides) to the FokI catalytic domain. Two ZFNs are required as FokI is only active as a dimmer. (b) Transcription activator-like effector nucleases (TALENs) are fusions of repeated TAL DNA-binding domains to the FokI catalytic domain. Two TALENs heterodimers separated by 15-20 nucleotides allow the dimerization of the FokI domains required for its activity. (c) Cas9 is a nuclease guided to a 20 nucleotides genome target by a single guide RNA. The Cas9 recognizes a specific protospacer associated motif (PAM) sequence on the DNA (NGG).

6. What could be expected from Plant Synthetic Biology in the future?

To date, the number of published examples illustrating the use of SynBio in plants is still very low. Nevertheless, it is expected that enlarging the collection of genetic elements performing new functions will exponentially increase the possibilities of genetic combinations to be assayed. Moreover, the increase in the efficiency of DNA assembly will facilitate the construction of such combinations.

An immediate application arising from the availability of an enriched palette of genetic elements, which includes many plant endogenous elements, is the possibility of recreating previous transgenic designs, now using exclusively genetic elements derived from the chassis, a strategy known as intragenesis (Rommens et al., 2007). In the context of Plant SynBio, intragenesis can be defined as a crop breeding strategy that limits itself to the use of endogenous "phytobricks" for genetic engineering. Intragenesis is conceived as a strategy for easing the current negative perception of technology crops and evidencing the need for regulating final products rather than processes. Chapter 4 illustrates the use of GoldenBraid to engineer intragenic tomatoes biofortified in flavonols.

Another interesting field of application is the design of sensing devices that enable plants to detect and react to environmental signals in unprecedented ways. Plant synthetic sensors can be designed to perceive an internal or external stimulus and they can be transcriptionally or post-transcriptionally controlled. Some examples of monitoring internal stimuli are detecting changes on the levels of auxins (Wend et al., 2013) or cytokinins (Zurcher et al., 2013). Some synthetic circuits have also been developed for sensing an external stimulus such as TNT (Antunes et al., 2011) or a bacterial pathogen (Liu et al., 2011, Liu et al., 2013). The TNT biosensor is based on a bacterial signal transduction pathway that was transferred to plants (Antunes et al., 2009) and connected to a de-greening circuit that provides a visual reporting signal when TNT is perceived (Antunes et al., 2006). The de-greening circuit has the re-set capacity when the levels of the stimulus are low, which is a desirable feature in any synthetic sensor.

Plant Metabolic Engineering is among the applications more likely to benefit from incorporating engineering principles. Plants have a wide range of metabolites and taking advantage of them it is possible to produce valuable compounds that are difficult to obtain by chemical synthesis or classical biotechnology. In addition to the value of the compounds themselves, metabolic engineering of plants is also highly demanded for their use as improved green factories (Xu et al., 2012). The introduction of new branches on the existing pathways and the redirection of the metabolic flux to reach higher yields of the desired compound is a challenge that SynBio is overcoming thanks to several approaches such as the use of repressible and inducible promoters (Yang et al., 2013) or the repression of non-desired branches of the pathway (Verpoorte & Memelink, 2002). Design of new circuits making use of the TALE or the CRISPR/Cas9 technology for activation or repression is being explored.

As described in this introduction chapter, there are many factors influencing the advancement of Plant SynBio that should be explored. We decided to investigate to what

extent was possible to catalog standard parts by developing a set of software-tools for multigene engineering and data collection adapted to the GoldenBraid standard (Chapter 1). In a next natural step we extended the generated tools to include functional characterization of any genetic element (Chapter 2) and subsequently, each new generated device was tested following standard approaches (Chapters 2 and 3). Furthermore, we proved the feasibility of the generated set of tools with a Metabolic Engineering example (Chapter 4). Altogether, this Thesis represents a significant contribution to the standardization in Plant Synthetic Biology.

2 | Objectives

The driven hypothesis in this Thesis is that the introduction of the Synthetic Biology principles of Standardization, Modularity and Abstraction of Function in the field of Plant Biotechnology will facilitate the achievement of increasingly challenging biotechnology objectives.

Accordingly, the General Objective of this Thesis was the design of key tools that enable the incorporation of those principles in Plant Genetic Engineering, testing their functionality and providing examples of their application in crop biotechnology. To fulfil this General Objective, the following specific objectives were set:

1. To introduce modularity and standardization in the DNA assembly tools for Plant Biotechnology, developing new utilities in the frame of the GoldenBraid cloning system, in particular:
    a. To develop new software tools assisting the assembly of synthetic DNA parts (Chapter 1).
    b. To propose and refine experimental standards for the characterization of synthetic DNA parts (Chapter 2).
    c. To incorporate functional specifications describing synthetic DNA parts in the form of associated datasheets (Chapter 2).

2. To use GoldenBraid capabilities to develop new biotech tools, in particular:
    a. To design small genetic circuits for Plant Genetic Engineering (Chapter 2).
    b. To develop CRISPR/Cas9 tools for Genome Engineering (Chapter 3).
    c. To generate intragenic tools for tomato transformation (Chapter 4).

3 | Chapter 1

Software-assisted stacking of gene modules using GoldenBraid 2.0 DNA-assembly framework

Vazquez-Vilar M, Sarrion-Perdigones A, Ziarsolo P, Blanca J, Granell A, Orzaez D.

Methods Mol Biol. 2015;1284:399-420. doi: 10.1007/978-1-4939-2444-8_20.

1. Introduction

Gene stacking is an important requirement in plant research, reaching its biggest applicability on Plant Metabolic Engineering and Plant SynBio. As custom DNA synthesis is still unaffordable for big constructs, a wide range of DNA assembly technologies appeared on recent years trying to offer alternatives for multigene engineering (Gibson et al., 2009, Geu-Flores et al., 2007, Hartley et al., 2000). The most interesting ones are the modular assembly strategies that approach genetic fragments as standard DNA building blocks that can be combined to create modules with a defined function (Zhang & Jiang, 2010). The standardization of genetic parts builds up on the idea of SynBio as a community effort, promoting exchange of genetic parts and speeding up engineering, since with standardization new composite parts will fit with the old ones (Muller & Arndt, 2012).

GoldenBraid is a modular DNA cloning system that relies on the use of Type IIS restriction enzymes (Sarrion-Perdigones et al., 2011, Sarrion-Perdigones et al., 2014). It makes use of previously-described Golden Gate multipartite assemblies of standard parts (Engler et al., 2009) to generate transcriptional units (TU). These TUs can be combined in a binary way following a double-loop iterative cloning strategy that allows the assembly of increasingly complex multigenic modules. A minimum set of only four destination vectors is required to complete the double-loop since TUs/modules assembled in two compatible vectors of the same level (α) can be combined together in any of the two destination vectors of the opposite level (Ω) and vice versa.

Despite the simplicity of GB, the multigene engineering design process can be tedious and time-consuming, requiring from computational approaches to manage standard parts and to perform *in silico* assemblies. GoldenBraid 2.0 (Sarrion-Perdigones et al., 2013) provides a set of online-tools that software-assists users during part standardization and multigene assembly with GB (www.gbcloning.upv.es).

This Chapter describes the GBcloning assembly procedure for the stacking of three gene modules within the same T-DNA, a simplified example of a Plant Metabolic Engineering approach. The software-assisted design process and the wet-lab methods for GB multigene assemblies are explained in detail. This example covers the assembly of transcriptional units carrying the *Solanum lycopersicum* transcription factors SlANT1 (Mathews et al., 2003), SlJAF13 (De Jong et al., 2004) and SlMYB12 (Adato et al., 2009) regulated by the constitutive 35s promoter. Transcription factors involved in the regulation of the flavonoids biosynthetic pathway, such as SlANT1 (MYB) and SlJAF13 (bHLH), *Solanum lycopersicum* orthologous of the *Antirrhinum majus* Rosea1 and Delila

proteins, have been demonstrated to activate the expression of several anthocyanin biosynthetic genes (Shimada et al., 2007). SlMYB12 over-expression was shown to increase the levels of naringenin-chalcone among other flavonoids (Ballester et al., 2010). To check the functionality of the assembled module, a transient expression assay in *Nicotiana benthamiana* leaves was performed.

2. Results

2.1. GoldenBraid collection

A new database was generated in the Django framework using Python as programming language to host the sequence information of all the parts, modules and vectors included in the increasingly populated GBcollection. All the building blocks in the GBcollection were generated using the GB assembly system and share the GB2.0 standard. The GBelements can be classified in the following categories:

- **GBparts and GBsuperparts** (also known as level 0 parts): fragments of DNA flanked by 4-nucleotides overhangs, which define their category within the TU, stored as inserts within a specially designed entry vector, the universal part domesticator (pUPD). There are eleven categories (in GB2.0 syntax), each of them with its pre-defined flanking sites (Figure 3a). Basic standard categories are named with numbers following a positional notation: positions 01, 02 and 03 comprise the *bona fide* promoter regions, excluding 5´ UTR. Next are positions 11 to 17, which cover the 5´UTR, the ORF and the 3´ UTR. Finally, position 21 is reserved to the 3´ non-transcribed elements. However, contiguous basic categories with a specific function can be grouped together defining a GBsuperpart, which is also a DNA fragment stored within the pUPD vector. Either a BsaI or a BtgZI digestion releases the GBpart from pUPD generating the same 4-nucleotides overhangs. This is possible since both, BsaI and BtgZI, are Type IIS restriction enzymes with different number of nucleotides between their recognition and cutting sites allowing the arrangement of their target sites to cut exactly at the same position.
- **GBdestination vectors (pDGBs)**: binary plasmids that receive the assembled TUs and/or modules in exchange for the LacZ cassette. The GBdatabase contains two sets of GBdestination vectors, the pGreen and the pCambia series, depending on their original backbone. The minimal number of GBdestination vectors for iterative cloning is 4, however 4 additional plasmids are required to enable binary assemblies in all possible orientations. GBdestination plasmids are classified in two levels according to the enzyme that releases the transcriptional unit/module and to their backbone resistance: α level plasmids have kanamycin resistance and release the insert upon BsmBI digestion; Ω

level plasmids contain spectinomycin resistance and BsaI sites. Furthermore, the plasmids of each level are named as 1 or 2 depending on the overhangs obtained after restriction. The compatible plasmids are α1-α2 and Ω1-Ω2.

- **GBtranscriptional units (TU)** (also known as level 1 parts): any DNA structure cloned into any GB destination vector as a result of a multipartite assembly reaction. The simplest way to assemble a new TU is by performing a multipartite (tripartite) assembly reaction with the following elements: a (01-02-03-11-12) GBsuperpart, which comprises all five contiguous standard GBparts and includes the promoter and the 5' untranslated region; a (13-14-15-16) GBsuperpart, which comprises the entire coding sequence and a (17-21) GBsuperpart comprising a 3'UTR and a transcriptional terminator element.
- **GBmodules:** two or more transcriptional units assembled together in one of the GBdestination vectors as a result of a binary assembly reaction.
- **GBpatches** (also known as level -1 parts): non-standard DNA fragments obtained by PCR amplification and employed for the removal of internal restriction sites during the construction of standard GBparts and GBsuperparts (a process known as domestication). GBpatches, not being standard elements, are usually not stored in the GBdatabase.

Besides the public GBdatabase hosting the GBcollection users can build their own private databases. In private databases users can store their own GBelements in order to combine them with those deposited in the public GBcollection.

2.2. GoldenBraid Software Tools

For software-assisted GoldenBraid cloning, a set of free-online software tools was created using Python as programming language.

a) The **GB Domesticator tool** (www.gbcloning.upv.es/do/domestication) takes a fasta or a GenBank DNA sequence file as input and provides the best PCR strategy to remove internal restriction sites and to add flanking nucleotides to it according to the specified category. The GB Domesticator takes into consideration the maintenance of the open reading frame in the coding sequences and the prevention of reconstitution of a restriction site when the mutations are performed.

b) The **GB TU Assembler tool** (www.gbcloning.upv.es/do/multipartite) assists in the *in silico* assembly of new transcriptional units by combination of standard parts or superparts. Only parts in the database with compatible flanking sites are displayed. The program output is a downloadable GenBank file containing the fully assembled DNA sequence (including the destination plasmid) together with a lab protocol describing the cloning procedure.

c) The **GB Binary Assembler tool** (www.gbcloning.upv.es/do/bipartite) performs *in silico* binary assemblies between single TUs or composite modules stored in the

GBdatabase. Only composite parts in compatible GBdestination vectors are displayed. After selection of the first element (position 1), the software will display all the compatible TUs/modules stored in the database. Once the election of the element in position 2 is made, a destination vector can be selected to carry out the assembly. As in the previous tool, output files are a GenBank file with the assembled sequence and a protocol to perform the reaction in the lab. A schematic view of the track followed by a DNA sequence along the three software tools and its interaction with the GBdatabase is shown in Figure 3b.

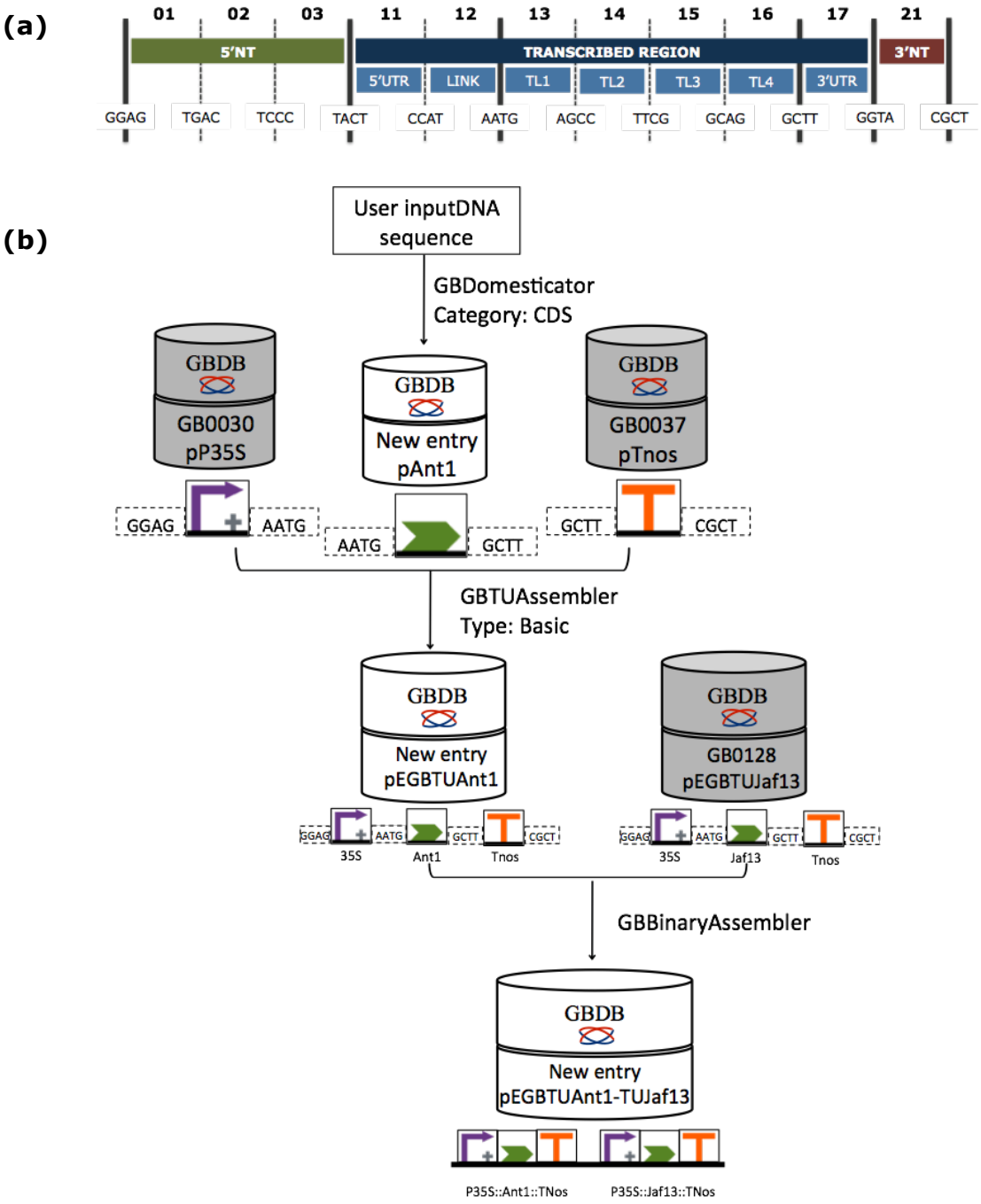


Figure 3. The GoldenBraid cloning software-assisted assembly procedure.

(a) Schematic overview of the GB2.0 syntax where the eleven basic standard categories are depicted: 01, 02 and 03 comprise the 5′ nontranscribed region (5′NT); positions 11 to 17 comprise the transcribed region including the 5′UTR (position 11), a linker (12), four divisions of the translated region (13 to 16) and the 3′UTR (17); and position 21 is the 3′ nontranscribed region (3′NT). (b) The GB Domesticator adapts the input DNA sequence provided by the user to the GBstandard according to the selected category. The GB TU Assembler performs the *in silico* multipartite assembly with any compatible GBpart stored in the database (depicted in grey) to create a transcriptional unit (TU). Finally, GB Binary Assembler allows the binary assembly of preformed single TUs or combinations of TUs (named modules) over the GoldenBraid loop to produce multigenic structures.

2.3. Exemplified GoldenBraid assembly procedure

2.3.1. GBparts domestication

The first step in the GoldenBraid assembly strategy is the adaptation of the new genetic elements to the GoldenBraid standard and their cloning into the pUPD vector, a process known as domestication. This adaptation process implies (i) the addition of specific flanking overhangs to each GBpart/superpart according to its expected position within the transcriptional unit, and (ii) the removal of the internal BtgZI (optional), BsaI and BsmBI restriction sites. For this purpose the sequence is split in *n+1* parts, known as GBpatches, where *n* is the number of internal restriction sites. For each GBpatch a pair of primers is designed. These primers incorporate the appropriate extensions for internal restriction sites mutation and grammar adaptation, as well as flanking BsmBI recognition sites for cloning into the pUPD vector. An overview of the domestication approach is represented on Figure 4a.

As an example, the software-assisted domestication of the coding region of the SlJAF13 transcription factor is described below:

1. Click the GB Domesticator link at www.gbcloning.upv.es/tools/domestication/.
2. Upload the DNA sequence to domesticate as a GenBank or fasta file. Once the file is uploaded, select the intended category for this new element (*see* **Note1**). For domestication of a coding region, as for example *SlJAF13*, select the CDS (13-14-15-16) option. Click "submit", download and open the resulting txt file in a text editor.
3. Order the primers indicated in the text file (*see* **Note 2**). For the domestication of *SlJAF13*, two pairs of primers are required since the sequence has one internal BsaI restriction site (*see* **Note3**). As the selected category is a CDS, the primers were designed to produce a synonymous mutation that removes the internal restriction site while maintaining an intact open reading frame. Note that the same procedure can be followed for the domestication of the remaining transcription factors (*ANT1* and *MYB12*) used in the multigenic construct described in this chapter. The primers employed for the domestication of all three genes are listed in Table 2.

Patch	Forward primer	Reverse primer
SlANT1 patch1	GCGC**CGTCTC**GCTCG*AATG*AACAGTACA TCTATGTCTTCATTG	GCGC**CGTCTC**GCTCG*AAGC*TCAATCAAGTAG ATTCCATAAGTC
SlJAF13 patch1	GCGC**CGTCTC**GCTCG*AATG*GCTATGGGA CACCAAGA	GCGC**CGTCTC**GGCCTCCAAAGGCTATTCTTT T
SlJAF13 patch2	GCGC**CGTCTC**GAGGCCAGAAGTTGATGA CATT	GCGC**CGTCTC**GCTCG*AAGC*TCAAGATTTCCA TACTACTCTCTG
SlMYB12 patch1	GCGC**CGTCTC**GCTCG*AATG*GGAAGAACA CCTTGTTGT	GCGC**CGTCTC**GGCGACCATCTGTTACCCAAA
SlMYB12 patch 2	GCGC**CGTCTC**GTCGCTTATAGCAGAACA TTTATCA	GCGC**CGTCTC**GCGTCGCTTGGTATCCTTAAG
SlMYB12 patch 3	GCGC**CGTCTC**GGACGAGAAGTTACCTAA AGCC	GCGC**CGTCTC**GCTCG*AAGC*CTAAGACAAAAG CCAAGATACAAT

Table 2. Software-designed primers for *SlANT1*, *SlJAF13* and *SlMYB12* domestication as CDS (13-14-15-16).

BsmBI recognition and cleavage sites are represented in bold and underlined, respectively. In italics, category specific flanking sites.

4. Amplify the different GBpatches using the primers specified in the protocol and a suitable DNA template. For instance, *Solanum lycopersicum* mature fruit cDNA was used as template for *SlJAF13* amplification.

5. Check whether the PCR reactions resulted in the expected GBpatches loading one-tenth of the reaction volume in an agarose gel electrophoresis. The size of each GBpatch can be directly calculated from its sequence, which is also provided as result after step 2. Figure 4b shows an electrophoresis gel with all the GBpatches needed for *SlANT1*, *SlJAF1*3 and *SlMYB12* domestication.

6. Purify the PCR fragments from remaining reaction volumes using the QIAquick PCR purification kit, as indicated by the manufacturer.

7. Set up the GBdomestication reaction to get the GBpart cloned into the pUPD plasmid by mixing in the same tube 40 ng of each purified GBpatch, 75 ng of pUPD, 5 U of BsmBI, 3 U of T4 ligase and 1 µL of 10X ligase buffer in a 10 µL reaction, according to the instructions produced by the GB Domesticator software.

8. Incubate the BsmBI restriction-ligation reaction in a thermocycler during 25 cycles (2min x 37ºC, 5min x 16ºC) (*see* **Note 4**).

9. Thaw 50 µL of *E. coli* electrocompetent cells on ice and mix with 1 µL of the reaction product. Pipet the mixture to a pre-chilled electroporation cuvette and carry out the electroporation immediately.

10. Add 500 µL of room temperature SOC media to the cuvette, recover the cells and grow them on a shaker at 37ºC for 1 hour.

11. Spread two different volumes (50 and 500 µL) of the cells onto LB/ampicillin/IPTG/X-Gal plates. Incubate the plates overnight at 37ºC.

12. Pick four white colonies (the blue ones carry the intact pUPD) and inoculate them in LB with ampicillin. Grow them overnight on a shaker at 37ºC.

13. Miniprep the cultures and check the correct domestication of each GBpart by restriction analysis. The restriction map can be obtained from the sequence of the domesticated GBsuperpart also provided as a result of step 2. Following the same procedure *SlMYB12* and *SlANT1* were also domesticated to GBsuperparts with a CDS B3-B4-B5 category. Figure 4c shows the restriction analysis of correctly domesticated GBparts pAnt1, pJaf13 and pMyb12.

14. Check those clones from step 13 showing a correct restriction pattern by Sanger sequencing using the T7 and the SP6 universal sequencing primers.

15. If the resulting sequence is correct, upload the GenBank file obtained from step 2 to the GBdatabase (*see* **Note 5**).

16. Store the GBpart as a DNA miniprep and the cells containing it in the form of a glycerol stock (15% glycerol).

The 35S promoter (GB0030_p35s) and the nopaline synthase terminator (GB0037_pTnos) had been previously domesticated to PROM+UTR+ATG (01-02-03-11-12) and to TER (17-21) standard categories respectively, and stored in the GBdatabase as fully reusable GBsuperparts. The reuse and exchange of previously adapted parts is one of the most important benefits of standardization.

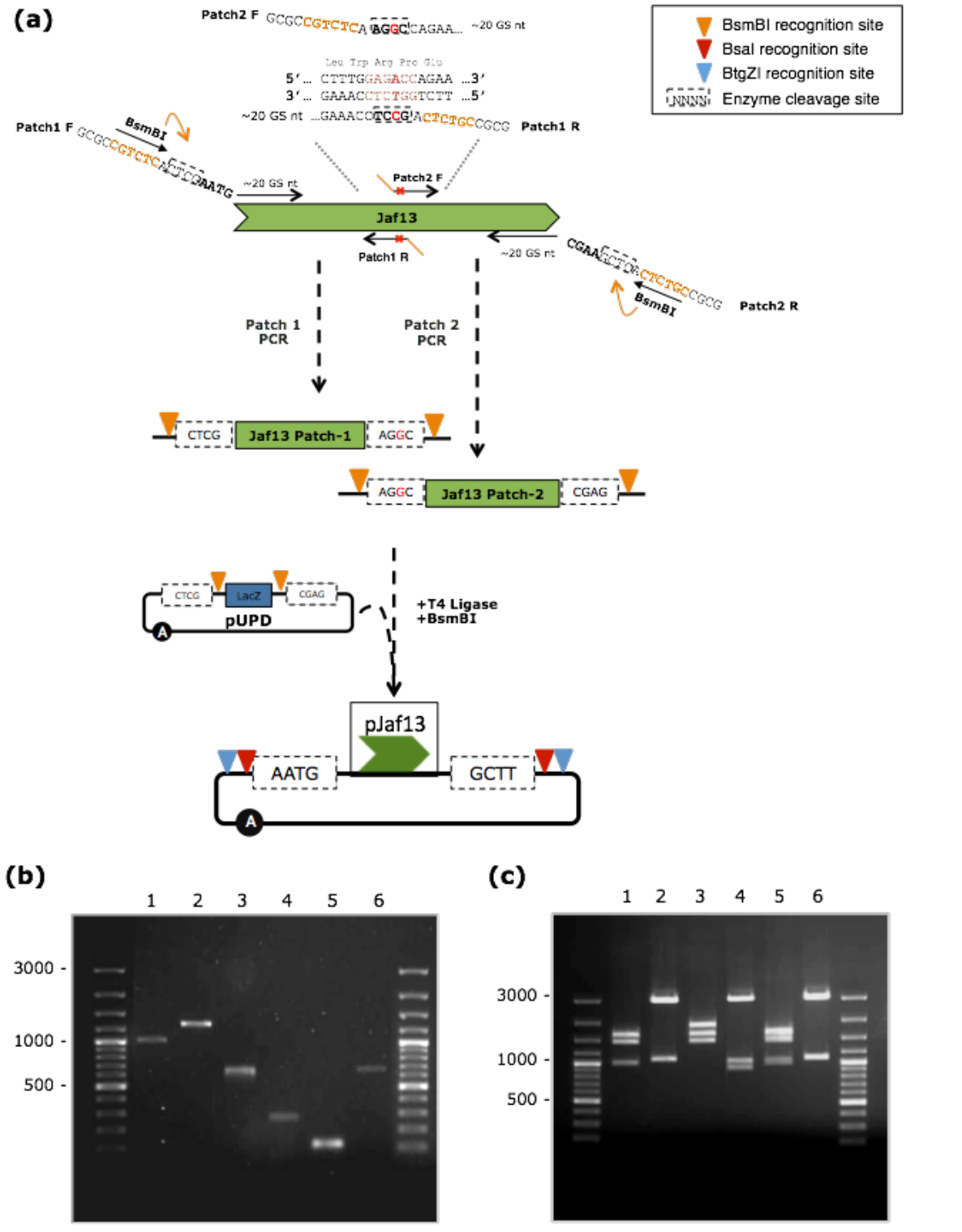


Figure 4. General strategy for GBparts domestication.

(a) Jaf13, having an internal BsaI recognition site (GAGACC), is domesticated by amplifying the target sequence in two separated PCR products. External GBadapted primers (Patch1 F and Patch2 R) are designed so they include a BsmBI recognition site, the cleavage site for cloning into pUPD, and the 4-nucleotide barcodes (AATG and GCTT) in addition to approximately 20 gene specific (GS) nucleotides. Internal GBadapted primers (Patch1 R and Patch2 F) have extensions that incorporate the BsmBI recognition sites and the single nucleotide change (A>G) to mutate the internal BsaI recognition site. After the amplification of both patches, a BsmBI restriction-ligation reaction will bring the domesticated sequence into the pUPD. The BsmBI recognition sequences are shown in orange, and their position is also indicated with orange arrows; BsaI and BtgZI are represented with red and blue triangles, respectively; the enzymes cleavage sites are boxed. Single nucleotide

mutations are marked in red in the DNA sequence (b) PCR amplification products of all the GBpatches required for part domestication of *SlANT1* (one patch of 1048 base pairs in lane 1), *SlJAF13* (two patches of 1309 and 633 bp in lanes 2 and 3, respectively) and *SlMYB12* (three patches of 298, 128 and 678 bp in lanes 4, 5 and 6, respectively). (c) Restriction analysis of correct clones of the GBparts pAnt1 with BsaI (1622-1433-1013 bp) and EcoRI (2997-1071 bp) in lanes 1 and 2 respectively; pJaf13 with BsaI (1882-1622-1433 bp) and EcoRI (2997-1025-915 bp) in lanes 3 and 4 respectively and pMyb12 with BsaI (1622-1433-1018 bp) and EcoRI (2997-1076 bp) in lanes 5 and 6 respectively.

2.3.2. Multipartite assemblies in α-level destination vectors

For the multipartite assembly of a new TU in an α-level destination plasmid, all the GB elements are incubated in a single-tube reaction in the presence of BsaI and T4 ligase. Individual GBparts are released from the pUPD by the BsaI activity. In parallel, the same enzyme opens the α-level destination plasmid releasing the LacZ-cassette. Next, the different GBparts are orderly assembled into the destination plasmid by the T4 ligase, creating a functional transcriptional unit. Only the correctly assembled structures will end up free of BsaI recognition sites, whereas wrong assemblies will remain susceptible to targeted digestion by the restriction enzyme during the entire reaction time.

In the following example, two new TUs, named pEGB1α1_35s::Ant1::Tnos and pEGB1α2_35s::Jaf13::Tnos (Figure 5a), will be constructed in parallel in complementary α-level plasmids, following the procedure described below:

1. Simulate the assembly of the TU on a computer with the GB TU Assembler tool. For basic protein expression in the nucleocytoplasm, select the 'BASIC' shortcut.
2. Select the GBparts to build the desired TU. In this example: the p35S as promoter element, the recently included pJaf13 as coding sequence; and the pTnos element as terminator. Select any of the α-GB destination vectors displayed in the screen, pDGB1α2 in this example (*see* **Note 6**). Click "Submit" to download the assembled DNA file and the assembly protocol in txt format.
3. Set the reaction by mixing 75 ng of each GBsuperpart (the ones selected on the previous step), 75 ng of the α-level destination vector, 5-10 U of BsaI, 3 U of T4 ligase and 1 μl of 10X ligase buffer in a 10 μl reaction, following the specifications of the GB TU Assembler protocol.
4. Incubate the BsaI restriction/ligation reaction in a thermocycler for 25 cycles (37ºC x 2min, 16ºC x 5min) (*see* **Note 4**).
5. Mix 1 μl of the reaction with 50 μl of electrocompetent cells previously thawed on ice. Transform them by electroshock in an electroporation cuvette, outgrow by adding 500 μl of SOC and shake for 1 hour at 37ºC. Plate two volumes (50 and 500 μl) of cells in LB/kanamycin/IPTG/X-Gal petri dishes. Incubate the plates overnight at 37ºC.

6. Pick four white colonies and grow them overnight in LB/kanamycin (*see* **Note 7**).
7. Miniprep the cultures for plasmid extraction and perform a restriction analysis. The choice of the restriction enzymes is based on the restriction map obtained from the sequence of the assembly retrieved as result of the online tool in step 2.

Following the same procedure, the TU carrying Ant1 was assembled into the pDGB1α1 vector. The restriction patterns of correct assemblies of the constructs pEGB1α1_35s::Ant1::Tnos and pEGB1α2_35s::Jaf13::Tnos are shown in Figure 5c (lanes 1 to 4).

8. Once the assembly has been found correct by restriction analysis, upload the GenBank files obtained in step 2 to the GBdatabase.
9. Store the GBtranscriptional unit as a DNA miniprep and the cells containing it in the form of a glycerol stock (15% glycerol).

2.3.3. Binary assembly into Ω-level destination vectors

Once the two TUs pEGB1α1_35s::Ant1::Tnos and pEGB1α2_35s::Jaf13::Tnos are assembled, the next step is to binarily combine them to create a module containing both genes. For this purpose, as they were assembled in compatible α GB destination vectors, they can be joined together with a single BsmBI reaction in any Ω-level GB destination vector with the steps described below. A representation of this assembly is shown at the top of the Figure 6a.

1. Perform the *in silico* assembly reaction using the GB Binary Assembler webtool (https://gbcloning.upv.es/do/bipartite/). Select the previously uploaded TUs and any Ω-level destination vector. In this example, the assembly is performed using the pDGB1Ω1 vector.
2. Set the reaction by mixing 75 ng of each TU (from step 7 of the previous section), 75 ng of the pDGB1Ω1 destination vector, 5-10 U of BsmBI, 3 U of T4 ligase and 1 µl of 10X ligase buffer in a 10 µl reaction.
3. Incubate the reaction in a thermocycler for 25 cycles (37ºC x 2min, 16ºC x 5min).
4. Transform 50 µl of electrocompetent cells with 1 µl of the reaction. Outgrow the cells by adding 500 µl of SOC and incubate shaking for 1 hour at 37ºC. Plate two volumes (50 and 500 µl) in LB/spectinomycin/IPTG/X-Gal petri dishes.
5. Pick four white colonies and grow them overnight in LB with spectinomycin (*see* **Note7**).

6. Miniprep the plasmids and check them by restriction analysis. The choice of restriction enzymes for the analysis is based on the restriction map that can be obtained directly from the sequence of the assembly retrieved in step 1. The restriction patterns of the correct pEGB1Ω1_35s::Ant1::Tnos-35s::Jaf13::Tnos assembly is depicted in Figure 6b (lanes 1 and 2).

7. Upload the GenBank file containing the *in silico* assembled module to the GBdatabase.

8. Store the GBmodule as a DNA miniprep and the cells containing it in the form of a glycerol stock (15% glycerol).

At this point a combination of two transcription factors that activate key enzymes of the anthocyanin biosynthetic pathway (Butelli et al., 2008) was assembled. However, to increase the content of other flavonoids in the plant, MYB12, a third transcription factor will be included in the construct.

2.3.4. Multipartite assemblies in Ω-level destination vectors

For multipartite assemblies in Ω-level plasmids, a BtgZI/BsmBI restriction-ligation reaction is required. In this reaction, BtgZI releases the parts from the pUPD leaving the same overhangs as BsaI, and BsmBI opens the Ω-level plasmids allowing the entry of the GBparts in the right order. BtgZI/BsmBI reactions are less efficient than BsaI reactions, so BsaI reactions are preferred. However, in this example, the assembly of the third TU in a Ω-level destination vector complementary to the pDGB1Ω1, will save us an extra binary reaction. The following steps were followed to assemble the TU named pEGB1Ω2_35s::Myb12::Tnos as it is shown in Figure 5b:

1. Perform the assembly of the TU *in silico* with the GB TU Assembler tool by clicking on the 'BASIC' shortcut.

2. Select the GBparts to conform the TU, namely p35S as promoter, pMyb12 as CDS, and pTnos as terminator. As destination plasmid, the Ω-level destination plasmid pDGB1Ω2, complementary to the pDGB1Ω1 that hosted the binary assembly 35s::Ant1::Tnos-35s::Jaf13::Tnos is selected.

3. Set the reaction by mixing 75 ng of each GBpart, 75 ng of the pDGB1Ω2, 5-10 U of BsmBI, 5 U of BtgZI, 3 U of T4 ligase and 1 µl of 10X ligase buffer in a 10 µl reaction, following the protocol produced by the webtool.

4. Incubate the reaction in a thermocycler for 25 cycles (37ºC x 2min, 16ºC x 5min).

5. Mix 1 µl of the reaction with 50 µl of electrocompetent cells. Transform them by electroshock into a cuvette and outgrow by adding 500 µl of SOC and shaking during 1 hour at 37ºC. Plate two volumes of cells (50 and 500 µl) in LB plates containing spectinomycin, IPTG and X-Gal.

6. Pick four white colonies and grow them overnight in LB with spectinomycin (*see* **Note 7**).

7. Miniprep the cultures for plasmid extraction and analyze them using restriction enzymes and gel electrophoresis. The choice of the restriction enzymes for the analysis is based on the restriction map obtained from the sequence of the assembly retrieved as result of the online tool in step 2. The restriction pattern of a correct colony of the pEGB1Ω2_35s::Myb12::Tnos assembly is shown in Figure 5c (lanes 5 and 6).

8. Upload the GenBank file of the new TU to the GBdatabase.

9. Store the GBtranscriptional unit as a DNA miniprep and the cells containing it in the form of a glycerol stock (15% glycerol).

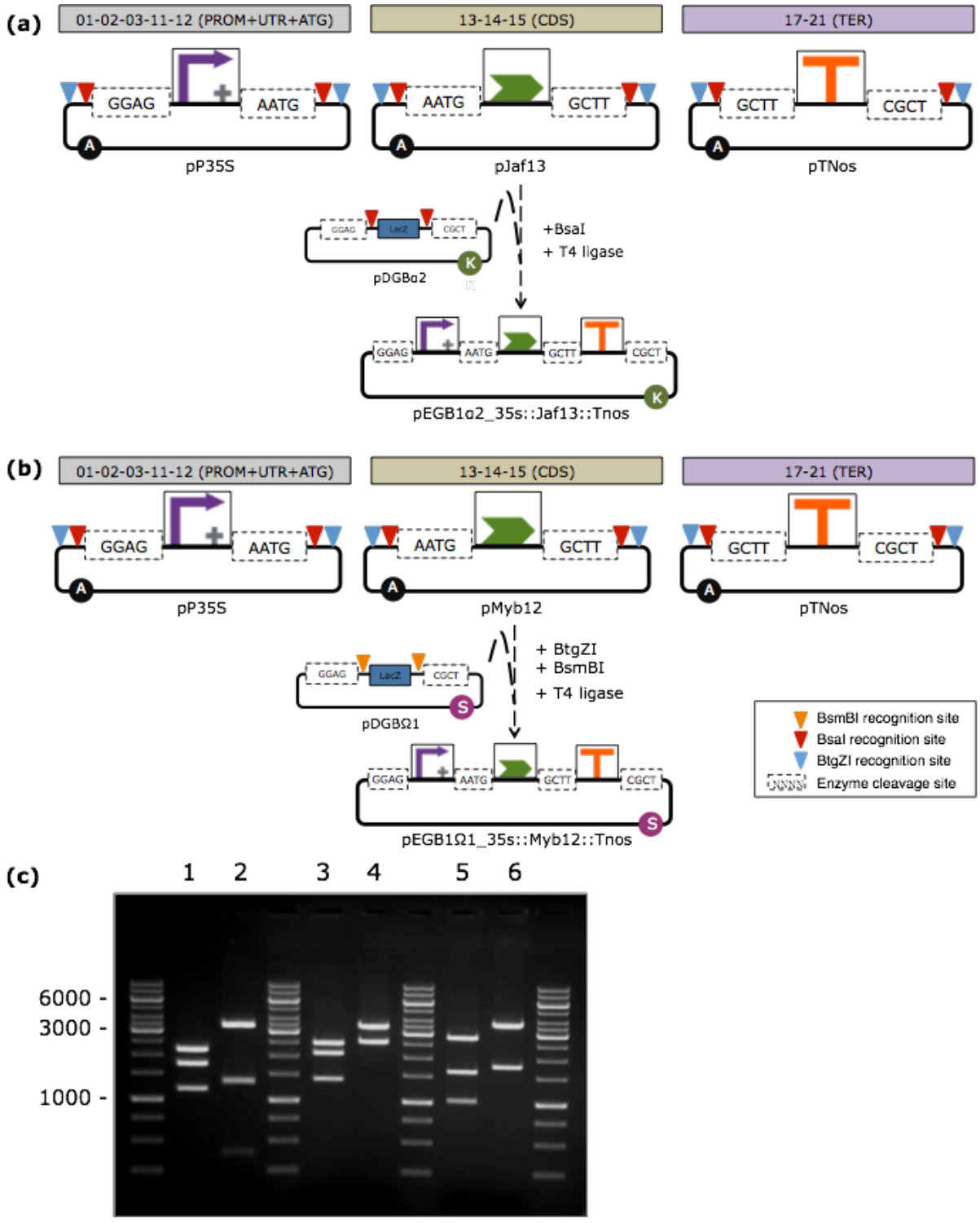


Figure 5. Multipartite assemblies in α and Ω destination vectors.

(a) Multipartite assembly of the GBparts GB0030_p35s, pJaf13 and GB0037_pTnos for the construction of the TU pEGB1α2_35s::Jaf13::Tnos in a α-level destination vector. (b) Multipartite combination of the GBparts GB0030_p35s, pMyb12 and GB0037_pTnos for the assembly of the TU pEGB1Ω2_35s::Myb12::Tnos in a Ω-level destination vector. (c) Restriction pattern of correct assemblies of the plasmids pEGB1α1_35s::Ant1::Tnos (in Lane 1 EcoRV+PvuI restriction bands 2188-1750-1202; in Lane2 NcoI bands 3404-1346-390), pEGB1α2_35s::Jaf13::Tnos (in Lane 3 BglII restriction bands 2489-2116-1405; Lane4: HindIII digestion bands 3437-2573) and pEGB1Ω2_35s::Myb12::Tnos (in Lane 5 BglII restriction bands 2818-1621-1037; Lane 6: PvuI digestion bands 3701-1775).

2.3.5. Final binary assembly in α-level destination vectors

The final assembly step is the binary combination between the Ant1/Jaf13 module (pEGB1Ω1_35s::Ant1::Tnos-35s::Jaf13::Tnos) and the Myb12 single transcriptional unit (pEGB1Ω2_35s::Myb12::Tnos) to reach a three-gene combination for polyphenol overproduction. This assembly implies a binary reaction with BsaI in any α-level destination plasmid, which is represented in Figure 6a (bottom).

1. Perform the software-assisted assembly with the GB Binary Assembler tool. Select the previously uploaded module pEGB1Ω1_35s::Ant1::Tnos-35s::Jaf13::Tnos, the TU pEGB1Ω2_35s::Myb12::Tnos and any α-level destination vector. In this example, the assembly is performed in the pDGB1α1 vector.
2. Set the reaction by mixing 75 ng of the module DNA, 75 ng of the TU DNA, 75 ng of the pDGB1α1 destination vector, 5-10 U of BsaI, 3 U of T4 ligase and 1 µl of 10X ligase buffer in a 10 µl reaction, according to the protocol obtained in step 1.
3. Incubate the reaction in a thermocycler for 25 cycles (37ºC x 2 min, 16ºC x 5 min).
4. Transform 50 µl of electrocompetent cells with 1 µl of the reaction. Outgrow the cells by adding 500 µl of SOC and shaking during 1 hour at 37ºC. Plate two cell volumes (50 and 500 µl) in LB plates containing kanamycin, IPTG and X-Gal.
5. Pick four white colonies and grow them overnight in LB with kanamycin (*see* **Note 7**).
6. Miniprep the plasmids and check them with a restriction analysis. The choice of restriction enzymes for the analysis is based on the restriction map that can be obtained from the sequence of the assembly retrieved as result of the online tool in step 1. The restriction patterns of a correct assembly pEGB1α1_35s::Ant1::Tnos-35s::Jaf13::Tnos-35S::Myb12::Tnos are shown in Figure 6b (lanes 4 and 5).
7. Upload the GenBank file containing the *in silico* assembled module to the database.
8. Store the GBmodule as a DNA miniprep and the cells containing it in the form of a glycerol stock (15% glycerol).

2.3.6. *Agrobacterium tumefaciens* transformation and transient expression in N. benthamiana leaves.

To verify the ability of the assembled construct to induce the polyphenol biosynthetic pathways, a transient expression experiment was performed in *N. benthamiana* leaves. In particular, the accumulation of anthocyanins is evidenced by an intense purple color observable with the naked eye (Bedoya et al., 2012).

1. Transform 10 ng of the plasmid containing the final multigene module into 50 µL of *A. tumefaciens* home-made electrocompetent cells carrying the pSoup vector (*see* **Note 8**); collect the cells from the cuvette with 500 µL of LB and grow them on a shaker at 28ºC for 2 hours.
2. Spread two cell volumes (20 and 100 µL) on LB/kanamycin/rifampicin/tetracycline plates (*see* **Note 9**). Incubate plates at 28ºC for two days.
3. Pick four colonies and inoculate them in 5 ml liquid LB containing kanamycin, rifampicin and tetracycline. Grow them for two days on a shaker at 28ºC.
4. Miniprep the cultures and check the colonies by restriction analysis.
5. Subculture (1/100 dilution) into a new tube (5 mL final volume) and grow overnight at 28ºC.
6. Pellet the cells by centrifugation (20 min, 2000 ×*g*).
7. Re-suspend cells in agroinfiltration buffer (10mM MES, pH 5.6, 10mM $MgCl_2$, and 200µM acetosyringone) and incubate them for 2 hours at room temperature on a horizontal rolling mixer in the dark.
8. Dilute the cell suspension with agroinfiltration buffer to a final optical density of 0.2 at 600 nm.
9. Infiltrate the cell suspension into the leaf intercellular spaces of 4-6 weeks old *Nicotiana benthamiana* plants using a needle-free syringe.
10. Ten days post-infiltration, anthocyanin over-production can be visualized in the infiltrated leaves. See Figure 6c.

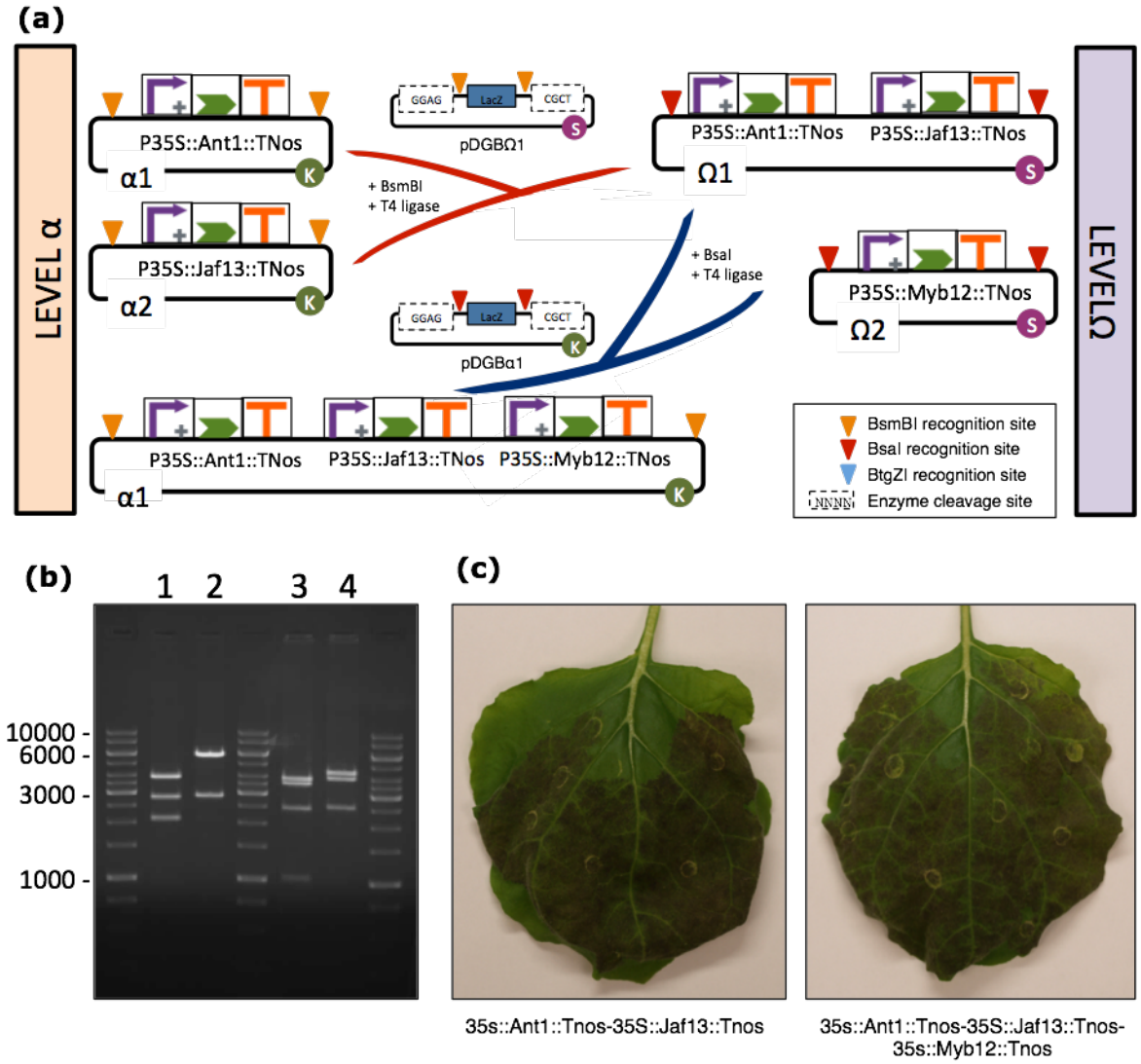


Figure 6. Binary assemblies in α and Ω levels for polyphenols overproduction.

(a) From α to Ω level: schema of a binary assembly between TUs pEGB1α1_35s::Ant1::Tnos and pEGB1α2_35s::Jaf13::Tnos in the Ω-level vector pDGB1Ω1 to create the module pEGB1Ω1_35s::Ant1:.Tnos-35s:Jaf13::Tnos. From Ω to α level: schema of the binary assembly between the module pEGB1Ω1_35s::Ant1:.Tnos-35s:Jaf13::Tnos and the TU pEGB1Ω2_35s::Myb12::Tnos to create the three TUs assembly pEGB1α1_35s::Ant1:.Tnos-35s:Jaf13::Tnos-35s::Myb12::Tnos. (b) BglII (3949-2818-2121) (Lane 1) and BamHI (5985-2903) (Lane 2) restriction patterns of a correct clone of the binary assembly pEGB1Ω1_35s::Ant1:.Tnos-35s:Jaf13::Tnos; BglII (3948-3633-2489-1040) (Lane 3) and EcoRI (4479-4057-2574) (Lane 4) restriction patterns of the final multigenic assembly pEGB1α1_35s::Ant1:.Tnos-35s:Jaf13::Tnos-35s::Myb12::Tnos. (c) Pictures of two *N. benthamina* leaves expressing the combination of the transcription factors, Ant1 and Jaf13 (left) and Ant1, Jaf13, Myb12 (right).

3. Discussion

Standardization in biology simplifies the assembly of transcriptional units from genetic parts as well as the stacking of several genes in the same T-DNA. Although sequence manager tools can help during the design of multigenic assemblies, it is important to

automate this process minimizing the user input to achieve the expected result. However, the coexistence of many assembly standards is a challenge for the creation of shared repositories of genetic parts and software tools for multigenic assemblies.

In recent years, many information-sharing platforms for data management have emerged to handle synthetic biology data and favor the communication among users (Ham et al., 2012, Mutalik et al., 2013b, Peccoud et al., 2008). These databases are the way to share and exchange, as well as keep organized all the data generated in a synthetic biology laboratory. However, most of them are poorly curated and the stored parts are not always well characterized and not suitable for plants (Peccoud et al., 2008). The design and development of the GBdatabase opens a new way to exchange biological parts for Plant Synthetic Biology and Plant Metabolic Engineering.

GoldenBraid offers, not only a repository of standard parts, but also an integrated framework for multigene engineering including software tools that facilitate the multigenic assemblies. Modular assembly methods allow the automation and make easy the development of computational tools to design the assemblies (Muller & Arndt, 2012). The set of software tools designed for GoldenBraid guides users from a raw DNA sequence to the assembly of a multigenic structure. One of the priorities during the development of these tools was to offer the user maximum simplicity. The assembly of a three-genes module described on this chapter shows that a minimal previous knowledge of the system is sufficient to start using these tools.

The population of the GBdatabase with our in-house collection of GBparts, GBsuperparts, GB transcriptional units and GBmodules showed the functionality and potential of the different software tools. The integration of the GBdatabase together with the software tools in a website established a framework from which we expect the plant community can benefit.

4. Notes

1. The GB category is given by the position of each part within the TU (see www.gbcloning.upv.es/do/multipartite). Decide the type of TU to assemble before starting part domestication. For *in silico* domestication of unprocessed coding sequences containing introns, use the free option of the GB Domesticator since the automatic one requires an uninterrupted open reading frame.
2. GB-adapted primers consist of 20-40 nucleotides and standard desalt purification is sufficient.

3. When internal restriction sites are less than 70 nucleotides away from the 3' or 5' end of the sequence or when internal restriction sites are too close within the sequence, they are domesticated with long primers. A long GB-adapted primer includes as many single nucleotide mutations as restriction sites to be removed and has a minimum length equal to the farther mutation position plus seven. The GB Domesticator automatically designs long primers when this option is preferred
4. Increasing the number of restriction/ligation cycles up to 50 can improve the efficiency of the reaction, especially in multipartite assemblies with more than three GBparts.
5. To upload a new sequence use the tab 'Add GB_element' and fulfill the form with the right data. To facilitate the upload of new entries, each results page in the tools section provides a direct link that can be used to upload the new GBelement to the database, saving time and minimizing the chances of introducing wrongly assigned GBelements in the database.
6. The choice of GBdestination vector depends on the number and order of TUs to be assembled. The *in silico* design of the final intended construct is highly recommended. Remember that vectors named as 1 and 2 of the same level (α1-α2 and Ω1-Ω2) are always compatible regardless of their backbone (pGreen or pCambia) and the TU orientation (forward or reverse).
7. Any *E.coli* cell transformed with an undigested GBpart will not grow due to antibiotic counterselection. Bacteria transformed with an intact α or Ω-level GBdestination plasmid will result in blue colonies, due to the presence of X-Gal and IPTG-LB in the plates.
8. Sinc*e som*e of the GBdestination vectors are based on pGreenII, *Agrobacterium* cells require a co-resident plasmid, the pSoup, for pGreen replication (Hellens et al., 2000). If *Agrobacterium* cells with pSoup are not available, the final construct can be co-transformed with the pSoup vector. The set of GBdestination vectors based on pCambia does not require pSoup.
9. Rifampicin is added to the media to reduce no *A. tumefaciens* bacteria contamination; tetracyclin is used for pSoup selection and kanamycin is used since the construct transformed in *A. tumefaciens* in this Chapter is assembled in an α-GBdestination vector.

4 | Chapter 2

Refinement of the Standardization of Phytobricks with the Plant Standard Syntax

Vazquez-Vilar M, Fernandez-del-Carmen A, Sarrion-Perdigones A, Ochoa-Fernandez R, Ziarsolo P, Blanca J, Granell A, Orzaez D.

1. Introduction

Synthetic Biology aims to applying the engineering principles of Standardization, Modularity and Abstraction of function to Biotechnology. It is expected that the adoption of these design principles will speed genetic engineering and will allow the accomplishment of increasingly complex goals. Synthetic Biology is influencing Plant Biotechnology primarily with the adoption of new cloning methods, now renamed as DNA assembly methods. A panoply of new assembly strategies have been developed for Plant Biotech, based either on site-specific recombination (Dalal et al., 2015), PCR-overlap (Cha-aim et al., 2009, Nour-Eldin et al., 2010) or Type IIS enzymes (Engler et al., 2008, Blake et al., 2010), which bring the efficiency required to facilitate complex multigene engineering. Type IIS systems based on the original Golden Gate strategy (Engler et al., 2009) are particularly interesting in the context of Synthetic Biology, as they open the way for the definition of assembly standards that, if widely adopted, will facilitate the exchange of DNA parts. In this respect, a common syntax for Golden Gate-based methods as MoClo (Weber et al., 2011a) and GoldenBraid (Sarrion-Perdigones et al., 2011) has been recently proposed and supported by developers and users of those technologies. This Standard Plant Syntax establishes the physical composition rules that govern the way in which individual DNA components (referred as phytobricks) are to be connected together to create higher order modules and devices, as for instance how to clone a promoter next to a CDS and a terminator to create a transcriptional unit.

The definition and the adoption of standard rules for physical assembly of genetic elements in plants is a first step forward in Synthetic Biology. However, a far more challenging objective ahead is the definition of functional rules that allow standard biological components (i.e. parts, modules or other devices) to be reliably and predictably assembled into higher order functional devices (Wang et al., 2013b) in a systematic way. Following a flow chart borrowed from mechanical engineering, standard specifications describing basic parts could be employed to anticipate the behavior of a composite part, and conversely, new composite parts could be designed by selecting their basic components, based solely on their standard specifications.

It can be anticipated that the complexity and the stochasticity at molecular level of biological networks will make the definition of even the simplest set of functional composition rules an extremely challenging task. However, it should be also noticed that rudimentary (intuitive) composition rules are underlying most of design decisions taken by today´s plant biotechnologists. For instance, the classical selection of a strong ethylene-regulated fruit-specific promoter to drive the expression of a transcriptional

factor upon fruit ripening (Butelli et al., 2008), is ultimately a non-systematic/non-standard exercise of implementing a functional composition rule.

A step further in Plant SynBio, beyond the adoption of common assembly standards, should be the establishment of comprehensive collections of parts covering a wide range of genetic activities and organized in databases where the experimental information required for functional description of each part is also integrated. The latter is a prerequisite for establishing any kind of functional composition rules. Although a few collections of standard parts for Plant Biotechnology have been recently created and deposited in repository databases (Engler et al., 2014), this effort is still very limited, and to date very little has been advanced in the integration of functional specifications in those databases.

Integrating experimental data into genetic databases can be of very little use unless certain uniformity in the experimental conditions is established. A strategy to partially circumvent this problem is the definition of standard experimental conditions for the description of parts belonging to the same category (Schaumberg et al., 2015). Performing quantitative characterization of biological parts and then summarizing their properties in the form of standard datasheets has been previously proposed as a way to maximize the usability of this information (Canton et al., 2008). Datasheets physically and functionally describe each element in a collection. Ideally, standard descriptions contained in datasheets should facilitate the creation of new assemblies and the anticipation of their behavior under different circumstances. This is conceivable specially with those collections whose elements are modular and reusable in biological sense, meaning that once created can be reassembled or replicated without changes (e.g. without introducing assembly seams or PCR-born errors).

In this Chapter we describe the development of GoldenBraid 3.0 (GB3.0), the first database of reusable genetic elements for Plant SynBio that incorporates functional descriptions of its synthetic parts. We have built GB3.0 database on top of the previously described GB2.0 assembly system. In its previous version, the GoldenBraid database stored only the sequence information of each DNA element. New genetic devices were assembled using software-assisted tools instructed with the so-called GB physical composition rules (Sarrion-Perdigones et al., 2013). The new GB3.0 assembly software adopts for the first time the new Plant Standard Syntax (PSS) (Patron et al., 2015) and registers the assembly history of each composite part, keeping record its genealogy. Most notably, GB3.0 enables the definition of standard experiments and the introduction of experimental results in the database. As a result, the new GB3.0 DNA elements are described by standard datasheets displaying their genealogy, their physical sequence information and their behavior under standard experimental conditions. Finally, to

illustrate the possibilities that GB3.0 offers for the design of novel genetic devices, we assembled a small genetic circuit where a chemically-inducible regulatory module was connected to a regulatory module controlling *in series* the activity of the dihydroflavonol 4-reductase DFR promoter, resulting in unprecedented glucocorticoid-dependent modulation of the DFR promoter and showing that descriptions of individual modules can led to anticipate the behavior of the whole circuit.

2. Results

2.1. Refinement of the GB assembly tools and rules: The GB3.0 software package.

A renovated website at www.gbcloning.upv.es was created to host the new features of GB3.0. As displayed in its front-page, the new GB3.0 web comprises four major sections, namely Design, Collection, Experiment and Genome Engineering.

The **Design** section contains improved software tools that facilitate *in silico* assembly of multigenic constructs. Briefly, the GB software, comprises three webtools, namely the domesticator, the multipartite assembler and binary assembler. The domesticator tool serves as entry point, converting raw DNA sequences in standard level 0 GBparts, typically promoters, coding regions, protein domains, terminators, etc. For domestication, raw DNA sequences are cloned into standard entry vectors while internal restriction sites are removed by PCR-mutagenesis. The multipartite assembler takes individual level 0 parts and clones them together to create level 1 elements, typically a full transcriptional unit. Later, level 1 elements can be assembled binarely (Binary assembler tool) in an iterative way to create level >1 elements, typically multigenic constructs. Although the basis for *in silico* assembly were developed in GB2.0 and described in Chapter 1, the new version incorporates a number of updates and improvements listed below (see Supplementary Figure 1):

- Two new part domestication tools (the 'Phylogeny Search' and the 'Synthetic Strategy') were introduced. As the domestication of coding DNA parts is often hampered by the presence of internal restriction enzyme sites (RES), the first tool searches plant genome databases to find homologous sequences from related species that contain minimal internal restriction sites as potential substitutes. With gene synthesis becoming increasingly affordable, it is often convenient to chemically synthesize RES-free DNA fragments and clone them directly in the Universal entry vector instead of removing internal RES by mutagenesis. The Synthetic Strategy tool was introduced to enable this domestication option.

- Domestication tools allow the removal of additional RES (e.g. BpiI) to enhance compatibility with other plant assembly methods (e.g. MoClo).
- In response to user´s requests, a domestication option for intron-containing coding regions was introduced, that avoids the disruption of the ORF.

The **Collection** section serves as interface for the GBdatabase. The GB3.0 database has been reshaped to adapt to the new PSS. Accordingly, all GB parts allowed in the database conform to PSS. A SBOL-inspired symbol was assigned to each basic GBpart, and to each relevant combination of parts. An interactive page was designed to browse GBparts in the database using GB symbols (Figure 7). Furthermore, the database was updated with new structural elements adapted to PSS:

- A new set of destination pCambia-based vectors (pDGB3 series) was introduced fixing bugs and increasing transformation efficiency.
- The ampicillin resistant Universal Entry Vector of GB2.0 was substituted by a new chloramphenicol resistant pUPD2 as the standard vector where all GB3.0 parts are stored. Adoption of pUPD2-hosted GBparts enhances compatibility with MoClo and with iGEM collections. Accordingly, all new GB parts are domesticated using pUDP2. Parts in old pUPD version can be used in new assemblies, but the use of pUPD as entry vector is no longer supported by GB software tools.
- Besides GBparts, the extended GB3.0 database harbors now experiments performed with standard GBparts. Experiments need to conform to one of the defined standard experimental types, and can be searched in the database attending to different strings (experiment type, parts involved, quantitative output range, etc.). Experimental data is used to feed quantitative descriptions of parts and modules displayed in standard datasheets (see below).

The new **Experiment** section contains templates for the introduction of experimental data. In GB3.0 database the information available for each DNA part is enriched with the introduction of experimental results. This is achieved with the introduction of a new type of element in the GB schema, the GBexperiment. A GBexperiment is defined as a set of data produced as the result of the transformation of one or more GB elements in a plant chassis. The experimental data comprises both the experimental conditions in which the experiment has been performed, as well as any observation (results), quantitative or qualitative, obtained as a consequence of the presence of the GB element(s) within the plant chassis. New experiments are introduced in the database by filling a standard questionnaire, where either qualitative (images, text descriptions) or numeric results (in the form of a standard datasheet) can be uploaded. An important feature of the GBexperiment is that it is always associated to a GB element, which corresponds to the complete GB composite part that was employed to perform the test. As the GBdatabase

keeps record of all individual parts that were combined to build the GB element, every new experiment enriches the description not only of its main GB element, but also of all individual components that make it up. Tentatively, we have defined five standard experiment types (SE_001 to SE_005), plus one non-standard template (SE_000). SE_001 and SE_002 templates were designed to accommodate transcriptional activity data obtained with the normalized Luciferase/Renilla reporter pair in agroinfiltrated *Nicotiana benthamiana* leaves. SE_002 is reserved to experiments with intact leaves whereas SE_001 accepts leaf disks incubated with chemical inducers (described in more detail in subsequent sections). SE_003 template describes stable transformation efficiencies, SE_004 is used to collect standardized levels of recombinant protein production, and SE_005 was defined to collect mutagenesis efficiency in CRISPR/Cas experiments. Each experiment type, except NS_000, is defined by a number of compulsory experimental conditions, accepts additional (declared) specific conditions, and it is expected to produce a limited number of predefined quantitative outputs. SE_001 for instance has strict rules as for the plant chassis, growing condition, disks size, harvesting and incubation timing, enzyme reaction conditions, internal references, etc., but it is agnostic towards the use of chemical inducers, optogenetic signals, temperature, etc. Detailed experimental conditions for each experiment type are shown in https://gbcloning.upv.es/add/experiment/ or in the Supplementary Table 1. Examples of all experiments can be consulted in GBdatabase (https://gbcloning.upv.es/search/experiment/) or representative ones in the Supplementary Table 2.

Finally, GB3.0 incorporates a new section devoted to **genome engineering**. This section integrates tools for construction of gRNA and CRISPR/Cas9 assemblies for multiplexing gene editing. The adaptation of CRISPR/Cas9 system to the GB3.0 standard is described in Chapter 3 and in Vazquez-Vilar et al. (2016).

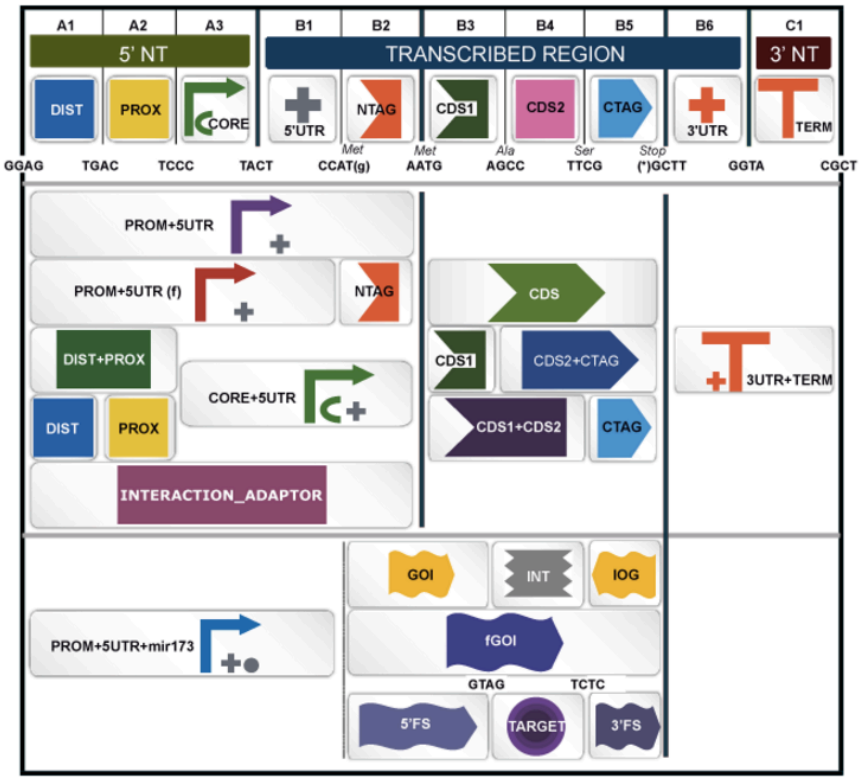


Figure 7. Parts database browser page.

Screenshot of the parts database browser search including SBOL-based images as shortlinks to the different part categories.

2.2. The GB3.0 standard datasheet: integrating experimental data.

A consequence of the introduction of GB experiments to the GB database is that the GB elements can now incorporate functional information to their descriptions. To handle this information we have designed a basic GB datasheet, a page that displays the most important information of each GB element. The basic datasheet information includes name, type of element (level 0, 1, >1; and the standard position in case of level 0), a short text description of the element, the GB plasmid where it is hosted, a graphical description adapted from the SBOL standard (Quinn et al., 2015) and, when available, the graphical output of a maximum of two representative experiments. The objective of the basic GBdatasheet is to bring maximum information about the GBelement in a single screenshot. In addition, more detailed information is linked to the basic page that comprises, among others, the nucleotide sequence in GenBank format, and a clickable list with all the experiments performed with that element.

An example of a GB3.0 datasheet is depicted in Figure 8a, and shows the standard specifications of the level 0 GB0030 phytobrick (see online entry at https://gbcloning.upv.es/feature/GB0030/). GB0030 was constructed based on the constitutive 35sCaMV promoter, as it is indicated in the Description section, and it is flanked by GGAG and AATG and therefore comprises basic parts A1-A2-A3-B1-B2 (see Figure 7). The datasheet includes, next a SBOL-like symbol representative of its standard position, sequence-related information as the plasmid where GB0030 is inserted, the presence of internal sites, selection marker, and a link to the complete DNA sequence in GenBank format. Level 0 phytobricks are introduced in the system directly via the Domesticator tool and therefore do not generate a genealogy section. Most remarkably, the datasheet contains the complete clickable list of experiments where GB0030 has taken part as a key element. Two representative experiments are depicted in the lower part of the datasheet, illustrating the behavior of GB0030 under standard conditions. In Figure 8a, GB0030 activity is reported next to the reporter system Luciferase/Renilla (as phytobrick GB0164 or GB1119), therefore conforming standard experiments SE_001 and SE_002.

An example of a datasheet describing a composite phytobrick is depicted in Figure 8b. In this case, GB0157 entry corresponds to a Transcriptional Unit comprising four standard basic parts, GB0552 (the A1-A2-B1 CaMV35S promoter), GB0531 (a B2 NT encoding the rat glucocorticoid receptor domain), GB0465 (a B3-B4-B5 CDS encoding a chimeric transcriptional factor comprising the binding domain of LacI and the transcription activation domain of Gal4) and GB0037 (a B6-C1 phytobrick derived from the transcriptional terminator of the Nopaline Synthase). As shown in the Figure 8b, GB0157 datasheet contains a "Component Elements" section describing this genealogy, which is automatically generated during *in silico* assembly within the GB3.0 frame. All the genealogy elements are clickable and linked to the standard datasheets of the phytobricks of the previous assembly level. As in the previous example, GB0157 datasheet contains a list of standard experiments performed with this composite part, and displays a representative chart showing the transcriptional activation that GB0157 confers to a minimal promoter containing a LacI operator in presence of dexamethasone, as evaluated using the SE_001 standard (see https://gbcloning.upv.es/feature/GB0157/ for an online version of GB0157 datasheet).

Figure 8. Examples of two GB3.0 datasheets.

(a) Datasheet of the level 0 phytobrick 35s (GB0030) including a SBOL-inspired symbol describing its category (top), assembly information such as part category, vector or information of compatibility with the MoClo standard (general information table) and experimental information with links to all experiments performed with devices including this phytobrick as key element (experiments table) and graphs of two experiments (bottom). (b) Datasheet of the composite phytobrick GB0157 including a compilation of the SBOL-inspired symbols describing each of its components (top) a list of these components with links to each of them facilitating the traceability of the assembly (component elements table), assembly information such as the vector and phytobrick category (general information table) and experimental information with links to all experiments performed with devices including this phytobrick as key element (experiments table) and the graphs of the last two experiments (bottom).

2.3. Refinement of the standard measurements of transcriptional activity in plant cells

Our group focuses in the design of transcriptionally-regulated genetic circuits in plants, and therefore we paid particular attention to the establishment of quantitative measurements for estimating transcriptional activity using phytobricks, as represented by the proposed standard SE_001. Relative Transcriptional Activities (RTAs) in SE_001 are measured in the *N. benthamiana* leaves making use of Agrobacterium-mediated transient expression, using Luciferase/Renilla system as indirect measurement of the steady-state transcript levels. Furthermore, the relative luciferase measurements are

normalized with an internal reference (phytobrick GB0166, a fixed Luciferase/Renilla device where Luciferase is driven by the weak promoter of the Nopaline synthase), agroinfiltrated in a different plant during the same experiment, and its Luciferase/Renilla readings are arbitrarily set as one rpu (relative promoter unit). The GB0166 reference was found necessary to allow comparison among experiments. The remaining SE_001 experimental conditions were carefully selected on the basis of detailed observations. An important parameter in transactivation experiments is the average number of active T-DNA copies that enter a leaf cell upon agroinfiltration and its dependence with the concentration of the Agrobacterium cultures. To estimate this parameter, a strategy similar to that used to estimate the Multiplicity of Infection (MOI) in viral infections was employed, assuming that the rate of co-transformation of two independent phytobricks (GB1287 and GB1288, carrying T-DNAs expressing Yellow and Red fluorescent proteins respectively, Figure 9a) follows a Poison distribution. As can be observed in Figure 9b, the number of transcriptionally-active T-DNA copies (defined here as Multiplicity of Transformation, MOT) adjusts well to a logarithmic function of the optical density of the agro culture, with a MOT=1 obtained at OD=0.002. MOT dependence with OD decreases as OD increases. We therefore decided to set a standard OD= 0.1 (estimated MOT= 7.4) for SE_001 and SE_002, as it balances maximum MOT stability (low dependence to OD variations) with low Agrobacterium input and acceptable copy numbers for transactivation activities.

The experimental data deposited in the GBdatabase was then used as an internal test of the reproducibility of the data obtained using SE_001 and SE_002. We took advantage of the repeated use GB0030 as internal control in several independent experiments to test the variability of the transcriptional activity of a CaMV35S constitutive promoter under standard conditions. The GB0030 datasheet was found to contain links to nine SE_001 and SE_002 experiments, and the relative transcriptional activity in each experiment was plotted and depicted in Figure 9c. As observed in the Figure 9c, average GB0030-confered transcriptional activity was maintained in a range between 9 and 13 relative units, this despite the use of two different backbones (either pGreen or pCambia) and that the depicted experimental conditions involved different plant batches, growth chambers, and individual researchers.

The remaining standard experiments were defined to illustrate the generation of an increasingly informative GBdatabase. We assembled and experimentally tested a number of new GB devices using the five pre-stablished experiment types and uploaded them to the database. Supplementary Table 3 shows a non-exhaustive list of characterized devices comprising protein-protein interactors, constructs for metabolic engineering,

recombinant protein production, tomato stable transformation, CRISPR/Cas9-based mutagenesis and transcriptional regulation, among others.

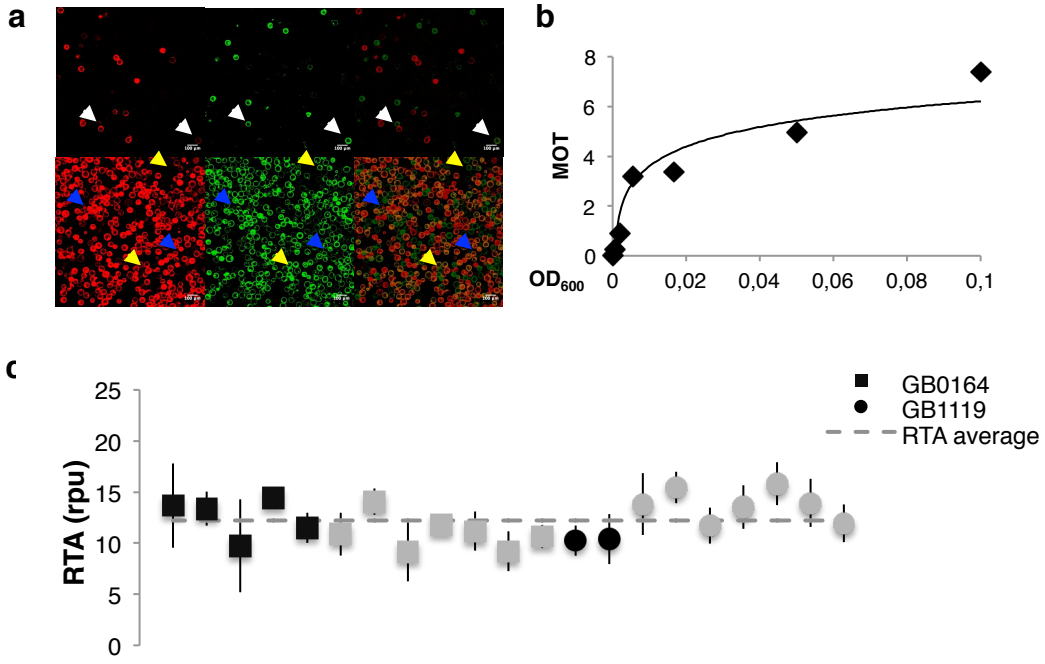


Figure 9. Relationship between the OD_{600} and the multiplicity of transformation in transient expression assays and behavior of the 35s promoter among experiments.

(a) MOT follows a logarithmic distribution in relation to the OD_{600}. Multiplicities of transformation were estimated based on the % of cotransformed protoplasts (see Materials and Methods for details). (b) Pictures of protoplasts cotransformed with GB1287 and GB1288 at OD_{600}= 0.000617 (top) and OD_{600} = 0.05 (down). From left to right Red channel, GFP channel and overlay of the two previous pictures. White arrows point to cotransformed protoplasts, blue arrows to protoplasts showing only red fluorescences and yellow arrows to protoplasts showing only yellow fluorescence. (c) Relative transcriptional activity of the 35s promoter (GB0030) tested as part of devices GB0164 and GB1119 over different experiments. GB0164 has a pGreen backbone while GB1119 has a pCambia backbone. Black squares and circles correspond to measurements at 4dpi while grey squares and circles are measurements determined at different time points. Error bars represent standard deviations of three different leaves expressing the same GBelement on the same experiment.

2.4. Functional composition guidelines obtained from datasheets.

To exemplify the use of GB3.0 in the construction of transcriptional gene circuits, we built and tested two independent transcriptional modules, and subsequently connected one to another, making the second one to take control on the first one.

The first transcriptional unit (DFRmodule, GB1160) was a simplified version of a plant endogenous module for the control of flavonoids biosynthesis in plants. This system comprised the DFR promoter (GB0606) and the transcriptional regulators Rosea1 (GB0026) and Delila (GB0079). Rosea1 and Delila are known to bind a WD40 factor forming quaternary activation complexes with DNA at the promoter regions of several enzymes of the flavonoid pathway, including dihydrofavonol-4-reductase (DFR) gene. In the absence of additional transcriptional factors, GB0606 showed negligible transcriptional activity in SE_001 conditions (Figure 10a). When constitutively-expressed Del (GB0126) or Ros1 (GB0125) factors were added to the DFR construct, transcription activity raised to low (0.07 rpu) or moderate (2.3 rpu) levels respectively (Figure 10a). Expectedly, the simultaneous inclusion of (GB0126) and (GB0125) resulted in a device with high (7.2 rpu) transcriptional levels (see Figure 10a). As it can be deduced, DFR functions as an imperfect AND gate, with Ros1 and Del acting *in series* in the activation of DFR, but with Ros1 having a stronger influence in DFR expression. Partial AND gate activation illustrates the problem of non-orthogonal systems and probably reflects the ability of Ros1 to recruit endogenous bHLH factors that compensate the absence of Del in the activation of DFR promoter.

The second transcriptional system is a conditional transactivator consisting of a constitutively-expressed glucocorticoid-responsive (GR) chimeric transcription factor (TF) next to an artificial promoter. The chimeric TF comprised a GR element fused to the Gal4 activation domain and the LacI DNA binding domain (GB0157), which binds the lacI operator present in GB1349. In the absence of chemical inducer, the device shows low transcriptional activity when coupled to Luciferase/Renilla reporter (average 1.7 rpu at t=24h). In the presence of 5μM dexamethasone GB1349 is activated to levels up to 16.0 rpu at t=24h (see Figure 10b).

Once both modules were characterized, it was anticipated that, by connecting them, the DFR promoter would turn responsive to dexamethasone. This exercise wanted to simulate the "hacking" of an endogenous regulatory module (flavonoid biosynthesis), making it responsive to a new stimulus. The new circuit admits at least two possible configurations, as either Ros1 or Del TFs can be directly connected to the Dexamethasone module. However, attending to the transcriptional activity reported by Ros1 and Del modules separately, which showed that DFR is more responsive to Ros1, it could be easily anticipated that connecting GB0157 to Ros1 would result in a better inducibility factor (defined as the RTA ratio at t=24h with and without dexamethasone). To test this model, we constructed the new circuit in both configurations (Figure 10) and tested it in presence and absence of inducer. As expected, both circuits resulted

dexamethasone-responsive devices, with the configuration that connected Ros1 to Dexa resulting in a higher inducibility factor (1,45 and 1,97 respectively).

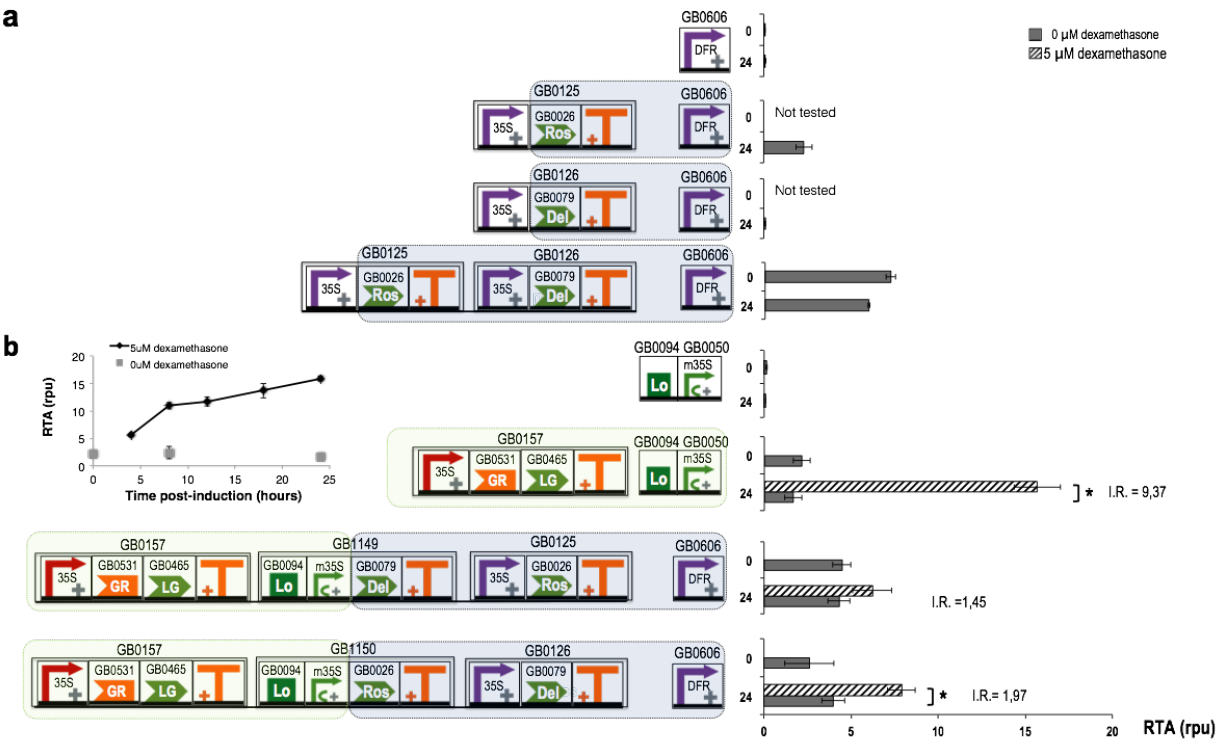


Figure 10. Connection of two transcriptional modules for conferring dexamethasone regulation to the DFR promoter.

(a) Transcriptional activity of the DFR promoter by Ros1 and Del, either alone or in combination at 72h post-infiltration (referred as 0) or 96 hours post-infiltration (referred as 24). (b) Transcriptional activity of the conditional transactivator GB0157 and the effect of its combination with Ros1 and Del turning the DFR promoter responsive to dexamethasone. A detailed time-course of the induction of the conditional transactivator is show on the graph on the left. All transcriptional activities are expressed as relative transcriptional units calculated normalizing relative transcriptional activities to GB0166. Error bars represent standard deviation of Luciferase/Renilla ratios determined on at least two independent leaves. I.R. are the induction ratios calculated by dividing values obtained in presence of dexamethasone by values obtained when dexamethasone was absent. The * indicate differences on the values significant in a T-Test with a p-value < 0.05.

3. Discussion

A new wave of innovative crop traits will be needed in an immediate future to respond with sustainable bio-production solutions to a rapidly changing environment (Ort et al., 2015). Plant Synthetic Biology proposes the rational and systematic design of genetic systems (i.e. regulatory networks, biosynthetic pathways, etc) as a new breeding strategy for obtaining radically new traits, especially those that are plausibly beyond the repertoire offered by natural variation. Undoubtedly, applying rational design to crop breeding is not devoid of phenomenal hurdles. Expectably, the strong context

dependency of many genetic elements will make them inoperative for any technically feasible rationally-designed solution. Even in cases where rational designs could conceivably be applied, it is expectable that the first iteration of a synthetic biological system will not match the breeding objectives, and multiple iterations of the design-build-test cycle will be needed to reach the desired trait. Moreover, defining even the simplest set of functional composition rules will face operational challenges, like the adoption of appropriate standards and the collection and reliable characterization of bioparts (Kelwick et al., 2014). These challenges are, however, worth addressing since the standard biological data is essential to improve the predictive capabilities of forward-design *in silico* models.

GB3.0 proposes a multigene design platform that connects DNA elements with biological data by means of both DNA assembly and experimental standards. This is to our knowledge the first attempt to create such an integrated platform for Plant Synthetic Biology. Previously, other initiatives as the BioBricks Foundation in the frame of the iGEM competition (Smolke, 2009), and the International Open Facility Advancing Biotechnology (BioFab) in a more-research-oriented scope (Cambray et al., 2011), have produced, characterized and validated large collections of standard biological parts including catalogues of promoters (Mutalik et al., 2013b), terminators (Cambray et al., 2013), Ribosomal Binding Sites (Mutalik et al., 2013a), as well as small regulatory devices to support bioengineering mainly in *E. coli* and *Saccharomyces cerevisiae*. To manage and reuse parts and devices, the SBOL (Synthetic Biology Open Language) standard aims to facilitate the exchange of information and to communicate designs in SynBio (Bartley et al., 2015). SBOL is an extensible standard created to encode additional information beyond an annotated sequence as required by synthetic biology, including measurements of performance characteristics, experimental context information, computational models of behavior etc.

The present particularities of Plant Biotechnology impose technical constrains that preclude the practical use of microbial-oriented catalogues as BioFab. In principle there is little chance for inter-kingdom exchange of bioparts due to differences in codon usage and to the lack of functional conservation in many regulatory elements (Angov, 2011). Besides, plant synthetic devices are very often delivered in the form of T-DNAs, and therefore genetic constructs need to be enclosed in Agrobacterium binary vectors. T-DNAs are randomly integrated into the plant genome, leading to positional effects. Also plant biotechnologists rarely make use of elements as RBSs, 5´-UTRs and 3-UTRs in their repertoire, as those elements are usually incorporated to promoter and terminator regions without functional dissection of their specific contributions. All in all, it seems more operative for the plant community to start developing platforms and standards that

are specially adapted for Plants, as reflected by the PSP initiative. Fortunately, there is a growing number of plasmid collections for Plant SynBio in the Addgene repository covering hundreds of standard elements (Engler et al., 2014, Weber et al., 2011a, Mutalik et al., 2013a). However, the development of information managing tools for those collections, such as software tools, protocols for exchange of information, experimental standards or automatization, etc is to a large extend still lacking. GB3.0 is an example in this direction. A first advantage of using a closed assembly system is that gaps or scars between elements are perfectly defined, opening the way to proper and reliable functional description of individual elements. Traditionally, the gaps between key components are almost never reported in Plant Biotechnology, presumably because they are not considered crucial. Yet, synthetic biology relies on the premise that synthetic DNA can be engineered with base-level precision. Lack of proper boundaries definition affects reproducibility, limiting reusability (Peccoud et al., 2011).

In its current form, GB3.0 provides a working solution that accommodates the actual needs of most Plant Biotechnologists, whose main hurdles come from the difficulties to assemble complex and combinatorial designs using available non-standard parts. In this sense GB3.0 is "cloning" oriented, meaning that does not make use of abstraction of function for the design of new devices. Instead, the syntactic category of each element is the actual driver of the new designs, as the usability of each phytobrick is largely conditioned by its flanking overhangs. This is in contrast with the general workflow proposed by SBOL, where physical constrains are not taken into account in the design phase (Galdzicki et al., 2014). It is expectable that progressive advancements in DNA synthesis will reduce physical assembly constrains in Plant Biotechnology as well, to such extend that they can be overlooked during the design phase. However this is not the case yet in the daily experience of most plant biotechnologists and therefore a design tool constrained by DNA-assembly is currently the most operative solution. This does not preclude, however, that Plant Synthetic Biologists tend to accommodate progressively to SBOL or similar standards to enhance compatibility with other design platforms. In this regard, there is a close match between the DNA elements defined in the Plant Standard and the elements proposed by SBOL that will facilitate future adaptations. File conversion tools between GenBank and SBOL have become recently available (Hillson et al., 2012), and it will be advisable to keep contact between both communities to ensure that Plant standards are developed in such way that confluence with general standards in Synthetic Biology is favoured.

A main novelty in GB3.0 is that experimental data is automatically associated to the genetic elements involved in its generation, enriching functional descriptions. We propose the definition of standard experiments as the most straightforward way to

populate SynBio databases with functional descriptions. Initially, we defined five standard experiments in GB3.0, measuring transcriptional activity (SE_001, SE_002), recombinant protein production (SE_003), transformation efficiency (SE_004) or mutagenesis efficiency for CRISPR/Cas9 constructs (SE_005). This is not an exhaustive list but rather an attempt to accommodate to GB3.0 the main activities in our lab. Thus, we proposed Firefly/Renilla luminescence ratios (Fluc/Rluc) measured in agroinfiltrated *N. benthamiana* leaves as a general standard for transcriptional activity specifications. Agroinfiltration is a widely used, rapid and straightforward transient expression method amenable for medium and high throughput analysis (Yang et al., 2000, Kapila et al., 1997). In this method, transgene expression takes place in differentiated leaf epidermal and parenchyma cells; however the high co-transformation efficiency facilitates the incorporation of trans-acting elements that simulate the molecular environment in other specialized tissues, therefore extending its applicability to a wide range of experimental setups. Absolute units for measurement of transcriptional and translational activities in Synthetic Biology have been proposed, as the number of polymerases per second (PoPs) or ribosomes per second (RiPs) (Canton et al., 2008); however in practical terms the use of relative units calculated with a constant reference, has turned to be more operative. We found that normalization of Fluc/Rluc ratios to a Nopaline synthase promoter-derived standard element resulted in a robust and reproducible method for estimation of the relative transcriptional activity in agroinfiltration experiments.

An important concern when setting up SE_001 standard conditions was the number of T-DNAs that simultaneously enter a plant cell in a typical agroinfiltration experiment, as this could affect reproducibility and interpretation in one hand, but also the ability to reliably perform co-transformation with non-linked T-DNAs on the other hand. Therefore we carefully investigated the T-DNA co-transformation levels and its dependence of the concentration of Agrobacterium in the infiltration culture. Although the levels of co-transformation in stable transgenics has been throughout investigated (Dafny-Yelin & Tzfira, 2007, Tzfira et al., 2004, De Buck et al., 2000), to our knowledge this question has not been convincingly addressed for the case of agroinfiltration. We followed a strategy similar to that described for estimation of MOI in viral infections and found that at low OD values the T-DNA copy number showed a strong dependence with Agrobacterium concentration, but this dependence sharply declined when OD was close to 0.1, corresponding to a MOT=7.4. We therefore set 0.1 as standard OD for agroinfiltration, although for certain analysis it would be advisable a lower MOT value. In the current work, we have not found evidence that T-DNA size could influence MOT (no bias has been observed in luciferase activity for larger constructs, data not shown). However the T-DNA constructs assayed so far are of average size range and therefore such dependence cannot be discarded for larger constructs. The characterization of

increasingly complex constructs will probably require a more comprehensive study of the dependence of MOT with T-DNA size.

We also show that the GB3.0 database itself serves as quality control for the reproducibility of a given experimental standard, as GB3.0 keeps record of all the experiments conducted with the same phytobrick. Thus, we found high consistency and reproducibility in all nine experiments conducted with phytobrick GB0030 (35Spromoter) in our lab, despite having been conducted by different personnel over a time period of two and a half years. Notwithstanding this, refinements of SE_002 could be proposed to increase amenability to high throughput analysis as e.g. whole tissue measurements that skip tissue homogenization steps. In its present form GB3.0 is flexible and allows introduction of new types of experiments covering a wide range of needs and including stable transgenics. In the future, an effort should be made in our opinion to find common grounds for experimental comparisons e.g. defining specific genomic landing paths for each species, agreeing standard growth conditions, etc, so that gene constructs can be reliably characterized in a whole plant genome context. For the time being, the standards proposed here could serve as a first step for functional phytobrick characterization. In addition, we have defined also NS_000 to serve as repository to those experiments not fitting any current standard. Non-standard experiments provide specific information of each phytobrick, but cannot be used as basis for comparisons or for creating functional composition rules.

To exercise genetic design in GB3.0, we constructed more than 150 level >1 elements (single TUs, gene modules), involving more than 50 level 0 phytobricks, and tested them in standard and/or non-standard conditions generating more than 80 experimental entries. As an example of use of GB3.0, we show here in more detail the composition of a small circuit that results from the connection of a MYB/bHLH regulatory module to a conditional activation module triggered by Dexamethasone. The Gal4 domain fused to a DNA binding domain has been previously employed as orthogonal transcriptional factor in plants, in enhancer trap strategies and for the spatial control of transgene expression; in addition, the fusion of Gal4 to the GR domain turns it into an efficient chemically inducible switch (see Moore et al. (2006) for a review). On the other hand, Rosea1 and Delila MYB/bHLH TFs are known to bind a WD40 factor forming quaternary activation complexes with DNA at the promoter regions of several enzymes of the anthocyanin biosynthesis pathway, including the dihydrofavonol-4-reductase (DFR) gene. In some tissues, like the tomato fruit, expression of downstream anthocyanin biosynthesis genes requires to a great extend the simultaneous presence of both Rosea 1 and Delila factors (Butelli et al., 2008), scoring close to a canonical AND gate with Ros1 and Del acting in series in the activation of DFR. In contrast, we show that in *Nicotiana benthamiana*,

Rosea1 and Delila function as an imperfect AND gate in the transcriptional activation of DFR promoter, with Ros1 having a stronger influence in DFR expression. Imperfect AND gate activation illustrates the problem of non-orthogonal systems and probably reflects the ability of Ros1 to recruit endogenous bHLH factors that compensate the absence of Del in the activation of DFR promoter. This lack of orthogonality is a very common situation in Plant SynBio, and it is not devoid of practical relevance e.g. for food crops as public opinion trends favour the use of intragenic elements rather than cross-kingdom elements for crop biotechnology. The activation of entire pathways by ectopic expression/repression of endogenous or homologous TFs has been shown a very successful strategy for metabolic engineering, including biofortification of food crops (Davuluri et al., 2005, Zhang et al., 2015, Butelli et al., 2008). Ectopic expression often results from overexpression of the TFs using constitutive promoters or, in a few examples, is made dependent of endogenously regulated factors as e.g. ripening (Davuluri et al., 2005), senescence (Ori et al., 1999), or tissue specific factors (Azuma et al., 2016, Paine et al., 2005, Houmard et al., 2007). A step forward in transcriptional control would involve connecting endogenous pathways (e.g. anthocyanin biosynthesis) with externally operable modules such as optogenetic or chemically inducible switches (Muller et al., 2014, Padidam, 2003, Moore et al., 2006, Kinkema et al., 2014). This would allow to externally operate the biochemical and/or physiological status of the plant, for instance anticipating biotic or abiotic threats or triggering the accumulation of target compounds immediately before harvesting. The constructs built here prototype this type of circuits by physically combining a chemical switch and a Ros/Del (endogenous) module in a single cloning step, and show that this results in a functional combination of both modules. Moreover we showed that, by attending to the standard descriptions of the different elements analysed separately, it was possible to anticipate the behavior of the combined system and to select the configuration that provides higher Dexamethasone induction rates to the DFR promoter, a first step towards implementing functional composition rules in plants.

4. Materials and methods

4.1. *Nicotiana benthamiana* transient expression

For transient expression, plasmids were transferred to *Agrobacterium tumefaciens* strain GV3101 by electroporation. *N. benthamiana* plants were grown for 5 to 6 weeks before agroinfiltration in a growing chamber with 24°C (light)/20°C (darkness) in a 16-h-light/8-h-dark photoperiod. Agroinfiltration was carried out as previously described by Orzaez et al. (Orzaez et al., 2009). Briefly, overnight Agrobacterium cultures were

pelleted and resuspended in agroinfiltration solution (10mM MES, pH 5.6, 10mM $MgCl_2$, and 200µM acetosyringone) to an optical density of 0.1 at 600nm. Bacterial suspensions were incubated for 2 hours at room temperature on a horizontal rolling mixer and they were mixed for experiments in which more than one GBelement was used. Agroinfiltrations were carried out through the abaxial surface of the three youngest leaves of each plant with a 1ml needle-free syringe.

4.2. Luciferase/Renilla assays

Samples of leaves coinfiltrated with GBdevices listed at Supplementary Table 3 for SE_001 and SE_002 were collected at 3dpi (SE_001) and 4dpi (SE_002). For the determination of the Luciferase/Renilla activity one disc per leaf (d=0.8cm, approximately 18-19mg) was excised. For SE_001 leaf discs were kept in plates with or without inductor and samples were frozen in liquid nitrogen at the standard timepoints. Inductions were performed with D1756-dexamethasone (Sigma Aldrich) diluted to the final concentrations listed on each experiment in 0.02% Tween-80. For SE_002 excised leaf discs were directly freeze in liquid nitrogen after excision.

Leaf discs were homogenized and extracted with 150µl of 'Passive Lysis Buffer', followed by 15 minutes of centrifugation (14000 x g) at 4ºC. Then, the supernatant was diluted 2:3 in Passive Lysis Buffer resulting in the working plant extract. Fluc and Rluc activities were determined following the Dual-Glo® Luciferase Assay System (Promega) manufacturer's protocol with minor modifications: 10µl of working plant extract, 40µl of LARII and 40µl of Stop&Glo Reagent were used. Measurements were made using a GloMax 96 Microplate Luminometer (Promega) with a 2-second delay and a 10-second measurement. Fluc/Rluc ratios were determined as the mean value of three samples coming from three independent agroinfiltrated leaves of the same plant and were normalized to the Fluc/Rluc ratio obtained for GB0166.

4.3. Protoplasts isolation and MOT calculation

N. benthamiana protoplasts were isolated from 5 days post-infiltrated leaves co-infiltrated with two Agrobacterium strains carrying GB1287 and GB1288 at 1:1 proportion at seven different OD600 (0.1, 0.05 and five 1:3 dilutions ranging from 0.05 to 0.00021). Protoplast isolation was performed as previously described by Sang-Dong Yoo et al. (Yoo et al., 2007) with minor modifications. Vacuum infiltration of cut leaves in enzyme solution was performed for 10 instead of 30 minutes. After filtration, intact protoplasts were further purified from dead protoplast and remaining cellular debris by the sucrose flotation method on 20% (w/v) sucrose. After the washing steps protoplasts

were kept in WI solution (4 mM MES (pH 5.7) containing 0.5 M mannitol and 20 mM KCl).

Expression of Yellow and/or the DsRed Fluorescent Proteins on the isolated protoplasts was detected with the photomultiplier of the LSM 780 (Zeiss) confocal microscope. Image processing was performed with ImageJ (Schneider et al., 2012) followed by manual counting of the no fluorescent (untransformed), yellow fluorescent, red fluorescent and yellow and red fluorescent (cotransformed) protoplasts.

By assuming that T-DNA distribution over cells of infiltrated leaves occur randomly and independently for T-DNAs of both fluorescent proteins, it can be considered that the number of T-DNAs of each type entering a cell follows a Poisson distribution. With this consideration, we calculated the multiplicity of transformation (MOT) based on the frequency of cotransformed protoplasts for each tested OD_{600} (see Supplementary Table 4) following the same approach that Gutiérrrez et al. used for MOI calculation (Gutierrez et al., 2010). At least 1000 protoplasts per OD_{600} were used for calculation.

4.4. Software tools development

The web site www.gbcloning.upv.es was implemented using Django, a Python web framework that supports rapid design and the development of Web-based applications (version 1.5; Django Software Foundation; http://djangoproject.com). The database management system PostgreSQL was chosen to host the schemas of the GBelements and the experiments databases and all software tools accessible on the web site were developed using Python. The software-package contains flexible modular blocks which are interconnected and can be classified in five main categories: (1) adaptation of raw DNA sequences to the GB standard, (2) creation of gene-cassettes from standard parts, (3) binary assembly of pre-made gene-cassettes, (4) generation of GBelement datasheets and (5) generation of experiment views. All tools run all functions behind the screen. For the cloning tools, the submitted data is directly passed by Python functions for sequence checking and for output generating by creating a GenBank file and a protocol to each assembly step which is sent back to web server for user download. For generation of experiment views, data submitted is passed by Python functions for plotting a graph with the quantitative values and for incorporating links to the experiments on datasheets of involved GBelements. The source code of all tools is available on the Github repository at https://github.com/pziarsolo/goldenbraid.

4.5. Website functionality

The gbcloning web site is organized in four different modules (Design, Collection, Experiments and Genome Engineering) as described on the Results section. The user-action workflow between the different cloning tools for the design of either multigene constructs or constructs for genome engineering is explained in detailed on Chapter 1 and Vazquez-Vilar et al. (2015) and on Chapter 3 and Vazquez-Vilar et al. (2016), respectively. Access and search to the Collection of GBelements is described at Chapter 1 and Vazquez-Vilar et al. (2015) and also at Sarrion-Perdigones et al. (2013). On the 'Experiments' section, standard and non-standard experimental information can be incorporated to the user-database (an user account is required) and associated to either user or public GBelements at https://gbcloning.upv.es/add/experiment/ following the guidelines specified on the same address. There are two options for users to access the experiments performed with a GBelement. First, on each GBelement datasheet a list of all public and user-associated experiments with links to them can be found. Second, experiments can be searched at https://gbcloning.upv.es/search/experiment/ following different criteria (experiment ID, keywords or words included on the experiment description, and GBelements used on the experiment). Additionally, for standard experiments the search can be filtered by the values of the quantitative outputs.

5 | Chapter 3

A modular toolbox for gRNA-Cas9 genome engineering in plants based on the GoldenBraid standard

Vazquez-Vilar M, Bernabé-Orts JM, Fernandez-Del-Carmen A, Ziarsolo P, Blanca J, Granell A, Orzaez D.

Plant Methods. 2016 Feb 1;12:10. doi: 10.1186/s13007-016-0101-2. eCollection 2016.

1. Introduction

Since its discovery, the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas immune bacterial system has rapidly become a powerful technology for genome editing in many organisms. This system is based on a guide RNA (gRNA) that directs the *Streptococcus pyogenes* Cas9 nuclease to its target site. The application of the RNA-guided Cas9 technology is being widely exploited by the scientific community in cell cultures (Ran et al., 2013), animals (Yang, 2015, Wang et al., 2013a) or plants (Bortesi & Fischer, 2015, Belhaj et al., 2015).

On the plant field, RNA-guided genome engineering via Cas9 has been employed in diverse approaches, from single and/or multiple gene knock-outs (Shan et al., 2013, Gao et al., 2015, Fauser et al., 2014) to targeted insertions of donor sequences (Schiml et al., 2014) or even targeted transcriptional regulation through the fusion of transcriptional activation or repressor domains to an inactivated Cas9 (Piatek et al., 2015). A remarkable feature of gRNA-Cas9 is that facilitates targeting multiple sequences simultaneously. While similar technologies such as the ZFNs (zinc finger nucleases) (Beerli & Barbas, 2002) or the TAL effectors (Bogdanove & Voytas, 2011) require recoding of a new protein for each target sequence, with the gRNA-Cas9 a change of 20 nts in the guide RNA is enough, paving the way for multiplex editing and design of complex regulatory circuits among other engineering possibilities (Nielsen & Voigt, 2014).

The direct transfection of Cas9 and guide RNAs into plant protoplasts followed by plant regeneration from single-cell has been shown effective for genome editing in rice and tobacco, however the efficiency remained relatively low, and besides, whole plant regeneration from protoplasts is not currently feasible for many crop species (Eeckhaut et al., 2013). A successful alternative for plants is the use of Agrobacterium mediated T-DNA transformation, followed by callus induction and organogenic plant regeneration (or floral dip transformation in the case of Arabidopsis). In this case, T-DNA-delivered gRNA-Cas9, besides acting transiently during callus formation, can also integrate in the genome and continue its activity in somatic tissues (Bortesi & Fischer, 2015). To exploit the full potential of the T-DNA strategy it is important to expand the ability to combine different gRNAs together with Cas9 within a single T-DNA, as it has been demonstrated that all-in-one plasmid approaches significantly increase editing efficiency (Mikami et al., 2015).

Modular cloning methods are being increasingly adopted by the plant research community as they greatly facilitate the combinatorial assembly of pre-made DNA

elements into multigene constructs (Patron et al., 2015, Liu & Stewart, 2015). GoldenBraid is a modular cloning standard that makes use of the Type IIS restriction enzyme BsaI for the assembly of basic, so-called "level 0" DNA elements (promoters, coding regions, terminators, etc) into transcriptional units (TUs), and then incorporates a second enzyme, BsmBI, to build higher level structures using a double-loop iterative strategy (Sarrion-Perdigones et al., 2013). Level 0 parts are flanked by 4 nucleotides overhangs, the sequence of which determines the relative position of each part in the transcriptional unit. To be usable in GB cloning, all level 0 parts need to be previously adapted with the incorporation of flanking BsaI recognition sites, the addition of flanking 4bp standard barcodes, and the removal of internal BsmBI and BsaI sites. The whole process of adaptation to the standard is often referred to as "domestication". Once domesticated, GB parts can be efficiently combined to create large multigenic constructs within binary destination plasmids ready to be used in Agrobacterium-mediated plant transformation. A key feature of GB is that all constructs can be reused in new combinations following the same cloning scheme, fostering the exchange of genetic elements. Interestingly, GB part reusability enables the unequivocal association of physical parts with experimental information, as no further modifications (i.e. subcloning, re-assembly or PCR re-amplification) are required to incorporate a GB part into different genetic modules. The GB webpage (https://gbcloning.upv.es/) offers a set of online tools for *in silico* multigenic assemblies and a database for the collection and exchange of GB standard parts (Vazquez-Vilar et al., 2015). Although Type IIS cloning methods have been employed for multi-gene assemblies with a wide range of applications in several organisms (Duportet et al., 2014, Guo et al., 2015), the GB framework is specially designed for plants since the GB destination plasmids are two sets of binary vectors (one based on pGreen and a second one based on pCambia) and all the GB standard parts including promoters and terminators are suitable for plant biotechnology.

The GB cloning strategy is especially suited for the construction of vectors incorporating Cas9 together with multiple guide RNAs in the same T-DNA. Here, we report the implementation of a GB-adapted gRNA-Cas9 toolbox for plants, which includes the domestication of gRNA/Cas9 elements, the definition of a CRISPR cloning workflow and incorporation of new online tools for building CRISPR-based genome engineering constructs in binary vectors.

2. Results

2.1. GB-adapted cloning strategy for CRISPR/Cas9 plant constructs

To facilitate the assembly of CRISPR/Cas9 constructs and the delivery of multiple guide RNAs in the same T-DNA, we designed the CRISPR cloning workflow depicted on Figure 11a. As a first step, twenty nucleotides sequences designed against a specific genomic target can be incorporated to the GoldenBraid scheme using the 'GB CRISPR domesticator' tool available at https://gbcloning.upv.es/do/crispr/. This tool generates a new target-specific GB element (D-Target /M-Target, syntax structure B3c-B4-B5c or B3c-B4-B5d), which can be used immediately or stored in the database for future assemblies. The D/M-Target comprises two partially complementary oligonucleotides yielding a double-stranded DNA fragment flanked by four nucleotides overhangs. In a next step, the D/M-Target is combined with a PolIII promoter (currently, Arabidopsis U6-26 and U6-1 and rice U3 promoters are available in the GB collection) and with the scaffold RNA in a cyclic digestion/ligation Golden Gate reaction (Engler et al., 2009) to build the complete gRNA expression cassette. This step is assisted by the 'CRISPR Assembler' tool available at https://gbcloning.upv.es/tools/crisprsassembler.

The conditions for gRNA assembly were optimized by checking three key parameters, namely primer concentration, primer dilution buffer and annealing conditions in a total of 12 combinations. The resulting assemblies were then transformed into *E.coli* and the efficiency assessed by the number of colonies obtained (Figure 11b and c). Two colonies of each of the 12 assembly reactions were selected for restriction analysis resulting in a 100% of positive clones (see Supplementary Figure 2a). Primer dilution was found the main factor affecting reaction efficiency, with best results obtained at low primer concentrations. Only minor effects were observed associated to buffer or denaturing condition (Figure 11b and c). Accordingly, recommended conditions for CRISPR assembly in multipartite GB reactions were set at 1µM primer concentration in water with a 30 minutes annealing step performed at room temperature.

Following the GB workflow, every gRNA expression cassette assembled in GB compatible vectors can be combined with each other and/or with a Cas9-encoding transcriptional unit (Figure 11a) with the 'GB Binary Assembler' web tool (https://gbcloning.upv.es/do/bipartite/). GB binary reactions were highly efficient as previously described Sarrion-Perdigones et al. (2011) and accurate since white colonies analyzed resulted in 100% correct assemblies in most cases (see Supplementary Figure

2b and Supplementary Table 7). The current GB-adapted gRNA-Cas9 toolbox incorporates seven different Cas9-encoding TUs which have been designed for gene editing, gene activation and gene repression projects. All Cas9 TUs described in this Chapter were created by combining only protein-coding GBparts, leaving constitutive plant expression elements invariant. The assembly of inducible and/or tissue-specific expression of Cas9 is also possible using other standard parts from the collection.

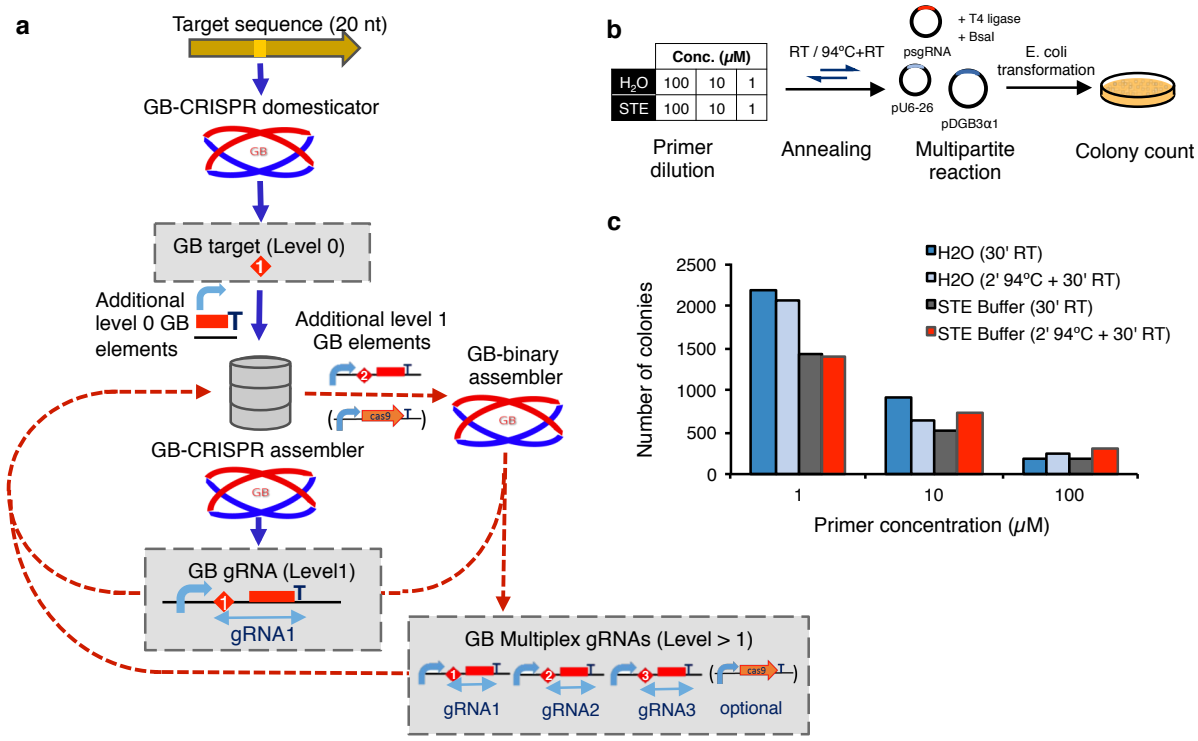


Figure 11. Multiple guide RNAs assembly with GoldenBraid.

(a) Software-assisted CRISPR cloning workflow. Targets are adapted to the GoldenBraid standard with the 'GB-CRISPR domesticator'. Then, these level 0 parts (D/M-Targets) are combined with other standard GBparts with the 'GB-CRISPR assembler' to create the guide RNA expression cassettes, which can be combined between them and/or with a Cas9 transcriptional unit with the 'GB-binary assembler'. (b) Optimization of GB-CRISPR multipartite reactions. Forward and reverse primers were diluted to different concentrations with different solvents; they were mixed and twelve independent multipartite reactions were set up. After transformation into *E. coli*, the number of colonies was estimated. (c) Number of colonies obtained on the twelve independent guide RNA multipartite assembly reactions.

2.2. Transient expression of GB-adapted Cas9 TUs provides efficient targeted mutagenesis in N. benthamiana leaves

To experimentally validate the different GB modules for gRNA-Cas9-mediated gene mutation, we tested them in *N.benthamiana* by targeting the endogenous xylosyltransferase (XT) gene. A BLAST search on the *N.benthamiana* genome with the GenBank accession ABU48858, resulted in scaffolds Niben101Scf04205Ctg025 and

Niben101Scf04551Ctg021 corresponding to predicted cDNAs Niben101Scf04205g03008 (XT1) and Niben101Scf04551g02001 (XT2) respectively. We decided to target the two of them using a specific guide RNA for each one. The 20-bp target sequences for each guide RNAs were designed with the CRIPSR-P online tool (Lei et al., 2014), imposing the requirement for a G at the 5' end of the sequence and minimizing off-targeting. An extra criterion for selection was the presence of a restriction site overlapping the Cas9 cleavage site to facilitate the detection of the mutations. The selected targets are depicted on Figure 12a.

GB-based gene targeting constructs carrying human-optimized (h) (Mali et al., 2013) and plant-optimized (pco) (Li et al., 2013) Cas9 variants directed to the single target of XT2 were transferred to Agrobacterium and infiltrated into *N. benthamiana* leaves. To test the mutation efficiency, genomic DNA was extracted from leaves, the targeted region amplified by PCR and the presence of mutated fragments estimated based on the elimination of the internal SpeI restriction enzyme (RE) site. The mutation efficiency for the hCas9 was estimated as 11% based on the intensity of the undigested band (Figure 12b Lanes 2 and 3) relative to the undigested DNA present on the negative control (Figure 12b Lane 1). For pcoCas9 mutation efficiency was below detection levels as it was not possible to visualize the undigested band on the agarose gel.

To test the efficiency of simultaneous targeting, we next assembled both gRNAs targeting XT1 and XT2 together with the hCas9 TU in a single T-DNA and transiently expressed them in *N. benthamiana* leaves. hCas9-induced mutations were detected as above with the restriction enzyme site loss method using BsmBI for XT1 and SpeI for XT2 (Figure 12c). The gRNA-guided Cas9 activity resulted in part of the DNA being resistant to RE digestion (see undigested band in Lanes 2 and 4) that was not detected when only hCas9 was expressed (Lanes 1 and 3). To corroborate the presence of mutations on the undigested PCR products, the undigested amplicons were cloned and individual clones were sequenced. The most prevalent mutations observed for XT1 were deletions of less than 10 nucleotides, while for XT2 a 32% of the mutated clones had single nucleotide insertions (C or T) (Figure 12d). Mutation rates of 17% (XT1) and of 14.5% (XT2) were observed for the new construct. Since 29% (XT1) and 32% (XT2) of the clones showed the wild type sequence, we included this correction factor to obtain a more accurate estimation of the mutation rate. As result, we obtained a mutation rate of 12.1% for XT1 and a mutation rate of 9.9% for XT2, consistent with the 11% obtained for the same target when a single gRNA was used.

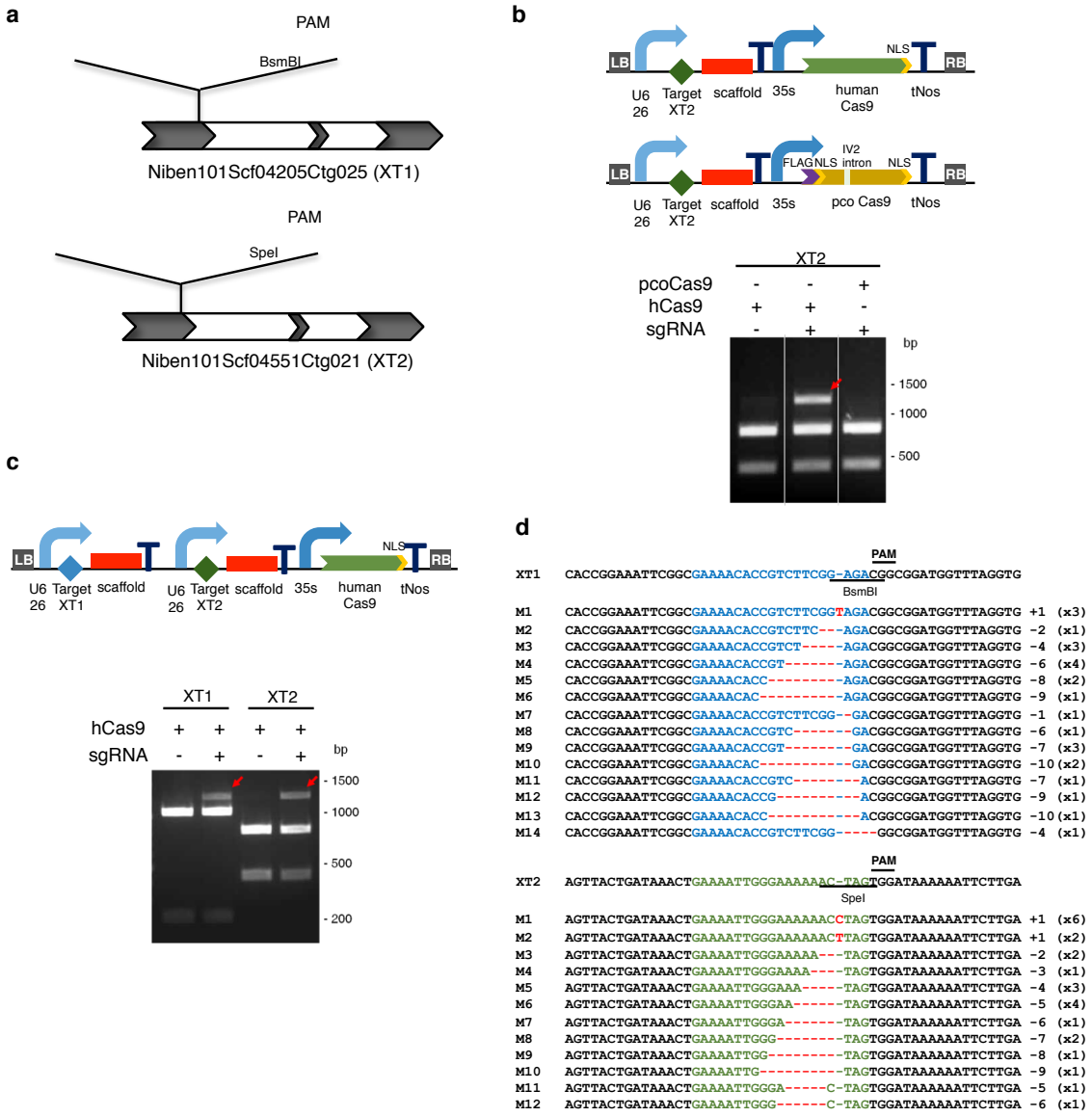


Figure 12. Targeted mutagenesis using the CRISPR/Cas9 system in transient expression in *N.benthamiana*.

(a) Schematic representation of the structure of Niben101Scf04205Ctg025 (XT1) and Niben101Scf04551Ctg021 (XT2) (exons in grey, introns in white) with the sequences of the target sites. Diagnostic restriction sites are underlined and the PAM sequence is shown in bold. (b) Comparison of the mutation efficiency of hCas9 and pcoCas9 targeting the XT2. Red arrow shows SpeI resistant PCR fragments only visible on the gRNA and hCas9 combination. (c) PCR/RE assay to detect simultaneous targeted mutations on XT1 and XT2. Red arrows show BsmBI and SpeI resistant PCR fragments amplified from *N.benthamiana* genomic DNA. (d) Alignment of XT1 and XT2 sequences obtained from different clones of uncleaved bands (see Figure 12c). XT1 target site appears in blue and XT2 target site in green. Red letters and dashes indicate insertions and deletions respectively.

2.3. GB-adapted dCas9 variants modulate transcriptional activity in *N. benthamiana* transient assays

The modularity of GoldenBraid assembly facilitates the design of Cas9 variants with novel functions as e.g. transcriptional activators, repressors, chromatin remodeling factors, etc, by incorporating additional coding modules as translational fusions to an inactive (dead) version of Cas9 (dCas9). To validate this option we built and tested a number of GB-based transcriptional regulators which were targeted to a Nopaline synthase promoter (pNOS) fused to a luciferase reporter.

Making use of level 0 standard genetic parts, we assembled five different transcriptional units (TUs) expressing either the dCas9 (D10A H840A) alone or C-terminus chimeric versions of it fused either to an activator (VP64 or EDLL) or a repressor (SRDX and BRD) (Supplementary Figure 3). These five chimeric transcriptional regulators were tested in combination with five gRNAs directed against different regions of pNOS on both sense and antisense strands (Figure 13a). Changes in the transcriptional activity in these constructs were estimated with the luciferase/renilla system using a reporter construct (REP) that included the firefly luciferase (Fluc) driven by the pNOS and the renilla luciferase (Rluc) driven by the 35S promoter as an internal reference. Transient co-transformations of REP with Cas9 and gRNA constructs were performed in order to test the ability of GB-built dCas9 chimeras to modulate transcription.

Since previous studies reported that dCas9 itself could act as a transcriptional repressor (Bikard et al., 2013), we first tested the repressor activity of the non-chimeric dCas9 TU. All five gRNAs targeting pNOS induced variable repression rates depending on their position (Figure 13b). The Fluc/Rluc ratios decreased as the position of the gRNA gets closer to the Transcription Start Site (TSS) whereas no repression was detected neither for gRNA4 (positions -161 to -142) nor for gRNA5 (positions -211 to -192). Co-expression of the two most effective gRNAs, gRNA 1 and 2, showed a nearly additive effect. However, the addition of a further gRNA, such as gRNA4, to one or both of them did not change the repression level.

Next, the dCas9 fusions to the BRD and the SRDX repressor domains were tested in combination with gRNAs 3, 4 and 5, all three designed to bind upstream the TATA-box. Figure 13c shows that only gRNA4, the gRNA designed on the sense strand, was capable of producing a significant repression on the transcriptional activity. A slight decrease in the Fluc/Rluc ratio was observed when gRNA4 was combined with the two additional

gRNAs. The repression levels found with the dCas9:BRD and dCas9:SRDX were similar (Figure 13c).

To determine whether the presence of the repressor domain modified the effect of the dCas9 itself, we compared the transcriptional activity obtained for the gRNAs 1, 2 and 4 in presence of the dCas9 with the ones obtained with the dCas9:BRD (Figure 13d). While in the case of the gRNA4 only dCas9:BRD had an effect on the reduction of the transcriptional activity, for the gRNAs overlapping the TATA-box and the TSS, both dCas9 and dCas9:BRD achieved similar repression levels.

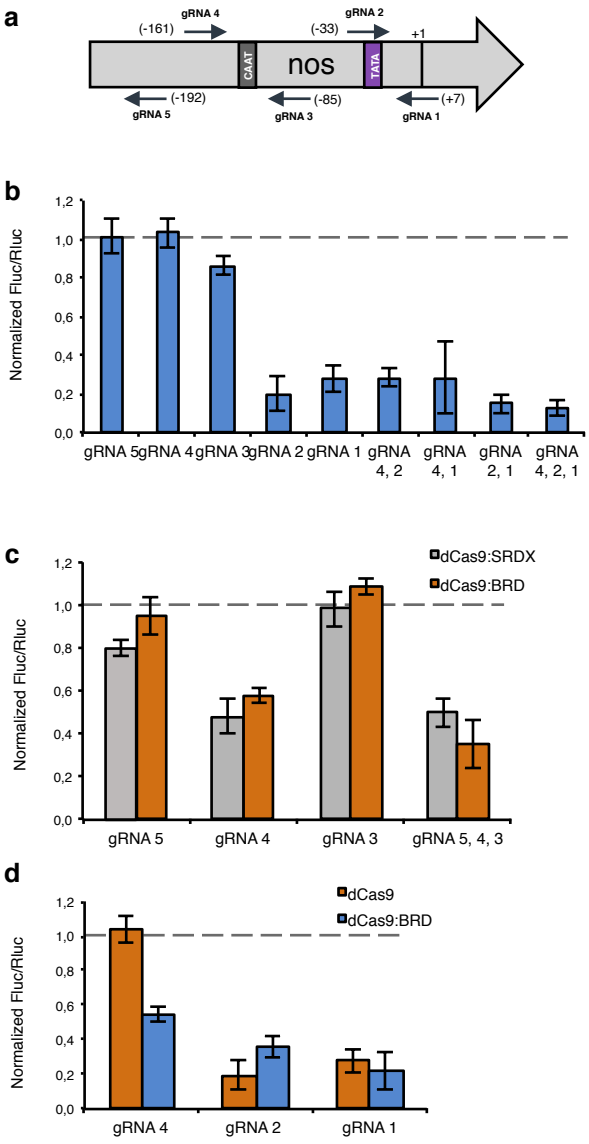


Figure 13. Transcriptional repression of the nopaline synthase promoter (pNOS) with different variants of the dead Cas9.

(a) Schematic representation of the gRNA target positions on the pNOS. The gRNAs were selected in both sense and antisense strands. In parenthesis the 5' position of each gRNA according to the pNOS transcription start site. (b) Comparison of the repression rates mediated by the different gRNAs combinations targeting the pNOS in combination with the dCas9. (c) Repression rates of the dCas9:BRD and dCas9:SRDX in combination

with gRNAs targeting different positions upstream the pNOS TATA-box. (d) Influence of the presence of the BRD domain fused to the dCas9 on the repression levels induced by gRNAs 1, 2 and 4. All values were normalized to the Fluc/Rluc ratios of a reference sample set as 1. Bars represent average values of three samples ± standard deviations.

Next, we decided to test whether the dCas9 fused to an activator domain could increase the transcriptional activity on the same reporter construct. The results showed that dCas9:VP64 and dCas9:EDLL raised the reporter levels in combination with gRNA4, while in combination with gRNA5 only a small induction rate was detected and no induction was observed with gRNA3, corroborating the functionality observed for the same gRNAs with dCas9:SRDX and dCas9:BRD (Figure 14a). Using both the dCas9:VP64 and the dCas9:EDLL variants in combination with 3x multiplexed gRNAs (gRNA 3, 4 and 5), the pNOS transcriptional activity was doubled.

These results demonstrated that it is possible to modulate the transcriptional activity driven by the pNOS using one or more gRNAs in combination with different chimeric versions of the dCas9. The maximum induction rate, calculated with the values of the best reported repression and activation Fluc/Rluc ratios, was 6.5x (Figure 14b).

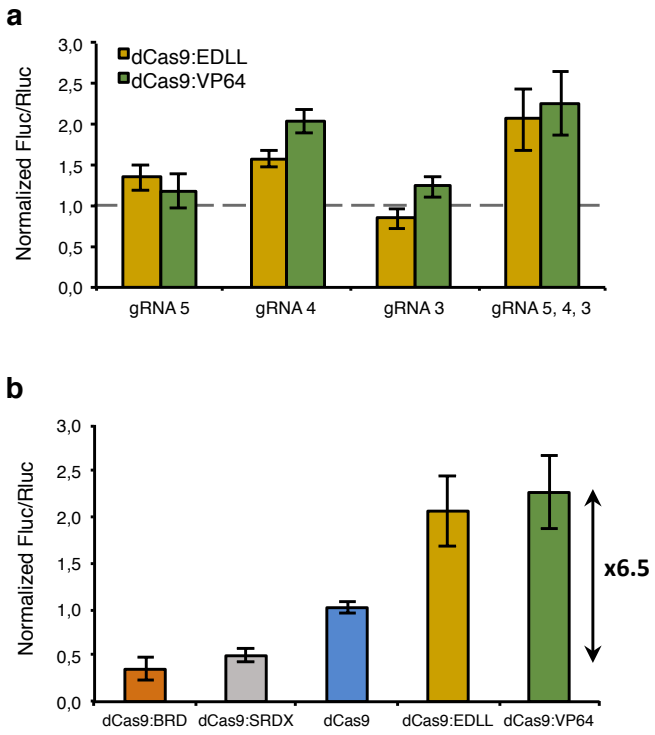


Figure 14. Transcriptional activation and modulation of the nopaline synthase promoter (pNOS).

(a) Fluc/Rluc ratios obtained with dCas9:VP64 and dCas9:EDLL in combination with gRNAs 3, 4 and 5. (b) Comparison of the Fluc/Rluc ratios obtained for gRNAs 3, 4 and 5 in combination with the different dCas9 variants reported on this work. All values were normalized to the Fluc/Rluc ratios of the reference sample set as 1. Bars represent average values of three samples ± standard deviations.

2.4. Second-dimension multiplexing using GoldenBraid.

To further increase the gRNA multiplexing capacity we decided to incorporate a polycistronic strategy to the GB pipeline. This strategy, which has been validated in rice (Xie et al., 2015), allows the simultaneous expression in a single transcript of multiple gRNAs, which are later processed by the endogenous tRNA ribonucleases P and Z to produce the individual gRNAs. To adapt the general GB cloning system to the polycistronic strategy we incorporated single tRNA-gRNA oligomers as level 0 GBparts, which are then multipartitely assembled on level 1 to create polycistronic tRNA-gRNAs (Figure 15a). To avoid using PCR reactions during the construction of each tRNA-gRNA oligomer, we designed new level -1 plasmids containing both the tRNA and the gRNA flanked by BsmBI restriction sites. The BsmBI assembly of level -1 plasmids with the D-target primers heteroduplex results in level 0 GB-oligomers. In turn, these level 0 elements are combined together with the level 0 PolIII promoter to create a level 1 polycistronic tRNA-gRNA in a software-assisted step available at https://gbcloning.upv.es/do/multipartite/free/. We validated the assembly efficiency of the 2-D multiplexing schema by assembling a level 2 construct targeting simultaneously *N. benthamiana* fucosyl and xylosyltransferase genes. As the two gRNAs targeting XTs have been previously tested in this work, we used the same targets (Supplementary Table 6) for the assembly of a polycistronic tRNA-gRNA combining two GBoligomers. Since the number of genes encoding fucosyltransferases in the *N. benthamiana* genome is very high, we decided in this example to target only five of them using a combination of three gRNAs (Supplementary Table 6), one of them targeting three genes and the remaining two gRNAs targeting a single gene. After assembling firstly all five level 0 oligomers and subsequently the two level 1 polycistronic structures, they were combined together in a GB binary reaction (Figure 15a) to generate a single binary plasmid containing all five gRNAs targeting a total of seven genes encoding fucosyl and xylosyltransferases. All the assembly steps resulted in 100% accuracy rates (at least 4 white colonies analysed in each step) demonstrating the efficiency of the proposed scheme for 2D multiplexing (Figure 15b). The whole process took just nine working days, and in three extra days the Cas9 was added to the assembly. To validate the functionality of the polycistronic strategy in N. benthamiana, genomic DNA was extracted from leaves agroinfiltrated with the plasmid containing the five gRNAs together with the Cas9 and the targets XT1 and XT2 were evaluated with the PCR/RE site loss method. The mutation efficiency for XT1 was estimated as 6% based on the intensity of the undigested band (Figure 15c Lane 3) relative to the undigested DNA present on the

negative control (Figure 15c Lane 1). For XT2, observed mutation efficiency was estimated as 8% (Figure 15c Lane 6) relative to the undigested DNA present on the negative control (Figure 15c Lane 4). On the same experiment, the monocistronic gRNAs targeting XT1 and XT2 previously tested were included as positive control (Figure 15c Lane 2 for XT1 and Lane5 for XT2), resulting in a calculated mutation efficiency of 5,5% for XT1 and 7% for XT2.

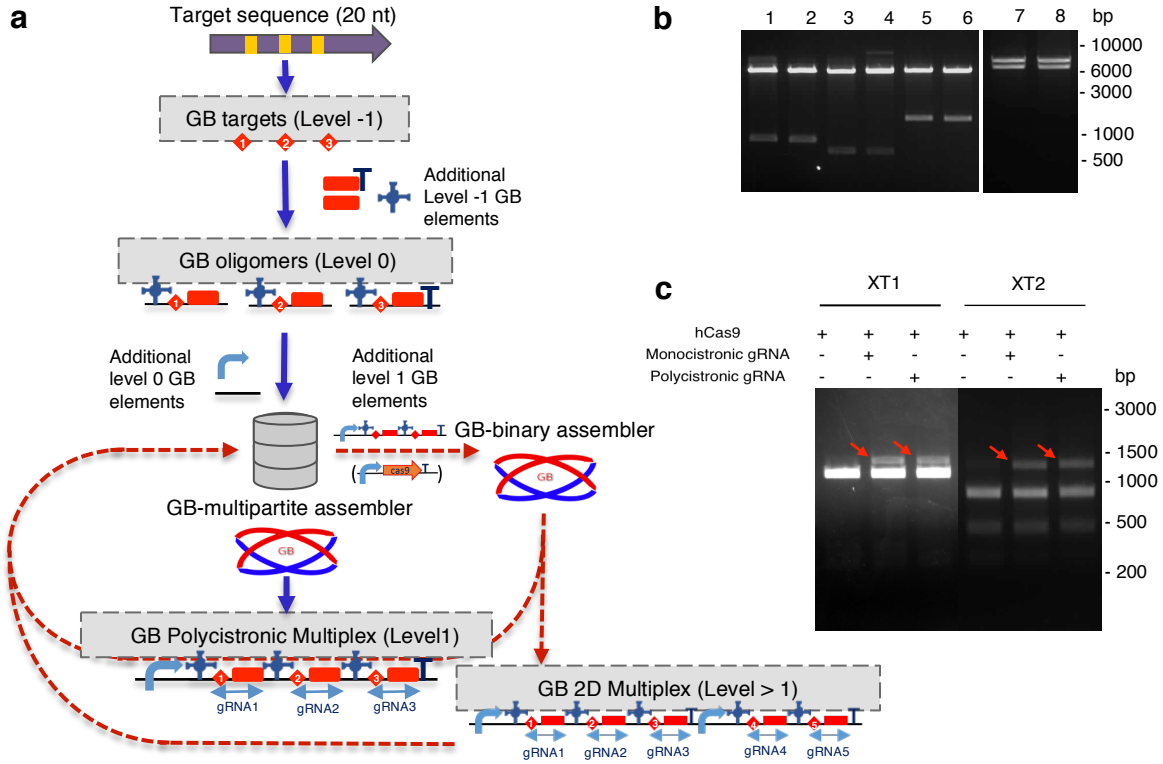


Figure 15. Second dimension Multiplexing with GoldenBraid.

(a) Pipeline of the 2D multiplexing strategy. Targets are designed as level 0 structures and combined with standard level -1 parts to create individual oligomers that are combined in level 1 polycistronic tRNA-gRNA structures. The binary combination of two polycistrons incorporates a 2D multiplexing step on the CRISPR cloning workflow. (b) Restriction analysis of two clones of level 1 polycistronic structures targeting fucosyl (Lanes 1 and 2; EcoRI expected bands: 6345-796) and xylosyltransferases (Lanes 3 and 4; HindIII expected bands: 6345-623), two clones of a level 2 construct derived from their binary assembly (Lanes 5 and 6; BamHI expected bands: 6674-1401) and two clones of its assembly with the hCas9 (Lanes 7 and 8; BsmBI expected bands: 7215-6367). (c) Comparison of the simultaneous targeted mutations on XT1 and XT2 with monocistronic gRNAs and a polycistronic gRNA. Red arrows show BsmBI and SpeI resistant PCR fragments amplified from *N.benthamiana* genomic DNA.

3. Discussion

The adoption of standard rules and modular design has promoted the expansion of many engineering disciplines from mechanics to electronics and is likely to have an

impact in genome engineering as well. Modular cloning methods based on TypeIIS restriction enzymes such as Golden Gate (Engler et al., 2009), MoClo (Weber et al., 2011a) and GoldenBraid (Sarrion-Perdigones et al., 2011), greatly facilitate the construction of large multigene assemblies enabling the concurrent delivery of multiple pieces of genetic information into the cell. Moreover, Type IIS cloning systems are especially well suited for the definition of standard assembly rules. Very recently, a common DNA assembly syntax for TypeIIS cloning has been agreed by 26 different Plant laboratories and research groups worldwide, constituting one of the first examples of a Bioengineering Standard adopted by the Scientific Community (Patron et al., 2015). We have introduced the necessary modifications in GB to make the gRNA-Cas9 toolbox fully compliant with the new standard.

The first step towards GB adaptation for gene targeted mutation consisted in the design of a GB-compatible assembly scheme that facilitates both gRNA multiplexing and Cas9 modification. We decided to build both gRNAs and Cas9 transcriptional units as level 1 structures to maximize their exchangeability while preserving the combinatorial potential. In the GB system, level 1 constructs grow only binarely, which poses a certain limitation in terms of cloning speed. Other systems growing multipartitely using Golden Gate assembly have been proposed for mammalian and plant systems, however this is at the cost of flexibility and reusability of the constructs (Sakuma et al., 2014, Ma et al., 2015, Lowder et al., 2015). Conversely, level 1 GB constructs are exchangeable, offering the possibility to reuse efficient gRNA constructs in new editing or regulatory combinations. Furthermore, this initial decision proved to be most adequate with the incorporation of polycistronic tRNA-gRNA constructs at level 1, which provides a new combinatorial dimension for multiplexing, and makes possible to hierarchically combine gRNAs using different assembly levels. Hence, in our 2D editing example we grouped homologous functions (either xylosyl or fucosyltransferases) in level 1, and later combined them in level 2 in a binary assembly step. Similarly, hierarchical assembly approaches can be used to build increasingly complex gRNA-Cas9-based transcriptional regulatory circuits in few days.

The assembly and functional validation of several gRNA-Cas9 constructs provides evidence of the efficiency of the process and the functionality of the elements that were incorporated to the GB toolkit. GB is based on Golden Gate typeIIS cloning which is an extremely efficient multipartite assembly method when parts are conveniently cloned within an entry plasmid. Whether the same high efficiency is maintained when one of the parts is made of two partially overlapping 23-25 mer oligonucleotides encoding the target sequence remained to be tested. Counterintuitively, the efficiency of the reaction was shown to be significantly higher when low concentrations of oligonucleotides (nM

range) were employed in the reaction mix. Also, it is worth to notice that in the proposed GB gRNA building scheme, the only variable input specific for each new construct are the two 25 mer oligonucleotides; all the remaining building elements are invariant and stored in the GB collection, a feature that significantly reduces gene synthesis costs for building gRNA-Cas9 constructs for plants.

The first functional characterization of the new GB targeted mutagenesis tools was the quantification of Cas9 nuclease activity in a *N. benthamiana* transient expression method (Nekrasov et al., 2013, Li et al., 2013). As shown, efficiencies up to 12% were observed using a human codon optimized Cas9 (hCas9) directed against two independent targets. Similar mutation rates were observed expressing the gRNAs as two monocistronic gRNAs or as a polycistronic tRNA-gRNA cassette. The variation on the mutation rate observed for the monocistronic gRNAs among experiments can be explained with the variability of the expression levels on *N. benthamiana* transient expression assays or with the unreliability of gel image analysis tools. In our hands hCas9 performed better than plant-optimized pcoCas9 in *N. benthamiana* transient assays, although it remains to be seen if the same differences are observed in other experimental systems. The mutation rate observed here with the hCas9 is consistent with those described when hCas9 and gRNAs were assembled in the same T-DNA (Upadhyay et al., 2013) and much higher than the rates obtained by Upadhyay et al. (2013) and Nekrasov et al. (2013) when the same were co-delivered in different plasmids by *in trans* co-transformation. The reported efficiency for the plant-optimized pcoCas9 when co-expressed with the gRNA on the same vector was substantially lower (4.8%) (Li et al., 2013). Therefore it is possible that our detection system based on the presence of an undigested band was not sensitive enough to detect this mutation rate.

The ability of GB-adapted gRNA/Cas9 elements to conduct RNA-guided transcriptional regulation was assessed by using the pNOS fused to luciferase as a reporter system. We observed that, by directing a nuclease-inactivated Cas9 to promoter regions around the transcription origin of the reporter gene, expression levels were severely reduced. These results were in line with previous reports showing an intrinsic repressor activity of a dCas9 without further modifications (Piatek et al., 2015, Bikard et al., 2013); however in our experimental conditions dCas9 intrinsic repression was almost completely abolished when paired to gRNAs targeting distal regions upstream of the -100 position. In the same upstream regions, however, the translational fusion of dCas9 with specific transcription modulating protein domains efficiently conducted the downregulation (BRD, SRDX) or upregulation (VP64, EDLL) of the reporter activity respectively. It was also observed that, by targeting several gRNAs towards the same promoter, the activation/repression effect was increased, highlighting the convenience of multiplex

targeting to achieve efficient transcriptional regulation. Altogether, the range of transcriptional activities that we were able to modulate using current GB gRNA-Cas9 tools was relatively modest, approximately seven times from the strongest repressor to the strongest activator. Further optimization of the system (e.g. improved fusion linkers, optimization of fusion sites, etc.) will be necessary to increase this efficiency. Nevertheless it should be noticed that, given that in the *N. benthamiana* agroinfiltration system several T-DNA copies of the reporter gene are co-delivered simultaneously in each cell there is probably a high demand for dCas9 fusions to achieve substantial activation/repression. In future experiments the quantification of the effect of dCas9 fusions on single copy genes stably integrated in the plant genome will be investigated.

Very recently, new gRNA-Cas9 toolkits for targeted mutagenesis or transcriptional regulation have been reported including animal (Senis et al., 2014, Port et al., 2014) and plant-dedicated (Lowder et al., 2015, Xing et al., 2014, Ma et al., 2015) systems, although none of them involved a standardized strategy. Interestingly, the toolbox reported by Lowder et al. incorporates gRNA-Cas9 elements for targeted mutagenesis and transcriptional regulation using a combination of type IIs and gateway recombination for multiplex assembly. In comparison, the GB toolbox showed here present a number of distinctive features. First, the GB toolbox includes a number of software tools that generate standardised protocols in each gRNA-Cas9 assembly step. The implementation of assembly software tools not only serves to facilitate construct-making for non-trained users, but most importantly, it turns GB into a self-contained, fully traceable assembly system, where all elements generated with GB software tools, now including also gRNA/Cas9 elements, are perfectly catalogued and their genealogy documented. Second, the modularity of GB facilitates combinatorial arrangements as e.g. between pre-set gRNA arrays and different Cas9 versions and enables the exchange of pre-made combinations. Finally, the GB cloning loop enables endless assembly of both monocistronic and polycistronic tRNA-gRNA expression cassettes, enhancing the multiplexing capacity of the system.

4. Conclusions

A modular gRNA-Cas9 toolbox conforming the GoldenBraid standard for Plant Synthetic Biology was developed and functionally validated. The GB-gRNA/Cas9 toolbox, comprising an adapted cloning pipeline, domesticated gRNA/Cas9 elements and a dedicated software tool, was shown to facilitate all-in-one-T-DNA cloning and gRNA multiplexing. The GB-adapted gRNA/Cas9 elements combined among them and/or with other GB elements were shown effective in targeting reporter genes for mutagenesis,

transcriptional activation and transcriptional repression in *N.benthamiana* transient assays. The GB adaptation enhances CRISPRs/Cas9 technology with traceability, exchangeability and improved combinatorial and multiplexing capacity.

5. Methods

5.1. GBparts construction

GBparts used in this work were created following the domestication strategy described in (Sarrion-Perdigones et al., 2013). For parts GB0575, GB1001 and GB1079, PCR amplifications with the primers obtained at https://gbcloning.upv.es/do/domestication/ were performed using the Phusion High-Fidelity DNA polymerase (Thermo Scientific). For level 0 parts GB0273, GB0645, GB1175, GB1185, GB1186, GB1187 and for level -1 parts GB1205, GB1206, GB1207 double-stranded DNA was synthesized using IDT gBlocks® Gene Fragments. GB1041 was amplified from GB0575 to incorporate the D10A and H840A mutations. For level 0 parts, 40ng of the PCR products or gBlocks® were cloned into the pUPD with a BsmBI restriction-ligation reaction. Level -1 parts were cloned into the pVD1 (GB0101) with a BsaI restriction-ligation reaction following the same protocol. A list of the level -1 and level 0 parts is provided in the Supplementary Table 7; their nucleotide sequences can be searched at https://gbcloning.upv.es/search/features/ with their corresponding ID numbers. All level -1 and level 0 GB parts were validated by restriction enzyme (RE) analysis and confirmed by sequencing.

5.2. Guide RNA assembly on level 0 and level 1

Assembly optimization reactions were performed as follows: primers gRNA_XT2_F/ gRNA_XT2_R were resuspended in water and STE buffer (10 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA) to final concentrations of 100, 10 and 1 µM. Equal volumes of forward and reverse primers were mixed. The mixture was split into two different tubes and one of them was incubated at 94ºC for 2 minutes prior to a 30 minutes incubation at room temperature while the other was directly incubated at room temperature for 30 minutes. The BsaI restriction-ligation reactions were set up in 10µl with 1µl of primers mix, 75ng of GB1001 (U626 promoter), 75ng of GB0645 (scaffold RNA) and 75ng of pDGB3α1 destination vector. One microliter of the reaction was transformed into *E. coli* TOP10 electrocompetent cells and the number of white colonies growing on agar plates counted. The selected conditions for the gRNA assemblies were dilution in water, incubation at room temperature for 30 minutes and set the restriction-ligation reaction with a final

primer concentration of 0.1µM. For gRNA assemblies on level 1, two complementary primers designed at www.gbcloning.upv.es/do/crispr/ and listed on Supplementary Table 6, were included in a BsaI restriction-ligation reaction following the selected conditions. For the assembly of guide RNAs on level 0, the primers listed on Supplementary Table 6 were included in a BsmBI restriction-ligation reaction following the selected conditions together with the pUPD2 and 75ng of the corresponding level -1 tRNA-scaffold plasmid depending on the desired position of each target on the level 1 assembly. All level 1 gRNA constructs were validated by RE-analysis, analyzed by sequencing and confirmed correct.

5.3. Cloning in α and Ω-level destination vectors

Multipartite BsaI restriction-ligation reactions from level 0 parts and binary BsaI or BsmBI restriction-ligation reactions were performed as described in Sarrion-Perdigones et al. (2013) to obtain all the level≥1 assemblies. A list with all the TUs and modules used in this work is provided on the Supplementary Table 7. All level≥1 were validated by restriction enzyme (RE) analysis. Furthermore, partial sequencing was carried out to check part´s boundaries. The sequences of all level≥1 constructs can be found entering their IDs (displayed at Supplementary Table 7) at https://gbcloning.upv.es/search/features/.

5.4. Nicotiana benthamiana agroinfiltration

For transient expression, plasmids were transferred to *Agrobacterium tumefaciens* strain GV3101 by electroporation. *N. benthamiana* plants were grown for 5 to 6 weeks before agroinfiltration in a growing chamber compliant with European legislation. Growing conditions were 24°C (light)/20°C (darkness) with a 16-h-light/8-h-dark photoperiod. Agroinfiltration was carried out with overnight-grown bacterial cultures. The cultures were pelleted and resuspended on agroinfiltration solution (10mM MES, pH 5.6, 10mM $MgCl_2$, and 200µM acetosyringone) to an optical density of 0.2 at 600nm. After incubation for 2 hours at room temperature on a horizontal rolling mixer, the bacterial suspensions were mixed in equal volumes. The silencing suppressor P19 was included in all the assays; in the same T-DNA for the transcriptional regulation experiments and co-delivered in an independent T-DNA for the targeted mutagenesis assays. Agroinfiltrations were carried out through the abaxial surface of the three youngest leaves of each plant with a 1ml needle-free syringe.

5.5. Genomic DNA extraction and PCR/ Restriction enzyme assay

Samples for genomic DNA extraction were collected from 5 days post infiltrated leaves. For genomic DNA extraction, 50mg of tissue powder coming from a pool of three leaves were mixed with in 500µl of DNA extraction buffer (200mM TrisHCl-pH 7.5, 250mM NaCl, 25mM EDTA, 0.5% SDS). The plant extract was mixed gently and it was spin at 14000 x *g* for 3 minutes. The supernatant was transferred to a new tube and an equal volume of isopropanol was added for DNA precipitation. The supernatant was removed after centrifugation (5min at 14000 x *g*) and the DNA was washed twice with 70% ethanol. The pellet was dried for half an hour and it was dissolved with 100µl of elution buffer (10mM TrisHCl-pH8, 1mM EDTA).

DNA amplicons covering the XT1 and XT2 target sites were obtained by PCR of genomic DNA using the Phusion High-Fidelity DNA polymerase (Thermo Scientific) and two pairs of gene specific primers: XT1_F/XT1_R for XT1 and XT2_F/XT2 _R for XT2 (Supplementary Table 5). The resulting PCR products were purified with the QIAquick PCR purification kit (QIAGEN) following the manufacturer's protocol and restriction reactions were set up with 500ng of purified DNA and the corresponding restriction enzyme; BsmBI (Fermentas) for XT1 and SpeI (Fermentas) for XT2. Band intensities were estimated using the 'Benchling Gels' (https://benchling.com) tool.

5.6. Gel band purification and BsaI-cloning

PCR products resistant to BsmBI and SpeI digestion were purified from a 1% agarose gel with the QIAEX II Gel Extraction Kit following the manufacturer's protocol. For sequence analysis, the purified PCR products were subsequently amplified with XT12BsaI_F/XT12BsaI_R primers (Supplementary Table 5) to incorporate BsaI sites for improving cloning efficiency. Finally, they were cloned into the pDGB3α1 with a BsaI restriction-ligation reaction and individual clones were sequenced.

5.7. Luciferase/Renilla activity determination

Samples of leaves coinfiltrated with the REP (GB1116), different activator/repressor TUs (GB1172 and GB1188 to GB1191) and the independent or combined gRNAs targeting the pNOS were collected at 4 days post infiltration. For the determination of the luciferase/renilla activity one disc per leaf (d=0.8cm, approximately 18-19mg) was excised, homogenized and extracted with 150µl of 'Passive Lysis Buffer', followed by 15 minutes of centrifugation (14000 x *g*) at 4ºC. Then, the supernatant was diluted 2:3 in

Passive Lysis Buffer resulting in the working plant extract. Fluc and Rluc activities were determined following the Dual-Glo® Luciferase Assay System (Promega) manufacturer's protocol with minor modifications: 10µl of working plant extract, 40µl of LARII and 40µl of Stop&Glo Reagent were used. Measurements were made using a GloMax 96 Microplate Luminometer (Promega) with a 2-second delay and a 10-second measurement. Fluc/Rluc ratios were determined as the mean value of three samples coming from three independent agroinfiltrated leaves of the same plant and were normalized to the Fluc/Rluc ratio obtained for a reference sample including the REP (GB1116) co-infiltrated with an unrelated gRNA (GB1221) and the corresponding activator/repressor TU.

6 | Chapter 4

Development of an intragenic selection marker for tomato transformation

Vazquez-Vilar M, Fernandez-del-Carmen A, Presa S, Granell A, Orzaez D.

1. Introduction

Cisgenesis and intragenesis were born as alternatives to ease public opposition to transgenic crops restricting the transfer of foreign DNA to a sexual compatibility group (Holme et al., 2013, Rommens, 2004, Rommens et al., 2007). In both strategies, DNA from the same or cross-related species is used for engineering crops with new agronomic characters. While cisgenesis involves the transfer of genes with their own regulatory regions (Schouten et al., 2006), intragenics aims to create new combinations of genetic elements (i.e. coding regions, promoters and other regulatory regions). Advances in DNA assembly technologies are enhancing the application of intragenics for generating complex traits in plants. With classical cloning strategies, the combinatorial potential of genetic elements was limited and time-consuming, preventing the generation intragenic constructs. However, modular methods give the flexibility required for creating new combinations of genetic elements from standard parts. Moreover, some of those methods such as GoldenBraid (Sarrion-Perdigones et al., 2011) also offer standardized strategies for stacking several genes in the same T-DNA.

To avoid the introduction of foreign DNA in intragenic crops, new transformation protocols resulting in plants without bacterial selection markers have been developed. These strategies include co-transformation with an antibiotic selection marker followed by segregation or site-specific recombination-mediated marker deletion (Darbani et al., 2007). However, the incorporation of intragenic selection markers in the same T-DNA where the gene of interest is contained would avoid high labor requirements of marker-removal techniques. Some selection markers derived from plants and conferring resistance to herbicides were developed in the last years including the use of mutated versions of the *Petunia hybrida* 5-enolpyruvylshikimate-3- phosphate synthase, the *Arabidopsis thaliana* acetolactate synthase and tryptophan synthase beta or the tobacco anthranilate synthase (Sundar & Sakthivel, 2008). The selection marker based on mutated forms of the acetolactate synthase gene (ALS) is one of the best established. ALS catalyzes the first step of the branched-chain amino acids biosynthesis pathway (Binder, 2010). Different amino acid substitutions on the ALS were reported to confer tolerance to different ALS-inhibiting herbicides, sulfonylureas, imidazolinones and sulfonanilides (Zhou et al., 2007). While other mutations cause resistance against imidazolinone, the Pro-197 mutation (in reference to AtALS) results in sulphonylureas resistance (Haughn et al., 1988). Pro-197 is located within the ALS substrate access channel and is important for herbicide binding, thus resulting in the impediment of the access of the substrate to the catalytic site when the herbicide is present (Zhou et al., 2007). Selection markers based on a mutated ALS were developed for some crops such

as apple (Yao et al., 2013) or tobacco (Haughn et al., 1988), but still unavailable for tomato transformation. Moreover, although some plant-based selection markers are already available, to our knowledge all of them use heterologous promoters such as the CaMV35s promoter (Yao et al., 2013) for driving gene expression. However, a full intragenic approach requires not only plant-CDS markers, but also plant regulatory regions. The availability of RNA sequencing data for most relevant crops makes possible the identification of regulatory regions driving strong and constitutive transcription levels that could replace heterologous promoters and terminators.

Intragenesis can be used in fruits to confer biofortified phenotypes, and several tools that might be useful for modifying fleshy fruits have been described (Molesini et al., 2012). Tomato is one of the most relevant crops in the world and their fruits have been engineer in many occasions to enrich their quality and nutritional and health value (Raiola et al., 2014). Flavonoids are a subclass of phenylpropanoids with confirmed health-promoting properties as result of their antioxidant activity. Based on it, it was proposed their role in prevention of heart coronary diseases or colon cancer (Martin et al., 2013). Tomato fruits have only small amounts of flavonoids which are mainly located in the peel (Raiola et al., 2014). The main flavonoids in tomato are naringenin chalcone and some flavonols including kaempferol and quercetin. Two different approaches have been followed in the past to increase the level of flavonoids in tomato fruits (Schijlen et al., 2004, Bovy et al., 2007). The first of them consisted in the expression of different biosynthetic enzymes, which led to a modest increment on the flavonoid quantity (Muir et al., 2001, Niggeweg et al., 2004). The expression of regulatory genes that activate the expression of several enzymes of the pathway was reported to be more effective, leading to the accumulation of different flavonoid compounds such as anthocyanins, flavonols or both, depending on the specific combinations of transcriptional factors employed (Luo et al., 2008, Bovy et al., 2002, Zhang et al., 2015). In particular, overexpression of Arabidopsis MYB12 (AtMYB12), a master regulator of the phenylpropanoid biosynthetic pathway, in tomato fruit, led to a dramatic increase in flavonol levels in fruits, as well as to the overaccumulation of caffeoyl quinic acids (Luo et al., 2008). The *Solanum lycopersicum* MYB12 homologous gene SlMYB12 binds the same promoter regions as its orthologous Arabidopsis counterpart (Zhang et al., 2015) and controls the accumulation of flavonoids in tomato peel (Ballester et al., 2010) suggesting a similar role in the flavonoid biosynthesis regulation.

In this chapter we report the generation of a mutated version of the *Solanum lycopersicum* acetolactate synthase (mSlALS) to be used as an intragenic selection marker for tomato transformation. As proof of principle, we combined the acetolactate synthase selection marker with a transcriptional unit that overexpresses SlMYB12 with

fruit-specific regulatory regions obtaining intragenic tomato fruit with increased levels of flavonols.

2. Results

2.1. Assembly of an intragenic selection marker for tomato transformation with GoldenBraid

We decided to clone and mutate the *S.lycopersicum* ALS (mSlALS) as earlier described for tobacco and apple (Yao et al., 2013). The mutation performed was a change proline to serine at amino acid 186 (position 197 in reference to *Arabidopsis thaliana* ALS) which had been described to confer the highest resistance to chlorsulfuron to the apple ALS (Figure 16a). Following the GoldenBraid standard domestication procedure, we generated the Level 0 part GB0816, consisting on the mutated SlALS (mSlALS). The mutant conversion of the conserved proline residue to serine was confirmed by sequencing (Figure 16a).

To create a completely intragenic selection marker, GB0816 was multipartite assembled with tomato regulatory regions. A search of highly expressed genes on tomato leaves was performed using the Tomato Functional Genomics Database (http://ted.bti.cornell.edu/). The top fifty highly expressed genes were further categorized on the basis of their expression levels in different tissues and the three ones showing both high expression and broad tissue distribution (Solyc01g099770, Solyc06g007510 and Solyc09g010800) were selected, and their regulatory regions were isolated and domesticated using the GoldenBraid procedure to create a minilibrary of regulatory 'phytobricks' conforming the Plant Standard Syntax. The relative transcriptional activity conferred by each promoter element was tested transiently in *N. benthamiana* following a standardized procedure (SE_001) as described in Chapter 2. Among the different assayed promoters, GBpart GB0080, corresponding to the metallothionein-like protein type 2B (Solyc09g010800) promoter, in combination with GB0142, the terminator region of the same gene was shown to confer highest relative expression levels, 6.94±0.95 RPUs (Sarrion-Perdigones et al., 2013). Based on this data we assembled the mSlALS with the Solyc09g010800 promoter and terminator in a multipartite reaction obtaining the TU of the intragenic selection marker (Figure 16b).

Next, the efficiency of tomato transformation procedure using the new selection marker was estimated using a reporter construct comprising a red fluorescent protein DsRed under the control of the standard CaMV35s promoter (GB0030) and the *Nos*

terminator (GB0037) elements available on the GBcollection at the moment this work was carried out. Making use of the modularity of GoldenBraid and the reusability of the GBelements, the two-gene reporter construct was assembled with a single binary reaction (Figure 16c).

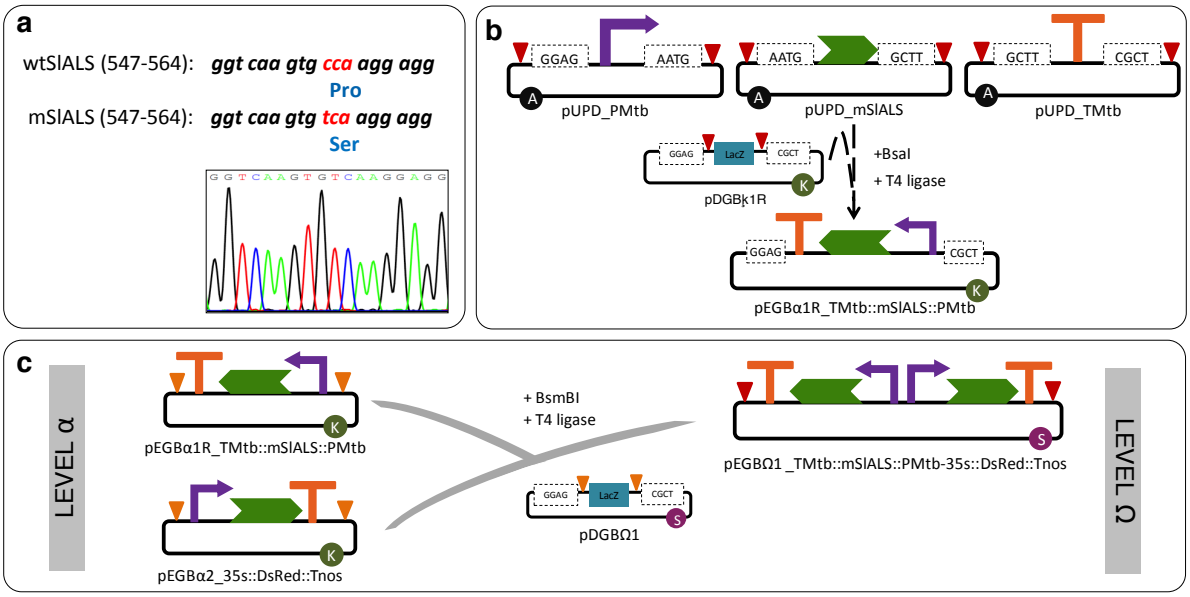


Figure 16. Intragenic selection marker assemblies with GoldenBraid.

(a) mSlALS sequence compared with the wtSlALS sequence: the Pro residue at position 186 (197 in reference to the AtALS) was changed to Ser to create the mutant version mSlALS confirmed by sequencing. (b) Multipartite assembly of the Level 0 parts GB0080_PMtb, GB0816_mSlALS and GB0142_TMtb for the construction of the transcriptional unit pEGBα1R_TMtb::mSlALS::PMtb in a α-level reverse destination vector (GB0818). (c) Schema of a binary assembly between TUs pEGB1α1R_TMtb::mSlALS::PMtb and pEGB1α2_35s::DsRed::Tnos in the Ω-level vector pDGB1Ω1 to create the reporter module pEGB1Ω1_TMtb::mSlALS:.TMtb-35s::DsRed::Tnos (GB0829).

2.2. Determination of the optimal chlorsulfuron dose for tomato transformation

For the determination of the optimal chlorsulfuron (CLS) dose for tomato transformation we first tested the ability of shoot regeneration of untransformed cotyledon explants at different CLS concentrations. Shoot regeneration was unaffected at 0.5 and 1 μg L^{-1} CLS, slightly reduced at 3 μg L^{-1}, highly reduced at 5 μg L^{-1} and resulted in death of explants at 10 μg L^{-1} (Figure 17).

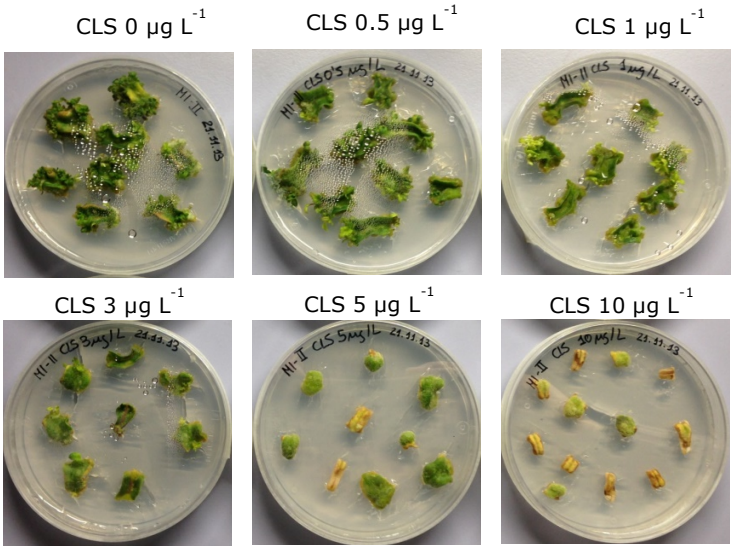


Figure 17. Untransformed explants kept for 5 weeks on induction medium with different chlorsulfuron concentrations.

Based on the tolerance results, explants infected with *Agrobacterium* carrying the construct GB0829 were transferred to medium containing CLS at 5, 7.5, 10, 15 and 20 µg L^{-1} for shoot regeneration. The same concentrations of CLS were maintained on the elongation and on the rooting media. Transformed explants were counted based on their DsRed fluorescence (Wenck et al., 2003). While concentrations below 10 µg L^{-1} in the shoot induction medium resulted in the regeneration of several untransformed plants, at higher concentrations the number of regenerated plants showed a severe decrease. The higher number of DsRed positive plants was obtained with 10 µg L^{-1} of CLS, resulting in a transformation efficiency of 6.7% (Table 3). Altogether, our results indicated that 10 µg L^{-1} of CLS is the most suitable concentration for tomato transformation with the GB0818 selection marker.

No. Explants	CLS concentration at induction medium (µg L^{-1})	DsRed+ plants	
		No.	%
40	0	0	0.0
60	5	1	1.7
60	7.5	1	1.7
178	10	12	6.7
135	15	2	1.5
141	20	2	1.4

Table 3. CLS resistant plants expressing DsRed.

2.3. Effects of SlMyb12 overexpression following an intragenic approach on tomato phenylpropanoids

AtMYB12 has been reported as a transcription factor capable of activating flavonol biosynthesis in Arabidopsis (Mehrtens et al., 2005), tobacco and tomato (Luo et al., 2008). To study whether an intragenic construct for SlMYB12 overexpression in tomato fruit could have the same effect, SlMYB12 was multipartite assembled with GB0914 and GB0144 derived from the E8 5' and 3' regulatory regions and the module TMtb::MSlALS::PMtb-E8::SlMYB12::TE8 (hereafter referred to as mSlALS-SlMYB12 or GB0830) was generated in a binary GoldenBraid reaction similar to the one performed to assemble the reporter construct (Figure 16c). Tomato transformation with mSlALS-SlMYB12 was performed with the optimal concentrations of CLS determined with the reporter construct and described above. With these conditions eleven independent herbicide-resistant plants rooted *in vitro* were obtained from 200 explants (5.5% transformation efficiency), which agrees with the transformation efficiency previously estimated with the reporter construct.

Further confirmation of the intragenic status was obtained by genotyping herbicide-resistant plants by PCR using specific primers listed on Supplementary Table 8 for the MYB12 transcriptional unit (Figure 18a). Confirmed intragenic plants were transferred to soil and grown to maturity. At the moment of preparing this chapter, four independent intragenic plants had produced ripen fruits. From them, two lines (lines 1A and 10) showed an orange color despite being fully ripen, a phenotype previously described for AtMYB12 by Luo et al. (Luo et al., 2008). In contrast, lines 2 and 9 showed a red color slightly darker than wild type (Figure 18b). To obtain preliminary data of flavonoid composition, samples from lines 1A and 2, representing each of the observed fruit color phenotypes were collected for further examination. To investigate which metabolites

were differentially accumulated in ripe fruits, peel and flesh samples of five fruits of each line together with a control wild type were separately analyzed by LC-ESI(+/-)-MS and the most relevant flavonoids were identified and quantified (see Table 4). The peel of both intragenic and control lines contained naringenin chalcone as the most abundant flavonoid detected, but this compound was 5-fold and 3.5-fold most abundant in wild type peel when compared to line 1A and line 2 respectively. Naringenin chalcone is always present in the peel and its content is highly variable between fruit samples, a variability that could partially explain this discrepancy. No drastic differences were found for the remaining flavonoids identified in the peel. A similar amount of flavonoids in the peel of transformed and control lines was expected since the ethylene-inducible E8 promoter drives gene expression in the flesh of ripe fruits (Kneissl & Deikman, 1996). In contrast, severe differences were observed in a number of compounds identified in the flesh samples. For some specific flavonol compounds as quercetin and kaempferol and their glycoside derivatives, over-accumulation levels above 100 fold in comparison with the control line were observed. Significant over-accumulations were found also for naringenin chalcone and its glucosides. The highest difference between intragenic lines 1A and 2 was observed in the naringenin chalcone glycoside content, which levels are 8-fold higher in line 1A than in line 2, a variation that correlated with the different color phenotypes observed in the fruits of these two lines. However, levels of other flavonoid glycosides such as quercetin-hexose-deoxyhexose-pentose, kaempferol-hexose-deoxyhexose-pentose and kaempferol-diglucoside are more than 2-fold higher in line 2 than in line 1A. As Luo et al. reported modifications on caffeoyl-quinic acids (CQAs) levels when AtMYB12 was expressed in tomato fruits (Luo et al., 2008), we decided to investigate whether this content was also modified in the mSlALS-SlMYB12 lines. On the peel we found that amounts of CQAs were similar to the control on both intragenic lines. In the flesh, we observed a modest 3-fold enrichment of chlorogenic acid and dicaffeoyl quinic acid for line 2, while for line 1A no substantial modifications on CQAs levels were detected.

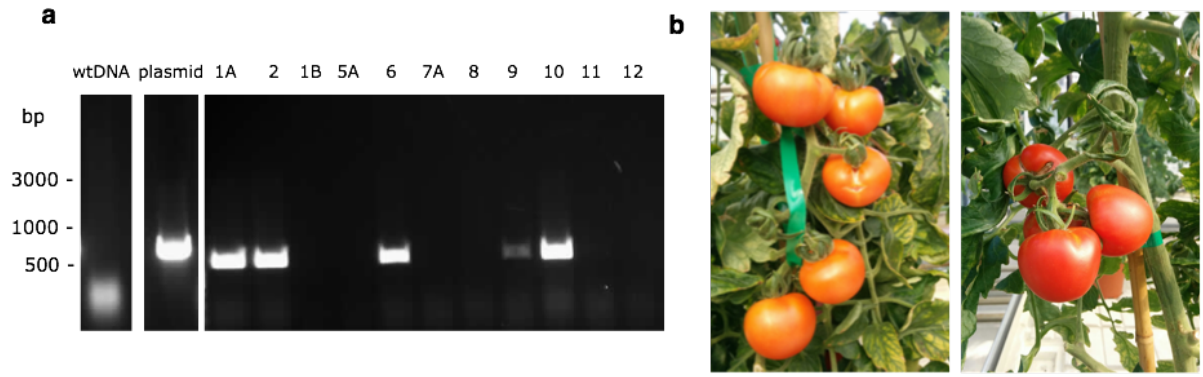


Figure 18. Phenotype of mSlALS-SlMYB12 tomato lines.

(a) Genotyping of 12 independent lines with specific primers covering the end of the mSlALS and the TMtb. (b) Tomato plants var. Moneymaker of mSlALS-SlMYB12 lines 1A (left) and 2 (right).

	Peel		Flesh	
	Line 1A	Line 2	Line 1A	Line 2
4-caffeoyl-quinic-acid	0.66 ± 0.08	0.61 ± 0.04	0.72 ± 0.11	3.30 ± 0.65
Dicaffeoylquinic-acid	0.56 ± 0.04	0.69 ± 0.23	0.37 ± 0.02	3.39 ± 0.87
Hydroxylated naringenin chalcone	2.30 ± 0.42	2.84 ± 0.70	57.98 ± 4.78	35.93 ± 8.30
Naringenin-chalcone	0.20 ± 0.03	0.29 ± 0.02	6.84 ± 0.13	3.40 ± 0.70
Naringenin-chalcone-dihexose	1.18 ± 0.20	2.26 ± 0.45	267.67 ± 64.92	172.81 ± 45.67
Naringenin-chalcone-glucoside	0.88 ± 0.13	1.28 ± 0.42	44.33 ± 11.42	5.84 ± 0.27
Naringenin-chalcone-hexose-I	0.59 ± 0.09	0.97 ± 0.28	58.30 ± 15.02	26.29 ± 8.76
Naringenin-chalcone-hexose-II	1.78 ± 0.45	2.57 ± 0.58	628.78 ± 104.16	212.15 ± 69.93
Naringenin-dihexose-I	0.59 ± 0.23	2.40 ± 0.36	0.76 ± 0.22	2.62 ± 0.39
Naringenin-hexose	1.78 ± 0.45	2.57 ± 0.58	567.32 ± 173.29	213.70 ± 70.44
Kaempferol-diglucoside	0.63 ± 0.10	1.72 ± 0.29	7.89 ± 2.00	17.47 ± 5.08
Kaempferol-dihexose	0.32 ± 0.05	1.11 ± 0.15	6.21 ± 0.76	9.08 ± 0.95
Kaempferol-dihexose-deoxyhexose	4.28 ± 0.68	7.13 ± 1.64	884.61 ± 255.30	1293.97 ± 291.59
Kaempferol-glucose-rhamnose	0.83 ± 0.10	2.48 ± 0.16	149.86 ± 28.03	244.58 ± 54.67
Kaempferol glucosyl-glucoside rhamnoside	4.98 ± 0.75	6.47 ± 1.30	690.93 ± 189.57	1186.32 ± 330.89
Kaempferol-rutinoside	0.79 ± 0.04	1.95 ± 0.46	663.98 ± 237.49	726.85 ± 52.84
Quercetin deoxyhexose-hexose-deoxyhexose	11.68 ± 1.52	23.93 ± 7.89	1730.45 ± 420.65	2648.28 ± 457.90
Quercetin-hexose	1.24 ± 0.36	1.29 ± 0.28	687.86 ± 158.51	103.31 ± 13.87
Rutin	0.32 ± 0.05	1.11 ± 0.15	6.21 ± 0.76	9.08 ± 0.94

Table 4. Relative CQAs and flavonoids content of mSlALS-SlMYB12 tomato lines 1A and 2.

Metabolites were quantified based on the amount of internal standards and data were normalized by setting the control line (non-transformed wild-type Moneymaker) as one. All data are presented with their standard deviations. In red and green, metabolites over- and down-accumulated respectively on the mSlALS-SlMYB12 line in reference to wild type (T-Test with a p-value < 0.05).

3. Discussion

The exponential decrease of genome sequencing costs is expanding the number of well-defined components arising from natural diversity that might be useful for plant genetic engineering. The search of new genetic elements within Solanum species provided the resources required to carry out intragenic approaches in tomato. Currently, intragenic/cisgenic crops are regulated as transgenics in most countries since regulation is process-based and the assessment requires, not only the evaluation of the new traits of the plant, but also the consideration of the process by which it was created (Russell & Sparrow, 2008). Currently only Canada and US have a product-based regulatory system, which allows a less strict control of intragenic/cisgenic plants compared to transgenics (Hou et al., 2014, Schaart et al., 2015).

Here, we showed how modular cloning could enhance the application of intragenesis as a new plant breeding technique in tomato by providing (1) a standard collection of coding sequences and regulatory elements from *Solanum lycopersicum* and (2) a well-designed strategy to expand the combinatorial possibilities of those elements, facilitating the assembly of multigenic constructs as those required for metabolic engineering. Thus, modularity and traceability are features exemplified in this work that speed up research and make the ready-to-transform constructs more reliable in terms of functionality.

On this chapter we reported the development of an ALS-based selection marker for tomato transformation. To our knowledge, the combination of a mutated ALS with regulatory regions of the same plant species had not been previously reported. Using GoldenBraid, the tomato regulatory regions were easily assembled with the mSlALS, and subsequently with the intragenic TU containing SlMyb12, illustrating the combinatorial capacity of the system. The 186 Pro>Ala conversion (197 in reference to AtALS) of SlALS was shown to confer resistance to chlorsulfuron, as previously demonstrated for AtALS in tobacco (Haughn et al., 1988) and apple ALS in tobacco and apple (Yao et al., 2013). The development of the ALS-based selection marker provides not only an intragenic marker for tomato transformation as was shown here, but also a useful alternative for pyramiding new genes in previously transformed tomato plants.

The transformation efficiencies of 5.5% shown here are far below those obtained with the use of the well established nptII antibiotic selection marker, (above 50% tested in a standard experiment SE_003 with equivalent transformation vectors, data not shown), but are high enough to be considered for routine transformations if final intragenic lines need to be generated, particularly if the intragenesis status were to confer a commercial advantage to the resulting product. Nevertheless, further optimization of the

transformation protocol would be advisable to reach higher transformation efficiencies. Several optimization strategies could be envisioned, for instance, to increase the expression levels of the herbicide resistance gene. In this regard, it has been reported that expression of the tomato SlMT1 gene, whose regulatory regions were used to create the intragenic marker, can be induced in root tips with zinc sulfate (Giritch et al., 1998). Addition of this compound to the media during tomato transformation could increase resistance to chlorsulfuron and subsequently enhance the transformation efficiency.

As proof of concept, we decided to generate intragenic tomato lines with enhanced flavonoid content, making use of the developed ALS-based selection marker. We aimed to induce the flavonoid biosynthetic pathway on tomato fruit by expressing SlMYB12 with fruit-specific E8 regulatory regions. At the moment of elaborating the results of this Chapter, only four T1 intragenic lines yielded mature fruits, and therefore the results presented here can only be considered as preliminary. However, both the trend as the magnitude of the changes observed in the metabolic profile of the analyzed fruits points toward a clear effect of the intragenic construct in the over-accumulation of specific compounds of the flavonol biosynthesis pathway in the fruit flesh, in line with previous results reported with transgenic lines overexpressing orthologous transcription factors (Luo et al., 2008). The two intragenic lines overexpressing SlMYB12 analyzed so far showed a drastic increase of naringenin, kaempferol and quercetin glycosides in the flesh in reference to wild type. The same differences were not observed in the peel, although this is not fully unexpected as the ethylene-inducible E8 promoter drives gene expression mainly in the flesh of ripe fruits. All results should be confirmed on T2 lines. Also, the total accumulation of antioxidant flavonoids in flesh and peel need to be carefully estimated to confirm that the intragenic strategy pursued here yields fruits that are substantially enriched in its antioxidant levels. In this respect, the data obtained so far indicates an important reduction of the content of naringening chalcone in the peel of intragenic tomatoes, which could result from the augmented consumption of precursors in the flesh. Although the flesh/peel weight relation is likely to compensate any reduction of antioxidant in the peel, quantitative data needs to be carefully evaluated in T2 to evaluate the real biofortification gains of this strategy.

In conclusion, by designing, building and testing mSlALS-SlMYB12 intragenic tomato plants here we illustrate how the elements and principles of Synthetic Biology, i.e. standard DNA parts, assembly rules and experimental results, can be used to generate applied plant biotechnology products. The characterization of the ALS-based selection marker adds a new device to the tomato-engineering toolbox that can be used for the generation of intragenic lines as well as for gene pyramiding in previously transformed

plants. Moreover, the combination of mSlALS with SlMYB12 intragenic construct seems to produce tomatoes with increased level of anti-oxidant flavonols.

4. Materials and methods

4.1. Plant transformation vectors construction

SlALS gene (SGN-U572742) was amplified from *Solanum lycopersicum* cv. Moneymaker genomic DNA. Three nucleotide changes were introduced on the ALS coding sequence, two synonymous mutations to remove one internal BsaI and one internal BsmBI restriction sites and a mutation that leads the Ala>Pro amino acid change at position 186. To introduce the mutations on the SlALS and create level 0 parts GB0816, GB0914 and GB0144 we followed the GB domestication standard procedure (Sarrion-Perdigones et al., 2013) with the primers depicted at Supplementary Table 8. GB0080 and GB0142 (Sarrion-Perdigones et al., 2013) and GB0075 (Sarrion-Perdigones et al., 2014) had been adapted to the GB standard in previous works. Level 0 parts making the selection marker transcriptional unit (GB0080, GB0816 and GB0142) were assembled in the pDGB1α1R GBvector (https://gbcloning.upv.es/feature/pDGB1_alpha1R/) with a Golden Gate reaction to create the level 1 transcriptional unit GB0818. With two additional Golden Gate reactions transcriptional units 35s::DsRed::tNos and E8::SlMyb12::TE8 were assembled in pDGB1α2 vectors. Finally, the selection marker was combined in two different binary reactions with each of the α2 TUs in the pDGB3Ω1 vector generating level>1 elements GB0829 and GB0830.

4.2. Tomato transformation

GB0829 and GB830 were transferred to *Agrobacterium tumefaciens* LBA4404 strain for stable tomato transformation. Tomato (var. MoneyMaker) transformation was carried out as described by Ellul et al. (2003) with minor modifications. Briefly, cotyledons of 10 days tomato plants were cut and explants were submerged in the Agrobacterium culture for half an hour. After that they were transferred to coculture medium and keep in the dark 48 hours. Then explants were transferred to the organogenic medium with different doses of chlorsulfuron (N11461 - Chlorosulfuron from Sigma-Aldrich). Chlorsulfuron was resuspended in a small volume of KOH 1M and diluted to the desired final concentration with distilled water. Individual shoots were excised and transferred to elongation medium prior to being transferred to the rooting medium for root regeneration.

4.3. LC-ESI(+/-)-MS analysis of tomato fruit phenylpropanoids

For phenylpropanoid extraction 10 mg of ground freeze-dried ripe fruit pericarp and peel powder were extracted with 0.75 mL cold 75% (v/v) methanol, 0.1% (v/v) formic acid, spiked with 10 µg/ml formononetin. Samples were vortexed for 30'', and shacked for 15' at 15 Hz using a Mixer Mill 300 (Qiagen) and kept at RT for 5' (twice). After centrifugation for 15' at 20,000 g at 4°C, 0.6 mL of supernatant were removed and transfer to HPLC tubes. For each genotype, 5 fruits and at least two independent extractions were performed. Separation was carried out using an Ultimate 3000 HPLC coupled to a Q-EXACTIVE mass spectrometer (ThermoFisher) equipped with a C18 Luna reverse-phase column (150 x 2.0 mm, 3 µm; Phenomenex, Macclesfield, UK) and a gradient system as follows: 95%A:5%B for one minute, followed by a linear gradient to 25%A:75%B over 40 minutes. LC conditions were kept for 2 more minutes, before going back to the initial LC conditions in 18 minutes. Ten µl of each sample were injected and a flow of 0.2 mL was carried out during the whole LC runs. Detection was performed continuously from 230 to 800 nm with an online Ultimate 3000 photodiode array detector (PDA, Thermo Fischer Scientific, Waltham, MA). All solvents used were LC-MS grade quality (CHROMASOLV® from Sigma-Aldrich). The Exactive Plus Orbitrap mass spectrometer was equipped with a heated electrospray probe (H-ESI). ESI and MS parameters were as follows: spray voltage −5.0 kV, sheath gas and auxiliary nitrogen pressures 30 and 10 arbitrary units, respectively; capillary and heater temperatures were set at, respectively 250 and 150 ºC, while tube lens voltage was 50 V. Data were acquired in profile mode. Identification was performed through comparison of chromatographic and spectral properties of authentic standards and reference spectra, and on the basis of the m/z accurate masses, as reported on Pubchem database (http://pubchem.ncbi.nlm.nih.gov/) for monoisotopic masses identification, or on Metabolomics Fiehn Lab Mass Spectrometry Adduct Calculator (http://fiehnlab.ucdavis.edu/staff/kind/Metabolomics/MS-Adduct-Calculator/) in case of adduct ion detection. Metabolites were relatively quantified on the basis of the internal standard amounts.

7 | General Discussion

Synthetic Biology aims to create biological systems through a systematic design. Towards this objective, Synthetic Biology adopted one of the fundamental principles of engineering, the Design-Build-Test cycle traditionally applied to the development of new products (Kelwick et al., 2014). This cycle can perfectly describe the iterative process followed in the design of a new biological circuit (Figure 19). The first step consist in the planning of the structure and components of the new circuit (DESIGN), and it is followed by the assembly of its components (BUILD) and finally by the analysis of its behavior (TEST). To ensure an iterative design on this cycle some considerations should be addressed, being the standardization the most relevant. Most importantly, standardization should reach all levels in the process, including (1) physical parts structure and assembly methodology, (2) experimental data acquisition and (3) display of functional specifications.

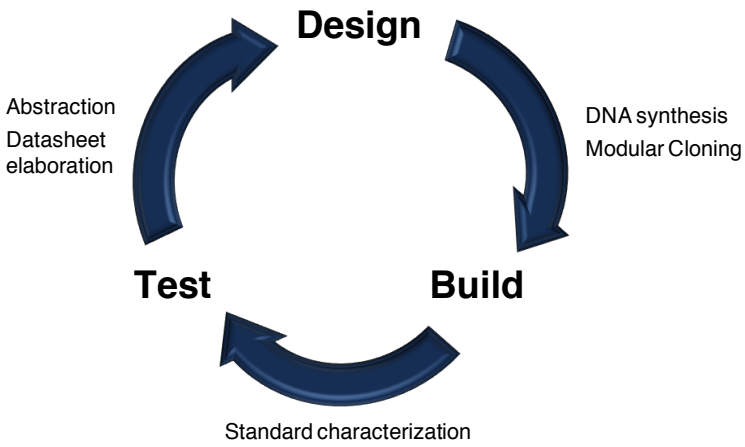


Figure 19. Iterative design in Synthetic Biology.

The Design-Build-Test cycle is one of the engineering principles adopted by SynBio and it describes the process of designing a biological system. After designing, DNA synthesis and Modular Cloning enable the building of genetic devices that are tested with a standard characterization process. After testing, the abstraction and systematic capture of retrieved data ensures the iteration of the cycle by providing the information required for a new Design step.

In this thesis we provide new tools for the assembly of genetic devices for plant biotechnology, and propose new strategies for the systematic acquisition and integration of experimental data within the DBT cycle, especially oriented to Plant Synthetic Biology. Here we will discuss the significance of the development of the integrated gbcloning web-platform described in Chapters 1 and 2 as well as its future perspectives. Furthermore, we will discuss how the work exposed in Chapters 3 and 4 have expanded the GoldenBraid toolbox with the adaptation of the gRNA/Cas9 tools to the GoldenBraid standard and the development of an intragenic selection marker for tomato transformation.

1. Facilitating DNA assemblies with software-tools (BUILD)

Recent advances in DNA assembly technologies have conducted to the standardization of physical DNA composition rules (Hillson, 2011). In Type IIS cloning, a DNA assembly standard defines the prefix and suffix of each part or composite part and establishes standard assembly rules (Casini et al., 2015). Biobricks, the assembly standard introduced by Tom Knight in 2003, constituted the first successful standard for Synthetic Biology (Shetty et al., 2008). However, it has several limitations as it implies only binary assembly steps, slowing down the process. Moreover, it is not suitable for plant transformation since most of the available vectors are only for bacteria transformation. In recent years other methods specially dedicated to plants have appeared. Among them, type IIS based methods such as MoClo and GoldenBraid are extensively being used for stacking several genes in a single T-DNA (Polturak et al., 2015, Marillonnet & Werner, 2015).

Chapter 1 shows how an assembly standard with very simple assembly rules such as GoldenBraid facilitates the development of a database and a group of software tools that assist the design of new multigene constructs from standard parts. With the GB software tools users can smoothly create multigenic assemblies from raw DNA sequences getting a detailed lab protocol for each assembly step. The performance of the gbcloning web-tools was illustrated in Chapter 1 with the step-by-step assembly of a three-gene construct for metabolic engineering. Since the launch of GB2.0 and the creation of the gbcloning website the number of users has grown up very quickly. We know that most of them decided to start using GoldenBraid instead of an alternative assembly strategy because of the software tools. The feedback we get from users is being really helpful for the website improvement. Their insights, concerns and questions helped to improve the website content and therefore the overall user experience. Together with GB users community we have made gbcloning a useful site for Plant Synthetic biologists, although it still requires some improvements in terms of likeability and aesthetics.

The gbcloning site also hosts the GBrepository, a collection that includes the most popular elements used in plant biotechnology. Full potential of modular methods is reached when parts are easily exchangeable and thanks to the GBrepository standard parts can be shared, speeding up the assembly of new devices. On the last three years we distributed around seventy starter kits (comprising destination GBvectors and some basic phytobricks) and we received requests for sixty additional phytobricks. These numbers show that GBusers are taking profit from parts already adapted to the GB standard in our lab for their own assemblies. To fully exploit the potential of standard

parts, the plant synthetic biology community should address the development of a common repository of physical DNA elements in which users can deposit their own parts. While gbcloning can be employed for the exchange of sequence information, our lab is not prepared to receive and store physical DNA parts. For this reason, we are starting to move our GBcollection to Addgene, a nonprofit global plasmid repository. Next step would be to encourage GB users to act in the same way to enlarge the number of publicly available GBparts. This should be done in parallel with the publication by the gbcloning administrator of the GBpart sequence, making it in this way accessible to other users in the community.

The potential of modular strategies relies not only on the facility to share standard parts among users, but also on the number of users. That is the reason why GoldenBraid and MoClo developers has recently reached a consensus in the four nucleotide overhangs that define each part in the transcriptional unit making GB and GG level 0 parts compatible in terms of grammar (Patron et al., 2015). Full compatibility between both systems can be accomplished with two extra features incorporated to the GBdomesticator. The first of them is the incorporation of a new universal entry plasmid with chloramphenicol resistance instead of ampicillin which makes possible the assembly of GB level 0 parts in MoClo level 1 entry plasmids. A second feature was the incorporation to the GBdomesticator tool the choice to domesticate level 0 parts not only to the GB enzymes (BsaI and BsmBI), but also to MoClo enzymes (BpiI). MoClo compatibility or incompatibility is depicted in level 0 part datasheet information. We hope these changes will enlarge the collection of domesticated level 0 parts allowing their exchange between MoClo and GoldenBraid communities.

We created a special section in GB webpage devoted to building devices for genome engineering, as we understand that these new tools will be an integral part of Plant Synthetic Biology in the near future. gRNA/Cas systems have multiple applications in crop breeding, and they are becoming especially attractive under the expectation that deregulation of crops bred with genome editing techniques in general, and with CRISPR/Cas9 in particular, could be reached under less stringent conditions (Jones, 2015), or even bypass regulation (Luo et al., 2015). The gRNA/Cas9 technology, as a targeted mutagenesis tool, can be applied to several breeding goals, as for improving yield potential, increasing the concentrations of some secondary metabolites, extending the life of fruits or enhancing plant pest and disease resistance (Xiong et al., 2015). Moreover, as targeted transcriptional regulators, modified gRNA/Cas9 can be used for hacking endogenous gene expression, an interesting new tool in the plant toolbox with

multiple potential applications. In Chapter 3 we described the adaptation to GoldenBraid of the gRNA-Cas9 tools both for targeted gene editing and targeted gene regulation.
We designed two alternative strategies for the assembly of guideRNAs with GoldenBraid, either as level 1 parts enhancing combinatorial possibilities or as level 0 parts following a polycistronic strategy. The second strategy reduces the combinatorial alternatives while enhances the potential for multiplexing gRNAs. Validation of both assembly strategies give users the choice of selecting one or another depending on their interests. Adaptation of the software tools described in Chapter 1 allowed the incorporation of level 1 gRNA-cassettes as part of multigenic assemblies with GoldenBraid. The generation of software tools assisting the assembly of polycistronic gRNA-cassettes (2-D multiplexing) should be considered in the future to make possible that any assembly performed in the lab could be simulated in the gbcloning site.

Functionality of gRNA-Cas9 GBelements was verified following two of the standard experiments described in Chapter 2, obtaining mutation efficiencies with our constructs of up to 12% with transient expression in *N.benthamiana* and a maximum induction/repression ratio of seven fold from the strongest activator to the strongest repressor. Despite results showed on Chapter 3 cannot be considered actually an example of genome editing, it is true that GB provide the tools for several genome engineering applications in plants such as conferring resistance to plant virus diseases (Ali et al., 2015, Baltes et al., 2015, Ji et al., 2015).

2. Standardizing genetic part characterization in plants (TEST)

A main goal of synthetic biologists is to develop a platform for the automated design of devices with a desired function (Rodrigo et al., 2012). On the design step of the DBT cycle, the behavior of a new biological device has to be predicted based on the functional specifications of its individual components (Kelwick et al., 2014). For the formulation of functional composition rules a modeling approach is required, and the creation of working models is heavily dependent on the quantity and quality of experimental data. Therefore making reliable working models relies on the ability to first obtain and then store functional data in a standard manner. Useful information can only be obtained by testing genetic parts following standard procedures. One of the aims of Chapter 2 is to propose standard procedures for part characterization in plants and provide functional specifications to the elements stored in the GBrepository described in Chapter 1.

The 'Registry of Standard Parts' (http://parts.igem.org) has provided some part characterization data for 10 years. However, there is no a consensus on how parts should be characterized and most of them have a vague characterization. We proposed five types of standard experiments in plants with different degrees of definition on their standard conditions. This is not a definitive not even an optimized list. With this selection we wanted to test the viability of experimental standardization, carefully choosing those situations that serve to characterize most of our in-house phytobricks, with special attention to transcriptional activity measurements. We also tried to reach a compromise between standardization and usability, knowing that the more flexible the experimental conditions are, the wider the range of elements that can be assayed using a single standard, but less comparable the specifications will be, and therefore less useful the standard will be for creating working models for quantitative design.

Most GBelements, including promoters, terminators, transcription factors and other genetic regulators can be characterized with the same standard measurement in the same standard experiment. To date, some measurement standards have been proposed for promoter characterization (Pasotti et al., 2012). Polymerase operations per second (PoPS) and ribosomes per second (RIPS) are abstract measurements whose major disadvantages are that they cannot be directly measured and that they do not capture the modifications at the post-translational level (Kelwick et al., 2014). An alternative measurement standard is the relative promoter units (RPU), which can be measured directly and it is less prone to variation than other measurements across different groups, equipment or slightly different protocols (Kelly et al., 2009). However, it requires a consensus on the choice of the promoter reference. Our proposal is to settle promoter units relative to the nopaline synthase promoter (PNos) as the standard measurement for plant part characterization. Verification of this standard measurement was done following a standard experiment in which we, as testers, specified a set of restrictive operating conditions. We postulated luciferase as reporter gene for data acquisition and the use of a second luminiscent protein as internal reference to decrease possible sources of variation. Several standard experiments performed with the 35s promoter resulted always in measurements of 11±1.5 rpu (relative to PNos) validating the measurement technique.

However, other GBelements such as gRNA/Cas9 constructs for targeted genome mutagenesis cannot be characterized with RPUs. For this reason we defined *ad hoc* parameters such as the 'percentage of overall mutation efficiency' in order to characterize the efficiency of each guide RNA on each target site. This measurement is less general than RPUs, but valid for extracting conclusions when experiments are

performed following the defined conditions. Data generated in any experiment can be useful for researchers in validating the functionality of a part or in providing relevant information for a subsequent design.

The gbcloning database was expanded to store data coming from experiments performed with GBelements. This will help not only to characterize genetic parts, but also to keep experimental information organized. As mentioned, the five standard experiments that we defined cover the range of experiments we usually perform on the laboratory. However, other labs may be interested in incorporating to gbcloning the results of other types of experiments more suitable to their parts´ characterization. Therefore we decided to keep the option of incorporating information obtained in non-standard experiments. Additionally, the set of standard experiments can be expanded to incorporate new ones if and when required.

To our knowledge this is the first proposal towards the establishment of the rules for part characterization in Plant Synthetic Biology. Whether these or similar standards are incorporated by the Plant SynBio community remains to be seen, but the exercise of proposing and testing the practical value of certain experimental conditions is a prerequisite worth to be explored.

3. Datasheets as the way to summarize DNA assembly experimental info (DESIGN)

After testing, abstraction principles are applied to model the behavior of any device, leading to redesign adjustments of a previous design or to a new design. The definition of conditions for testing genetic parts is required but not sufficient to move to the design step of the DBT cycle; first an intermediate step of information organization is required. Some proposals of file structures that can capture sequence and experimental data have been outlined for bacteria (Canton et al., 2008, Lee et al., 2011). Useful datasheets should include vector and assembly standard information, chassis, experimental methodology and experimental data for predictive modeling (Kelwick et al., 2014).

In Chapter 2 we adapted the datasheet fields proposed by Canton et al. (Canton et al., 2008) to plant requirements. This allowed the formulation of the GBdatasheets, which summarize the most relevant assembly and experimental information and include links to all the data generated with a given phytobrick. The use of symbols associated with each level 0 category, a feature based on the SBOL visual standard (Quinn et al.,

2015), together with the traceability offered by the GBtools, makes possible the automatic generation of a graphical display for each GBelement. This traceability also permits to automatically link each datasheet to all the experimental data involving not only the full GBelement, but all its standard components. In the future will be advisable the adaptation of the GBdatasheet to more general Synthetic Biology, as the XML standard. SBOL XML format files are required for interoperation with emerging Synthetic Biology software tools for gene circuit design such as GenoCAD (Cai et al., 2010). Since all the datasheet information is stored in the GBdatabase, it is expected that the conversion to other output formats will be possible without great effort.

4. Final Remarks

The practical implementation of this thesis involved the design and building of countless genetic devices using the GoldenBraid standard. Although most of them were created for transient expression experiments, we wanted to provide also a practical example of the application of cloning standards to plant breeding. To this goal, we designed, built and experimentally tested an 'all-tomato' genetic device comprising an intragenic selectable marker linked to an activator of flavonol biosynthesis. Despite presenting only preliminary T1 results, and although it does not strictly adhere to SB engineering rules in all aspects, we decided to include this example in this Thesis as it illustrates the essence of PSB, successfully completing a full Design-Build and Test cycle. First, although the **design** of the device was not the result of a predictive but rather of an 'intuitive' modeling exercise, we indeed used available gene expression data obtained with standard experiments SE_001 for the selection of appropriated regulatory sequences directing ALS expression. Moreover, the selection of phytobricks Myb12, ALS, and fruit-specific E8 promoter and terminator was based on non-standard but reliable data generated elsewhere. Only plant standard, reusable phytobricks were used for **building** this device, conforming to the Standard Plant Syntax. Consequently we are fully confident that the phytobricks involved in the design will function equally reliably when reused to create any new device. In other words, the functional characterization of our intragenic device serves as a functional specification (e.g. quality assurance) for each of its individual components. Finally, the intragenic device was stably transferred to the plant genome and its performance was functionally **tested**. Transformation efficiency conferred by the ALS selectable marker was tested using a predefined standard (SE_003). This enables comparisons with other selectable markers employed in the same context (tomato transformation using pDGB3 plasmid series). Although it is not possible to establish experimental standards covering all possible metabolic engineering designs, quantitative characterization of the resulting plants involving secondary

metabolite profiles was still possible using a non-standard questionnaire. Eventually, a new experimental setup could be established and adopted if comparisons among similar devices with high reliability had to be conducted in the frame of a dedicated project. In this standardization context, the results obtained with the characterization of T1 flavonol-rich intragenic tomatoes could be easily used to refine the design, therefore entering a new DBT cycle. Alternatively, new accelerated tests could be carried out in a relevant experimental environment (e.g. Myb12 transient transactivation of target tomato promoters of the flavonoid pathway tested in *N. benthamiana*), so that the resulting specifications could serve to feed new designs with quantitative data.

We expect that the work presented in this Thesis will contribute to fuel the engine that keeps the Design-Build-Test wheel moving for the advance of Plant Biotechnology.

8 | Conclusions

C1. The GB2.0 web resource was first created, comprising a database of DNA elements and a package of software tools. GB2.0 enables *in silico* assembly of standard biological parts (Phytobricks) for Plant Synthetic Biology using the GoldenBraid cloning format. The resources generated within GB2.0 were made publicly available and hosted at www.gbcloning.upv.es.

C2. A new improved version of the GB resource (GB3.0) was created next, conforming to the new Standard Plant Syntax. The GB3.0 database was adapted to host experimental data generated by transformation of GB standard biological parts into the plant chassis. To ensure operability, a number of experiment types were defined and proposed as standards for the physical characterization of phytobricks.

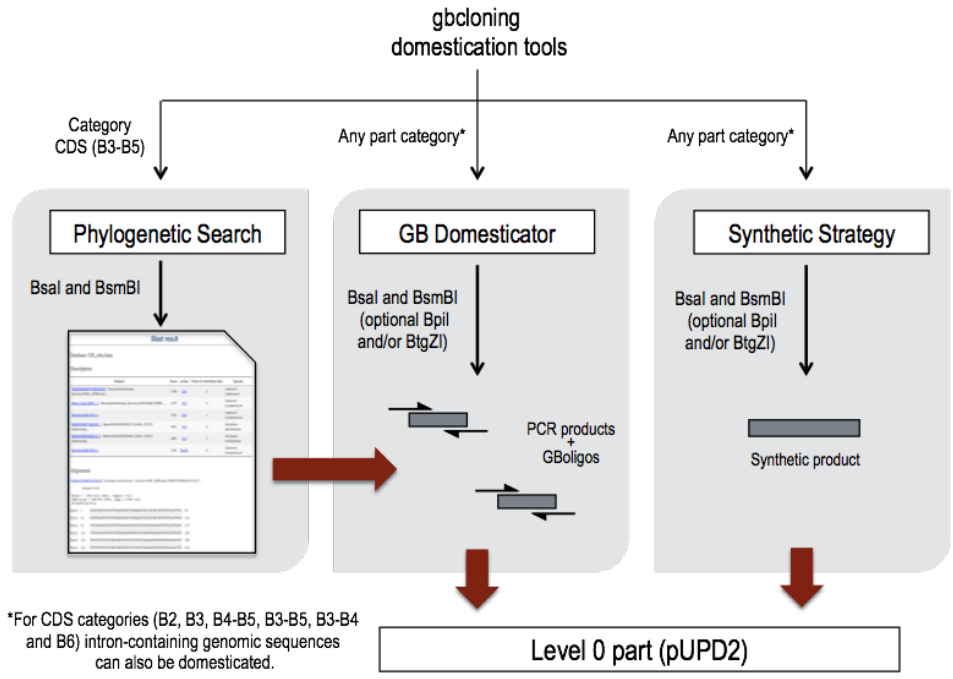
C3. The GB3.0 software automatically generates physical and functional quantitative specifications describing each phytobrick, which are collected and displayed as datasheets. GB datasheets are proposed as basic elements for creating functional composition rules in Plant Synthetic Biology.

C4. Using GB3.0 tools, a simplified Ros1/Del transcriptional regulator prototype for flavonoid biosynthesis was conferred dexamethasone responsiveness by physically connecting it to a dexamethasone-dependant conditional transactivator. The new transcriptional device showed a qualitative behavior that could have been anticipated from the standard specifications displayed at the datasheets of its basic components.

C5. All elements required for the use of the CRISPR/Cas9 technology in plants were adapted to the GB standard. Transient expression assays conducted in *N. benthamiana* following GB standard experiments proved the functionality of the generated devices for gene editing, gene activation and gene repression. To enable *in silico* multigene assemblies of sgRNA-Cas9 devices, the GB3.0 resource was updated to support sgRNA-cassette assemblies.

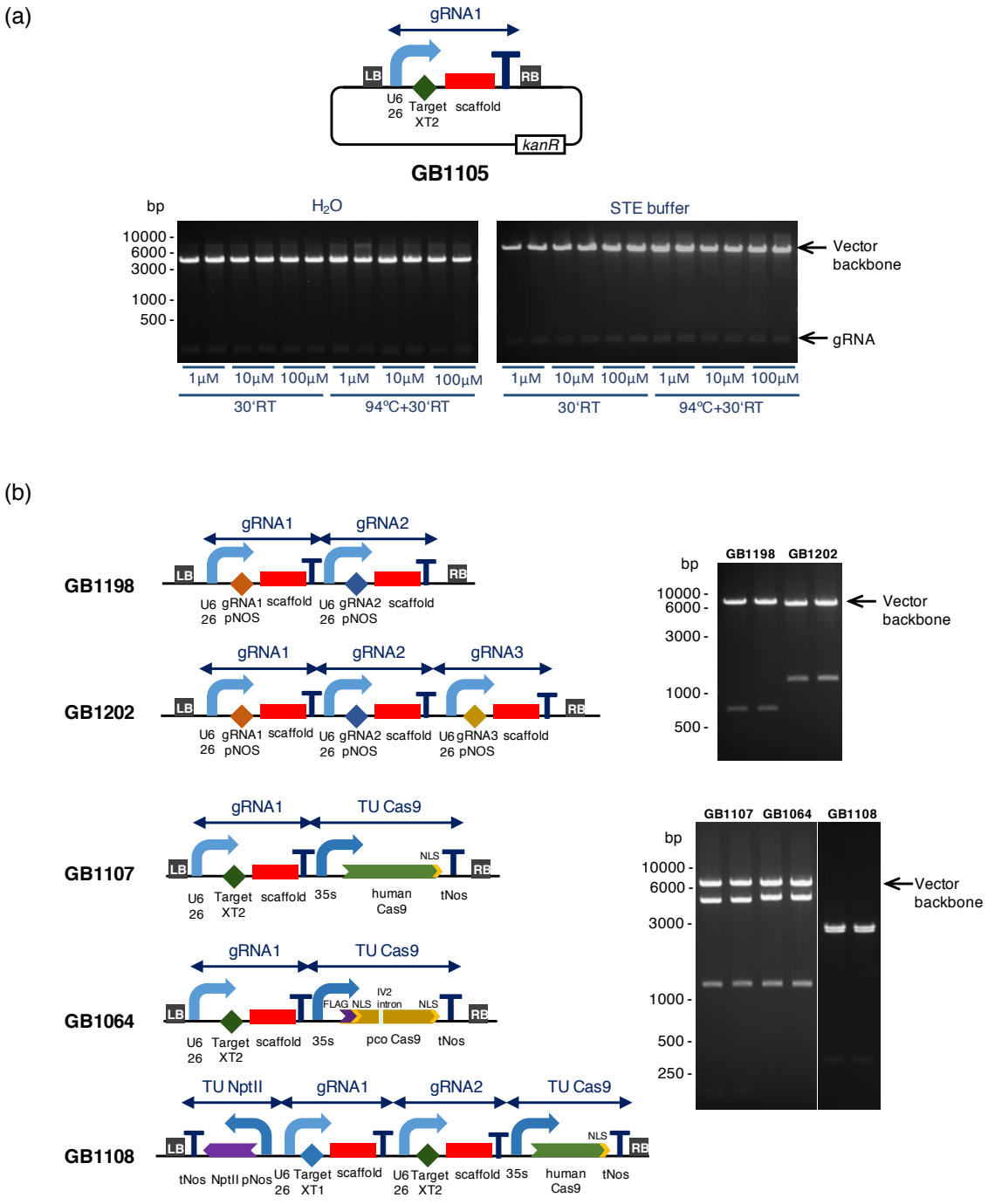
C6. An intragenic tomato selection marker based on a mutant Acetolactate Synthase gene was designed, built and assembled to an intragenic construct directing the ripening-dependant activation of flavonol biosynthesis in fruit following GB3.0 standard procedures. The resulting all-tomato intragenic T1 plant lines showed phenotypically distinctive tomatoes that, according to preliminary analysis, accumulate high levels of anti-oxidant flavonols, illustrating the application of GB3.0 to advanced Plant Breeding.

9 | Supplementary Figures



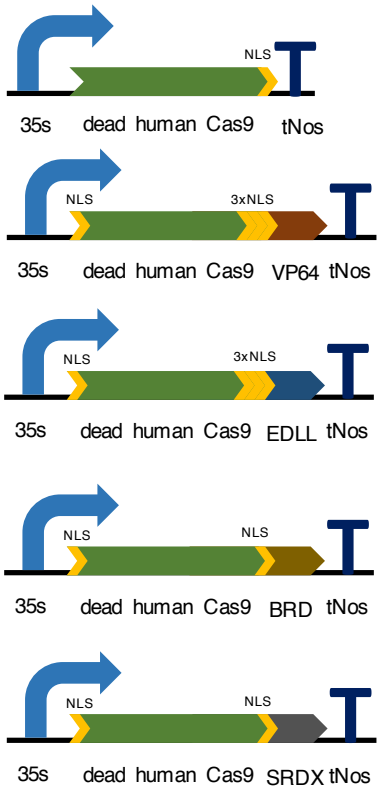
Supplementary Figure 1. GB3.0 domestication tools workflow.

The Phylogenetic Search Tool offers sequence retrieval of the best hits with minimal internal restriction sites from a Blast search in a coding sequences database. Upon sequence election, it offers conection with the GBDomesticator. The GB Domesticator provides the list of the PCR products and GB oligos required for part domestication while the Synthetic Strategy Tool provides directly the synthetic product to be ordered. The GB Domesticator and the Synthetic Strategy Tool offer in this new version the possibility of domesticating intron containing genomic sequences when the introns are written in lowercase in the input file. They also offer the choice to domesticate the sequences not only for BsaI and BsmBI but also for BpiI and/or BtgZI. All domestication tools generate a GenBank file with the level 0 part cloned into the pUPD2.



Supplementary Figure 2. Cloning efficiency of representative Level≥1 GBelements.

(a) Restriction analysis of 24 plasmids obtained from randomly selected colonies to test the accuracy of gRNA assemblies under different conditions (HindIII expected bands: 6345-374). (b) Restriction analysis of GB binary reactions combining gRNAs with each other and/or with the Cas9-encoding TU. All tested plasmids resulted in the expected restriction patterns. GB1198 BamHI expected bands: (6674-728); GB1202 EcoRI expected bands: (6345-1242); GB1107 EcoRV expected bands: (6674-4620-1236-194); GB1064 EcoRV expected bands: (6674-5117-1236); GB1108 EcoRI expected bands: (6987-6345-828).



Supplementary Figure 3. Schema of the architecture of the dead Cas9 transcriptional units tested on the repression and activation experiments.

10 | Supplementary Tables

Supplementary Table 1. Experimental conditions.

Experiment type	Mandatory experimental conditions	Recommended experimental conditions	Quantitative output
SE_001 Promoter strength tested in discs	- Plant species: *N.benthamiana*. - Chassis: Agroinfiltrated leaves of 5-6 weeks old *N.benthamiana* plants. - Constructs: tested constructs must include the Luc gene as reporter (GB0096) and the 35s:Renilla:Tnos transcriptional unit in cis. The silencing supressor P19 must be coinfiltrated either in cis or in trans. - Sampling: collect discs (0.8 cm in dimater) at 3 days post infiltration. - Analysis: keep the discs on plates with/without chemical inductor and take samples at different times. - Data normalization: normalize the data with the values obtained for GB1116 in the same experimental conditions.	- Plant growth conditions: 24ºC/20ºC 16h light / 8h darkness. - Sampling: collect replicas from 3 independent agroinfiltrated leaves. - Analysis: use the Dual-Glo© Luciferase Assay System (Promega) for the luminiscence assay.	- RTA at 0h. - RTA at 4h. - RTA at 8h. - RTA at 12h. - RTA at 18h. - RTA at 24h. - RTA at 36h. - RTA at 48h. All relative transcriptional activities (RTA) have to be expressed in rpu (relative promoter units to GB1116).
SE_002 Promoter strength tested in leaves	- Plant species: *N.benthamiana*. - Chassis: Agroinfiltrated leaves of 5-6 weeks old *N.benthamiana* plants. - Constructs: tested constructs must include the Luc gene as reporter (GB0096) and the 35s:Renilla:Tnos transcriptional unit in cis. The silencing supressor P19 must be coinfiltrated either in cis or in trans. - Sampling: collect discs (0.8 cm in dimater) at 4 days post infiltration for analysis. - Data normalization: normalize the data with the values obtained for GB1116 in the same experimental conditions.	- Plant growth conditions: 24ºC/20ºC 16h light / 8h darkness. - Sampling: collect replicas from 3 independent agroinfiltrated leaves. - Analysis: use the Dual-Glo© Luciferase Assay System (Promega) for the luminiscence assay.	- RTA (in rpu) Relative transcriptional activity (RTA) has to be expressed in rpu (relative promoter units to GB1116).
SE_003 Transformation efficiency	- Analysis: calculate the transformation efficiency by dividing the number of obtained transgenic plants by the number of inoculated explants.	Not defined	- % transformants
SE_004 Recombinant protein production	- Plant species: *N.benthamiana*. - Chassis: Agroinfiltrated leaves of 6-7 weeks old *N.benthamiana* plants. - Sampling and analysis: extract your protein, purify it and determine the amount of recombinant protein relative to the total amount of protein or weight of plant tissue.	- Plant growth conditions: 24ºC/20ºC 16h light / 8h darkness.	- µg/gDW (dry weight) - µg/gFW (fresh weight) - %TSP (total soluble protein)
SE_005 CRISPR target efficiency	- Constructs: tested constructs must include a Cas gene together with at least one single guideRNA (sgRNA). - Analysis: determine the efficiency calculating the percentage of mutated genomic DNA (for transient expression) or the number of plants with mutations in reference to the total number of transgenic plants (for stable transformation).	Not defined	- Overall mutations efficiency (%)
NS_000 Non-standard experiment	Not defined	Not defined	Not defined

Supplementary Table 2. Non-exhaustive list of GBexperiments.

*All listed experiments can be consulted at https://gbcloning.upv.es/search/experiment/ by introducing the 'Experiment IDs'. Extra experiments can be searched on the same link by using different search criteria.

Title	Short description	Experiment type	GB elements	Experiment IDs
Dexamethasone dose-response	Study of the inducibility factor of a genetic element inducible by dexamethasone at different dexamethasone concentrations.	SE_001	GB0162	GB_EXP_3F, GB_EXP_40, GB_EXP_41, GB_EXP_42, GB_EXP_43, GB_EXP_45, GB_EXP_47, GB_EXP_49, GB_EXP_4A, GB_EXP_4B
Dexamethasone time-course	Determination of the induction profile of a genetic element inducible by dexamethasone.	SE_001	GB1254	GB_EXP_7A, GB_EXP_7B
Estradiol dose-response	Study of the inducibility factor of a genetic element inducible by estradiol with different β-estradiol concentrations	SE_001	GB1132	GB_EXP_36, GB_EXP_37, GB_EXP_38, GB_EXP_3A, GB_EXP_3B, GB_EXP_3C, GB_EXP_3D, GB_EXP_3E
Transactivation with a synthetic transcription factor time-course	Determination of the levels of activation of a regulated promoter with either the constitutive or self-regulated expression of a synthetic transcription factor	SE_001	GB1118, GB1121, GB1122, GB1124	GB_EXP_1A, GB_EXP_17, GB_EXP_18, GB_EXP_19
Regulated transactivation of the SlDFR promoter time-course	Determination of the transcriptional activity induced by the *Solanum lycopersicum* DFR promoter when it is coexpressed with a MYB and a bHLH transcription factors (Rosea1 and Delila) either constitutively expressed or regulated.	SE_001	GB1160, GB0129, GB1156, GB1157	GB_EXP_87, GB_EXP_90, GB_EXP_8C, GB_EXP_8D, GB_EXP_8E, GB_EXP_8F
Expression of the 35s constitutive promoter	Determination of the transcriptional activity induced by the 35s promoter over different experiments to test the stability of its expression and the reproducibility of the employed experimental method	SE_001	GB0164	GB_EXP_2D, GB_EXP_2E, GB_EXP_2F, GB_EXP_33, GB_EXP_26
Expression of the 35s constitutive promoter	Determination of the transcriptional activity induced by the 35s promoter over different experiments to test the stability of its expression and the reproducibility of the employed experimental method	SE_001	GB1119	GB_EXP_34, GB_EXP_35, GB_EXP_16
Transactivation induced by the TEV protease	Test of the ability of the TEV protease constitutively expressed to reléase a synthetic transcription factor from a transmembrane protein.	SE_002	GB0588 GB0594	GB_EXP_24 GB_EXP_25
Protein-protein interaction determined with the split-TEV system	Test of the interaction of two proteins fused to the N-term and C-term domains of the TEV protease by measuring expression from a promoter that is activated by a synthetic transcription factor released from a transmembrane protein upon reconstitution of the TEV protease.	SE_002	GB0592 GB0593	GB_EXP_22, GB_EXP_23

Supplementary Table 2 (continuation). Non-exhaustive list of GBexperiments.

Title	Short description	Experiment type	GB elements	Experiment IDs
Transactivation of two DFR promoters with plant transcription factors	Determination of the transcriptional activity induced by the *Solanum lycopersicum* and the *Antirrhinum majus* DFR promoters when they are coexpressed with two MYB transcription factors (Rosea1 and Ant1) either alone or in combination with two bHLH transcription factors (Delila and Jaf13).	SE_002	GB1160, GB1161, GB0125, GB0126, GB0127, GB0128, GB0129, GB0130	GB_EXP_4D, GB_EXP_51, GB_EXP_54, GB_EXP_55, GB_EXP_57, GB_EXP_58, GB_EXP_59, GB_EXP_5B, GB_EXP_5C, GB_EXP_5D, GB_EXP_5E, GB_EXP_60, GB_EXP_61, GB_EXP_62
Transcriptional activation using the CRISPR/Cas9 technology	Comparison of the transcriptional activation of the nopaline synthase promoter by targeting to it the dCas9 fused to the EDLL or to the VP64 activation domains with different sgRNAs either alone or combined.	SE_002	GB1116, GB1189, GB1190, GB1221, GB1192, GB1197, GB1195, GB1220	GB_EXP_A3, GB_EXP_A5, GB_EXP_A6, GB_EXP_A7, GB_EXP_A8, GB_EXP_A9, GB_EXP_AB, GB_EXP_AF, GB_EXP_B0, GB_EXP_B1
Transcriptional repression using the CRISPR/Cas9 technology	Transcriptional repression of the nopaline synthase promoter by targeting to it the dCas9 fused to the BRD or to the SRDX repressor domains with with different sgRNAs either alone or combined.	SE_002	GB1116, GB1188, GB1172, GB1221, GB1192, GB1197, GB1195, GB1220	GB_EXP_B2, GB_EXP_B3, GB_EXP_B4, GB_EXP_B5, GB_EXP_B6, GB_EXP_B8, GB_EXP_B9, GB_EXP_BA, GB_EXP_BB, GB_EXP_BC
Tomato transformation with a intragenic selection marker	Determination of the transformation efficiency obtained using a mutated version of the tomato acetolactate synthase as selection marker.	SE_003	GB0830	GB_EXP_BD
Recombinant antibody production	Comparison of the expression levels of three monoclonal antibody formats against the human tumor necrosis factor.	SE_004	GB_UA_BD1, GB_UA_C27, GB_UA_C29	GB_EXP_95, GB_EXP_98, GB_EXP_99
Gene editing with the CRISPR/Cas9 technology	Mutagenesis efficiency of the Cas9 in combination with two sgRNAs targeting each of them a different locus of the *N.benthamiana* xylosyltransferase gene.	SE_005	GB0639, GB1108	GB_EXP_83, GB_EXP_85, GB_EXP_86
Anthocyanins production	Quantification of anthocyanins produced in transient expression by the expression of two MYB transcription factors (Rosea1 and Ant1) either alone or in combination with two bHLH transcription factors (Delila and Jaf13).	NS_000	GB0125, GB0126, GB0127, GB0128, GB0129, GB0130	GB_EXP_BF, GB_EXP_C0, GB_EXP_C1, GB_EXP_C2, GB_EXP_C3, GB_EXP_C4

Supplementary Table 3. List of GBelements tested under experimental conditions described in Chapter 2.

*Datasheets of all listed GBelements can be consulted at https://gbcloning.upv.es/search/features/ by introducing their GB IDs.

Protein-protein interaction	
GB0592	Device for testing the interaction of Ros1 and SOC1 using the Split TEV system
GB0593	Device for testing the interaction of FUL and SOC1 using the Split TEV system
Reporter devices for transcriptional regulation studies	
GB1160	Reporter device including a transcriptional unit for the expression of the Luciferase gene driven by the *Solanum lycopersicum* DFR promoter
GB1161	Reporter device including a transcriptional unit for the expression of the Luciferase gene driven by the *Antirrhinum majus* DFR promoter
GB0178	Reporter device including a transcriptional unit for the expression of the Luciferase gene driven by a synthetic promoter including the LacI operon and the minimal 35s promoter
GB1130	Transcriptional unit for the expression of the Luciferase gene driven by a synthetic promoter including the LexA operon and the minimal 35s promoter
GB1116	Reporter device including a transcriptional unit for the expression of the Luciferase gene driven by the nopaline synthase promoter
Constitutive transcriptional regulation	
GB1120	Transcriptional unit for the constitutive expression of a synthetic transcription factor conformed by the LacI DNA binding domain and the Gal4 activation domain
GB0129	Module for the constitutive expression of the MYB transcription factor Rosea1 and the bHLH transcription factor Delila
GB0130	Module for the constitutive expression of the MYB transcription factor Ant1 and the bHLH transcription factor Jaf13
GB1189	Transcriptional unit for the constitutive expression of the dCas9 fused to the VP64 activation domain. In combination with any module expressing one or more sgRNA targeting a promoter (i.e. GB1202) can activate expression from it.
GB1188	Transcriptional unit for the constitutive expression of the dCas9 fused to the SRDX repressor domain. In combination with any module expressing one or more sgRNA targeting a promoter (i.e. GB1202) can repress expression from it.
Conditional transcriptional regulation	
GB1111	Transcriptional unit for the regulated expression of a synthetic transcription factor conformed by the LacI DNA binding domain and the Gal4 activation domain
GB1156	Module for the conditional expression of the MYB transcription factor Rosea1 and the constitutive expression of the bHLH transcription factor Delila including a synthetic transcription factor responsive to dexamethasone constitutively expressed.
GB1157	Module for the conditional expression of the bHLH transcription factor Delila and the constitutive expression of the MYB transcription factor Rosea1 including a synthetic transcription factor responsive to dexamethasone constitutively expressed.
GB0157	Transcriptional unit for the constitutive expression of a synthetic transcription factor conformed by the GR glucocorticoid receptor domain fused to the LacI DNA binding domain and the Gal4 activation domain
GB1129	Transcriptional unit for the constitutive expression of a synthetic transcription factor conformed by the ER estradiol receptor domain fused to the LexA DNA binding domain and the Gal4 activation domain

Metabolic engineering	
GB0830	Device for tomato transformation expressing the *S.lycopersicum* MYB12 under the E8 fruit promoter. MYB12 is a master regulator of the flavonoids biosynthetic pathway.
GB0130	Device for overproduction of anthocyanins in transient expression in *N.benthamiana* leaves comprising the *S.lycopersicum* MYB and bHLH transcription factors Ant1 and Jaf13.
CRISPR/Cas9 based gene editing	
GB1108	Device including two monocistroninc sgRNAs targeting each of them one locus of the N.benthamiana xylosyltransferase gene and the constitutively expressed Cas9.
GB1222	Device including two polycistroninc sgRNAs, one with two targets for the two genes of the N.benthamiana xylosyltransferase and the second one with three sgRNAs targeting five N.benthamiana fucosyltransferase genes. It also includes the TU for the constitutive expression of the Cas9.
Recombinant antibody production	
GB_UA_BD1	Transcriptional unit for the constitutive expression of the human scFv-Fcgamma1 antibody format against the human TNF-alpha.
GB_UA_C27	Transcriptional unit for the constitutive expression of a monoclonal antibody with the gamma1 heavy chain and the lambda light chain against the human TNF-alpha.
GB_UA_C29	Transcriptional unit for the constitutive expression of a monoclonal antibody with the gamma1 heavy chain and the kappa light chain against the human TNF-alpha.

Supplementary Table 4. Frequencies of plants cells untransformed and transformed with two different T-DNA upon agroinfiltration.

OD	Frequency of untransformed protoplasts	Frequency of cotransformed protoplasts	Calculated MOT
0,0002058	0,887	0,006	0,027
0,000617	0,598	0,065	0,265
0,00185	0,281	0,213	0,890
0,0055	0,053	0,652	3,190
0,0167	0,033	0,675	3,360
0,05	0,004	0,834	4,950
0,1	0,000	0,945	7,400

Supplementary Table 5. Primers used for the amplification of the N.benthamiana xylosyltransferases XT1 (Niben101Scf04205Ctg025) and XT2 (Niben101Scf04551Ctg021) regions.

XT1_F	5´-AACCACTTTTCCTCGTCGGAAA-3'
XT1_R	5´-TAACTATTCAACTAAAGCTTCAAACAG-3'
XT2_F	5'-AACCACTTTTCCTTGTCGGAAA-3'
XT2_R	5'-GGAATGAAATTAACCACTTCAGG-3'
XT12BsaI_F	5'-GCGGGTCTCAGGAGCCCTAATGTTGCTTGGAGATC-3'
XT12BsaI_R	5´-GCGGGTCTCAAGCGCCGTCTAAGGTTCAATTTGAGTAGC-3'

Supplementary Table 6. List of forward and reverse primers used to construct the targets.

gRNA XT1_F	5´-ATTGAAAACACCGTCTTCGGAGA-3’
gRNA XT1_R	5´-AAACTCTCCGAAGACGGTGTTTT-3’
gRNA XT2_F	5´-ATTGAAAATTGGGAAAAAACTAG-3’
gRNA XT2_R	5´-AAACCTAGTTTTTTCCCAATTTT-3’
gRNA1 pNOS_F	5’-ATTGAGACTCTAATTGGATACCG-3’
gRNA1 pNOS_R	5’-AAACCGGTATCCAATTAGAGTCT-3’
gRNA2 pNOS_F	5’-ATTGACGTTCCATAAATTCCCCT-3’
gRNA2 pNOS_R	5’-AAACAGGGGAATTTATGGAACGT-3’
gRNA3 pNOS_F	5’-ATTGACTTTTGAACGCGCAATAA-3’
gRNA3 pNOS_R	5’-AAACTTATTGCGCGTTCAAAAGT-3’
gRNA4 pNOS_F	5’-ATTGCCACTGAGCCGCGGGTTTC-3’
gRNA4 pNOS_R	5’-AAACGAAACCCGCGGCTCAGTGG-3’
gRNA5 pNOS_F	5’-ATTGGGACAAGCCGTTTTACGTT-3’
gRNA5 pNOS_R	5’-AAACAACGTAAAACGGCTTGTCC-3’
gRNA FT5447tRNA_F*	5’-GTGCACCCAAAAGAAATGGTCCAAT-3’
gRNA FT5447tRNA_R*	5’-AAACATTGGACCATTTCTTTTGGGT-3’
gRNA FT1272tRNA_F*	5’-GTGCACCAATAAGCAATGGCGCAAT-3’
gRNA FT1272tRNA_R*	5’-AAACATTGCGCCATTGCTTATTGGt-3’
gRNA FT2631tRNA_F*	5’-GTGCACCGATAAACAATGGCGCAAT-3’
gRNA FT2631tRNA_R*	5’-AAACATTGCGCCATTGTTTATCGGT-3’
gRNA XT4205tRNA_F*	5’-GTGCAGAAAACACCGTCTTCGGAGA-3’
gRNA XT4205tRNA_R*	5’-AAACTCTCCGAAGACGGTGTTTTCT-3’
gRNA XT4551tRNA_F*	5’-GTGCAGAAAATTGGGAAAAAACTAG-3’
gRNA XT4551tRNA_R*	5’-AAACCTAGTTTTTTCCCAATTTTCT-3’

*Primers used on the assembly of GBoligomers on level 0.

Supplementary Table 7. List of GBelements generated in Chapter 3.

Level -1 GBparts		
GBdatabase ID	**Name**	
GB1205	tRNA-gRNA position [D1_2]	
GB1206	tRNA-gRNA position [2_n-1]	
GB1207	tRNA-gRNA position [n]	
GB1208	tRNA-gRNA position [D1_n-1]	
GB1209	tRNA-gRNA position [M1_2]	
GB1210	tRNA-gRNA position [M1_n-1]	
Level 0 GBparts		
GBdatabase ID	**Name**	**Category**
GB0273	ppcoCas9	B3-B4-B5
GB0575	phCas9	B3-B4-B5
GB0645	psgRNA	B6b-C1
GB1001	pAtU6-26	A1-A2-A3-B1-B2c
GB1041	pdCas9	B3-B4-B5
GB1079	pdCas9	B3-B4
GB1175	pNLS-BRD	B5
GB1184	pOsU3	A1-A2-A3-B1-B2d
GB1185	pNLS-SRDX	B5
GB1186	p3xNLS-VP64	B5
GB1187	p3xNLS-EDLL	B5
GB1204	pAtU6-1	A1-A2-A3-B1-B2c
GB1211	ptRNA-target5447_5494_17626FT-gRNA[D1_2]	Other
GB1212	ptRNA-target1272FT-gRNA[2_n-1]	Other
GB1213	ptRNA-target2631FT-gRNA[n]	Other
GB1214	ptRNA-target4205XT-gRNA[D1_n-1]	Other
GB1215	ptRNA-target4551XT-gRNA[n]	Other
Level 1 GB TUs		
GBdatabase ID	**Name**	**% accuracy***
GB0639	pEGB2α2 35s:hCas9:tNOS	100%
GB0576	pEGB3α2 35s:pcoCas9:tNOS	50%
GB1104	pEGB3α1 U626:gRNAXT4551:sgRNA	100%
GB1105	pEGB3α2 U626:gRNAXT4205:sgRNA	100%

GB1172	pEGB3α2 35s:hCas9:BRD:tNOS	100%
GB1177	pEGB3α1 U626:gRNA1pNOS:sgRNA	100%
GB1178	pEGB3α2 U626:gRNA1pNOS:sgRNA	100%
GB1179	pEGB3α1 U626:gRNA2pNOS:sgRNA	100%
GB1180	pEGB3α1 U626:gRNA4pNOS:sgRNA	100%
GB1188	pEGB3α2 35s:hCas9:SRDX:tNOS	100%
GB1189	pEGB3α2 35s:hCas9:VP64:tNOS	100%
GB1190	pEGB3α2 35s:hCas9:EDLL:tNOS	100%
GB1191	pEGB3α2 35s:dCas9:tNOS	100%
GB1192	pEGB3α1 U626:gRNA3pNOS:sgRNA	100%
GB1193	pEGB3α1 U626:gRNA5pNOS:sgRNA	100%
GB1194	pEGB3α2 U626:gRNA3pNOS:sgRNA	100%
GB1195	pEGB3α2 U626:gRNA5pNOS:sgRNA	100%
GB1196	pEGB3α2 U626:gRNA2pNOS:sgRNA	75%
GB1197	pEGB3α2 U626:gRNA4pNOS:sgRNA	100%
GB1221	pEGB3α1 U626:gRNA1pDFR:sgRNA	100%
Level >1 GB Modules		
GBdatabase ID	**Name**	
GB1064	pEGB3Ω2 U6-26:target4551XT:sgRNA-35s:pcoCas9:tNOS	75%
GB1106	pEGB3Ω1 tNOS:nptII:pNOS-U626:target4205XT:sgRNA	100%
GB1107	pEGB3Ω2 U6-26:target4551XT:sgRNA-35s:hCas9:tNOS	100%
GB1108	pEGB 3α1 tNOS:nptII:pNOS-U626:target4205XT:sgRNA-U626:target4551XT:sgRNA-35s:hCas9:tNOS	100%
GB1116	pEGB3α1 pNOS:Luciferase:tNOS-SF-35S:Renilla:tNOS-35S:P19:tNOS-SF	100%
GB1198	pEGB3Ω1 U626:gRNA1pNOS:sgRNA-U626:gRNA2pNOS:sgRNA	100%
GB1199	pEGB3Ω1 U626:gRNA1pNOS:sgRNA-U626:gRNA4pNOS:sgRNA	100%
GB1200	pEGB3Ω1 U626:gRNA2pNOS:sgRNA-U626:gRNA4pNOS:sgRNA	100%
GB1201	pEGB3Ω2 U626:gRNA4pNOS:sgRNA-SF	100%
GB1202	pEGB3α1 U626:gRNA1pNOS:sgRNA-U626:gRNA2pNOS:sgRNA-U626:gRNA4pNOS:sgRNA	100%
GB1216	pEGB3α1 U626:tRNA-target1FT-gRNA:tRNA-target2FT:gRNA:tRNA-	100%

	target3FT-gRNA	
GB1217	pEGB3α2 U626:tRNA-target1XT-gRNA:tRNA-target2XT:gRNA	100%
GB1218	pEGB3Ω1 U626:tRNA-gRNA(x3withFTtargets)-U626:tRNA-gRNA(x2withXTtargets)	100%
GB1219	pEGB3Ω1 U626:gRNA3pNOS:sgRNA-U626:gRNA5pNOS:sgRNA	100%
GB1220	pEGB3α1 U626:gRNA3pNOS:sgRNA-U626:gRNA5pNOS:sgRNA-U626:gRNA4pNOS:sgRNA	100%
GB1222	pEGB3α1 U626:tRNA-gRNA(x3withFTtargets)-U626:tRNA-gRNA(x2withXTtargets)-35s:hCas9:tNOS	100%

* Cloning accuracy expressed as the % of colonies showing the correct restriction pattern (number of colonies assayed 2-4).

Supplementary Table 8. Primers used on the domestication of GB0816 (mALS), GB0914 (E8 promoter) and GB0144 (E8 terminator).

M13MAY01ALS1aOF	GCGCCGTCTCACTCGAATGGCGGCTGCTGCCTCACC
M13MAY02ALS1aOR	GCGCCGTCTCATGACACTTGACCTGTAATAGCAACAATCGG
M13MAY03ALS1bOF	GCGCCGTCTCAGTCAAGGAGGATGATTGGTAC
M13MAY04ALS1bOR	GCGCCGTCTCAGCCTCAGCTCCTCACTTGATTG
M13MAY05ALS2OF	GCGCCGTCTCAAGGCGATTTGTGGAGCTTACAGG
M13MAY06ALS2OR	GCGCCGTCTCACCTCCCAACAGCCGCACCTAT
M13MAY07ALS3OF	GCGCCGTCTCAGAGGCCGGGTGAGATTGTGG
M13MAY08ALS3OR	GCGCCGTCTCACTCGAAGCTCAATAGGAACATCTCCCGTCGCC
M13OCT01_PE8F1	GCGCCGTCTCACTCGGGAGTCCCTAATGATATTGTTCATG
M13OCT02_PE8R1	GCGCCGTCTCACTCGCATTCTTCTTTTGCACTGTGAATGATTAG
M12MAY03TermE8F1	GCGCCGTCTCGCTCGGCTTGAATAAGAATAATAATG
M12MAY04TermE8R1	GCGCCGTCTCGCTCGAGCGCGTAAATTAGATAAGGAAAAC
M12ENE24TmtbR1	GCGCCGTCTCGCTCGAGCGTCGCAAAAACTATATGCTCTC

* Primers M13MAY07ALS3OF and M12ENE24TmtbR1 were used for genotyping transformed plants.

11 | References

Adato A, Mandel T, Mintz-Oron S*, et al.*, 2009. Fruit-surface flavonoid accumulation in tomato is controlled by a SlMYB12-regulated transcriptional network. *PLoS Genet* **5**, e1000777.

Agrawal PK, Kohli A, Twyman RM, Christou P, 2005. Transformation of Plants with Multiple Cassettes Generates Simple Transgene Integration Patterns and High Expression Levels. *Molecular Breeding* **16**, 247-60.

Ali Z, Abulfaraj A, Idris A, Ali S, Tashkandi M, Mahfouz MM, 2015. CRISPR/Cas9-mediated viral interference in plants. *Genome Biol* **16**, 238.

Alonso JM, Stepanova AN, Leisse TJ*, et al.*, 2003. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* **301**, 653-7.

Altaras NE, Cameron DC, 1999. Metabolic engineering of a 1,2-propanediol pathway in Escherichia coli. *Appl Environ Microbiol* **65**, 1180-5.

Andrianantoandro E, Basu S, Karig DK, Weiss R, 2006. Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* **2**, 2006 0028.

Angov E, 2011. Codon usage: nature's roadmap to expression and folding of proteins. *Biotechnol J* **6**, 650-9.

Antunes MS, Ha SB, Tewari-Singh N*, et al.*, 2006. A synthetic de-greening gene circuit provides a reporting system that is remotely detectable and has a re-set capacity. *Plant Biotechnol J* **4**, 605-22.

Antunes MS, Morey KJ, Smith JJ*, et al.*, 2011. Programmable ligand detection system in plants through a synthetic signal transduction pathway. *PLoS One* **6**, e16292.

Antunes MS, Morey KJ, Tewari-Singh N*, et al.*, 2009. Engineering key components in a synthetic eukaryotic signal transduction pathway. *Mol Syst Biol* **5**, 270.

Arkin A, 2008. Setting the standard in synthetic biology. *Nat Biotechnol* **26**, 771-4.

Azuma M, Morimoto R, Hirose M*, et al.*, 2016. A petal-specific InMYB1 promoter from Japanese morning glory: a useful tool for molecular breeding of floricultural crops. *Plant Biotechnol J* **14**, 354-63.

Ballester AR, Molthoff J, De Vos R*, et al.*, 2010. Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit color. *Plant Physiol* **152**, 71-84.

Baltes NJ, Hummel AW, Konecna E*, et al.*, 2015. Conferring resistance to geminiviruses with the CRISPR–Cas prokaryotic immune system. *Nature Plants* **1**, 15145.

Bartley B, Beal J, Clancy K*, et al.*, 2015. Synthetic Biology Open Language (SBOL) Version 2.0.0. *J Integr Bioinform* **12**, 272.

Bedoya LC, Martinez F, Orzaez D, Daros JA, 2012. Visual tracking of plant virus infection and movement using a reporter MYB transcription factor that activates anthocyanin biosynthesis. *Plant Physiol* **158**, 1130-8.

Beerli RR, Barbas CF, 3rd, 2002. Engineering polydactyl zinc-finger transcription factors. *Nat Biotechnol* **20**, 135-41.

Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V, 2013. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* **9**, 39.

Belhaj K, Chaparro-Garcia A, Kamoun S, Patron NJ, Nekrasov V, 2015. Editing plant genomes with CRISPR/Cas9. *Curr Opin Biotechnol* **32**, 76-84.

Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA, 2013. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res* **41**, 7429-37.

Binder S, 2010. Branched-Chain Amino Acid Metabolism in Arabidopsis thaliana. *Arabidopsis Book* **8**, e0137.

Blake WJ, Chapman BA, Zindal A, Lee ME, Lippow SM, Baynes BM, 2010. Pairwise selection assembly for sequence-independent construction of long-length DNA. *Nucleic Acids Res* **38**, 2594-602.

Bogdanove AJ, Voytas DF, 2011. TAL effectors: customizable proteins for DNA targeting. *Science* **333**, 1843-6.

Bortesi L, Fischer R, 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol Adv* **33**, 41-52.

Bovy A, De Vos R, Kemper M*, et al.*, 2002. High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes LC and C1. *Plant Cell* **14**, 2509-26.

Bovy A, Schijlen E, Hall RD, 2007. Metabolic engineering of flavonoids in tomato (): the potential for metabolomics. *Metabolomics* **3**, 399-412.

Boyle PM, Burrill DR, Inniss MC*, et al.*, 2012. A BioBrick compatible strategy for genetic modification of plants. *J Biol Eng* **6**, 8.

Brophy JA, Voigt CA, 2014. Principles of genetic circuit design. *Nat Methods* **11**, 508-20.

Butelli E, Titta L, Giorgio M*, et al.*, 2008. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat Biotechnol* **26**, 1301-8.

Cai Y, Wilson ML, Peccoud J, 2010. GenoCAD for iGEM: a grammatical approach to the design of standard-compliant constructs. *Nucleic Acids Res* **38**, 2637-44.

Callaway E, 2013. Glowing plants spark debate. *Nature* **498**, 15-6.

Cambray G, Guimaraes JC, Mutalik VK*, et al.*, 2013. Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res* **41**, 5139-48.

Cambray G, Mutalik VK, Arkin AP, 2011. Toward rational design of bacterial genomes. *Curr Opin Microbiol* **14**, 624-30.

Canton B, Labno A, Endy D, 2008. Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol* **26**, 787-93.

Casini A, Storch M, Baldwin GS, Ellis T, 2015. Bricks and blueprints: methods and standards for DNA assembly. *Nat Rev Mol Cell Biol* **16**, 568-76.

Castilho A, Neumann L, Daskalova S*, et al.*, 2012. Engineering of sialylated mucin-type O-glycosylation in plants. *J Biol Chem* **287**, 36518-26.

Cha-Aim K, Fukunaga T, Hoshida H, Akada R, 2009. Reliable fusion PCR mediated by GC-rich overlap sequences. *Gene* **434**, 43-9.

Chang MC, Keasling JD, 2006. Production of isoprenoid pharmaceuticals by engineered microbes. *Nat Chem Biol* **2**, 674-81.

Chao R, Yuan Y, Zhao H, 2014. Recent advances in DNA assembly technologies. *FEMS Yeast Res*.

Cook C, Martin L, Bastow R, 2014. Opportunities in plant synthetic biology. *J Exp Bot* **65**, 1921-6.

Dafny-Yelin M, Tzfira T, 2007. Delivery of multiple transgenes to plant cells. *Plant Physiol* **145**, 1118-28.

Dalal J, Yalamanchili R, La Hovary C*, et al.*, 2015. A novel gateway-compatible binary vector series (PC-GW) for flexible cloning of multiple genes for genetic transformation of plants. *Plasmid* **81**, 55-62.

Darbani B, Eimanifar A, Stewart CN, Jr., Camargo WN, 2007. Methods to produce marker-free transgenic plants. *Biotechnol J* **2**, 83-90.

Davey MR, Anthony P, Power JB, Lowe KC, 2005. Plant protoplasts: status and biotechnological perspectives. *Biotechnol Adv* **23**, 131-71.

Davuluri GR, Van Tuinen A, Fraser PD*, et al.*, 2005. Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. *Nature Biotechnology* **23**, 890-5.

De Buck S, De Wilde C, Van Montagu M, Depicker A, 2000. Determination of the T-DNA transfer and the T-DNA integration frequencies upon cocultivation of Arabidopsis thaliana root explants. *Mol Plant Microbe Interact* **13**, 658-65.

De Jong WS, Eannetta NT, De Jong DM, Bodis M, 2004. Candidate gene analysis of anthocyanin pigmentation loci in the Solanaceae. *Theor Appl Genet* **108**, 423-32.

Depaoli HC, Borland AM, Tuskan GA, Cushman JC, Yang X, 2014. Synthetic biology as it relates to CAM photosynthesis: challenges and opportunities. *J Exp Bot* **65**, 3381-93.

Duportet X, Wroblewska L, Guye P*, et al.*, 2014. A platform for rapid prototyping of synthetic gene networks in mammalian cells. *Nucleic Acids Res* **42**, 13440-51.

Eeckhaut T, Lakshmanan PS, Deryckere D, Van Bockstaele E, Van Huylenbroeck J, 2013. Progress in plant protoplast research. *Planta*.

Ellul P, Garcia-Sogo B, Pineda B, Rios G, Roig LA, Moreno V, 2003. The ploidy level of transgenic plants in Agrobacterium-mediated transformation of tomato cotyledons (Lycopersicon esculentum Mill.) is genotype and procedure dependent [corrected]. *Theor Appl Genet* **106**, 231-8.

Engler C, Gruetzner R, Kandzia R, Marillonnet S, 2009. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* **4**, e5553.
Engler C, Kandzia R, Marillonnet S, 2008. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647.

Engler C, Youles M, Gruetzner R*, et al.*, 2014. A golden gate modular cloning toolbox for plants. *ACS Synth Biol* **3**, 839-43.

Fauser F, Schiml S, Puchta H, 2014. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. *Plant J* **79**, 348-59.

Fu F, Sander JD, Maeder M*, et al.*, 2009. Zinc Finger Database (ZiFDB): a repository for information on C2H2 zinc fingers and engineered zinc-finger arrays. *Nucleic Acids Res* **37**, D279-83.

Gaj T, Gersbach CA, Barbas CF, 3rd, 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**, 397-405.

Galdzicki M, Clancy KP, Oberortner E*, et al.*, 2014. The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nat Biotechnol* **32**, 545-50.

Gao J, Wang G, Ma S*, et al.*, 2015. CRISPR/Cas9-mediated targeted mutagenesis in Nicotiana tabacum. *Plant Mol Biol* **87**, 99-110.

Gardner TS, Hawkins K, 2013. Synthetic biology: evolution or revolution? A co-founder's perspective. *Curr Opin Chem Biol* **17**, 871-7.

Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA, 2007. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res* **35**, e55.

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO, 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-5.

Giritch A, Ganal M, Stephan UW, Baumlein H, 1998. Structure, expression and chromosomal localisation of the metallothionein-like gene family of tomato. *Plant Mol Biol* **37**, 701-14.

Gorbunova VV, Levy AA, 1999. How plants make ends meet: DNA double-strand break repair. *Trends Plant Sci* **4**, 263-9.

Goulet C, Khalf M, Sainsbury F, D'aoust MA, Michaud D, 2012. A protease activity-depleted environment for heterologous proteins migrating towards the leaf cell apoplast. *Plant Biotechnol J* **10**, 83-94.

Guo Y, Dong J, Zhou T*, et al.*, 2015. YeastFab: the design and construction of standard biological parts for metabolic engineering in Saccharomyces cerevisiae. *Nucleic Acids Res* **43**, e88.

Gutierrez S, Yvon M, Thebaud G, Monsion B, Michalakis Y, Blanc S, 2010. Dynamics of the multiplicity of cellular infection in a plant virus. *PLoS Pathog* **6**, e1001113.

Ham TS, Dmytriv Z, Plahar H, Chen J, Hillson NJ, Keasling JD, 2012. Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic Acids Res* **40**, e141.

Hartley JL, Temple GF, Brasch MA, 2000. DNA cloning using in vitro site-specific recombination. *Genome Res* **10**, 1788-95.

Haughn G, Smith J, Mazur B, Somerville C, 1988. Transformation with a mutant Arabidopsis acetolactate synthase gene renders tobacco resistant to sulfonylurea herbicides. *Molecular and General Genetics MGG* **211**, 266-71.

Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM, 2000. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant Mol Biol* **42**, 819-32.

Hillson N, 2011. DNA Assembly Method Standardization for Synthetic Biomolecular Circuits and Systems. In: Koeppl H, Setti G, Di Bernardo M, Densmore D, eds. *Design and Analysis of Biomolecular Circuits.* Springer New York, 295-314.

Hillson NJ, Rosengarten RD, Keasling JD, 2012. j5 DNA assembly design automation software. *ACS Synth Biol* **1**, 14-21.

Holme IB, Wendt T, Holm PB, 2013. Intragenesis and cisgenesis as alternatives to transgenic crop development. *Plant Biotechnol J* **11**, 395-407.

Hou H, Atlihan N, Lu ZX, 2014. New biotechnology enhances the application of cisgenesis in plant breeding. *Front Plant Sci* **5**, 389.

Houmard NM, Mainville JL, Bonin CP, Huang S, Luethy MH, Malvar TM, 2007. High-lysine corn generated by endosperm-specific suppression of lysine catabolism using RNAi. *Plant Biotechnol J* **5**, 605-14.

Hwang I, Sheen J, 2001. Two-component circuitry in Arabidopsis cytokinin signal transduction. *Nature* **413**, 383-9.

Jarboe LR, Grabar TB, Yomano LP, Shanmugan KT, Ingram LO, 2007. Development of ethanologenic bacteria. *Adv Biochem Eng Biotechnol* **108**, 237-61.

Ji X, Zhang H, Zhang Y, Wang Y, Gao C, 2015. Establishing a CRISPR–Cas-like immune system conferring DNA virus resistance in plants. *Nature Plants* **1**, 15144.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E, 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-21.

Jones HD, 2015. Regulatory uncertainty over genome editing. *Nature Plants* **1**, 14011.

Joung JK, Sander JD, 2013. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* **14**, 49-55.

Kapila J, Derycke R, Vanmontagu M, Angenon G, 1997. An Agrobacterium-mediated transient gene expression system for intact leaves. *Plant Science* **122**, 101-8.

Kelly JR, Rubin AJ, Davis JH*, et al.*, 2009. Measuring the activity of BioBrick promoters using an in vivo reference standard. *J Biol Eng* **3**, 4.

Kelwick R, Macdonald JT, Webb AJ, Freemont P, 2014. Developments in the tools and methodologies of synthetic biology. *Front Bioeng Biotechnol* **2**, 60.

Khalil AS, Collins JJ, 2010. Synthetic biology: applications come of age. *Nat Rev Genet* **11**, 367-79.

Kinkema M, Geijskes RJ, Shand K*, et al.*, 2014. An improved chemically inducible gene switch that functions in the monocotyledonous plant sugar cane. *Plant Mol Biol* **84**, 443-54.

Kneissl ML, Deikman J, 1996. The Tomato E8 Gene Influences Ethylene Biosynthesis in Fruit but Not in Flowers. *Plant Physiol* **112**, 537-47.

Lee TS, Krupa RA, Zhang F*, et al.*, 2011. BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *J Biol Eng* **5**, 12.

Lei Y, Lu L, Liu HY, Li S, Xing F, Chen LL, 2014. CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. *Mol Plant* **7**, 1494-6.

Leonard E, Nielsen D, Solomon K, Prather KJ, 2008. Engineering microbes with synthetic biology frameworks. *Trends Biotechnol* **26**, 674-81.

Lewens T, 2013. From bricolage to BioBricks: Synthetic biology and rational design. *Stud Hist Philos Biol Biomed Sci* **44**, 641-8.

Li JF, Norville JE, Aach J*, et al.*, 2013. Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nat Biotechnol* **31**, 688-91.

Li MZ, Elledge SJ, 2012. SLIC: a method for sequence- and ligation-independent cloning. *Methods Mol Biol* **852**, 51-9.

Liu W, Mazarei M, Rudis MR*, et al.*, 2013. Bacterial pathogen phytosensing in transgenic tobacco and Arabidopsis plants. *Plant Biotechnol J* **11**, 43-52.

Liu W, Mazarei M, Rudis MR, Fethe MH, Stewart CN, Jr., 2011. Rapid in vivo analysis of synthetic promoters for plant pathogen phytosensing. *BMC Biotechnol* **11**, 108.

Liu W, Stewart CN, Jr., 2015. Plant synthetic biology. *Trends Plant Sci* **20**, 309-17.

Liu Y, Zeng Y, Liu L*, et al.*, 2014. Synthesizing AND gate genetic circuits based on CRISPR-Cas9 for identification of bladder cancer cells. *Nat Commun* **5**, 5393.

Lowder LG, Zhang D, Baltes NJ*, et al.*, 2015. A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiol* **169**, 971-85.

Luo J, Butelli E, Hill L*, et al.*, 2008. AtMYB12 regulates caffeoyl quinic acid and flavonol synthesis in tomato: expression in fruit results in very high levels of both types of polyphenol. *Plant J* **56**, 316-26.

Luo S, Li J, Stoddard Thomas j*, et al.*, 2015. Non-transgenic Plant Genome Editing Using Purified Sequence-Specific Nucleases. *Mol Plant* **8**, 1425-7.

Ma X, Zhang Q, Zhu Q*, et al.*, 2015. A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Mol Plant*.

Mali P, Yang L, Esvelt KM*, et al.*, 2013. RNA-guided human genome engineering via Cas9. *Science* **339**, 823-6.

Marillonnet S, Werner S, 2015. Assembly of Multigene Constructs Using Golden Gate Cloning. In: Castilho A, ed. *Glyco-Engineering.* Springer New York, 269-84. (Methods in Molecular Biology; vol. 1321.)

Martin C, Zhang Y, Tonelli C, Petroni K, 2013. Plants, diet, and health. *Annu Rev Plant Biol* **64**, 19-46.

Mathews H, Clendennen SK, Caldwell CG*, et al.*, 2003. Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport. *Plant Cell* **15**, 1689-703.

Mehrtens F, Kranz H, Bednarek P, Weisshaar B, 2005. The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiol* **138**, 1083-96.

Mikami M, Toki S, Endo M, 2015. Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. *Plant Mol Biol*.

Molesini B, Pii Y, Pandolfini T, 2012. Fruit improvement using intragenesis and artificial microRNA. *Trends Biotechnol* **30**, 80-8.

Moore I, Samalova M, Kurup S, 2006. Transactivated and chemically inducible gene expression in plants. *Plant Journal* **45**, 651-83.

Muir SR, Collins GJ, Robinson S*, et al.*, 2001. Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols. *Nat Biotechnol* **19**, 470-4.

Muller K, Siegel D, Rodriguez Jahnke F*, et al.*, 2014. A red light-controlled synthetic gene expression switch for plant systems. *Mol Biosyst* **10**, 1679-88.
Muller KM, Arndt KM, 2012. Standardization in synthetic biology. *Methods Mol Biol* **813**, 23-43.

Mutalik VK, Guimaraes JC, Cambray G*, et al.*, 2013a. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat Methods* **10**, 354-60.

Mutalik VK, Guimaraes JC, Cambray G*, et al.*, 2013b. Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat Methods* **10**, 347-53.

Naqvi S, Farre G, Sanahuja G, Capell T, Zhu C, Christou P, 2010. When more is better: multigene engineering in plants. *Trends Plant Sci* **15**, 48-56.

Nekrasov V, Staskawicz B, Weigel D, Jones JD, Kamoun S, 2013. Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 691-3.

Nielsen AA, Voigt CA, 2014. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol Syst Biol* **10**, 763.

Niggeweg R, Michael AJ, Martin C, 2004. Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nat Biotechnol* **22**, 746-54.

Nikel PI, Martinez-Garcia E, De Lorenzo V, 2014. Biotechnological domestication of pseudomonads using synthetic biology. *Nat Rev Microbiol* **12**, 368-79.

Nour-Eldin HH, Geu-Flores F, Halkier BA, 2010. USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories. *Methods Mol Biol* **643**, 185-200.

Oldroyd GE, Dixon R, 2014. Biotechnological solutions to the nitrogen problem. *Curr Opin Biotechnol* **26**, 19-24.

Ori N, Juarez MT, Jackson D, Yamaguchi J, Banowetz GM, Hake S, 1999. Leaf senescence is delayed in tobacco plants expressing the maize homeobox gene knotted1 under the control of a senescence-activated promoter. *Plant Cell* **11**, 1073-80.

Ort DR, Merchant SS, Alric J*, et al.*, 2015. Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc Natl Acad Sci U S A* **112**, 8529-36.

Orzaez D, Medina A, Torre S*, et al.*, 2009. A visual reporter system for virus-induced gene silencing in tomato fruit based on anthocyanin accumulation. *Plant Physiol* **150**, 1122-34.

Padidam M, 2003. Chemically regulated gene expression in plants. *Current Opinion in Plant Biology* **6**, 169-77.

Paine JA, Shipton CA, Chaggar S*, et al.*, 2005. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat Biotechnol* **23**, 482-7.

Pasotti L, Politi N, Zucca S, Cusella De Angelis MG, Magni P, 2012. Bottom-up engineering of biological systems through standard bricks: a modularity study on basic parts and devices. *PLoS One* **7**, e39407.

Pasotti L, Zucca S, 2014. Advances and computational tools towards predictable design in biological engineering. *Comput Math Methods Med* **2014**, 369681.

Patron NJ, 2014. DNA assembly for plant biology: techniques and tools. *Curr Opin Plant Biol* **19**, 14-9.

Patron NJ, Orzaez D, Marillonnet S*, et al.*, 2015. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytol* **208**, 13-9.

Peccoud J, Anderson JC, Chandran D*, et al.*, 2011. Essential information for synthetic DNA sequences. *Nat Biotechnol* **29**, 22; discussion -3.

Peccoud J, Blauvelt MF, Cai Y*, et al.*, 2008. Targeted development of registries of biological parts. *PLoS One* **3**, e2671.

Piatek A, Ali Z, Baazim H*, et al.*, 2015. RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol J* **13**, 578-89.

Polturak G, Breitel D, Grossman N*, et al.*, 2015. Elucidation of the first committed step in betalain biosynthesis enables the heterologous engineering of betalain pigments in plants. *New Phytol*.

Port F, Chen HM, Lee T, Bullock SL, 2014. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proc Natl Acad Sci U S A* **111**, E2967-76.

Posfai G, Plunkett G, 3rd, Feher T*, et al.*, 2006. Emergent properties of reduced-genome Escherichia coli. *Science* **312**, 1044-6.

Puchta H, Fauser F, 2013. Gene targeting in plants: 25 years later. *Int J Dev Biol* **57**, 629-37.

Quan J, Tian J, 2009. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One* **4**, e6441.

Quinn JY, Cox RS, 3rd, Adler A*, et al.*, 2015. SBOL Visual: A Graphical Language for Genetic Designs. *PLoS Biol* **13**, e1002310.

Raiola A, Rigano MM, Calafiore R, Frusciante L, Barone A, 2014. Enhancing the health-promoting effects of tomato fruit for biofortified food. *Mediators Inflamm* **2014**, 139873.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F, 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-308.

Rao AQ, Bakhsh A, Kiani S*, et al.*, 2009. The myth of plant transformation. *Biotechnol Adv* **27**, 753-63.

Rastogi Verma S, 2013. Genetically modified plants: public and scientific perceptions. *ISRN Biotechnol* **2013**, 820671.

Ro DK, Paradise EM, Ouellet M*, et al.*, 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-3.

Rodrigo G, Carrera J, Landrain TE, Jaramillo A, 2012. Perspectives on the automatic design of regulatory systems for synthetic biology. *FEBS Lett* **586**, 2037-42.

Rommens CM, 2004. All-native DNA transformation: a new approach to plant genetic engineering. *Trends Plant Sci* **9**, 457-64.

Rommens CM, Haring MA, Swords K, Davies HV, Belknap WR, 2007. The intragenic approach as a new extension to traditional plant breeding. *Trends Plant Sci* **12**, 397-403.

Russell AW, Sparrow R, 2008. The case for regulating intragenic GMOs. *Journal of Agricultural and Environmental Ethics* **21**, 153-81.

Sainsbury F, Lomonossoff GP, 2014. Transient expressions of synthetic biology in plants. *Curr Opin Plant Biol* **19**, 1-7.

Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T, 2014. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci Rep* **4**, 5400.

Sarrion-Perdigones A, Falconi EE, Zandalinas SI*, et al.*, 2011. GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One* **6**, e21622.

Sarrion-Perdigones A, Palaci J, Granell A, Orzaez D, 2014. Design and construction of multigenic constructs for plant biotechnology using the GoldenBraid cloning strategy. *Methods Mol Biol* **1116**, 133-51.

Sarrion-Perdigones A, Vazquez-Vilar M, Palaci J*, et al.*, 2013. GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol* **162**, 1618-31.

Schaart JG, Van De Wiel CC, Lotz LA, Smulders MJ, 2015. Opportunities for Products of New Plant Breeding Techniques. *Trends Plant Sci*.

Schaumberg KA, Antunes MS, Kassaw TK*, et al.*, 2015. Quantitative characterization of genetic parts and circuits for plant synthetic biology. *Nat Methods*.

Schijlen EG, Ric De Vos CH, Van Tunen AJ, Bovy AG, 2004. Modification of flavonoid biosynthesis in crop plants. *Phytochemistry* **65**, 2631-48.

Schiml S, Fauser F, Puchta H, 2014. The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J* **80**, 1139-50.

Schneider CA, Rasband WS, Eliceiri KW, 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-5.

Schouten HJ, Krens FA, Jacobsen E, 2006. Do cisgenic plants warrant less stringent oversight? *Nat Biotechnol* **24**, 753.

Senis E, Fatouros C, Grosse S*, et al.*, 2014. CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. *Biotechnol J* **9**, 1402-12.

Shan Q, Wang Y, Li J*, et al.*, 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol* **31**, 686-8.

Sharma N, Jung CH, Bhalla PL, Singh MB, 2014. RNA sequencing analysis of the gametophyte transcriptome from the liverwort, Marchantia polymorpha. *PLoS One* **9**, e97497.

Shetty RP, Endy D, Knight TF, Jr., 2008. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng* **2**, 5.

Shimada S, Otsuki H, Sakuta M, 2007. Transcriptional control of anthocyanin biosynthetic genes in the Caryophyllales. *J Exp Bot* **58**, 957-67.

Silver PA, Way JC, Arnold FH, Meyerowitz JT, 2014. Synthetic biology: Engineering explored. *Nature* **509**, 166-7.

Slocombe SP, Cornah J, Pinfield-Wells H*, et al.*, 2009. Oil accumulation in leaves directed by modification of fatty acid breakdown and lipid synthesis pathways. *Plant Biotechnology Journal* **7**, 694-703.

Smolke CD, 2009. Building outside of the box: iGEM and the BioBricks Foundation. *Nat Biotechnol* **27**, 1099-102.

Strasser R, Stadlmann J, Schahs M*, et al.*, 2008. Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol J* **6**, 392-402.

Sundar IK, Sakthivel N, 2008. Advances in selectable marker genes for plant transformation. *J Plant Physiol* **165**, 1698-716.

Taniguchi Y, Ohkawa H, Masumoto C*, et al.*, 2008. Overproduction of C4 photosynthetic enzymes in transgenic rice plants: an approach to introduce the C4-like photosynthetic pathway into rice. *J Exp Bot* **59**, 1799-809.

Tzfira T, Li J, Lacroix B, Citovsky V, 2004. Agrobacterium T-DNA integration: molecules and models. *Trends Genet* **20**, 375-83.

Unbottling the genes. 2009*Nat Biotechnol* **27**, 1059.

Upadhyay SK, Kumar J, Alok A, Tuli R, 2013. RNA-guided genome editing for target gene mutations in wheat. *G3 (Bethesda)* **3**, 2233-8.

Vazquez-Vilar M, Bernabe-Orts JM, Fernandez-Del-Carmen A*, et al.*, 2016. A modular toolbox for gRNA-Cas9 genome engineering in plants based on the GoldenBraid standard. *Plant Methods* **12**, 10.

Vazquez-Vilar M, Sarrion-Perdigones A, Ziarsolo P, Blanca J, Granell A, Orzaez D, 2015. Software-assisted stacking of gene modules using GoldenBraid 2.0 DNA-assembly framework. *Methods Mol Biol* **1284**, 399-420.

Verpoorte R, Memelink J, 2002. Engineering secondary metabolite production in plants. *Curr Opin Biotechnol* **13**, 181-7.

Villalobos A, Ness JE, Gustafsson C, Minshull J, Govindarajan S, 2006. Gene Designer: a synthetic biology tool for constructing artificial DNA segments. *BMC Bioinformatics* **7**, 285.

Von Caemmerer S, Quick WP, Furbank RT, 2012. The development of C(4)rice: current progress and future challenges. *Science* **336**, 1671-2.

Wang H, Yang H, Shivalila CS*, et al.*, 2013a. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-8.

Wang S, Tiwari SB, Hagen G, Guilfoyle TJ, 2005. AUXIN RESPONSE FACTOR7 restores the expression of auxin-responsive genes in mutant Arabidopsis leaf mesophyll protoplasts. *Plant Cell* **17**, 1979-93.

Wang YH, Wei KY, Smolke CD, 2013b. Synthetic biology: advancing the design of diverse genetic systems. *Annu Rev Chem Biomol Eng* **4**, 69-102.

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S, 2011a. A modular cloning system for standardized assembly of multigene constructs. *PLoS One* **6**, e16765.

Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S, 2011b. Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One* **6**, e19722.

Wenck A, Pugieux C, Turner M*, et al.*, 2003. Reef-coral proteins as visual, non-destructive reporters for plant transformation. *Plant Cell Rep* **22**, 244-51.

Wend S, Dal Bosco C, Kampf MM*, et al.*, 2013. A quantitative ratiometric sensor for time-resolved analysis of auxin dynamics. *Sci Rep* **3**, 2052.

What's in a name? , 2009*Nat Biotechnol* **27**, 1071-3.

Xia B, Bhatia S, Bubenheim B, Dadgar M, Densmore D, Anderson JC, 2011. Developer's and user's guide to Clotho v2.0 A software platform for the creation of synthetic biological systems. *Methods Enzymol* **498**, 97-135.

Xie K, Minkenberg B, Yang Y, 2015. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc Natl Acad Sci U S A* **112**, 3570-5.

Xing HL, Dong L, Wang ZP*, et al.*, 2014. A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol* **14**, 327.

Xiong J-S, Ding J, Li Y, 2015. Genome-editing technologies and their potential application in horticultural crop breeding. *Horticulture Research* **2**, 15019.

Xu J, Dolan MC, Medrano G, Cramer CL, Weathers PJ, 2012. Green factory: plants as bioproduction platforms for recombinant proteins. *Biotechnol Adv* **30**, 1171-84.

Yanagisawa S, Yoo SD, Sheen J, 2003. Differential regulation of EIN3 stability by glucose and ethylene signalling in plants. *Nature* **425**, 521-5.

Yang F, Mitra P, Zhang L*, et al.*, 2013. Engineering secondary cell wall deposition in plants. *Plant Biotechnol J* **11**, 325-35.

Yang X, 2015. Applications of CRISPR-Cas9 mediated genome engineering. *Mil Med Res* **2**, 11.

Yang YN, Li RG, Qi M, 2000. In vivo analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. *Plant Journal* **22**, 543-51.

Yao JL, Tomes S, Gleave AP, 2013. Transformation of apple (Malus x domestica) using mutants of apple acetolactate synthase as a selectable marker and analysis of the T-DNA integration sites. *Plant Cell Rep* **32**, 703-14.

Yoo SD, Cho YH, Sheen J, 2007. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat Protoc* **2**, 1565-72.

Zhang F, Maeder ML, Unger-Wallace E*, et al.*, 2010. High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases. *Proc Natl Acad Sci U S A* **107**, 12028-33.

Zhang H, Jiang T, 2010. Synthetic circuits, devices and modules. *Protein Cell* **1**, 974-8.

Zhang Y, Butelli E, Alseekh S*, et al.*, 2015. Multi-level engineering facilitates the production of phenylpropanoid compounds in tomato. *Nat Commun* **6**, 8635.

Zhang Y, Werling U, Edelmann W, 2012. SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res* **40**, e55.

Zhou Q, Liu W, Zhang Y, Liu KK, 2007. Action mechanisms of acetolactate synthase-inhibiting herbicides. *Pesticide Biochemistry and Physiology* **89**, 89-96.

Zurbriggen MD, Tognetti VB, Fillat MF, Hajirezaei MR, Valle EM, Carrillo N, 2008. Combating stress with flavodoxin: a promising route for crop improvement. *Trends Biotechnol* **26**, 531-7.

Zurcher E, Tavor-Deslex D, Lituiev D, Enkerli K, Tarr PT, Muller B, 2013. A robust and sensitive synthetic sensor to monitor the transcriptional output of the cytokinin signaling network in planta. *Plant Physiol* **161**, 1066-75.