UNIVERSITAT POLITÈCNICA DE VALÈNCIA

/ Design and development of modular DNA assembly tools for multigene engineering and Synthetic Biology in plants

Autor

Manuel Alejandro Sarrión Perdigones

Director

Dr. Diego Orzáez Calatayud

Co-Director

Dr. Antonio Granell Richard

Enero de 2014

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Design and development of modular DNA assembly tools for multigene engineering and Synthetic Biology in plants

Memoria presentada por

Manuel Alejandro Sarrión Perdigones

Para optar al Grado de

Doctor en Biotecnología

Director

Dr. Diego Orzáez Calatayud

Co-Director

Dr. Antonio Granell Richard

Enero de 2014

El Dr. Diego Orzáez Calatayud, Científico Titular del Consejo Superior de Investigaciones Científicas y Profesor Asociado de la Universidad Politécnica de Valencia, y el Dr. Antonio Granell Richart, Profesor de Investigación del Consejo Superior de Investigaciones Científicas, ambos pertenecientes al Instituto de Biología Molecular y Celular de Plantas,

CERTIFICAN que el Licenciado en Biotecnología MANUEL ALEJANDRO SARRIÓN PERDIGONES ha realizado bajo su dirección en el Instituto de Biología Molecular y Celular de Plantas el trabajo que lleva por título "Design and development of modular DNA assembly tools for multigene engineering and Synthetic Biology in plants", y autorizan su presentación para optar al grado de Doctor en Biotecnología.

Y para que así conste, expiden y firman el presente certificado en Valencia, a 9 de diciembre de 2013

Dr. Diego Orzáez Calatayud

Dr. Antonio Granell Richart

A mi abuelo José, el que más
habría disfrutado con todo esto

Agradecimientos

En principio, como algunos sabéis, no quería incluir una sección de agradecimientos en la Tesis pues aquellas personas a las que tengo que dar gracias por su ayuda durante estos años, ya saben de sobra lo que les querría decir. Finalmente, me he decidido por hacerla pues creo que os merecéis que este reconocimiento sea público y que no quede como algo entre nosotros.

La primera mención es a mi Director de Tesis. Diego confió en mi sin conocerme, me invitó (sí, me invitó... ¡él pagó el viaje!) a Valencia a hacer una entrevista en enero de 2009 y de esa reunión salió el compromiso de hacer la tesis a su laboratorio. En Septiembre de 2009 hice las maletas y me vine para Valencia. Pronto supe que no me había equivocado al hacer esta elección y no me he arrepentido de ello en ningún momento de estos años. Diego ha sido un magnífico Director. No son muchos los doctorandos que acaban su tesis con la sensación de cercanía, de confianza y de amistad con su Director que yo me llevo. Hemos discutido y nos hemos enfadado pero siempre con la tranquilidad de saber que siendo claros con el otro, no iba a haber ningún problema. De lo que más valoro es que me ha escuchado cuando pensaba que no lo estaba haciendo bien y en muchas ocasiones ha rectificado, siendo un jefe muy justo. Diego, no puedo estar más agradecido por todo lo que he aprendido aquí contigo. Me has permitido hacer una tesis de la que me siento muy orgulloso, hemos decidido juntos los pasos a dar y creo que hemos conseguido hacer un trabajo redondo. GoldenBraid fue un diseño que hicimos juntos en tu despacho y cuatro años después es una herramienta que se está utilizando en muchos laboratorios por todo el mundo; aún queda mucho camino pero podemos estar contentos. Aquí he tenido muchas experiencias más allá del trabajo de laboratorio: redactar patentes, organizar una escuela de verano con gente de toda Europa, dar clases en la Licenciatura de Biotecnología, haber co-dirigido cuatro proyectos de fin de carrera... todo eso ha sido posible por tu cercanía, por tu disponibilidad y por la enorme confianza que depositaste en mi desde el primer día. Gracias por todo, Diego.

A Toni Granell, mi co-director, agradecerle su confianza y las buenas palabras que siempre he escuchado hacia mi trabajo. Espero que las cosas que tenemos abiertas, lleguen a buen puerto.

A los miembros de mi laboratorio, agradecerles los buenos ratos compartidos, las discusiones científicas y no científicas. Entre todos vosotros, quiero agradecérselo de forma especial a tres personas sin que eso signifique que a los demás no os pudiera dedicar mis mejores palabras. Primero a Silvia Presa, que hace un trabajo indispensable para que el laboratorio funcione en el día a día. Gracias por tu compañía en todos estos años, por tus consejos y por las innumerables conversaciones en la terraza o la salita del café. La segunda persona que a la que quiero hacer una mención especial es Marta Vázquez pues sin su colaboración y buenas ideas, GoldenBraid2.0 no sería lo que es y sobre todo no tendría las herramientas informáticas en las que se ha dejado la piel. Eres una trabajadora increíble, lo vi con claridad cuando viniste al laboratorio como alumna de verano en 2010, y seguro que vas a hacer una tesis impresionante. Gracias por contribuir, opinar, criticar y dar ideas. Es un gusto trabajar a tu lado. Por último, agradecerle a José Manuel Julve todas las risas que me ha sacado con sus ideas locas. Gracias por los consejos que me has dado en innumerables ocasiones.

Aparte de los miembros *permanentes* del laboratorio, hay 4 personas que han pasado por el 2.10 en estos años que han tenido un papel destacado en esta tesis: Erica Falconi, Sara Izquierdo, Jorge Palací y Rocío Ochoa. No he podido tener mejor suerte con los alumnos de Proyecto Fin de Carrera que he tenido y a todos vosotros os debo un agradecimiento enorme. Trabajabais por vuestro proyecto… pero también lo hacíais por mi tesis. Todos tenéis una parte importante de culpa en los distintos capítulos que la componen y es justo que mi gratitud quede por escrito. De una forma especial, destacar el esfuerzo que hacía Erica Falconi después de su jornada laboral en el Hospital de la Villajollosa y los 250 kilómetros que recorría para venir al laboratorio. Tu ayuda fue imprescindible para que GoldenBraid fuera una realidad y debes estar muy orgullosa del trabajo que hiciste con nosotros en el IBMCP. A Rocío Ochoa quiero agradecerle el enorme trabajo que ha hecho este otoño, que me ha permitido desentenderme un poco de lo que estaba pasando en la poyata y centrarme en darle forma a esta tesis. Todos los demás han tenido su publicación como recompensa del trabajo en nuestro laboratorio y tú no vas a ser menos, llegará en no mucho tiempo. Gracias, Rocío, por tu enorme trabajo, tu disposición y tus ganas de aportar, por tu visión crítica y por no estar conforme nunca con ningún resultado, sino que siempre querías buscar uno mejor… aunque eso implicara transformar dos bandejas de plantas una vez más.

Fuera de las puertas del laboratorio 2.10, a la persona a la que más líneas quiero dedicarle es Mª Ángeles Argomaniz. Sin ti, los 4Labos no funcionarían como lo hacen. Tienes una

paciencia infinita con todos nosotros y no siempre los compañeros saben agradecerte lo indispensable que es tu labor (y personalmente, eso me da muchísima rabia). Desde el primer día, me has regalado una sonrisa en todo momento, me has escuchado cuando necesitaba un desahogo y has estado continuamente pendiente de mí, siempre preguntándome en cuanto me veías algo bajo de ánimos… y has conseguido que cada vez que salga de tu laboratorio me vaya contento. Cuando cumpliste los 65 te dije que por favor no te jubilaras hasta que no leyera yo la tesis… y aunque evidentemente tu decisión de seguir trabajando más allá de la edad de retiro no depende de lo que yo te haya pedido, quiero agradecerte muchísimo que lo hayas hecho. Ojalá cuando llegue la hora de mi jubilación tenga la mitad de la vitalidad, de la ilusión y las ganas de hacer cosas que tú tienes. Gracias por cuidarme todos estos años, de verdad.

Muchos otros miembros del grupo 4Labos me han aportado mucho en estos años, como son MD Gómez, Paco Vera, Carolina Gallego o Juan Carbonell. De entre los que siguen y los que no están, a quién más cosas tengo que agradecer es a Berta Sotillo. Berta, has sido una estupenda compañera de desahogos, de críticas y de risas. Es un placer trabajar contigo ya no solo en el IBMCP sino como estamos haciendo ahora en la Federación Española de Biotecnólgos. Mucho ánimo con lo que te queda por delante, termina de escribir tu tesis y preséntala que yo estaré muy orgulloso de ello. También quiero mencionar a Clara Fuster, por los ratos de café que hemos compartido por la mañana temprano, cuando casi nadie más ha llegado al laboratorio. Gracias por tu ayuda para cuidar a mis *Arabidopsis* en este último año pues lo has hecho sin que eso fuera tu tarea.

Del resto del IBMCP hay muchas personas a las que quiero dedicar unas líneas. A unas por las sonrisas y conversaciones en los pasillos (sobre todo a Sandra Fresquet, aunque podría mencionar a muchos otros como María *la limpiadora*, Mapi López, Cristina Martí, Liliam Martínez o Teresa Caballero) y a otras por su profesionalidad y saber hacer (en especial a Eugenio Grau y a Ana Marín del Servicio de Secuenciación, a Rafa de los Invernaderos y a Marisol del Servicio de Microscopía). A Lynne Yenush, Luis Cañas y Cristina Ferrándiz agradecerles que hayan probado GoldenBraid en sus laboratorios, como también lo están haciendo Miguel Pérez y Eugenio Gómez en los más cercanos al 2.10. A los miembros de la Comisión de Divulgación del IBMCP agradecerles su confianza en mi criterio y el apoyo recibido a mis propuestas (sobre todo a ti Luis, que me has dejado hacer con total libertad). Ha sido un placer desarrollar *PLANTéatelo* con vosotros y espero que la labor que hacemos tenga continuidad en los próximos años, que siempre habrá profesores interesados en reciclarse y sacar nuevas ideas para aplicarlas con sus alumnos.

Un lugar especial en todos estos años lo han tenido numerosos compañeros involucrados en la Federación Española de Biotecnólogos. Muchos os habéis convertido en grandes amigos, como es el caso de Joaquina Delás, Saray Soldado, Marcos Fernández o María José Conde. Entre todos, quiero destacar de forma especial a Ángela Bernardo, que ha sido mi compañera de fatigas los últimos cuatro años en múltiples proyectos e ideas. Eres una trabajadora incombustible y haces que desarrollar cualquier tarea contigo sea un placer. A otros miembros de FEBiotec como son Diego Balboa, Alex Fernández, Sandra Ciudad, Alba Timón, Julio Ríos, Rodrigo García, Roberto Flores, Roi Villar, Alba Olivares, Rocío Cánovas, David Gallardo o Marc Parrilla (y otros muchísimos más), decirles que me siento muy orgulloso del trabajo que hemos hecho y de las grandes acciones en defensa y promoción de nuestro perfil profesional que se siguen llevando a cabo. Quiero tener un recuerdo muy especial para Clara García y Albert Almasqué, que siempre estarán conmigo.

Después de todos estos agradecimientos relacionados con el mundo científico quiero tener unas palabras para varios de mis mejores amigos, que han estado ahí a lo largo de estos años. A Laura del Río y Carlos Bernal, mis amigos de siempre. Gracias por hacer que cada vez que he ido a Sevilla me sintiera como si nunca me hubiera marchado, como si el tiempo no hubiera pasado. Por otro lado a Alicia Jiménez y Tere Ciudad, mis compañeras en varias locas aventuras en China, Croacia y Grecia. Es un placer compartir con vosotras conversaciones de Ciencia… y de lo que no es Ciencia… y espero que pronto volvamos a viajar juntos!

Y por último, a mi familia. Gracias a mis padres por su constante e incondicional apoyo. Gracias por vuestros consejos, por vuestro trabajo incansable y por haberme dado la educación que me ha hecho ser quien soy y me ha permitido llegar al lugar en el que estoy. A mi abuela Concha porque aunque no entienda nada de lo que hago, ella siempre me va a apoyar y a estar orgullosa a más no poder. Y por último a Elías, que ha elegido el estar a mi lado y compartir conmigo todas las aventuras que nos vengan por delante. Has tenido una paciencia infinita cuando he tenido que ir al laboratorio los fines de semanas, cuando he estado trabajando horas y horas… Sin tu apoyo, muchas cosas habrían sido mucho más difíciles y esta tesis es una de ellas (y tampoco habría quedado tan bonita!)… ¿nos vamos ya al próximo país?

Summary

One of the aims of Plant Synthetic Biology is the engineering of entire biosynthetic or signaling pathways using plant crops as “chassis”. A major technological challenge for achieving this involves transferring large amounts of genetic information to the plant genome in the form of multigene constructs. To facilitate multigene engineering in plants we have developed a DNA assembly platform that we named GoldenBraid (GB). GB is a standardized DNA assembly system for Plant Synthetic Biology based on the use of Type IIS restriction enzymes that allows the indefinite growth of reusable gene modules made of standardized DNA parts. The GB toolbox includes eight destination vectors and a universal domesticator plasmid designed to incorporate multipartite assemblies made of standard DNA parts and to combine them binarily in increasingly complex multigene constructs. The use of the GoldenBraid framework is facilitated by a number of web resources, which include a publicly available database, tutorials and a software package that provides *in silico* simulations and tailored laboratory protocols for part domestication and multigene assembly. The GB toolkit is completed with a repository of standard DNA parts, the GBcollection. This collection contains more than 300 basic GBparts, more than 500 intermediate modules and pre-made transcriptional units, including a set of basic logic gates that will be the basis of future digitally-regulated gene networks. The most basic one-input operations and an initial set of two-input Boolean gates were developed. Finally a NOT logic gate, the first inducible basic inverter in plants, was also implemented.

Resumen

La Biología Sintética de Plantas tiene como objetivo la construcción de rutas de biosíntesis o señalización, utilizando plantas cultivables como "chasis". Para alcanzar este objetivo, uno de los principales desafíos tecnológicos es la transferencia de gran cantidad de información genética al genoma de la planta en forma de construcciones multigénicas. Para facilitar la ingeniería multigénica en plantas, se ha desarrollado una plataforma de ensamblado de DNA que hemos denominado GoldenBraid (GB). GB es un sistema estandarizado de ensamblaje de DNA para la Biología Sintética de Plantas basado en el uso de enzimas de restricción de Tipo IIS, que permite el crecimiento indefinido de módulos genéticos reusables ensamblados a partir de piezas de DNA estandarizadas. La "caja de herramientas" de GB incluye ocho vectores de destinos y un plásmido domesticador universal, diseñados para incorporar ensamblajes multipartitos a partir de piezas de DNA estandarizadas y para combinar estos de forma binaria en construcciones multigénicas de complejidad creciente. El uso del sistema GB se ve facilitado con la puesta en marcha de una serie de recursos web que incluyen una base de datos pública, tutoriales y herramientas software que permiten realizar simulaciones *in silico* y proporcionan protocolos de laboratorio a medida para la domesticación de las piezas de DNA y la construcción de estructuras multigénicas. Todo esto se completa con la Colección GB, un repositorio de piezas de DNA estandarizadas que contiene más de 300 GBparts y más de 500 módulos intermedios y unidades transcripcionales pre-ensambladas. Además, incluye un set de puertas lógicas básicas que serán el punto de partida para la creación de redes génicas regulables. Se han desarrollado las operaciones básicas de una entrada y un conjunto inicial de puertas Booleanas de dos entradas. Finalmente se ha implementado la puerta lógica NOT, el primer ejemplo de puerta lógica de inversión en plantas

Resum

La Biologia Sintètica de Plantes té com a objectiu la construcció de vies de biosíntesi o senyalització, utilitzant plantes cultivables com "xassís". Un dels principals desafiaments tecnològics per a arribar a aquest objectiu és la transferència de gran quantitat d'informació genètica al genoma de la planta en forma de construccions multigèniques. Per a facilitar l'enginyeria multigènica en plantes, hem desenvolupat una plataforma d'assemblatge amb el nom de GoldenBraid (GB). GB és un sistema d'assemblatge de DNA estàndard per a la Biologia Sintètica de Plantes basat en l'ús d'enzims de restricció de tipus IIS, que permet el creixement indefinit de mòduls genètics reutilitzables assemblats a partir de peces de DNA estandarditzades. La "caixa d'eines" de GB inclou vuit vectors de destinació i un plàsmid domesticador universal dissenyat per a incorporar assemblatges multipartits a partir de peces de DNA estandarditzades i per a combinar aquestes de forma binària en construccions multigèniques de complexitat creixent. L'ús del sistema GB es veu facilitat amb l'ús d'una sèrie de recursos web que inclouen una base de dades pública, tutorials i eines de software que permeten realitzar simulacions *in silico* i proporcionen protocols de laboratori per a la domesticació de les peces de DNA i per a la construcció d'estructures multigèniques. Tot això es completa amb la Col·lecció GB, un repositori de peces de DNA estandarditzades que conté més de 300 GBparts i més de 500 mòduls intermedis i unitats transcripcionals pre-assemblades. A més a més, inclou un set de portes lògiques bàsiques que seran el punt de partida del futur desenvolupament de xarxes gèniques regulables. S'han desenvolupat les operacions més bàsiques d'una entrada i un conjunt inicial de portes Booleanes de dues entrades. Finalment, es va implementar la porta lògica NOT, el primer exemple d'una porta lògica d'inversió desenvolupat en plantes.

Abbreviations

A.tumefaciens *Agrobacterium tumefaciens*
AD Activation Domain
AmpR Ampicillin resistance gen
BD Binding Domain
BiBAC Binary bacterial artificial chromosome
BiFC Bimolecular Fluorescence Complementation
BFP Blue Fluorescent Protein
bp base pairs
ccdB bacterial lethal gene
cDNA cyclic Deoxyribonucleic Acid
CDS Coding sequence
CH 1 Immunoglobulin Heavy Chain 1 constant domain
CH 2 Immunoglobulin Heavy Chain 2 constant domain
CmR Chloramphenicol resistance gene
C Immunoglobulin Light Chain Lambda constant domain
C Immunoglobulin Light Chain Kappa constant domain
DNA Deoxyribonucleic Acid
d.p.i. Days post infiltration
DsRed *Discosoma sp.* Red Fluorescent Protein
DV Destination Vector
E.coli *Escherichia coli*
EV Expression Vector
ELISA Enzyme-Linked Immunosorbent Assay
FOI Fragment of interest
FUS Frequently Used Structures
GB GoldenBraid
GB2.0 GoldenBraid2.0 Framework
GFP Green Fluorescent Protein
GmR Gentamicin resistance gen
GOI Gene of interest
h Hours
HEN Homing Endonuclease
h-IgA Human Immunoglobulin A
HRP Horseradish peroxidase
HygR Hygromicin resistance gene
IgA Immunoglobulin A
IgH Immunoglobulin Heavy chain
Ig Immunoglobulin Light Chain Kappa
Ig Immunoglobulin Light Chain Lambda
IP Intellectual Property
KanR Kanamycin resistance gen
kb kilo base
kDa kilo Dalton
LacZ *-galactosidase gene*
LB Luria broth
MAR Matrix attachment region
MES 2-(N-morpholino)ethanesulfonic acid
min minutes
ml milliliters
mM milli Molar
MoClo Modular cloning system

MRG MultiRound Gateway
nm nanometer
N.benthamiana *Nicotiana benthamiana*
N.tabaccum *Nicotiana tabaccum*
NptII Neomycin phosphotransferase II gene
NTR Non Transcribed Region
Op Operator
P19 Tomato Bushy Stunt Virus P19 silencing suppressor
PBS Phosphate buffered saline
PCR Polymerase Chain Reaction
pDGB GoldenBraid Destination plasmid
pEGB GoldenBraid Expression/Entry plasmid
POI Promoter of interest
PR Promoter
PVDF Polyvinylidene Fluoride
pUPD Universal Parts Domesticator Plasmid
rpm revolutions per minute
SB Synthetic Biology
SDS-PAGE Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
S.lycopersicum *Solanum lycopersicum*
SP Pectate lyase signal peptide
SpmR Spectinomycin resistance gen
TAC Transformation Competent Artificial Chromosome
TAE Tris-acetate-EDTA
TAL Transcription activator-like effectors
TM Terminator
TNos Nopaline synthase terminator
TOI Terminator of interest
TU Transcriptional Unit
T-DNA Transfer DNA
UAS Upstream Activation Sequence where Gal4 binds to
UTR Untranslated Region
UV Ultraviolet
V Volts
VH Immunoglobulin Heavy Chain Variable domain
VL Immunoglobulin Light Chain Variable domain
v/v volume/volume
v/w volume/ weight
w/v weight/volume
YFP Yellow Fluorescent Protein
ZFN Zinc Finger Nuclease
l microliters
ºC Degree Celsius
⊤ Tautology
⊥ Contradiction
P Proposition
¬ Logical Negation
∨ Logical Conjunction
∧ Logical Disjunction
35S Cauliflower Mosaic Virus 35s promoter

Index

Summary XI

Resumen XIII

Resum XV

Abbreviations XVI

Index XIX

Figure Index XXII

Table Index XXIII

Supplemental Figure Index XXIV

Supplemental Table Index XXV

Introduction 1

1. Multigene assembly is out of necessity and not on a whim 3

1.1. Adapt or die. Multigene engineering is possible with traditional Type II cloning enzymes 5

1.2. New in town (for Plant Biotech): don't let them out of your sight! 7

1.3. Revisiting site-specific recombination tools for multigene engineering 7

1.4. The use of Zinc Finger Nucleases and Homing Endonucleases 11

1.5. Building multigenic constructs using Type IIs restriction enzymes 13

2. The fine-control of the genetic expression 15

2.1. Logic gates 17

Objectives 21

Chapter 1. GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules 25

1. Introduction 27

2. Results 29

2.1. Part standardization and multipartite assembly of simple devices 29
2.2. The Double Loop design of the GoldenBraid system 31
2.3. Multigenic constructs for Plant Biology 35
2.4. Construction and combination of therapeutic and biosafety gene modules by GoldenBraid 39
3. Discussion 41
4. Materials and methods 46
4.1. Strains and growth conditions 46
4.2. Cloning and assembly of modular pieces 46
4.3. GB-Domestication of destination plasmids for plant biology 47
4.4. Plant transient transformation 47
4.5. Western Blot and ELISA analysis 47
Chapter 2. GoldenBraid2.0: a comprehensive DNA assembly framework for Plant Synthetic Biology 51
1. Introduction 53
2. Results 55
2.1. The GB2.0 cloning strategy 55
2.2. GBparts: words and phrases. Definition of the GB2.0 grammar 56
2.3. GBpart domestication: creating words and phrases 58
2.4. The GB2.0 destination plasmids kit 58
2.5. The composing strategy: from single words to full compositions 60
2.6. Innovative features in the GB2.0 cloning strategy 61
2.7. GB collection and software tools 68
3. Discussion 69
4. Material and methods 72
4.1. Strains and growth conditions 72
4.2. Restriction-ligation assembly reactions 73
4.3. GBpart domestication 73
4.4. pDGB construction 73
4.5. *N. bentamiana* transient transformation 74
4.6. *A. thaliana* stable transformation 74

4.7. Luciferase/Renilla expression assays 74
4.8. Glucocorticoid Receptor induction and Heat shock treatments 75
Chapter 3. Engineering modular and orthogonal logic gates for Plant Synthetic Biology 77
1. Introduction 79
2. Results 80
2.1. Simple operations 80
2.2. Binary operations 89
3. Discussion 93
4. Materials and methods 94
4.1. Bacterial strains and growth conditions 94
4.2. Restriction-ligation assembly and domestication reactions 95
4.3. *N. bentamiana* transient transformation 95
4.4. *A. thaliana* stable transformation and growth conditions 95
4.5. Glucocorticoid and estradiol receptor induction 96
4.6. Heat-shock treatments 96
4.7. Luciferase/Renilla expression assays 96
General Discussion 99
1. The Goldenbraid Framework: an insight look 101
1.1. A binary loop fed by a multipartite standard 101
1.2. The generation, selection and characterization of the GBparts 103
1.3. Size and speed considerations 104
1.4. The software tools 104
2. GoldenBraid perspectives: an outlook 106
2.1. Compatibility between different multigene assembly methods 106
2.2. The creation of a community of users 108
3. The development of Logic Gates adapted to Plant Biotechnology 108
Conclusions 113
References 117
Supplemental Figures 137
Supplemental Tables 152

Figure Index

Figure 1. Simple graphical overview of the multigene assembly methods described in this chapter ... 6
Figure 2. Detailed diagram of MultiRound Gateway (MRG) ... 9
Figure 3. Detailed diagram of MISSA ... 10
Figure 4. Detailed diagram of pSAT+ZFNs vectors ... 12
Figure 5. Functioning of the Golden Gate Assembly System ... 14
Figure 6. Detailed diagram of MoClo ... 16
Figure 7. Truth tables of the Boolean Logic Gates ... 17
Figure 8. Part standardization and multipartite assembly of single devices ... 30
Figure 9. Structure of the LacZ cassettes in the GoldenBraid system ... 32
Figure 10. The mechanism of GoldenBraid system ... 34
Figure 11. GoldenBraid-assisted co-transformation of fluorescent devices ... 37
Figure 12. GoldenBraid-assisted selection of plant-made IgA isotypes ... 38
Figure 13. New multigene assemblies using reusable composite parts ... 40
Figure 14. Comparison of the topology of MoClo and GoldenBraid ... 44
Figure 15. Analogies between GB2.0 and English grammar ... 56
Figure 16. The complete GB2.0 grammar and its most frequently used structures ... 57
Figure 17. Standardized domestication of GBparts ... 59
Figure 18. GB2.0 cassettes and assembly rules ... 62
Figure 19. Characterization of regulatory regions for basic expression cassettes ... 65
Figure 20. GoldenBraid versus GB2.0 ... 71
Figure 21. Truth tables for the one-input one-output models ... 80
Figure 22. Tautology and contradiction ... 83
Figure 23. Propositions ... 84
Figure 24. Transcriptional strength of the engineered versions of the CaMV 35S promoter and effect of the synthetic transcriptional repressors ... 86
Figure 25. NOT logic gate ... 88
Figure 26. Logical Disjunction ... 91
Figure 27. Logical Conjunction ... 92
Figure 28. Pathways the generation of a 5-gene construct using three different assembly systems ... 105
Figure 29. Wiring options for the development of Boolean Logic gates ... 109

Table Index

Table 1. Logic Operations for Plant Synthetic Biology 81
Table 2. Bacterial Resistances and Type IIS Enzymes used in MoClo and GoldenBraid for the different restriction-ligation reactions 103
Table 3. Multigene Assembly methods for Plant Biotechnology 107

Supplemental Figure Index

Supplemental Figure 1. GB Domestication with the removal of internal restriction sites .. 138
Supplemental Figure 2. Combinatorial antibody engineering using the GBpatch assembly level ... 139
Supplemental Figure 3. Theoretical and experimental transcriptional activity of different promoter/terminator combinations ... 140
Supplemental Figure 4. Functional characterization of heat-shock promoters ... 141
Supplemental Figure 5. Functional characterization of transactivation constructs ... 142
Supplemental Figure 6. Frequently Used Structures (FUS) for the protein-protein interaction analysis ... 143
Supplemental Figure 7. The Frequently Used Structures (FUS) for endogenous gene silencing ... 144
Supplemental Figure 8. Adapting Gateway (GW) technology to GB2.0 ... 145
Supplemental Figure 9. GoldenBraid Building pathways for the (A) ᴛLACI::GAL4AD and (B) ⊥UAS gates ... 146
Supplemental Figure 10. GoldenBraid Building pathways for the (A) P GR (B) P ER and (C) P HS logic gates ... 147
Supplemental Figure 11. GoldenBraid Building pathway for the ¬GR.LexA::SRDX version of the logic gate ... 148
Supplemental Figure 12. GoldenBraid Building pathways for the (A) GR∨HS (B) GR∨ER gates ... 149
Supplemental Figure 13. GoldenBraid Building pathways and results for the AND tests constructs (I) ... 150
Supplemental Figure 14. GoldenBraid Building pathways and results for the AND tests constructs (II) ... 151

Supplemental Table Index

Supplemental Table 1 GBparts used in Chapter 2 152
Supplemental Table 2 GBparts used in Chapter 3 154

/IN

INTRODUCTION

MULTIGENE ASSEMBLY IS OUT AND NOT ON A WHIM

THE FINE-CONTROL OF THE GENETIC EXPRESSION

Introduction

The nascent discipline of Plant Synthetic Biology aims to apply engineering principles to the plant genetic design. Two strategic requirements to achieve this goals are (i) the adoption of common standardized technologies that facilitate the construction of increasingly complex multigenic structures at the DNA level while enabling the exchange of genetic building blocks among plant bioengineers, and (ii) the generation of genetic components, such as logic gates, to fine-control genetic expression and to generate gene expression profiles beyond the spectrum of natural promoters.

1. Multigene assembly is out of necessity and not on a whim.

Plant Biotechnology is currently swift-moving from the traditional single-gene target approach towards the engineering of complex multigene traits **[1]**. The majority of valuable agronomic traits are polygenic, therefore the introduction of novel genetic pathways or the modification of the existing ones (either by overexpressing or down-regulating the native genes) may require the simultaneous incorporation of several transgenes to the plant **[2-4]**. Biotech challenges requiring a multigene approach go beyond quantitative multigene traits such as in the engineering of metabolic routes **[4-5]** to the modification of the photosynthetic carbon cycle, turning C3 plants into C4 or CAM **[6-7]**, or even the construction of whole-organism biosensors **[8]**. The recombinant production of added-value heterologous multimeric proteins also involves the transformation of several transcriptional units (TUs) into the plant and perfectly exemplifies the goals for which a standardized multigene assembly technology would be demanded **[9]**.

Stacking several genes from different origins in one plant is often tedious and time-consuming. It has traditionally been implemented through different strategies: (i) gene pyramiding by sexual crossing of plants containing the individual transgenes **[5]**, (ii) co-transformation

of the desired genes in a single step **[9-10]**, and (iii) the random **[11]** or site-specific **[12]** retransformation of lines already containing a transgene. Recently an interesting strategy for facilitating the cotransformation of un-linked genes has been reported. It consists on the co-bombardement of the constructs with tungsten or gold particles to generate transgenic plants carrying all the input genes with high cotransformation efficiency, but at the cost of generating complex transgenic loci, which are believed to be prone to instability and silencing **[13-14]**. This strategy has been successfully employed in the biofortification of soybean, rice **[13]** and maize, including different vitamins **[16]** and elements of the carotenoid biosynthetic pathway **[17]**. If the proteins or enzymes need to be expressed in a balanced molar ratio, alternative strategies are (i) the expression of polycistronic transcripts in the nucleocytosolic compartment including intrinsic cleavage sites **[18-21]**, (ii) the engineering of internal ribosome binding sites **[18-19]**, or (iii) the transformation of the chloroplast genome with synthetic multigene operons **[22-25]**, which is still restricted to a small number of species **[26]**. The integration of several transgenes into the same locus in the target plant genome brings important advantages: (i) a single marker gene is needed for the selection of the whole set of transgenes (ii) typically a smaller number of plants needs to be generated as all genes will be jointly inherited **[27-28]**, and (iii) it will simplify the IP registration and the legislative authorization as only one transformation event has to be described **[29]**.

The emergent discipline of Synthetic Biology (SB), which aims at applying engineering principles to biological design, strongly relies on the ability to introduce multiple transgenes within a target genome (often referred as “chassis” in SB nomenclature). Although random co-integration of unlinked genes by particle bombardment is an excellent strategy for the transfer of multiple traits in a plant genome, some sophisticated genetic designs are likely to ultimately require the construction of well-defined (rationally) designed genetic structures and, eventually, the orderly combination of multiple genetic elements within a single lineal DNA string. Consequently, there is a growing interest among plant biotechnologists in the development of such DNA assembly methods for SB and metabolic engineering, preferably (but not exclusively) through the adaptation of binary vectors as final destination vectors for multigenic constructs.

Agrobacterium-mediated linked co-transformation involves the transfer to the plant of large DNA constructs harboring large amounts of information. This strategy is not devoid of anticipated difficulties as (i) the need of high-capacity binary vectors such as BiBACs or TACs that are usually low-copy and therefore difficult to handle, (ii) the technical complexity involving the assembly of several genes in the same vector with the everyday methods, and (iii) the chances of construct instability in some *Agrobacterium tumefaciens* strains and/or in the plant genome. There is an increasing demand to overcome those technical

hurdles in order to achieve complex plant genetic engineering goals. In the context of SB, the development of new, highly efficient DNA assembly technologies is already enabling the adoption of standardized DNA building protocols and facilitating the assembly of multiple genes, linked together in a lineal DNA molecule **[30]**. However, further efforts are required in this direction, especially in the field of Plant Biotechnology. These efforts should involve not only the development of increasingly efficient DNA assembly methods, but also the design of modular frameworks, the definition and adoption of common standards, and ultimately, the development of structural scaffolds that ensure genomic stability. Recent reviews provide an overview on the available methods for transgene stacking **[31-32]**, on the analysis of the different strategies for multigene transformation in plants **[33]**, and on the strategies for heterologous co-expression of proteins under the control of single promoters **[34]**. During the preparation of this introduction, a review analyzing different advanced genetic tools for Plant Biotechnology, including the description in broad strokes of multigene engineering tools, has been published **[35]**. In this introduction chapter we will discuss the advantages and disadvantages most recent strategies for multigene assembly in plants. A simple graphical overview of the function of the discussed methods is depicted in Figure 1.

1.1. Adapt or die. Multigene engineering is possible with traditional Type II cloning enzymes.

Traditional Type II restriction enzymes are widely used for making combinatorial and multigene engineering in the prokaryote community especially due to the success of standardized platforms such as BioBricks or BglBricks **[36-37]**. These standards have some disadvantages for plant biotechnologists because they are not scar benign (leaving DNA fingerprints from the cloning strategy), said to be relatively efficient **[37]** and the actual collections do not include neither vectors nor plant parts. On top of everything, current assembly methods based on Type II restriction enzymes are unsuitable for the assembly of multiple DNA fragments in one step and obligates the use of a sequential cloning scheme **[30]**. In spite of the disadvantages that traditional cloning may have for the assembly of large constructs, researchers have succeed in building multiple TUs using non-standardized alternatives **[38-39]**. The main disadvantage is that it involves lots of sub-cloning and intermediate steps to produce a single-use non-reusable solution.

Two type of standardized solutions have been described **[28, 40]**, which employ iterative digestion and ligation reactions to combine TUs into a multigene structure. Each type of solution uses a pair of isocaudomer restriction enzymes (*Xba*I/*Avr*II and *Asc*I/*Mlu*I) to permit iterative cloning, restoring the original enzymes after ligation of the assembled units. Using

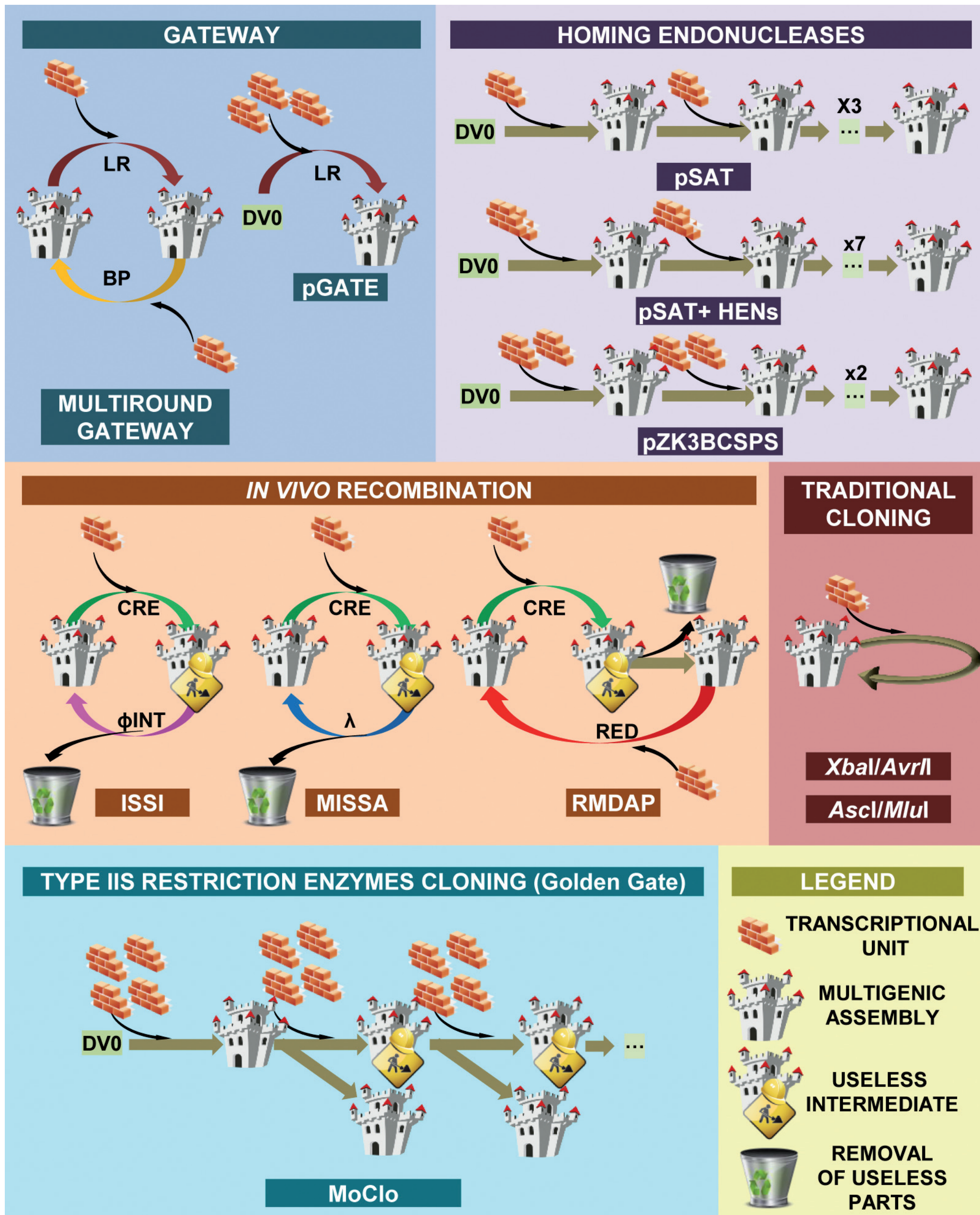


Figure 1. Simple graphical overview of the multigene assembly methods described in this chapter.

LR and BP are Gateway LR and BP Reactions; DV is Destination Vector; x3, x7, x2 are the number of times an operation can be repeated; CRE is Cre-loxP recombineering; ϕINT is the bacteriophage ΦC31 integrase; λ is phage λ site-specific recombination; RED is Red-mediated recombineering.

these simple systems, up to five transgenes were combined in one binary vector. The set of vectors described by Kuroda *et al.* has been used to assemble and transform several TUs into different plant species **[41-43]**.

1.2. New in town (for Plant Biotech): don't let them out of your sight!

There are a number of assembly methods that, being suitable for modular cloning and multigene engineering, have not been adapted yet for Plant Biotechnology **[44-47]**. Particularly interesting are those in the so-called the CBA group (for chew-back and anneal, the two basic reactions for these assembly systems), which include USER **[48]**, In-Fusion **[49]**, and Gibson Assembly **[50-51]**.

The most remarkable of the three, according to its larger assembly capacity and extensive use in different fields, is Gibson Isothermal Assembly. Gibson Assembly permits the simultaneous combination of numerous DNA fragments provided that they share a 20-40 bp overlapping region. Briefly, the process starts with the T5 exonuclease-mediated generation of single-stranded overhangs in the overlapping regions, followed by the pair-base annealing of the resulting overhangs. Junctions are then sealed with the action of the Phusion polymerase and the T4 DNA ligase. Using a simple cocktail that includes all three enzymes, multiple DNA fragments can be ligated in a 15 minutes to 1 hour isothermal 50ºC reaction.

Remarkable features of Gibson and other CBA methods are that they are sequence-independent and permit the simultaneous assembly of many PCR fragments in a one-step-one-pot reaction, thus simplifying and speeding the process. Important disadvantages are that they are neither modular nor suitable for the assembly of fragments with repetitive sequences.

1.3. Revisiting site-specific recombination tools for multigene engineering.

The use of site-specific recombinases for multigene engineering overcomes the main issue concerning the use of restriction enzymes, which is the presence of internal target sites. Different homologous recombination strategies have been proposed as suitable standardized methods for multigene engineering, including Gateway and different *in vivo* approaches.

1.3.1. Giving another turn of the screw to Gateway technology.

Gateway technology is a highly efficient and versatile ligase-free cloning system, which has been easily adapted for high-throughput approaches. Gateway is the preferred assembly technology for many plant biotechnologists especially due to the work of many labs that have developed interesting ready-to-use toolkits for different proposes and plant species **[52-57]**. Briefly, this DNA assembly method relies on the use of the λ phage site-specific recombination reactions and two Invitrogen proprietary enzyme mixtures named BPClonase and LRClonase. Any sequence flanked by *att*B sites can be cloned into a Donor Vector with *att*P sites resulting on an Entry Clone in which the sequence of interest is flanked by *att*L sites (BP reaction). Any Entry Clone containing *att*L sites can be transferred into a Destination Vector with compatible *att*R sites, resulting in an Expression Clone (LR reaction).

An evolution to the Gateway Cloning, based on the incorporation of modified *att*L and *attR* sites is the MutiSite Gateway Technology. It permits the simultaneous introduction of up to four DNA fragments to the destination vector in a single reaction and enables the construction of modular DNA parts collection to facilitate the combinatorial assembly of genetic elements **[53, 58-59]**. Despite the fact that the use of MutiSite Gateway enhanced the versatility of the initial series of vectors, the inability to produce reusable units burdens the applicability of this technology for multigene engineering.

A first Gateway-based platform specifically designed for multigene engineering, named MultiRound Gateway (MRG), was established in 2006 **[60]**. In MRG, single TUs are initially built using traditional digestion/ligation procedures. Next, multiple TUs can be sequentially and indefinitely assembled into a set of Gateway compatible destination vectors in consecutive LR reactions, by combining different *att*L and *att*R sites and two selectable markers. This initial platform was used in a later work where four TUs, all comprising the same promoter and terminator regions, were co-transformed into *Arabidopsis thaliana* to engineer a regulatory pathway for K^+ uptake **[61]**.

An improved version of the MRG technology was developed later (Figure 2) included a TAC based destination vector and some recommendations to facilitate larger assemblies **[62]**. A final construct of eight genes plus the selection marker was built using this method. The whole assembly process involved seven LR reactions, and led to the stable expression of all genes in *Nicotiana tabacum* plants for at least two generations.

A number of simplified platforms combine traditional cloning for the construction of the basic

expression units and take advantage of MultiSite Gateway for the multigene assembly step **[4, 58, 63]**. The pGATE platform **[63]**, that permits the assembly of a maximum of three TUs in a one LR reaction, is very easy to use since developers have made an effort to simplify and standardize the building scheme.

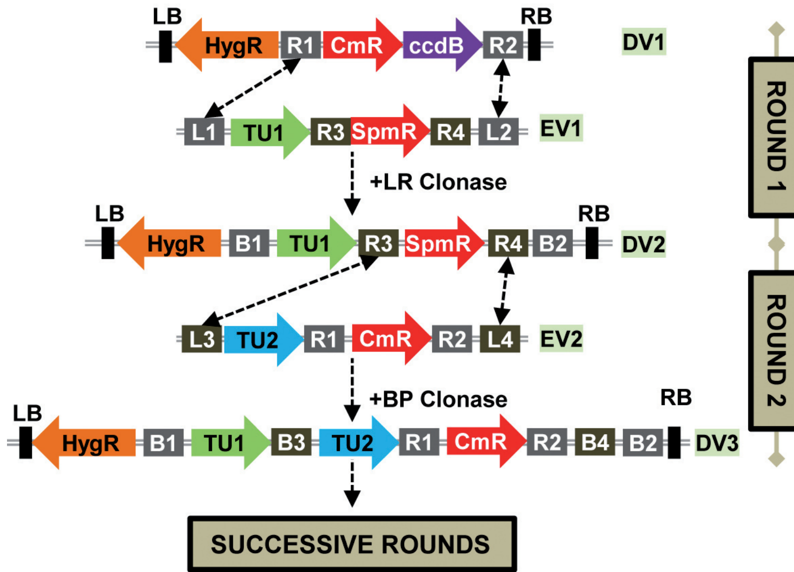


Figure 2. Detailed diagram of MultiRound Gateway (MRG).

The two first rounds of assembly are represented. The first round incorporates TU1 by LR recombination between the *att*L1 and *att*L2 of the DV1 and the *att*R1 and *att*R2 of the EV1. The second round incorporates TU2 by an BP recombination between the *att*R3 and *att*R4 sites of the DV2 and the *att*L3 and *att*L4 of the EV2. The loop continues with a third round similar to Round 1. Different bacterial resistance genes (CmR and SpmR) are used in the successive rounds to ensure counterselection. LB, T-DNA left border; RB, T-DNA right border; HygR, Hygromicin resistance; CmR, Chloramphenicol resistance; SpmR, Spectinomycin resistance; ccdB, bacterial lethal gene; DV, Destination Vector; EV, Entry Vector; B1, B2, B3, B4, *att*B1-2-3-4 sites; R1, R2, R3, R4, *att*R1-2-3-4 sites; L1, L2, L3, L4, *att*L1-2-3-4 sites. TU1 and TU2 are two transcriptional units.

1.3.2. In vivo recombination also makes it possible.

Going back a decade, a first attempt to develop a multigene assembly platform exploited the Cre–loxP recombination and homing endonucleases I-*Sce*I and PI-*Sce*I, to create endless alternating cycles with both elements **[64]**. A similar solution named 'Iterative Site Specific Integration' (ISSI) was described a couple of years later. ISSI makes use of the bacteriophage ΦC31 integrase instead of the Cre–loxP system **[65]**. Although the ISSI system was not

specifically created for Plant Biotechnology, it has been exploited as a tool for the iterative integration of transgenic DNA after *Agrobacterium*-mediated floral-dip transformation in *A. thaliana*, opening a new way for the site-directed multigene insertion by retransformation **[66-67]**.

A powerful method, according to the size of the achieved constructs, is the multiple-round ***in vivo*** site-specific assembly system (MISSA) **[68]**. MISSA combines inducible Cre recombination and phage λ site-specific recombination with conjugational transfer in two different recipient strains (Figure 3). The larger multigenic constructs achieved with this technology combined up to 9 TUs plus several matrix attachment regions (MARs), assembled together into BiBAC and TAC destination vectors. A particularly interesting construct comprised four genes involved in the SOS pathway for salt tolerance (SOS1/2/3 and CBL10) and included also four MARs intercalated among the TUs to increase the stability of the whole assembly. The construct was transformed into *Festuca arundinacea* and the analysis of the resulting transgenic plants suggests that the co-expression of the four genes cooperatively enhanced salt tolerance **[69]**.

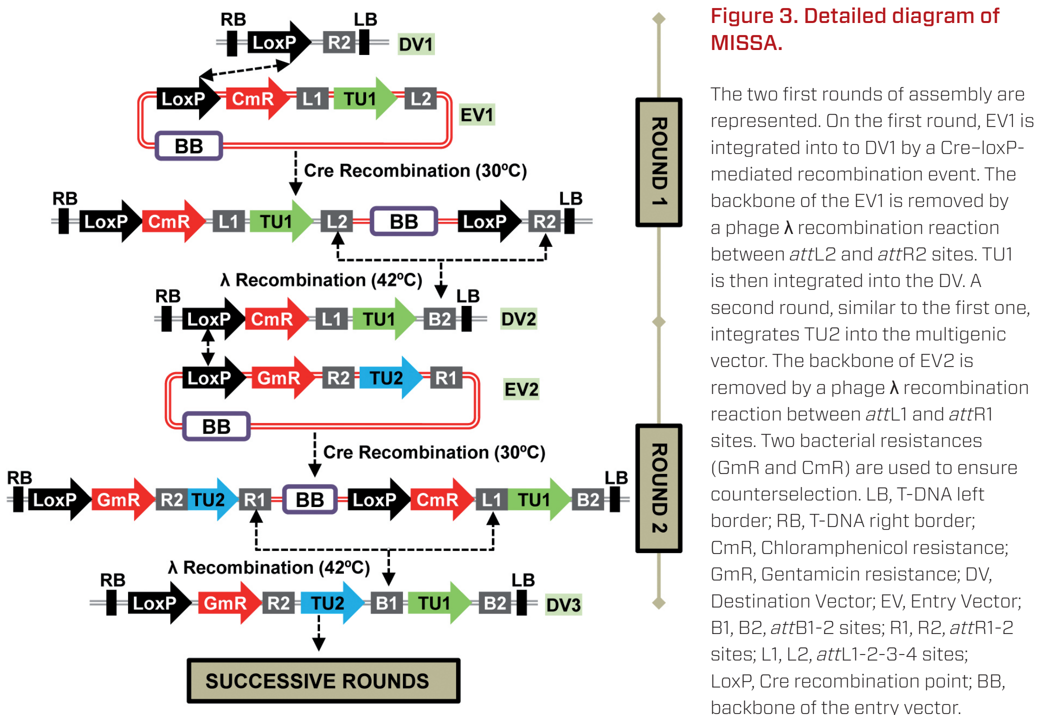


Figure 3. Detailed diagram of MISSA.

The two first rounds of assembly are represented. On the first round, EV1 is integrated into to DV1 by a Cre–loxP-mediated recombination event. The backbone of the EV1 is removed by a phage λ recombination reaction between *att*L2 and *att*R2 sites. TU1 is then integrated into the DV. A second round, similar to the first one, integrates TU2 into the multigenic vector. The backbone of EV2 is removed by a phage λ recombination reaction between *att*L1 and *att*R1 sites. Two bacterial resistances (GmR and CmR) are used to ensure counterselection. LB, T-DNA left border; RB, T-DNA right border; CmR, Chloramphenicol resistance; GmR, Gentamicin resistance; DV, Destination Vector; EV, Entry Vector; B1, B2, *att*B1-2 sites; R1, R2, *att*R1-2 sites; L1, L2, *att*L1-2-3-4 sites; LoxP, Cre recombination point; BB, backbone of the entry vector.

A last method applying *in vivo* recombination for multigene engineering is the Recombination-assisted Multifunctional DNA Assembly Platform (RMDAP) **[70]**. It uses Gateway technology for the TU construction and Cre–loxP for the first round of TU assembly. Furthermore, RMDAP includes the option of introducing new fragment(s) using Red-mediated recombineering **[71]**, which is an absolute novelty among all multigene engineering platforms. RMDAP includes a basic ready-to-use toolbox that provides all requirements for the construction of any vector of interest.

1.4. The use of Zinc Finger Nucleases and Homing Endonucleases.

Homing endonucleases (HENs) are restriction enzymes characterized by long recognition sequences, which make them convenient tools for multigene engineering. Binary plasmids constructed with polylinkers containing HEN target sites can easily accept large multigene constructs, as HEN target sites very rarely appear within the sequences to be combined. Following the same rationale, zinc finger nucleases (ZFNs), which are artificial restriction enzymes resulting from the fusion of the nonspecific cleavage domain of the Type IIS *Fok*I enzyme and an engineered zinc finger DNA domain, are also powerful tools for cloning. ZFNs are not only interesting because of their longer DNA recognition sequence but also because they can be tailored to digest a given sequence of interest **[72-73]**.

The use of HENs for multigene engineering was first introduced in the pAUX vector series **[74]**, which included a multicloning site (MCS) with 5 HENs among other target sites for classic Type II restriction enzymes. This was the basis for the development of the pSAT vector series, a first example of modular cloning vectors for plant transformation, which established a small collection of standardized DNA parts (basically promoters, terminators and reporter genes) **[75-76]**.

Since first published in 2005, the pSAT vectors have been used in several works **[77-79]**, but the ability to incorporate multiple genes using HENs in the same binary vector has not been fully exploited (authors state that this is possibly due to the very small number of commercially available HENs **[73]**). To overcome this limitation and to broaden the assembly capacity of the system, a new version of this technology was reported in 2012 **[73]**. The new system incorporated three ZFNs sites as additional entry points, allowing the assembly of up to nine transgenes following the rules of the system. Novel ZFNs can be easily incorporated to expand the system beyond the current assembly capacity. The building process begins with the individual cloning of the functional expression cassettes into the nine pSAT plasmids

which are flanked by different HENs and ZFNs. The flanking sites determine the position of each TU within the final construct. TUs are assembled into pRCS11.1, a destination binary vector that carries a polylinker with HENs and ZFNs recognition sites (Figure 4). To test the system, a nine-gene expression vector was built and successfully transformed into *A. thaliana*. Interesting features of this building system are (i) the possibility to assemble TUs in inverted orientations using the same standardized DNA parts, and (ii) the ability to remove and replace TUs during the construction process.

The list of multigene assembly methodologies based in rare cutters is completed with a system developed by Fujisawa et al **[80]**, based on HENs. Their strategy makes use of the rare cutter *Sfi*I to build up to two TUs in a single vector. Later, TUs are successively assembled using HENs into the binary destination vector pZK3BCSPS, which can harbor a maximum of eight genes.

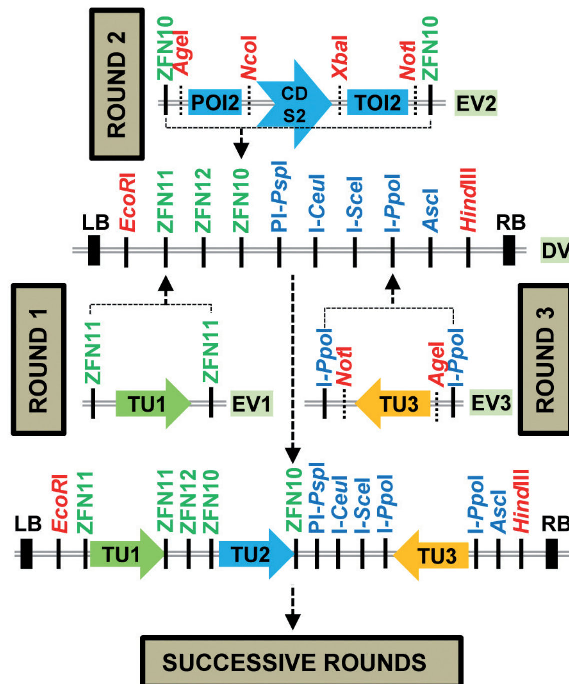


Figure 4. Detailed diagram of pSAT+ZFNs vectors.

The modular cloning scheme of a TU is exemplified by the combination of POI2-CDS2-TOI2 to build TU2 through the unique *Age*I, *Nco*I, *Xba*I and *Not*I sites. The reverse orientation assembly of a TU is exemplified by TU3. The DV incorporates restriction sites for five HENs (in blue) and four ZFNs (in green) so up to nine expression cassettes can be transferred into the T-DNA, in successive rounds of assembly. The incorporation of three TUS using ZFNs (TU1 and TU2) and HENs (TU3) is represented. LB, T-DNA left border; RB, T-DNA right border; DV, Destination Vector; EV, Entry Vector; POI, promoter of interest; CDS, coding sequence of interest; TOI, Terminator of interest.

1.5. Building multigenic constructs using Type IIS restriction enzymes.

Golden Gate (GG) is an ingenious cloning system that facilitates the assembly of multiple DNA fragments in a one-tube-one-reaction fashion with extremely high efficiency **[81-82]**. It is based on the ability of Type IIS restriction enzymes to cut at a defined distance from their non-palindromic asymmetric recognition sites, making possible scarless and combinatorial assembles in a one-step reaction (see Figure 5 for a detailed overview on the functioning of GG). Despite GG's short life, its high efficiency and versatility has secured many GG users from different fields, who have made this cloning system their *motto* **[83-88]**. As originally described, GG permits the simultaneous assembly up to 9 DNA parts in one step. Despite the impressive cloning capacity, the resulting units contain no more target sites for Type IIS restriction enzymes and therefore cannot be used for further constructs.

Chapter 2 in this thesis describes the development of GoldenBraid, a DNA assembly strategy that enables the building of multigene structures using a GG-based cloning schema. In parallel to the development of GoldenBraid in our lab, GG developers created MoClo, a different strategy that also enables multigene engineering based in GG. MoClo **[89]** and GoldenBraid **[90-91]** are therefore two solutions for the same problem. Both are modular systems that facilitate the conversion of multipartite constructs into reusable composite parts. These two strategies make use of a second Type IIS restriction enzyme flanking the assembly point so the composite parts can be released and used in subsequent assemblies. The basis of MoClo are discussed next in this introduction chapter, whereas the development of GB and its updated version GB2.0 will be described in detail in Chapters 1 and 2.

1.5.1. MoClo: a modular cloning system for multigenic constructs.

MoClo proposes a strategy for the combinatorial assembly of multigenic constructs based on the use of three Type IIS enzymes (*Bsa*I, *BsmB*I and *Bpi*I) and three basic assembly levels, named 0, 1 and 2.

The purpose of Level 0 is the generation of the basic DNA parts. In this step, the sequences of interest are PCR-amplified to include the appropriate flanking cleavage sites. The position of the DNA part within the TU determines the destination vector to be chosen (among 10 different available). Level 0 parts are used for the assembly of TUs in a *Bsa*I GG reaction using

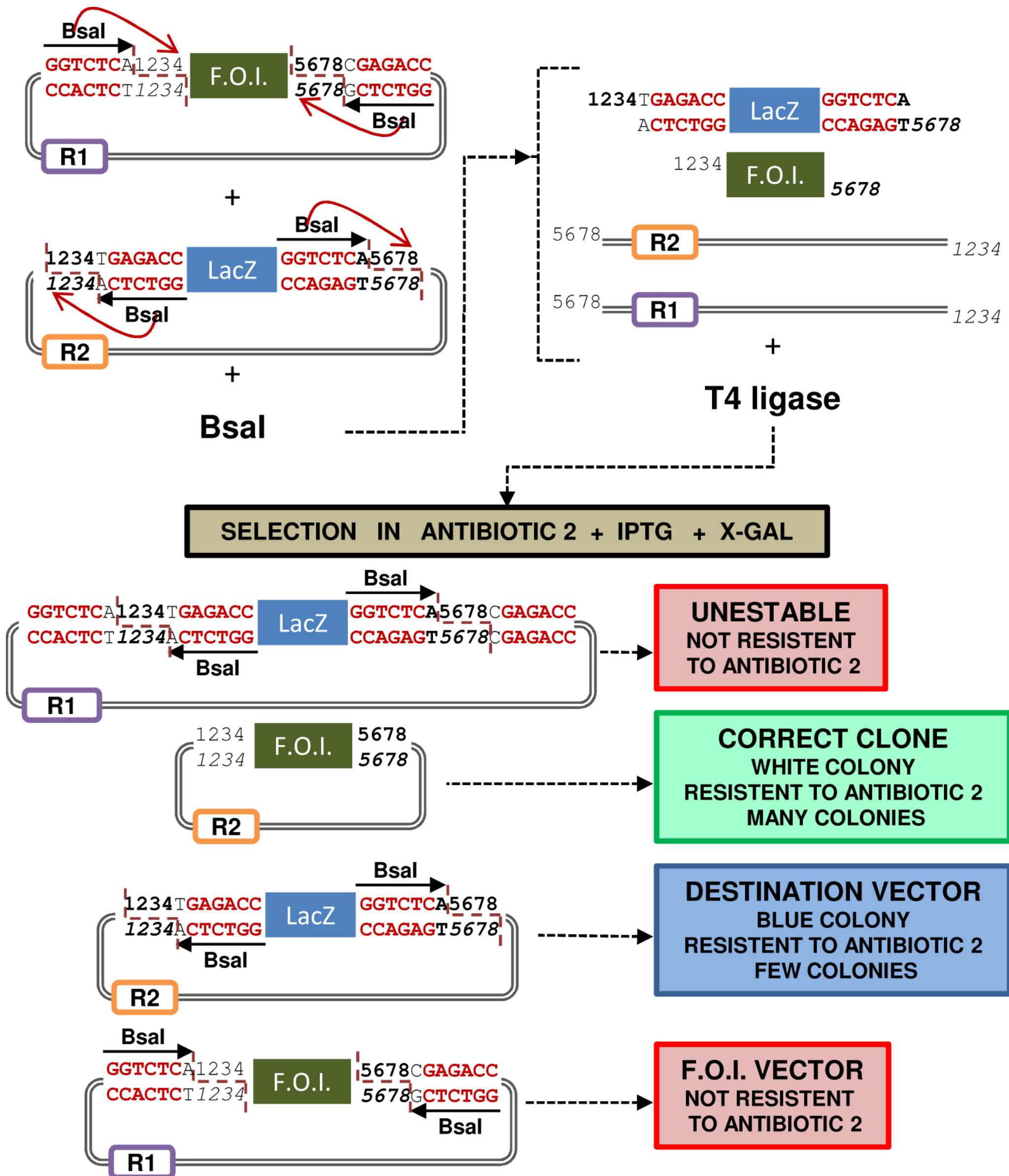


Figure 5. Functioning of the Golden Gate Assembly System.

An entry clone containing the DNA of interest (F.O.I.) and the expression vector containing the LacZ selection marker are mixed in one tube together with *Bsa*I and ligase. Of the 4 possible ligation products, two of them will not resist to the selection in antibiotic 2. The intact vector will produce blue colonies. Only the correct clone is stable and will produce white colonies. Numbers 1 to 8 denote any nucleotide of choice, and numbers in italics denote the complementary nucleotides. R1 and R2 are two antibiotic resistance genes. Adapted from Engler *et al.* [81]

any of the 14 different Level 1 destination vectors. This will fix the position and orientation of the assembled unit in the multigenic construct (Figure 6A and B).

At this point, MoClo proposes the user two alternative paths to continue with the assembly system. The simplest one is used when less than 6 TUs are to be combined resulting in a Level 2-1 construct. Those are end terminal products and do not permit the incorporation of additional TUs (Figure 6C). In order to enable the further growth of the construct, the second path has to be used. End linkers consisting on a blue or red selection cassette (as defined by the authors) can be linked after the last TU on the level 2 assembly step, making possible the creation of a Level 2i-1, which permits the additional incorporation of more TUs in successive reactions (Figure 6D).

As a result of all these levels and linkers, the final MoClo toolkit is composed of 28 different vectors and 21 end-linkers. Using this set of vectors, a 33kb construct containing 11 TUs was built in only three successive cloning steps. The functionality of the construct was tested by transient transformation of in *Nicotiana benthamiana* **[89]**. MoClo released a second version in 2012. This new version incorporated a loop (named levels M and P) into their cloning design. This resulted in an improvement in MoClo performance for subsequent multipartite assemblies that are functionally similar to our GoldenBraid α and Ω levels (see Chapter 1 in this thesis) but allowed the subcloning of more than two TUs at a time **[92]**. To fulfill this, 14 additional end-likers had to be created, making the toolbox even more complex. With this solution up to 17 genes were assembled in one vector, illustrating the power of this cloning method.

2. The fine-control of the genetic expression.

The new DNA assembly technologies are paving the way for the accomplishment of increasingly sophisticated genetic designs. The improved multigene assembly capacity offers the technological basis for applying Synthetic Biology principles to Plant Biotechnology. Synthetic Biology was born with the goal of engineering or “wiring” biological circuits for controlling gene expression. The engineering-driven approaches started on the 70s but have rapidly progressed since the description of the early examples in Synthetic Biology. The development in year 2000 of a bacterial toggle-switch was the first breakthrough. This bistable synthetic gene network was constructed by the connection of two mutually inhibitory transcriptional repressors **[93]**. In the same year, an artificial oscillating network

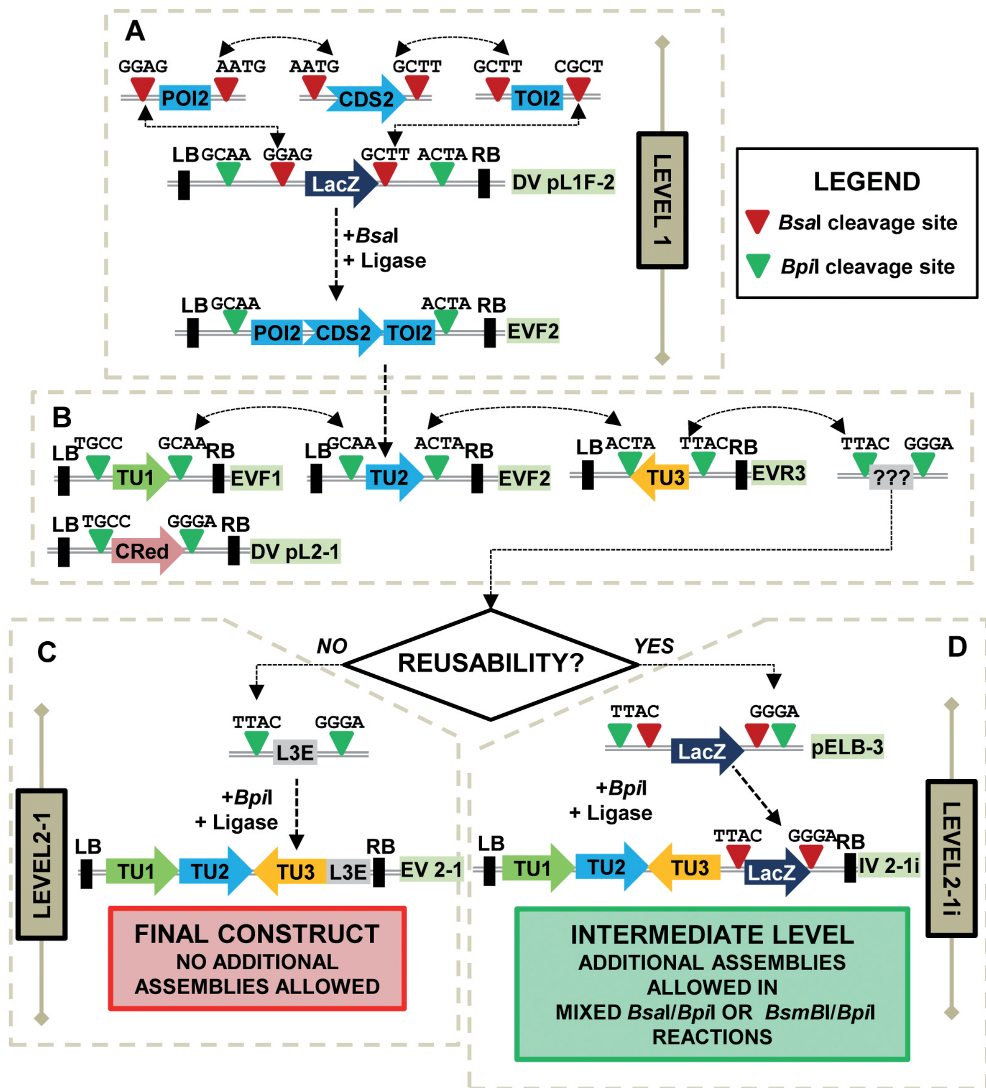


Figure 6. Detailed diagram of MoClo.

(A) The modular cloning scheme of a TU into Level1 plasmids is exemplified by the combination of POI2-CDS2-TOI2 to build TU2 into the pL1F-2 vector using a *Bsa*I GG Reaction. **(B)** The multipartite assembly of three TUs into the pL2-1 vector is depicted. Up to 7 TUs can be assembled in one step therefore and end-linker has to be chosen among two different possibilities. **(C)** If the multipartite assembly is the final construct and no more TUs have to be incorporated to the binary vector, the pL3E end-linker has to be chosen (decision represented with "???" on the scheme). This will close the assembly and will result in a multigenic level 2-1 construct with 3 TUs after a *Bpi*I GG reaction. **(D)** If the construct has to be further reused to incorporate additional TUs, the pELB-3 end-linker should be incorporated to the *Bpi*I assembly reaction. This will keep the Level 2-1i construct opened and additional TUs can be incorporated in a *Bsa*I/*Bpi*I (or *BsmB*I/*Bpi*I, depending on the intermediate level) assembly reaction. *Bpi*I and *Bsa*I recognition sequences are marked with green and red triangles respectively. LB, T-DNA left border; RB, T-DNA right border; DV, Destination Vector; EVF, Entry Vector in forward orientation; EVR, Entry Vector in reverse orientation; IV, intermediate construct; POI, promoter of interest; CDS, coding sequence of interest; TOI, Terminator of interest. LacZ and CRed are blue and red selection markers, respectively.

was engineered. It periodically induces the synthesis of GFP as readout of its state **[94]**. Many interesting developments followed these initial examples, including memory devices **[95-96]**, cell-to-cell communications **[97]**, biosensors **[98]** or the optimization of biosynthetic pathways for cost-effective and high-yield microbial production of compounds of interest **[99-100]**.

2.1. Logic Gates.

To exploit the full potential of SB in plants, it is mandatory to develop higher order modular devices like sensors, logic gates and memory switches. These genetic devices are analogous to the components of electronic circuits, and constitute the building blocks of the artificial gene networks required to reprogram gene expression in plants. One of the indispensable elements for the circuit-like connectivity are logic gates, that describe the biologic regulation schemes with mathematical models and help to control the genetic expression *à la carte* **[101]**.

Logic gates are elementary building blocks of a digital circuit that implement a logic operation on one or more inputs and produce a single output. There are sixteen different binary operations, also called Boolean functions (see Figure 7) and most of them can be assembled by the combination of simple one-input one-output operations. Logic gates are important tools to combine different inputs can be directly used to generate new genetic expression profiles (i.e. a combined response to two inducers that is not achievable using the standard promoters) or to be the basis of the construction of complex genetic circuits.

INPUT A	INPUT B	TRUE	FALSE	PROPOSITION		NOT			
				A	B	A	B	OR	AND
0	**0**	1	0	0	0	1	1	0	0
1	**0**	1	0	1	0	0	1	1	0
0	**1**	1	0	0	1	1	0	1	0
1	**1**	1	0	1	1	0	0	1	1

INPUT A	INPUT B	NAND	NOR	A IMPLY B	B IMPLY A	A NIMPLY B	B NIMPLY A	XOR	XNOR
0	**0**	1	1	1	1	0	0	0	1
1	**0**	1	0	0	1	1	0	1	0
0	**1**	1	0	1	0	0	1	1	0
1	**1**	0	0	1	1	0	0	0	1

Figure 7. Truth tables of the Boolean Logic Gates.

Truth Tables are composed of one column for each input variable and one column for each of the possible results of the logical operations included in the table. Each row of the truth table contains one possible configuration of the input and the result of the operation for those values (1 is for *true* and 0 is for *false*).

Several research groups have engineered and characterized the full set of logic operations in bacteria, yeast of mammalian cells **[102-104]**. The control of the transcription is the most usual mechanism to design logic gates **[105-107]** translation control elements are also used **[108-110]**. Two recent publications included a set of logic gates that could be engineered by DNA flipping in bacterial cells **[104, 111]**, representing a new paradigm in synthetic circuit design **[112]**. Once developed, these components are of outstanding interest for the further engineering of circuits **[113]**, since they are agnostic to towards their final use and therefore they can be combined to create increasingly complex networks. This is a complete new field for plant biotechnologists. In Chapter 3 of this thesis, we describe the firsts steps towards the characterization of an initial set of logic gates adapted to the plant cells.

/OB

OBJECTIVES

Objectives

The overall objective of this work was to establish a technological framework for the standardized assembly of DNA parts and to create a basic collection of modular building blocks for Synthetic Biology in plants. To fulfill this general objective, the following specific objectives have been developed:

01. The design and development of a modular DNA assembly method for multigene engineering in plants based on the use of Type IIS restriction enzymes.

02. The definition of an assembly standard and the generation of a collection of interchangeable DNA parts for genetic engineering in plants that conform to this standard.

03. The development of modular and orthogonal genetic logic gates adapted to Plant Biotechnology, using the assembly standard and some of the DNA part created in the objective 2.

/C1

CHAPTER 1

GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules

Chapter 1: GoldenBraid: an interactive cloning system for standardized assembly of reusable genetic modules.

1. Introduction

Synthetic Biology adapts the general engineering principle of assembling standard components, dating back to the Industrial Revolution, to biological components. This discipline aims at the design of artificial living forms displaying new traits not existing in nature **[114-115]**. This objective can be pursued following a bottom-up strategy, by creating new living forms from its basic components; however, a more straightforward option consists of integrating new genetic circuits within the genome of a current living organism or "chassis". In this top-down tinkering approach, the construction of new versions of an existing organism can be conducted following a modular hierarchical approach, by combining well defined basic DNA "parts" (e.g. promoters, coding sequences, terminators, etc.) into genetic devices (e.g. transcriptional units), those devices into basic genetic modules (e.g. biochemical pathways, genetic circuits, etc.), and those into higher order modules, which integrated in a natural genome or "chassis" will configure a redesigned organism displaying new traits. Modularity is not only an engineering strategy; multiple high-throughput genetic interaction studies have provided substantial evidence of modularity in the genetic organization of cellular systems **[116]**. In view of this fundamental modular structure of genetic networks, many key design solutions are likely to involve intermediate hierarchical levels, entailing structures ranging from a few devices to complex modules and comprising between five and a few hundred basic genetic parts. In recent years the ability to manufacture synthetic DNA molecules has increased exponentially. Chemical synthesis ordinarily produces *de novo* sequences in the size range of a genetic "part" (up to 0.5 -5 Kb) **[117-118]**. On the opposite side, increasingly efficient homologous recombination methods have enormously facilitated the assembly of large DNA sequences up to the genome range **[30]**, with the synthesis of a complete bacterial genome serving as best example **[50, 119]**. Despite these technical advances, many critical engineering issues as the exhaustive characterization of new genetic modules, their re-adaptation for additional purposes or their combination with other devices to produce combined traits still require from increasingly efficient and versatile DNA assembly methods operating at intermediate range. Moreover,

to facilitate engineering at this level, basic pieces (parts) need to be assembled following standard rules, which can be applied independently of the identity of the parts. Standardization is therefore a crucial feature that allows the exchange of pieces among laboratories and facilitates automation. Standardization also favors reusability, as any standard pieces can be exchanged for assembling different constructs following common rules of assembly.

When adopting standardization, it is highly preferable that the rules of assembly are kept to a minimum. Simplicity facilitates the adoption of the technology by the potential users, reduces the elements in the engineer's tool box and simplifies the automation process. The maximum expression of simplicity in assembly standards is idempotency, occurring when any new composite part can be assembled following the same rules used to generate its original components. Idempotency is at the basis of the success of the BioBricks, a community effort to build a standardized collection of genetic parts for Synthetic Biology **[120]**. BioBricks standards are binary assembly rules where two pieces flanked by a set of restriction sites, result, upon assembly, in a composite piece flanked by identical restriction sites than their predecessors. The simplicity of the idempotency has boosted the interest in BioBricks standards, which have evolved to deal with engineering drawbacks as those derived from the presence of assembly scars **[36]**.

BioBricks assemblies are strictly binary, meaning that only two elements can be assembled together in each assembly step. This feature slows down the engineering process, this being apparently an obligate penalty for idempotency. Oppositely, multipartite systems have been developed allowing the assembly of multiple DNA fragments in a single step. Among them, Golden Gate, a cloning system based on the use of Type IIS restriction enzymes, has a number of interesting features for operating at the level of genetic devices and modules **[81-82]**. Unlike other multipartite methods, which are often based on overlapping flanks and in vitro recombination, Golden Gate cloning does not require PCR amplification of each part prior to the assembly. Since amplification of self-complementary or repetitive parts can be problematic, Golden Gate is more permissive than other methods for the assembly of repetitive elements. Despite being based on restriction/ligation, its all-in-one-tube design avoids inconvenient gel extraction procedures that often reduce cloning efficiency; most interestingly, it allows seamless assembly by careful design of the restriction sites. This feature is particularly important when DNA fragments comprise coding sequences for sensitive applications (e.g. in the design of therapeutic proteins). Despite its obvious advantages, Golden Gate, as many multipartite systems, is limited in standardization and reusability. Hence, Golden Gate multipartite assemblies, as originally designed, cannot be reused to generate higher order devices and modules following standardized rules of assembly, limiting its use in Synthetic Biology.

Here we present GoldenBraid, a new modular assembly system that allows the binary combination of multipartite assemblies using an extremely simple set of rules, very close to idempotency. GoldenBraid makes use of the multipartite Golden Gate cloning method to generate a modular assembly of standardized basic parts, which are then incorporated to a double loop ("braid") cloning design that allows binary assembly of multipartite constructs. In this way, GoldenBraid technology enables the standardization of Golden Gate for its use in Synthetic Biology. Moreover, this is achieved with a small toolbox consisting of only four destination plasmids and a limited number of assembly rules.

Multigene engineering has an enormous potential in crop design, as for metabolic engineering, biofortification, molecular farming or for combination of traits of agronomic value via gene stacking **[33]**. Plant Synthetic Biology is a nascent discipline where the use of standard assembly rules has not yet rooted, and there is therefore room for efficient and innovative assembly methods to be adopted by the plant research community. Based on the features of GoldenBraid, here we propose its adoption as a common assembly standard for Plant Synthetic Biology. To substantiate this proposal we show here three examples of GoldenBraid-assisted multigene engineering in plants. In a first example we demonstrate the advantages of in-*cis* multigenic designs for *Agrobacterium*-mediated transient co-transformation. In a second example, we show the versatility of the system to assay recombinant antibody expression in a combinatorial way. Finally, we combine different modules to produce two alternative 14.3 Kb constructs each involving the assembly of 19 basic parts grouped in five different transcriptional units.

2. Results

2.1. Part standardization and multipartite assembly of simple devices.

GoldenBraid is an adaptation of Golden Gate to Synthetic Biology. Golden Gate is a multipartite assembly system based on the use of Type IIS restriction enzymes. These enzymes digest DNA at a defined distance few nucleotides away from its recognition site, not requiring any specific sequence in the actual cleavage site, and often leaving a short overhang. This feature makes them extremely useful in seamless cloning strategies: by carefully positioning recognition and digestion sites in opposite directions in entry and destination vectors, it is possible to design and obtain multipartite assemblies where all recognition sites in the final expression vectors have disappeared. Since there are no sequence requirements in the cleavage sites, these can be user-defined, and therefore accommodated to serve as assembly boundaries

for standard DNA parts. Following this rationale, we initially considered three categories of basic parts, namely promoters (PROM), coding sequences (CDS) and terminators (TERM). All parts are cloned as *Bsa*I fragments in entry clones. The inclusion in a category is defined by the flanking *Bsa*I digestion sites. A schematic view of a standardized multipartite assembly of a transcriptional unit is depicted in Figure 8. To facilitate the interpretation, we gave a label to each 4 bp cleavage site producing the corresponding overhang (e.g. numbers 1, 2, 3, IV, etc., to those sites digested by *Bsa*I enzyme). Therefore a promoter is a "part" flanked by sites 1 and IV, whereas CDSs are flanked by sites IV and III, and terminators are flanked by sites III and 2. In our approach, nucleotide boundaries were conveniently fixed to accommodate the nature/sequence of the different parts: site IV, defining PROM-CDS boundary, was designed GATG, conveniently containing an ATG start codon, whereas site III, that forms CDS-TERM boundary was designed to contain a TGA stop codon (namely TGAG). Parts are ordinarily created by PCR amplification of suitable templates, adding appropriate *Bsa*I extensions to the primers. Once amplified, parts can be used directly as PCR fragments and/or cloned and stored in a collection for future assemblies.

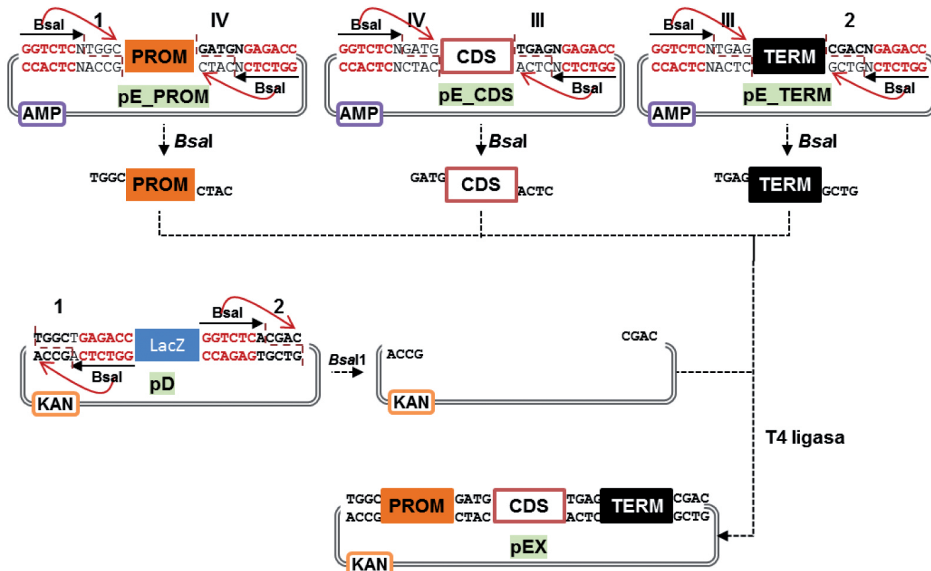


Figure 8. Part standardization and multipartite assembly of single devices.

PCR products of entry plasmids (pE) containing basic parts such as promoters (PROM), coding sequences (CDS) and terminators (TERM) are flanked by fixed convergent *Bsa*I recognition-cleavage sites. To facilitate the visualization of the design, we assigned each 4 bp cleavage sequence a different label: those produced by *Bsa*I digestion are labeled with Arabic and Latin numbers (1,2,3, III, IV, etc.). In assembling a single device, constituent parts (pEs) are incubated together with a destination plasmid (pD) containing a LacZ cassette flanked by *Bsa*I sites in divergent orientation. As a result, an expression plasmid (pEx) is created where all *Bsa*I recognition sites have disappeared. Boxed AMP and KAN represent ampicillin and kanamycin resistance genes.

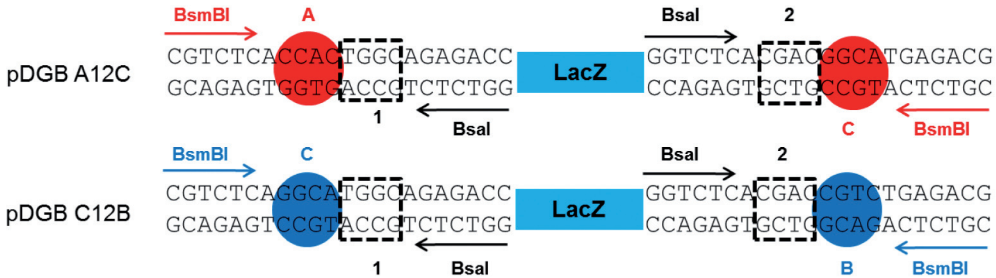
2.2. The double loop Design of the GoldenBraid system.

So far, the described method allows standardization, but the resulting units (expression vectors), lacking restriction sites, cannot be re-used in subsequent assembly reactions. A possible solution to this constraint would be the addition of restriction sites for a second Type IIS enzyme (e.g. *BsmB*I) in the backbone of the destination plasmid, so that *Bsa*I-assembled devices (first order assembly) could similarly be assembled in second order destination plasmids. However, in order to allow multipartite second order assemblies, this solution would require the design of a large number of destination plasmids, as the flanking *BsmB*I sites of the destination plasmids need to be different depending on the number of elements to be assembled in the second level. Moreover, in order to make the resulting composite parts fully reusable, an indefinite number of additional destination plasmids for subsequent hierarchy levels would be required.

A simple solution to this limitation, described here as GoldenBraid, is to insert a loop (braid) in the cloning design, so that the expression plasmids from first level become entry plasmids for second level assemblies and vice versa. In order to do this, two types of destination plasmids were designed, namely level **α** and level **Ω**. The key in GoldenBraid design is that, while all plasmids contain two restriction/recognition sites corresponding to two different Type IIS enzymes, level **α** and level **Ω** plasmids are designed to have their sites in inverted orientations (Figure 9). They also differ in the resistance marker associated to each of them, allowing counterselection. According to this strategy, only four destination plasmids are required to conform the loop cloning topology of GoldenBraid: plasmids pDGB_A12C and pDGB_C12B for assembling at level **α** and pDGB_1AB3 and pDGB_3AB2 for assembling at level **Ω**, where 1, 2 and 3 correspond to sequences of four nucleotide-overhangs produced by *Bsa*I and A, B and C refer to the four nucleotide-overhangs produced by *BsmB*I.

The cloning methodology used in GoldenBraid is shown in Figure 10. Standard parts are normally assembled in level **α** plasmids (Figure 10A). Those composite parts built into pDGB_A12C as destination vector can be merged with other structures assembled in pDGB_C12B, yielding two possible results depending on which of the two level-**Ω** plasmids is used as destination vector: a new structure flanked by 1-3 sites and/or a structure flanked by 3-2 sites (Figure 10B). In a second assembly round, composite parts assembled using level **Ω** plasmid can be assembled together using level **α** destination plasmids. As can be observed in Figure 10, GoldenBraid works as endless iteration of binary assemblies where the only limitations would be those imposed by the host on the size/composition of the DNA that can be stably propagated in a given destination vector backbone.

LEVEL α PLASMIDS



LEVEL Ω PLASMIDS

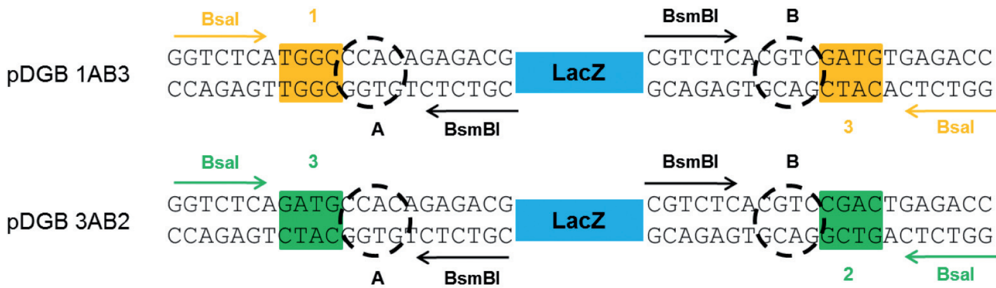


Figure 9. Structure of the LacZ cassettes in the GoldenBraid system.

GB plasmid set comprises four destination plasmids (pDGBs), two of them act as destination plasmids for level α assembly and the remaining two function as destination plasmids for level Ω. All pDGB vectors incorporate a LacZ selection cassette flanked by four Type IIS restriction sites (*Bsa*I, *BsmB*I), but positioned in inverted positions and orientations. To facilitate the visualization of the design, we assigned each 4 bp cleavage sequence a different label: those produced by *Bsa*I digestion are labeled with squares and named with Arabic numbers (1,2,3), whereas *BsmB*I 4 bp cleavage sites are encircled and named with capital letters (A,B,C).

GoldenBraid assembly can be formally described with a simple system of four assembly rules:

1. pE [1 (X_i) 3] + pE [3 (X_j) 2] + pD (A12C) = pE [A (X_i+X_j) C]
2. pE [1 (X_i) 3] + pE [3 (X_j) 2] + pD (C12B) = pE [C (X_i+X_j) B]
3. pE [A (X_i) C] + pE [C (X_j) B] + pD (1AB3) = pE [1 (X_i+X_j) 3]
4. pE [A (X_i) C] + pE [C (X_j) B] + pD (3AB2) = pE [3 (X_i+X_j) 2]

where,

- (X_i) and (X_j) are any DNA pieces, including Golden Gate assembled composite parts.
- (X_i+X_j) is a composite part of (X_i) and (X_j) that follows the same assembly rules than (X_i) and (X_j).
- Numbers 1, 2, and 3 are four-nucleotide sequences, which flank (X) pieces, and which are made protuberant ends upon *Bsa*I digestion.
- Letters A, B and C are four-nucleotide sequences, which flank (X) pieces, and which are made protuberant ends upon *BsmB*I digestion.
- pE[ ] is any plasmid (entry plasmid) hosting a piece (X), such piece flanked by sites as indicated by flanking numbers or letters.
- pD( ) is any plasmid (destination plasmid) hosting a LacZ cassette, such LacZ cassette flanked by two sites, as indicated by flanking numbers or letters.

As deduced from these rules, in order to be GB-assembled together each DNA fragment needs to be cloned in a different plasmid from the same GB level. A careful design of the assembly strategy will ensure in most cases that two pieces to be assembled are correctly positioned. For those cases where this is not possible (e.g. two devices designed independently in different labs), we have constructed four "twister" plasmids containing a small stuffer fragment that facilitate moving pieces from one level to the next in a single GB reaction (Figure 10C). The twister plasmids are indeed four entry plasmids hosting a "fixed" tomato intergenic region flanked by one of the four possible enzyme combinations each (A-C, C-B, 1-3 or 3-2). Using these plasmids, any GB-cloned composite part can be easily and conveniently GB-twisted into next level plasmids, allowing its assembly with parts located at the opposite level.

It is highly desirable that all the components in the GoldenBraid system are free of internal *Bsa*I and *BsmB*I sites. For part domestication, internal sites are removed using standard methodology as overlapping-PCR, directed mutagenesis, or direct DNA synthesis. For plasmid adaptation to GB system, we followed a general procedure using a third Type IIS enzyme (*Bbs*I). The original binary plasmid was deconstructed in pieces; the number of pieces depends on the number of internal sites to be removed and the functional structures that need to be kept as independent pieces. Usually, basic pieces involve the LacZ cassette, antibiotic resistance, and two additional pieces containing replication origins and each of the T-DNA borders. Four LacZ pieces (A12C, C12B, 1AB3 and 3AB2) and two different antibiotic resistance pieces (e.g. KanR and SpmR) are to be produced to generate a complete GB plasmid set. Additional pieces may be required to mutagenize internal Type IIS sites.

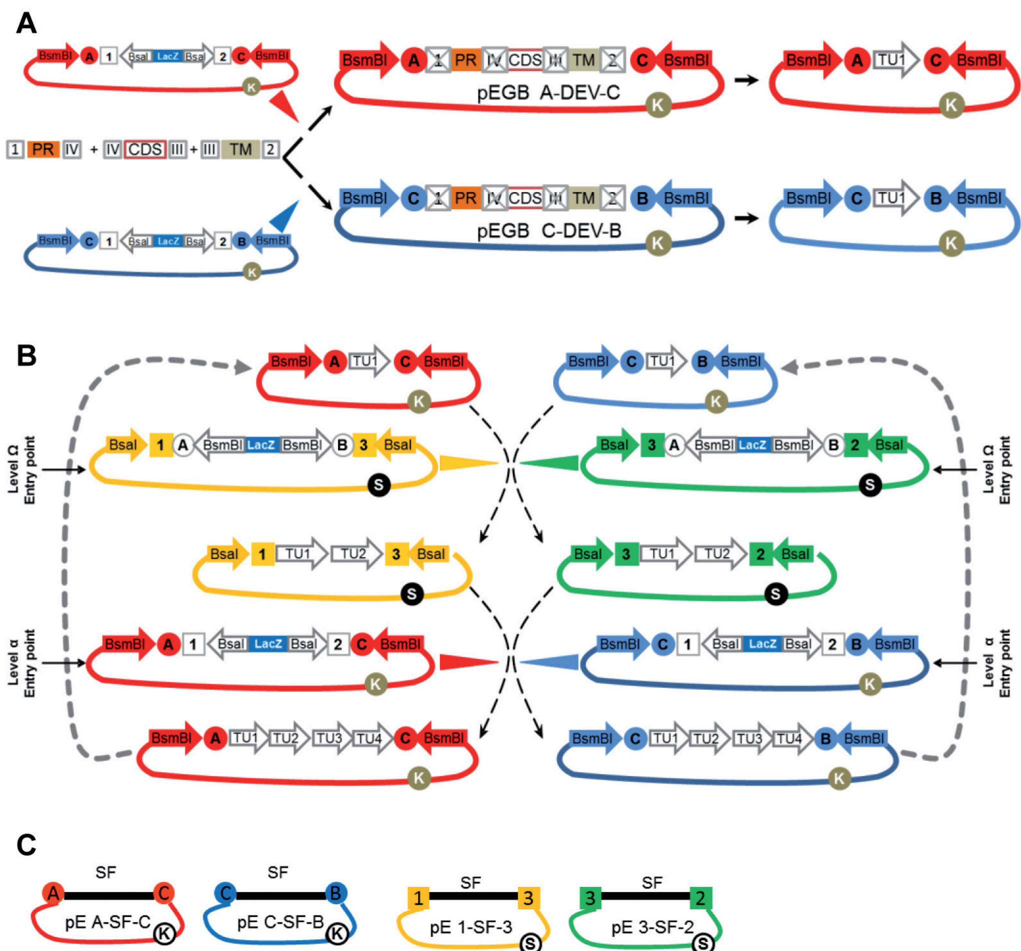


Figure 10. The mechanism of GoldenBraid system.

(A) Standard parts as promoters (PR), coding sequences (CDS) and terminators (TM), flanked by fixed *Bsa*I cleavage sites (represented as Arabic and Latin numbers) are ordinarily assembled using level α plasmids (pDGBA12C or pDGBC12B). As a result of multipartite assembly, *Bsa*I recognition sites disappear and the resulting boundary is not cleavable anymore (represented as a crossed label). Nevertheless, the newly assembled device (DEV, represented for simplification as an arrow) remains flanked by *BsmB*I cleavable sites (represented as encircled capital letters). **(B)** Two devices assembled in complementary α plasmids can be reused as entry vectors (pEGB) for a subsequent level Ω binary assembly, provided that they share a *BsmB*I sticky end (labeled as encircled C). Similarly, constructs assembled using opposite Ω plasmids can be reused as entry vectors for a subsequent level α binary assembly, provided that they share a *Bsa*I sticky end (labeled as squared 3). Level α and level Ω can alternate indefinitely creating increasingly complex structures, as depicted by the arrows closing the double loop. Encircled K and S represent KanR and SpmR respectively. **(C)** Representation of the four "twister" plasmids that can be eventually used to assist GoldenBraid cloning design. SF is a 150 bp stuffer fragment containing an intergenic region from Solanum lycopersicum (S.lycopersicum).

2.3. Multigenic constructs for Plant Biology.

2.3.1 GoldenBraid-assisted co-transformation ensures the coordinated expression of multiple genes in transient expression experiments.

Agrobacterium-mediated transient gene expression (agroinfiltration) in *N. benthamiana* is an efficient technology for recombinant protein production in plants. An interesting feature of this system is the high co-transformation efficiency obtained by simply combining two or more independent *Agrobacterium* cultures each carrying one of the genes of interests (this called *in trans* co-transformation). The cumbersome and inefficient assembly of multiple transcriptional units in a single T-DNA has often led many labs to rely on *in trans* co-transformation when the coordinated or simultaneous expression of two or more proteins in a single cell/tissue was pursued. The GoldenBraid strategy here described makes the cloning of multigene constructs a straightforward task. To test whether an *in cis* co-transformation approach outperforms the *in trans* approach, three different fluorescent devices were GB-assembled and its performance compared with that of an *in trans* approach.

As starting point for the assembly, we used a small collection of basic parts (pEs), namely promoters, CDS and terminators. Fluorescent devices (transcriptional units) were *BsaI*-assembled into GoldenBraid Level α vectors (Figure 11A). Three basic parts were assembled in each case: pE_35S (CaMV 35s promoter) and pE_TNos (Nopaline synthase terminator) were used in all the constructions and assembled to CDS parts carrying either a yellow fluorescent protein (pE_YFP), a blue fluorescent protein (pE_BFP), a Tomato Bushy Stunt Virus P19 silencing suppressor **[121]** (pE_P19) or *Discosoma sp*. red fluorescent protein (pE_DsRed) respectively. Two of the resulting devices (YFP and BFP transcriptional units) were assembled into pDGB_A12C and the two others (DsRed and P19 transcriptional units) were assembled into pDGB_C12B, generating four expression vectors: pEGB_A-YFP-C, pEGB_A-BFP-C, pEGB_C-P19-B and pEGB_C-DsRed-B. These reactions were extremely efficient with an average of 64000 colonies obtained in each transformation (generally 4 colonies were selected for mini-prep resulting in 100% correct colonies). Next, pEGB_A-YFP-C and pEGB_C-P19-B were assembled together into pDGB_1AB3, whereas pEGB_A-BFP-C and pEGB_C-DsRed-B were assembled into pDGB_3AB2, generating the expression vectors pEGB_1-YFP-P19-3 and pEGB_3-BFP-DsRed-2 respectively with the same high efficiency and accuracy.

Taking advantage of the different selection markers of the plasmids in levels α and Ω, we also tested the possibility of building double-device constructs directly from its basic parts in a single in vitro experiment. Ordinarily, devices are *BsaI*-assembled in one-tube multipartite

reactions using α level destination plasmids, and the resulting mix is used to transform *Escherichia coli* (*E. coli*). In this case, the double-device constructs were attempted by combining two independent single-device reactions (e.g. pEGB_A-BFP-C and pEGB_C-DsRed-B) in a new tube and incubating with *BsmB*I and ligase for additional 25 cycles. As a result the two functional devices were assembled in one T-DNA (pEGB_3-BFP-DsRed-2) with 1/10 efficiency of the two-step assembly, but in a single day experiment and without requiring intermediate *E. coli* transformation.

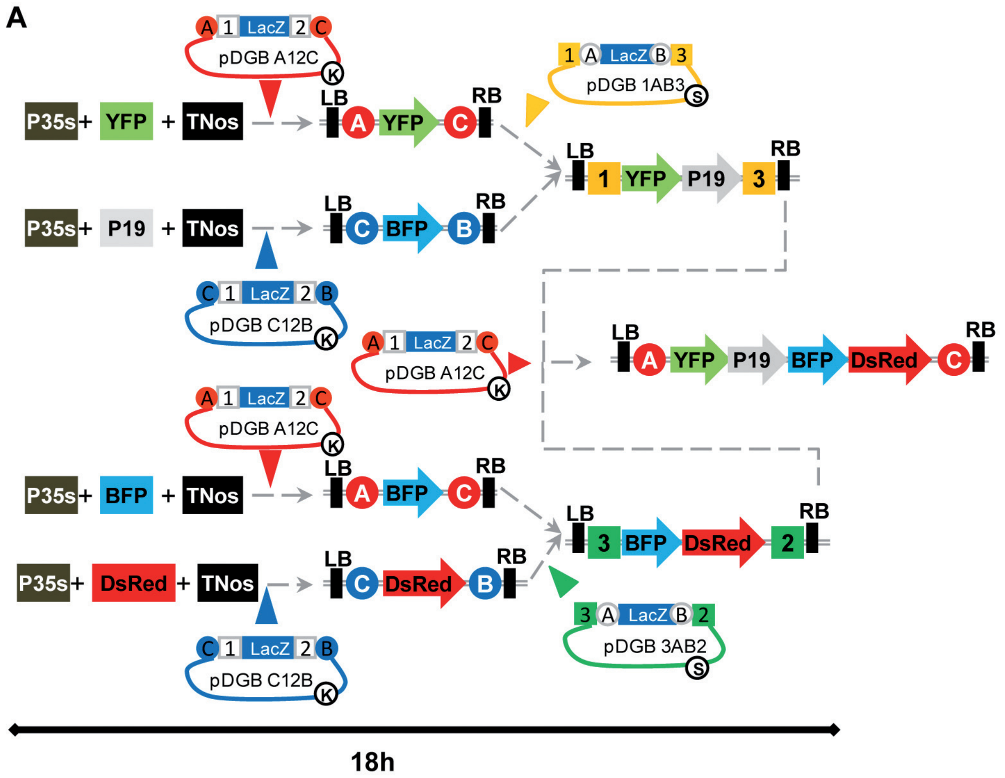
Finally, pEGB_1-YFP-P19-3 and pEGB_3-BFP-DsRed-2 vectors were assembled in a *Bsa*I reaction into the destination vector pDGB_A12C. This final multigenic construction pEGB_A-YFP-P19-BFP-DsRed-C, comprising 11.4 Kb and 12 parts, was functionally validated by agroinfiltration into *N. bentamiana* leaves. In parallel, single-assembled fluorescent proteins and P19 were also co-transformed *in trans* by mixing their respective *Agrobacterium* cultures. As can be observed in Figure 11B, GoldenBraid assembled fluorescent proteins showed coordinated expression in *N. bentamiana*, as deduced by the similar fluorescence intensity observed in all three channels. In contrast, when the fluorescent devices were agroinfiltrated *in trans*, each channel showed a different intensity distribution, evidencing heterogeneous expression levels of the different proteins.

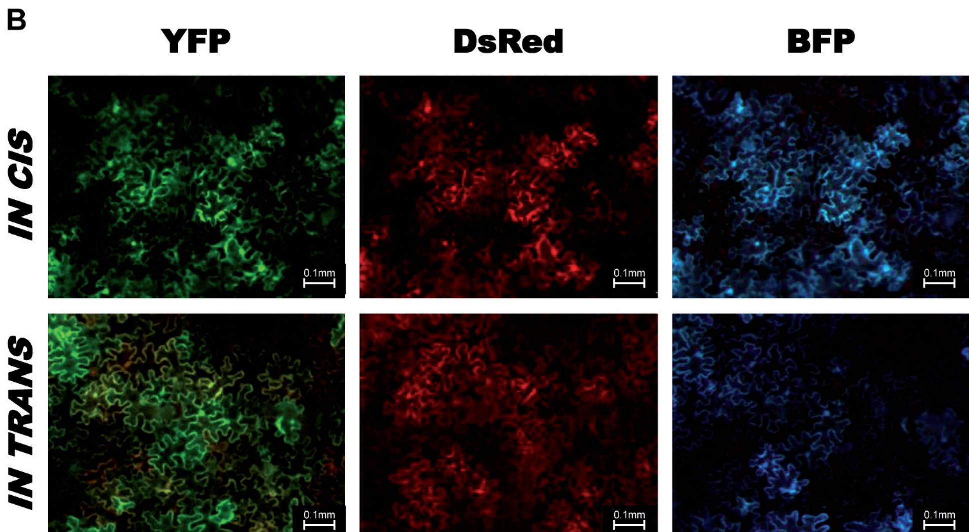
2.3.2 GoldenBraid-assisted antibody chain shuffling facilitates selection of antibody isotype.

The plant-based production of therapeutic antibodies is a field that requires flexible multigene cloning strategies. Therefore we evaluated our GoldenBraid-assisted cloning to build different "antibody devices" and compared the results obtained after expressing the proteins *in planta*. In the previous experiment with fluorescent proteins, parts were *Bsa*I-assembled into level α plasmids (entry point α in Figure 10). The loop design of GoldenBraid system should allow the use of both level α and level Ω plasmids for multipartite assembly of basic parts. In this second experiment we made use of entry point Ω to build and assemble basic parts for therapeutic devices.

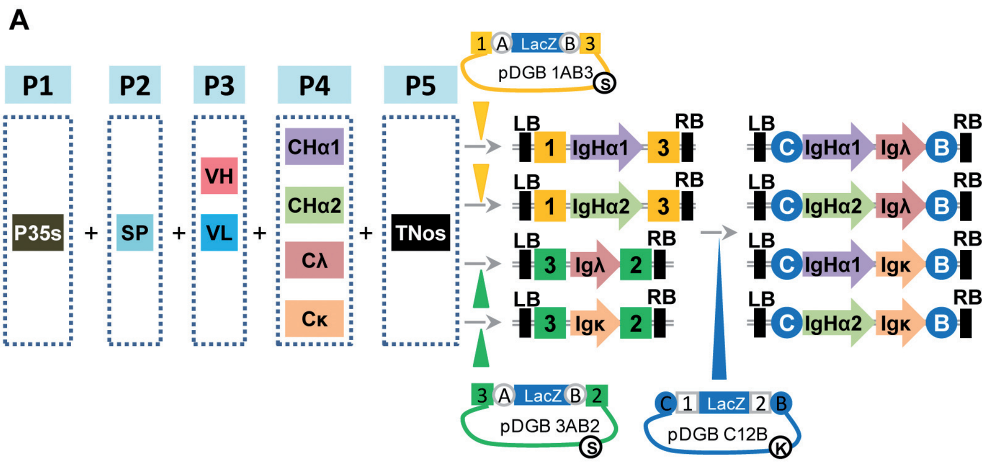
Figure 11. GoldenBraid-assisted co-transformation of fluorescent devices.

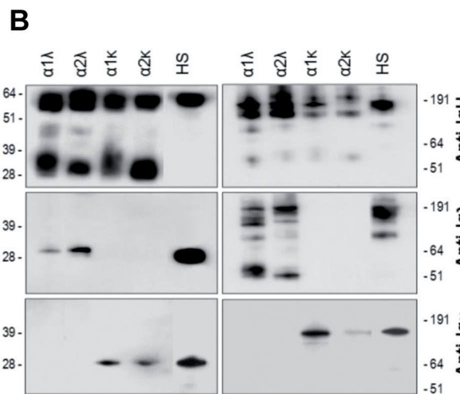
(A) GoldenBraid cloning path for the assembling of YFP, P19, BFP and DsRED transcriptional units in a single T-DNA. **(B)** Spatial expression patterns of BFP, YFP and DsRed in *N. benthamiana* leaves agroinfiltrated with pEGB_A-YFP-P19-BFP-DsRed-C- (left captures, 1, 2 and 3) or with a mixture of the individual devices pEGB_A-YFP-C, pEGB_C-P19-B, pEGB_A-BFP-C and pEGB_C-DsRed-B (right captures 4, 5 and 6). 35S is CaMV 35s promoter; YFP is yellow fluorescent protein; P19 is TBSV Silencing Suppressor; BFP is blue fluorescent protein; DsRed is *Discosoma sp.* red fluorescent protein; TNos is Nopaline synthase terminator.

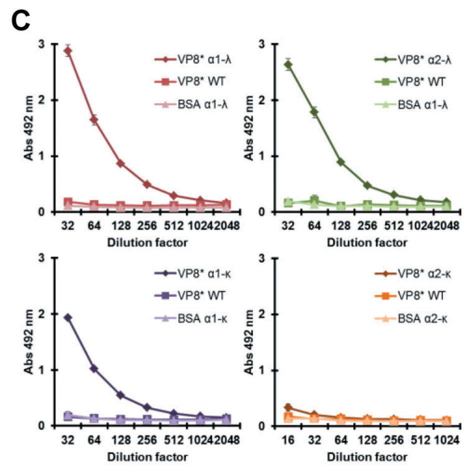

A
A 1 LacZ 2 C
pDGB A12C
K
P35s + YFP + TNos
LB
A YFP C
RB
1 A LacZ B 3
pDGB 1AB3
S
LB
1 YFP P19 3
RB
P35s + P19 + TNos
LB
C BFP B
RB
C 1 LacZ 2 B
pDGB C12B
K
A 1 LacZ 2 C
pDGB A12C
K
LB
A YFP P19 BFP DsRed C
RB
A 1 LacZ 2 C
pDGB A12C
K
P35s + BFP + TNos
LB
A BFP C
RB
LB
3 BFP DsRed 2
RB
P35s + DsRed + TNos
LB
C DsRed B
RB
3 A LacZ B 2
pDGB 3AB2
S
C 1 LacZ 2 B
pDGB C12B
K
18h


B
YFP
DsRed
BFP
IN CIS
IN TRANS
0.1mm

A versatile strategy was designed to assemble any desired human IgA (h_IgA) isotype. To gain flexibility, parts were classified in five categories, namely promoter, signal peptide, variable antibody regions, constant antibody regions and terminators. Next, five-part *BsmB*I reactions were performed to assemble the individual heavy and light antibody chains. The experiment showed here was aimed at selecting the best IgA isotype for *in planta* production of an anti-rotavirus antibody. For this purpose, two heavy chains (pEGB_1-IgHα1-3 and pEGB_1-IgHα2-3) and two light chains (pEGB_3-IgK-2 and pEGB_3-Igλ-2) were *BsmB*I-assembled into level Ω plasmids. Next, heavy and light chain devices were combined in a *Bsa*I-GoldenBraid reaction, generating the four different isotypes of human IgA (Figure 12A). The four h_IgA isotypes produced (separately) in agroinfiltrated leaves were compared by western blot (Figure 12B) and ELISA (Figure 12C), with the version combining IgHα1 and Igλ (pEGB_C-IgHα1-Igλ-B) showing best performance *in planta*.







2.4. Construction and combination of therapeutic and biosafety gene modules by GoldenBraid.

One of the strengths of GoldenBraid cloning is the reusability of pieces, so once assembled and tested for one purpose they can be easily included in further multigenic structures aimed at similar or different purposes. To illustrate this ability, we show the use of some of the devices described above to make two additional multigenic structures (highlighted with an asterisk in Figure 13A). In this case a "therapeutic" module (anti-rotavirus IgA) initially aimed at transient expression is reused for a different purpose, the engineering of a biosafe plant biofactory for anti-rotavirus IgA. For this goal, IgA "therapeutic" module is combined with a "selection" device for plant stable transformation (KanR) and two alternative biosafety modules, both comprising an "identity preservation" device and a "pollen-sterility" device.

Single-device constructs were assembled as follows (Figure 13A): first a Kanamycin resistance device was built in a multipartite *Bsa*I reaction into level **α** plasmid pDGB_A12C. Next, two alternative "Identity Preservation" devices were considered: the previously described pEGB_C-DsRed-B conferring red fluorescence to the plant, and the newly constructed Rosea1, consisting of a CaMV 35s, Nos terminator and the *Antirrhinum majus* Rosea1 transcription factor that confers purple color to the cells **[122]**. Finally a male sterility "device" was constructed, combining barnase-barstar CDS under pTA29 anther-specific promoter **[123-124]**. From here, the assembling of multigene structures was conducted as follows: the device pEGB_A-KanR-C was assembled to the IgA "therapeutic" module in a *BsmB*I reaction into pDGB_1AB3. Next, two alternative "biosafety" modules, namely pEGB_3-Barnase-Rosea-2 and pEGB_3-Barnase-DsRed-2 were assembled into level **Ω** plasmids as shown in Figure 13A. Finally, biosafety

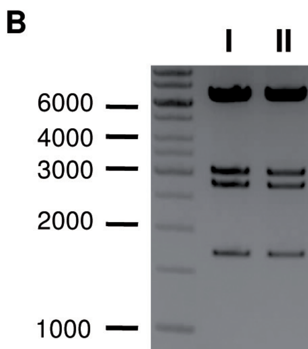
Figure 12. GoldenBraid-assisted selection of plant-made IgA isotypes.

(A) GoldenBraid cloning strategy followed in the assembly of different IgA isotypes. Multipartite assembly involved the combination of different basic parts each occupying a fixed position in the assembly (P1-P5). Individual antibody chains were assembled in pDGB_C12B plasmid to yield four IgA isotypes. Promoter and terminator pieces were flanked by the same 4 nucleotide extensions as in Figure 8. Signal peptides incorporated a GATG extension at its 5´ end, whereas constant antibody regions ended in TGAG extensions to match terminators. The remaining boundaries were designed to produce benign junctions within coding sequences. **(B)** Western Blot analysis of IgA transient expression in *N. benthamiana*. Leaves were infiltrated with the four previous combinations. Samples were resolved under either reducing (left) or non-reducing (right) conditions and decorated using anti-heavy chain antibody, anti-λ light chain antibody or anti-κ light chain antibody. HS lane contains control human serum. **(C)** End-point antigen-ELISA tittering of four IgA combinations tested by transient expression in *N. benthamiana* leaves. All samples were tittered against VP8* or against BSA and compared with equivalent samples derived from wild type leaves (WT). 35S is CaMV 35s promoter; SP, pectate lyase signal peptide; CHα1 and CHα2, are heavy chain constant domains; TNos, is Nopaline synthase terminator; Cλ and Cκ, are light chain constant domains; VH and VL are heavy and light variable regions of an antibody against rotavirus VP8* peptide; IgHα1 is immunoglobulin A heavy chain α1; IgHα2 is immunoglobulin A heavy chain α2; Igλ is immunoglobulin A light chain λ; Igκ is immunoglobulin A light chain κ.

modules were assembled to the IgA_KanR module in a final *Bsa*I reaction resulting in two alternative five-device constructs of 14.3 Kb and 19 pieces made of reused devices (Figure 13B).

A

B



3. Discussion

GoldenBraid is a tool that converts single-use Golden Gate multipartite assemblies into reusable composite parts. In this sense GoldenBraid assembly is an attempt to extend the capabilities of the previously described Golden Gate cloning system to the requirements of Synthetic Biology. There are no preconditions on the type of DNA pieces involved in the initial multipartite assembly, which can be basic parts, transcriptional units or even small pathways. However, we think that multipartite assemblies of basic DNA parts are most interesting, particularly when this is made in a standardized, community-based fashion. To do so, we propose (i) the creation of a standardized collection of basic parts flanked by Type IIS sites, (ii) the multipartite assembly of DNA parts into GB destination plasmids to generate simple genetic devices; (ii) the use of GB plasmids and GB rules to grow increasingly complex genetic modules and pathways.

Part standardization is pivotal for genetic engineering. The small junctions used by Type IIS-based cloning and the high efficiency of GoldenBraid procedure greatly favors standardization. We currently use a small collection of basic parts structured in promoters, CDS, and terminators, however, a more elaborated category list could be considered. It is important to notice that the relative position of a DNA fragment in a multipartite assembly, and therefore its identity, is determined by its 4-nucleotide flanking sequences. Adoption of common sequences by different labs would be required for taking full advantage of the system.

We think GoldenBraid has a number of characteristics that encourage its adoption by scientific community. One of them is reusability/exchangeability: all GoldenBraid composite parts can be either transformed directly into cells or used as a piece to build more complex structures. No PCR amplification or further modifications of the piece are required. Error-born and/or lengthy adaptation methodologies hamper the engineering processes, whereas full reusability ensures the reproducibility of the built-in genetic devices. A second advantage is speed: as the starting point of GoldenBraid scheme is a multipartite assembly, the overall

Figure 13. New multigene assemblies using reusable composite parts.

(A) GoldenBraid strategy for the assembly of two alternative 5-gene T-DNA constructs. **(B)** PvuI digestion of one colony of each final constructs pDGB_A-KanR-IgHα1-Igλ-Barnase-Rosea-C (lane I) and pDGB_A-KanR-IgHα1-Igλ-Barnase-DsRed-C (lane II). Asterisks highlight those GB-assembled transcriptional units that were reused in the assembly of new multigenic structures. PNos is Nopaline synthase promoter; KanR is neomycin phosphotransferase II gene; TNos is Nopaline synthase promoter; IgHα1 is immunoglobulin A heavy chain α1; Igλ is immunoglobulin A light chain λ; pTA29 is anther-specific promoter; Barnase is barnase-barstar CDS; DsRed is *Discosoma sp.* red fluorescent protein; Rosea is *A.majus* Rosea1 transcription factor.

engineering process is considerably accelerated when compared with purely binary systems as BioBricks. Moreover, we have shown that two expression cassettes can be assembled together in less than 24h starting from basic parts. A third comparative advantage is accuracy: Type IIS cloning allows the building of assemblies containing short "benign" seams, as earlier demonstrated in Golden Gate cloning. Finally, a distinctive characteristic of the GoldenBraid scheme is its simplicity: GoldenBraid can theoretically build indefinite assemblies with the only use of four destination plasmids and four basic assembling rules.

Plant genetic engineering currently relies on assembly methodologies poorly adaptable to Synthetic Biology. In an attempt to facilitate versatile cloning into plant binary vectors, we and others have developed plasmid collections based on Gateway technology **[53, 60, 125]**. Gateway cloning, based on site-specific recombination, is a highly efficient cloning technique; however it leaves long scars between pieces (*att*B sites) and the reusability of pieces is limited. A number of additional techniques, based on site-specific recombination, the use of rare cutters or homing endonucleases have been developed **[27, 68, 74, 80, 126]**, however in our opinion GoldenBraid compares favorably with most of them in terms of standardization, simplicity and reusability.

In view of this need, we have adapted GoldenBraid scheme to plant biotechnology by domesticating four binary plasmids, and demonstrated in a number of examples the feasibility of the methodology. In a first example, using fluorescent proteins, it was demonstrated that GoldenBraid is permissive with the repetition of single pieces in multiple assemblies. At least as long as transient expression is concern, the introduction of 4 copies of CaMV 35s promoter in a single T-DNA does not affect the transient expression of the fluorescent proteins. Just on the contrary, *in cis* co-transformation favors the coordinated expression of the transgenes. *In trans* co-agroinfiltration is currently used as a fast–track tool for e.g. plant glyco-engineering or metabolic engineering, both approaches often relying on coordinated expression of the different transgenes in each cell **[127]**. In the light of the results showed here, GB-assisted assembling would improve the outcome of these transient approaches, as it would do so if the same engineered T-DNAs were to be stably transformed in plants.

In a second example we illustrate the use of GB in antibody engineering by exchanging in a combinatorial way all the alternative constant regions of a human IgA against rotavirus. Moreover, this design also allows the exchange of variable regions, facilitating conversion of antibody idiotype. In this particular example we chose to build parts that enter the GB loop at the Ω level, therefore demonstrating the symmetry of the braid. Although this possibility remains open, it seems more reasonable for a general strategy the use a single entry level, as this facilitates part standardization. It is important to notice that, in its current design,

GB uses different entry sequences for level α (sites 1 and 2) and level Ω (sites A and B). It could be conceived a system where A=1 and B=2, which would allow standard pieces to be assembled indistinctly at level α or Ω. This would increase the exchangeability of the pieces, reducing the eventual need for twister plasmids. However, this would also require the use of an additional Type IIS restriction enzyme for the cloning of basic "parts". By doing so, parts could be multi-partite assembled at any level by using an "extra" enzyme that does not destroy the restriction sites to be used at the next level. In this case, the increased reusability would pay the toll of extra domestication requirements introduced by a third enzyme. We calculate that, by using our current two-enzyme design, 29% of tomato cDNAs would require domestication, whereas the use of a third enzyme (e.g. *Bbs*I) would increase this figure up to 51%. Considering the simplicity and efficiency of helper-assisted twists, we tend to favor current design over a three-enzyme design.

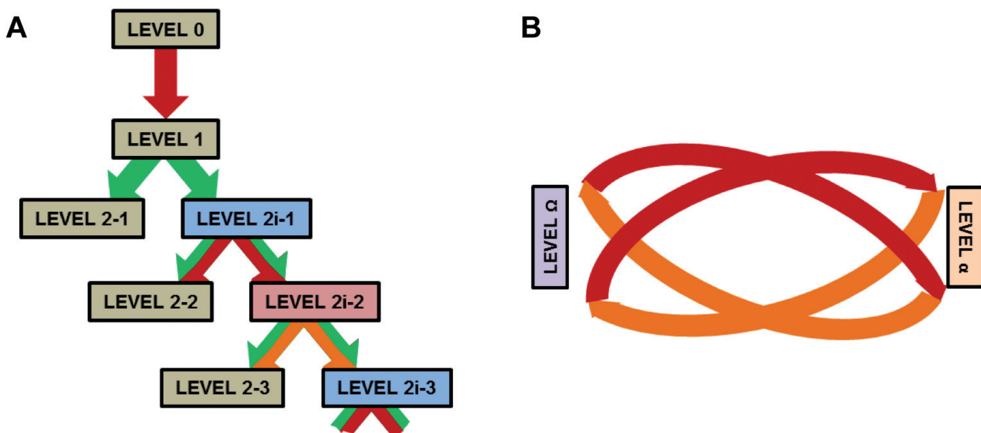
In a final example we demonstrate the reusability of GB constructs with the assembly of two alternative constructs comprising five transcriptional units. A "therapeutic" module (IgA) is combined with a "selection" module and two alternative "biosafety" modules. Biosafety modules are made of a "male sterility" device and two alternative "identity preservation" devices. In our opinion, this example fully illustrates the principles of modularity, standardization and reusability that drive Synthetic Biology aims.

Given the indefinite design of GB, the obvious limitation to GB assemblies is that imposed by the maximum insert size that can be harbored by binary plasmids. Although initially designed using binary plasmids, GB assemblies, as fully reusable units, can be easily transferred to newly domesticated structures such as BiBACs **[128]** suitable to host larger T-DNAS, or other devices for direct DNA transfer. Moreover, at any time GB constructs can be added new pieces that facilitate its conversion to alternative assembling methods. This may include, among other elements, *att*B cassettes for Gateway cloning, overlapping regions for in vitro or in vivo recombination, or recombination sites (e.g. loxP) for *in planta* gene stacking. We consider that standardized in vitro gene assembling methods as GB may become an important tool in engineering of complex traits, which lays at the horizon of modern Plant Biotechnology.

During the preparation of this manuscript, an alternative methodology for the standardization of Golden Gate cloning for Synthetic Biology (named MoClo) was published **[89]**. In their paper, Weber *et al.* show the construction of a 33 Kb multigenic structure with the only use of successive Golden Gate reactions, a result that demonstrates that Type IIS technologies (including GoldenBraid) can successfully be used for the assembly of complex genetic modules. MoClo proposes an elegant strategy for the cloning of "subparts" (level 0) that was not contemplated in GB strategy. This interesting strategy enhances the flexibility and

the combinatorial power of any part collection. Also, similarly to GB, MoClo proposes the use of a second enzyme in destination plasmids as a way to extend Golden Gate cloning to a second assembly level. The use of a second enzyme for extended cloning has been also very recently proposed by different authors as a tool to facilitate modular assembling of TAL effectors **[85-86, 129-130]**; however MoClo brings this idea to a general scheme for multigene assembling. In MoClo strategy a first enzyme (*Bsa*I) is used to assemble "parts" into devices (level 1, equivalent to GB level α), and a second enzyme (*Bbs*I) is used to combine devices into multigene structures (level 2, equivalent to GB level Ω). However at this point the solutions provided by MoClo and GB to achieve the indefinite growth of multigene structures become completely different. As the use of two enzymes limits the level of successive assembling levels to two, MoClo proposes the creation of intermediate assembly levels (2i-1, 2i-2, etc.), where an "extra" piece (end-linker) consisting of a selection cassette (LacZ or Red) is introduced as a way to leave the assembly "open" to the addition of new pieces. Further additions will involve the exchange of LacZ and Red cassettes by new "true" pieces in successive assembly levels.

GB has a number of features that differentiate it from the solution proposed by Weber *et al.* : (i) GB makes use of only two restriction enzymes whereas MoClo requires a third enzyme and an additional selection cassette to ensure indefinite growth; (ii) GB pieces are fully reusable, whereas in MoClo intermediate structures need to be assembled to allow further growth of the construct; (iii) GB assemblies are always binary, whereas MoClo allows multipartite assemblies at level 2; (iv) the topology of MoClo system is basically lineal, with successive assembly levels and lateral branches corresponding to intermediate levels. In contrast, GB has a circular topology, with pieces growing by alternating level α and Ω. A comparison of the topology of the two systems can be observed in Figure 14.



In synthesis, we consider that GB has two main distinctive features that can make it a useful alternative to MoClo for certain applications: its simplicity and the reusability of its composite parts. Conversely, MoClo main advantage is the possibility of building multipartite assembles at level 2. Both groups of features are probably mutually exclusive: MoClo multipartite assemblies at level 2 come at the expenses of the incorporation of a number of additional destination plasmids and end-linker plasmids to the system, which further increases its complexity. Analogously, additional destination and end-linker plasmids could be added to GB level α to allow multipartite assemblies at level Ω (e.g. A12D, D12C and C12B to obtain tripartite assemblies). However we doubt that the possible advances in speed could compensate the increased complexity of this solution provided that (i) indefinite growth of GB assemblies is ensured without the use of additional elements, (ii) intermediate binary assemblies are in itself useful as reusable entities (see last example of results section); (iii) in our experience multipartite cloning of large fragments has low efficiency, making often advisable to advance large constructs in binary form; (iv) speed in GB is satisfactory, as we show that 2-device assemblies can be constructed from its basic parts in a single in vitro 18h experiment; (v) the adoption of the technology by the community as well as its automation will be facilitated if simplicity is maintained.

It needs to be pointed out that both MoClo and GB are based on the same enzymatic reactions, and therefore, it can be expected that both should perform similarly in terms of construct size. The ability to assemble complex constructs will most likely depend on other factors not covered in this paper as the host plasmid (copy number, replication origin), the presence of repetitive regions, the host bacteria (whether *Agrobacterium*-mediated transformation is needed), etc. Either as GB or as MoClo, the extension of Golden Gate method to the standardized assembly of higher order genetic pieces as devices and pathways is an important step that will facilitate genetic engineering, particularly in the plant field. In our opinion, it would be highly beneficial to establish community-shared standards in aspects as piece identity and entry sites in order to facilitate the exchange of genetic pieces between labs and to facilitate further development of Plant Synthetic Biology.

Figure 14. Comparison of the topology of MoClo and GoldenBraid.

(A) Hierarchical topology of MoClo assembly. Level 0 hosts the flexible assembly of subparts into basic parts, allowing also part domestication. Level 1 hosts multipartite assembly of basic parts into transcriptional units. Level 2-1 hosts multipartite assembly of transcriptional units, yielding a non-reusable structure. Alternatively, level 1 can be branched into level 2-1i (intermediate) by adding an end-linker, yielding an open structure (albeit nonfunctional), which can host new transcriptional units (level 2-2). Successive intermediate levels ensure the indefinite structure of the cloning system. **(B)** Double loop topology of GoldenBraid. Level-α plasmids host the multipartite assembly of basic parts into transcriptional units. Two level-α transcriptional units can be assembled together yielding two alternative level-Ω constructs, which themselves can be assembled into level-α constructs. The overall structure is a double iterative loop that ensures the indefinite growth of the assembly system. *Bsa*I, *BsmB*I and *Bbs*I reactions are symbolized with red, orange and green arrows, respectively.

4. Materials and Methods

4.1. Strains and growth conditions.

Escherichia coli DH5α was used for gene cloning and *A. tumefaciens* strain GV3101 was used for plant agroinfiltration and transformation experiments. Both strains were grown in LB medium under agitation (200 rpm) at 37ºC and 28ºC respectively. Ampicillin, kanamycin and spectinomycin were used for *E. coli* at 50 µg ml^{-1}. Rifampicin, tetracycline and gentamicin were also used for *A. tumefaciens* at 50, 12.5 and 30 µg ml^{-1} respectively.

4.2. Cloning and assembly of modular pieces.

PCR amplification was performed by using the Advantage® 2 DNA Polymerase Mix (Clontech, California, USA) following the manufacturer's instructions. PCR was analyzed by agarose 1% gel electrophoresis and purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). Amplified parts were TA Cloned using the pGEM®-T Easy Vector System (Promega, Madison, USA) and 1 µl of the ligation was transformed into DH5α electrocompetent cells. Plasmid DNA preparations were obtained by using The E.Z.N.A. Plasmid Mini Kit I (Omega Bio-Tek, Norcross, USA). Plasmid DNA concentration was measured using a Nano Drop Spectrophotometer 2000 (Thermo Scientific, Rockford, USA). Positive clones were selected in ampicillin-containing plates and confirmed by plasmid restriction analysis (*EcoR*I, *Not*I) and by sequencing.

Assembly reactions were performed basically as described by Engler *et al.* **[81]** using *Bsa*I, *BsmB*I and *Bbs*I as restriction enzymes in 25 cycle digestion/ligation reactions. Restriction enzymes were purchased from New England Biolabs (Ipswich, USA). T4 DNA ligase was purchased from Promega.

One µl of the reaction was transformed into DH5α electrocompetent cells. Positive clones were selected in kanamycin or spectinomycin-containing plates. Plasmid DNA preparations were made by using The E.Z.N.A. Plasmid Mini Kit I (Omega Bio-Tek). Plasmid DNA concentration was measured using a Nano Drop Spectrophotometer 2000 (Thermo Scientific). Constructs were confirmed by plasmid restriction analysis and by sequencing. Constructs for plant functional assays were transferred to *A. tumefaciens* electrocompetent strain GV3101.

4.3. GB-Domestication of destination Plasmids for Plant Biology.

With some adaptations, domestication of pDGB plasmids was performed basically as earlier described by Engler *et al.* **[81]**. A third Type IIS enzyme was used (*Bbs*I) for domestication. All the components in the GoldenBraid system were made free of internal *Bsa*I and *BsmB*I sites. The original binary plasmid (pGreen II) **[131]** was deconstructed in four pieces involving the LacZ cassette, antibiotic resistance, and two additional pieces containing replication origins and each of the T-DNA borders. Four lacZ pieces (A12C, C12B, 1AB3 and 3AB2) and two different antibiotic resistance pieces (e.g. KanR and SpmR) were produced to generate a complete GB plasmid set. To assemble pDGB plasmids set, four *Bbs*I Golden Gate reactions between backbone pieces and LacZ cassettes were set up, yielding the four pDGB plasmids, each containing a different LacZ cassette and the kanamycin or spectinomycin resistance genes.

For the construction of twister plasmids, a small intergenic region (150 bp) was PCR-amplified from tomato gDNA, using *Bsa*I and *BsmB*I primer extensions that match the cloning sites of each pDGB (i.e. 1-2 for *Bsa*I and A-B for *BsmB*I). PCR fragments were purified and subsequently GB-cloned in each of the four destination plasmids.

4.4. Plant transient transformation.

For transient plant transformations plasmids were transferred to *A. tumefaciens* strain GV3101 by electroporation. Agroinfiltration was performed as previously described **[132]**. Briefly, overnight grown bacterial cultures were centrifuged and the pellets resuspended in agroinfiltration medium (10 mM MES pH 5.6, 10 mM MgCl2, 200 μM acetosyringone) to an optical density at 600 nm = 0.4. Co-infiltrations were performed by mixing equal volumes of the corresponding bacterial suspensions. Inoculations were carried out by syringe-agroinfiltration in leaves of 4-5 weeks old *N. benthamiana* plants (growing conditions: 24ºC day / 20ºC night in a 16 h light / 8 h dark cycle). Samples were collected 5-6 days post-infiltration and examined for transgene expression.

4.5. Western Blot and ELISA Analysis.

Detection of individual antibody chains and IgA complexes was carried out by western blotting. Leaf proteins were extracted in 3 volumes (v/w) of PBS (phosphate buffer

saline, pH7.4). Protein separation was carried out by SDS-PAGE on NuPAGE 10% Bis-Tris polyacrylamide gels (Invitrogen, Paisley, UK). Proteins were transferred to PVDF membranes (Amersham Hybond-P, GE Healthcare, UK) by semi-wet blotting (XCell IITM Blot Module, Invitrogen) following manufacturer instructions. Membranes were blocked with a 2% (w/v) solution of ECL AdvanceTM Blocking agent (GE Healthcare, UK) in PBS-T (0.1% (v/v) Tween 20 in PBS). For the detection of IgH_α1 and α2 heavy chains membranes were incubated with 1:20000 Anti-Human IgA (α-chain specific) peroxidase conjugate (SIGMA, St. Louis, USA); the Igλ and Igk light chains were detected by incubation with 1:10000 Anti-Human lambda light chain (Sigma) and 1:10000 anti-human-kappa chain (Pierce - Thermo Scientific) as primary antibodies, followed by an incubation with 1:10000 ECL Rabbit IgG, HRP-Linked (GE Healthcare) and 1:10000 Anti-Goat IgG-peroxidase (Sigma) respectively, as secondary antibodies. Blots were developed with ECL Plus Western Blotting Detection System (GE Healthcare) following manufacturer instructions and visualized by exposure to X-ray film (Fujifilm Corporation, Tokyo, Japan).

The binding activity of the recombinant IgA was determined by ELISA. Plates (CORNING, New York, USA) were coated overnight with 10 μg/ mL of recombinant VP8* in coating buffer (50 mM carbonate buffer pH 9,8) at 4º C. Plates were then washed 4 times in PBS and blocked with a 2% (w/v) solution of ECL AdvanceTM Blocking agent (GE Healthcare) in PBS-T (0.1% (v/v) Tween 20 in PBS). Samples were diluted in PBS as required for each assay and incubated for 1 hour at room temperature. After incubation, plates were washed 4 times in PBS and the anti-human IgA α specific-HRP 1:5000 (Sigma-Aldrich) in 5% blocking buffer (GE Healthcare) in PBS-T was added and incubated for 1 h at room temperature. After 4 PBS washes, the substrate (o-phenilenediamine from Sigma-Aldrich) was added and the reactions were stopped with 3M HCl. Absorbance was determined at 492 nm.

/C2

CHAPTER 2

GoldenBraid2.0: a comprehensive DNA assembly framework for plant synthetic biology

Chapter 2: GoldenBraid2.0: A comprehensive DNA assembly framework for Plant Synthetic Biology.

1. Introduction

Synthetic Biology is producing a paradigm shift in Biotechnology based on the introduction of engineering principles in the design of new organisms by genetic modification **[114-115]**. Whereas Synthetic Biology has rapidly permeated microbial biotechnology, the engineering of multi-celled organisms following Synthetic Biology principles is now emerging, and is mainly driven by the so-called top-down approaches where newly engineered genetic circuits are embedded into naturally-existing organisms used as a "chassis". The plant chassis offers an extraordinarily fertile ground for Synthetic Biology-like engineering. However, technology still faces the huge challenge of performing engineering-driven genetic designs. One of the main technological challenges of Plant Synthetic Biology requires the construction and transfer of multigene structures to the plant genome. This is putting pressure on developing a DNA assembly and transformation technologies adapted to plants. One main trend is the use of modular cloning, an engineering-inspired strategy consisting in the fabrication of new devices by combining prefabricated standard modules. In a modular strategy, pre-defined categories, the so-called "parts", are assembled together following a number of rules known as the "assembly standard". Modular DNA building has been enthusiastically adopted by microbial Synthetic Biologists because it offers a number of advantages such as speed, versatility, lab autonomy, combinatorial potential, and often lower cost **[30]**. Modular methods acquire full potential when parts are easily interchangeable, and when one or a few assembly standards are shared by many manufacturers.

A number of features define the value of a modular cloning method. Speed and efficiency are important characteristics, as are also its simplicity and the ability to produce scar-less or scar-benign assemblies. Moreover, any cloning strategy for Synthetic Biology should enable endless reusability; that is, it should ensure that new composite parts themselves can take part in new assemblies, therefore allowing unlimited growth. Several modular cloning strategies have been proposed in the literature, and each

presents advantages and shortcomings. For instance, the original BioBricks standard widely used in microbial Synthetic Biology scores a maximum for simplicity because a single rule governs all the assemblies (a property known as idempotency). However, it is not scar-benign and is only relatively efficient **[37]**. LIC **[133]**, USER's **[134]**, and specially Gibson Assembly **[50]**, are highly efficient DNA assembly methods, although they are neither strictly modular nor widely adopted by plant biotechnologists. In sharp contrast, Gateway Cloning (Hartley, Temple et al. 2000) is of widespread use in plant laboratories **[53, 58, 125]**. Recently, MultiRound Gateway technologies opened Gateway capabilities to the sequential delivery of multiple transgenes by multiple rounds of recombination reactions **[60, 63]**. In general, Gateway-based technologies are highly efficient. Unfortunately, they are not always scare-benign as they leave 21 bp scars between building blocks. Other technologies involving rare cutters or homing endonucleases-based strategies have also been developed and adapted to plant transformation **[27, 80, 126]**, including combinations of homing endonucleases and engineered zinc finger nucleases **[73]**, and iterative *in vivo* assembly rounds of Cre recombinase and phage1 site-specific recombination **[68]**. Many of these techniques can serve as efficient assembly methods for multigene engineering. Nonetheless, a pre-requisite to become a standard for Plant Synthetic Biology is the development of a set of rules and tools based on those technologies which can be shared by as many labs as possible.

Recently, a very powerful DNA assembly method named Golden Gate was described **[81-82]**. Golden Gate uses Type IIS restriction enzymes to generate four-nucleotide sticky ends flanking each DNA piece, which can be subsequently joined together efficiently by T4 ligase. The assembly reaction is multipartite and is performed in a single tube reaction to yield highly efficient scar-less or scar-benign assemblies. This is because Type IIS recognition sites are eliminated upon ligation, leaving only four nucleotides seams, which can be user-defined. These features make the Golden Gate technology an excellent candidate to set up a standardized Modular Cloning system. However, as originally conceived, Golden Gate is not a reusable system and cannot, therefore, be used efficiently for multigene engineering.

Most recently, two strategies were described to enable the reusability of the Golden Gate cloning scheme: MoClo **[89]** and GoldenBraid **[90]**. Both methods use the multipartite Golden Gate property to build transcriptional units (TUs) starting from basic standard building blocks, and both create specially-designed destination vectors to enable Golden Gate-built TUs to be assembled among them. Whereas the GoldenBraid minimalist cloning strategy allows multigene growth by enabling binary

assemblies between TUs, the MoClo destination vectors offer the interesting possibility of performing multipartite assemblies at the TU level, be it at the cost of the higher complexity of its vectors toolkit.

The Golden Gate-based strategies MoClo and GoldenBraid are ideal to serve as modular assembly systems in Plant Synthetic Biology as they are efficient, reusable and scare-benign. To realize their full potential, it is very important to: (i) advance in adopting common standards; so building blocks can be shared by as many users as possible; (ii) further optimize the design of cloning strategies to improve speed and efficiency; (iii) improve users´ experience by generating new hardware (building blocks and modules) and software (databases and assembly programs) tools which simplify and facilitate the engineering process.

To facilitate the implementation of Plant Synthetic Biology approaches, we present GoldenBraid 2.0 (GB2.0), a new version of the GoldenBraid cloning strategy. In this new version, we defined, in concert with MoClo developers, a common assembly standard by establishing arbitrary, yet scar-benign, assembly seams within a TU which facilitates part exchangeability. In addition, we optimized the versatility of the GB strategy by enhancing its minimalist design, creating a universal part entry vector and simplifying the cloning setup. Finally, we generated a collection of pre-made genetic modules and new software tools for the purpose of facilitating the building of frequently used genetic structures. In short, we present a new *grammar* for Plant Synthetic Biology and we introduce a comprehensive toolkit to facilitate the use of GB2.0 in composing new genetic designs.

2. Results

2.1 The GB2.0 cloning strategy.

To describe the GB2.0 assembly strategy, we follow an analogy with a natural language because we believe this comparison closely describes the GB2.0 cloning strategy structure and facilitates its understanding. This is because the hierarchical manner in which the different building blocks in GB2.0 are combined to form a multigenic structure become analogous to the way grammar elements (morphemes, words, phrases and sentences) are combined hierarchically to create a composition. Figure 15A provides an equivalence table between the elements of English grammar and the elements of the GB2.0 system.

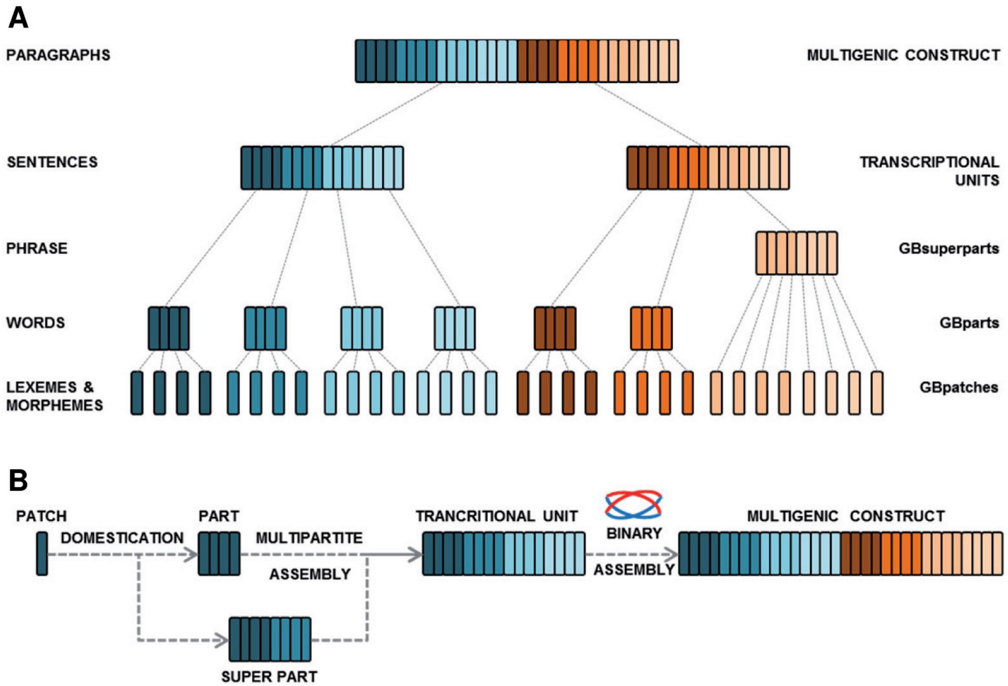


Figure 15. Analogies between GB2.0 and English grammar.

(A) GB2.0 elements can be compared with those of a natural language. In English grammar (left), morphemes are joined together to make words; words are combined together to make phrases and sentences, which are further joined to make a composition. In GB2.0 (right), the simplest units are GBpatches, used to build any of the 11 standard GBparts. GBpatches can be also combined in GBSparts to facilitate cloning (e.g. a whole promoter). GBparts and GBSparts are combined in a multipartite reaction to build TUs, which can be used for plant transformation, or can be reused and combined with other TUs to build multigene modules. **(B)** Flow chart of the GoldenBraid assembly steps. It starts with the GoldenBraid domestication of GBpatches into GBparts or GBSparts; GBparts are multipartitely combined to build up TUs; finally, TUs are binarily assembled to build modules and multigene constructs.

2.2 GBparts: words and phrases. Definition of the GB2.0 grammar .

The first task in upgrading GoldenBraid was to define the minimal standard building blocks in GB2.0, the so-called GBparts, which can be considered the "words" of the GBgrammar. GBparts are fragments of DNA flanked by four nucleotide overhangs. They are stored as inserts within a specially designed entry vector (pUPD), from where they are released by cleavage with BsaI or *BtgZ*I restriction enzymes to generate the corresponding flanking overhangs. GBparts are classified into different classes or categories according to their specific function. Each GB class is defined by its flanking four nucleotides which will overhang upon enzyme digestion

and will determine its position within the TU. We defined eleven standard classes (Figure 16A), which correspond to the basic functional categories in a typical TU. The first three categories (01, 02 and 03) were orderly set in the 5' non-transcribed region, and correspond to operators and promoter regions. Next, we defined seven categories in the transcribed region: one corresponded to the 5' UTR (11); one related to the 3' UTR (17); four were reserved to the coding region (13-14-15-16); an additional class was set as a buffer zone to facilitate, among other designs, the construction of non-coding TUs intended for gene silencing. Lastly, we set a final class (21) for standard 3' un-transcribed GBparts.

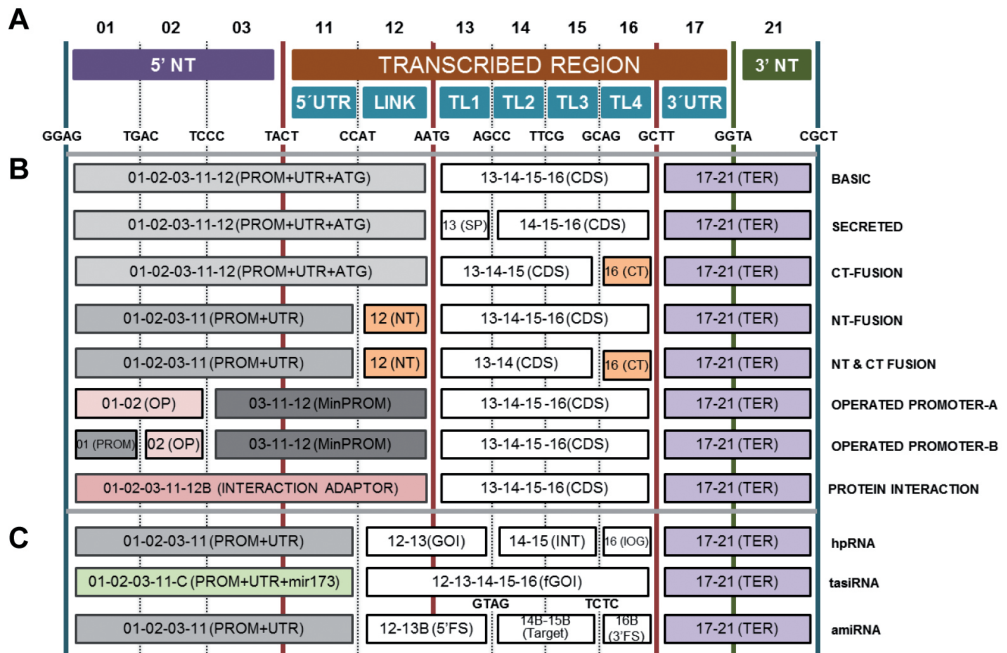


Figure 16. The complete GB2.0 grammar and its most frequently used structures.

(A) Schematic overview of a TU structure where the 11 standard GoldenBraid classes are depicted: 01, 02, and 03 GBparts form the 5′ non-transcribed region (5′NT); position 11 is the 5′ UTR of mRNA; 12 is a linker region; 13 to 16 (TL1–TL4) are four divisions of the translated region; 17 is the 3′ UTR of mRNA; and 21 is the 3′ non-transcribed region of the TU (3′NT). **(B)** Frequently used structures for the protein-coding TUs. The elements forming each frequently used structure and the class that they belong to are depicted. **(C)** Frequently used structures for RNA silencing, including amiRNA, hpRNA, and tasiRNA. 5′NT and 3′NT as well as 5′ UTR and 3′ UTR are defined above; LINK represents a region between the 5′ UTR and the coding sequence where tags or fused proteins can be placed; PROM is a promoter; CDS is the coding DNA sequence; TER represents the terminator; SP is signal peptide; NT and CT are N- and C-terminal tags or fusion proteins; OP is a promoter operator; MinPROM is a minimal promoter; 5′FS and 3′FS indicate the flanking sequences of the amiRNA precursor sequences; Target represents the region of the amiRNA structure comprising the loop and the complementary target sequences; GOI and IOG are the fragments of the gene of interest in an inverted orientation; INT is the intron for hpRNA processing; mir173 represents the mir173 target site for tasiRNA processing; fGOI indicates the fragment gene of interest to be silenced.

Besides the basic classes, GB2.0 also employs "superclasses". For practical purposes, it is convenient to group several contiguous basic GBparts which, together, perform a defined function (e.g., a complete promoter or a full coding region) in a single DNA element (a GB superpart, abbreviated to GBSpart) instead of splitting it into its basic standard parts. This is analog to an English phrase, which comprises a group of words that functions as a single unit within the hierarchical structure of the sentence syntax (e.g., a subject or a direct complement). As with GBparts, GBSparts are ultimately DNA fragments stored within the pUPD vector. Upon digestion with BsaI or *BtgZ*I, the whole phrase is released as a solid indivisible unit flanked by four nucleotides barcodes. In practice, GBSparts are very convenient as they reduce the number of elements that need to be assembled to produce a TU and, therefore, they enhance efficiency. Frequently used superclasses are depicted in Figure 16B and C. For example, the promoter regions normally employed in traditional cloning correspond to the superclass (01-12). GBparts and GBSparts are components of the GB collection, and their sequence information is stored in the GBdatabase.

2.3. GBpart domestication: creating words and phrases.

The process of adapting a DNA building block (GBparts or GBSparts) to the GBgrammar is referred to as domestication. GB domestication usually involves the PCR amplification of the target DNA (word or phrase) using GB-adapted primers (see Figure 17 for details), and the subsequent cloning of the resulting PCR fragment into the pUPD vector using a *BsmB*I restriction-ligation reaction. Occasionally, domestication may involve the removal of internal BsaI, *BsmB*I or *BtgZ*I restriction sites. In order to facilitate an eventual automation of the cloning process, the GB2.0 system includes a standard procedure for internal site removal. This procedure, described in detail in Supplemental Figure 1, involves the amplification of the target DNA in separated fragments (named GBpatches) using GB-adapted primers, which incorporate single mismatches to disrupt the enzyme target sites. Once amplified, GBpatches are re-assembled together in a single-tube *BsmB*I restriction-ligation reaction into pUPD to yield a domesticated GBpart or GBSpart.

2.4. The GB2.0 destination plasmids kit.

GoldenBraid destination vectors (pDGBs) are binary vectors that function as recipients of new assemblies. Each pDGB contains a GBcassette (the selection lacZ gene flanked by two restriction/recognition sites corresponding to two different Type IIS enzymes; see

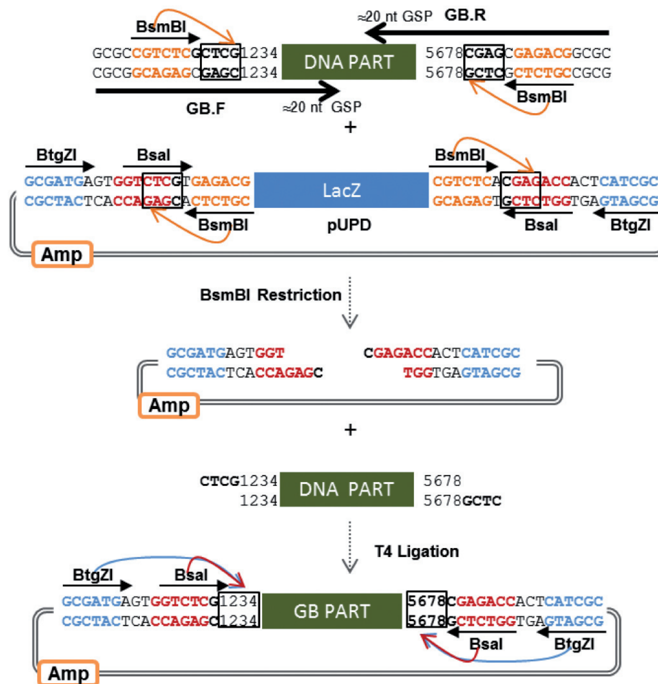


Figure 17. Standardized domestication of GBparts.

GBparts are domesticated by amplifying the desired sequence with standard GBprimers (GB.F and GB.R). GBprimers include approximately 20 nucleotides of the gene-specific primer (GSP) and a tail region that includes a *BsmB*I recognition site, the cleavage site for cloning into pUPD, and the four-nucleotide barcode (1234 and 5678). The amplified DNA part is cloned into pUPD in a restriction-ligation reaction, with *BsmB*I as the restriction enzyme. The resulting GBpart is cleavable by *Bsa*I and *BtgZ*I to produce 1234 and 5678 flanking overhangs. *BsmB*I recognition sequences are depicted in orange in the DNA sequence; *Bsa*I and *BtgZ*I are labeled in red and blue, respectively. Enzyme cleavage sites are boxed.

Figure 4A). In addition, GB2.0 plasmids include a watermark (i.e., a distinctive restriction site flanking the GBcassette) to help plasmid identification. Detailed information about the sequence of the different GBcassettes is also provided in Figure 18A. The special orientation and arrangement of the restriction enzymes defines two levels of pDGBs; the α-level and Ω-level plasmids; which are used for the *Bsa*I and *BsmB*I-GB reactions, respectively. Plasmids also differ in the resistance marker that is associated with each level (kanamycin for level α and spectinomycin for level Ω, allowing counter-selection). To ensure an endless cloning design, a minimum set of four pDGBs is required (pDGBΩ1, pDGBΩ2, pDGBα1 and pDGBα2). Additionally, this set can be expanded to eight plasmids to enable assemblies in different orientations (pDGBΩ1R, pDGBΩ2R, and pDGB1αR and pDGBα2R). For GB2.0, we constructed

two complete sets of pDGBs, one based on the pGreen-II backbone and another set based on the pCAMBIA backbone. The sequence information of all 16 pDGBs in GB2.0 is uploaded in the GBdatabase.

2.5. The GB2.0 destination plasmids kit.

The GB2.0 cloning strategy comprises two types of assemblies (see the GB2.0 chart in Figure 1B): multipartite assemblies and binary assemblies. Multipartite assemblies are performed to create single TUs. The different GBparts and GBSparts required to produce a well-constructed TU are mixed together in a single tube in the presence of a pDGB, the corresponding Type IIS restriction enzyme/s, and the T4 ligase and they are incubated in cyclic restriction-ligation reactions. If all the elements are correctly set in the reaction, they orderly assemble within the destination vector and generate a so-called expression vector, which harbors the assembled composite part. Our pDGBs are binary vectors; therefore the resulting expression clone is ready to be used directly for *Agrobacterium*-mediated plant transformation.

After building a new TU using a multipartite assembly, the resulting new expression clone can be binarily combined with another expression clone to produce increasingly complex multigene structures analogously to how sentences are combined to create a written composition. The solution provided by GB cloning relies on the special design of GB destination vectors, which introduces a double loop (braid) into the cloning strategy. A composite part (a TU or a group of TUs) cloned in a given entry vector can be combined only with a second composite part cloned in the complementary entry vectors at the same level. This is done in the presence of a destination vector of the opposite level and generates a new expression vector at the opposite level. A formal notation describing the rules for multipartite and binary assemblies is shown in Figure 18B and C.

By choosing appropriate combinations of expression and destination vectors, it is possible to create increasingly complex structures, and the only limits are the capacity of the vector backbone or the biological restrictions imposed by bacteria. Moreover, all the new composite parts are fully reusable (they can be used directly for part transformation or can be employed in new assemblies) and exchangeable (can be combined with the GB modules that are produced separately in different labs by following the same assembly rules).

2.6. Innovative features in the GB2.0 cloning strategy.

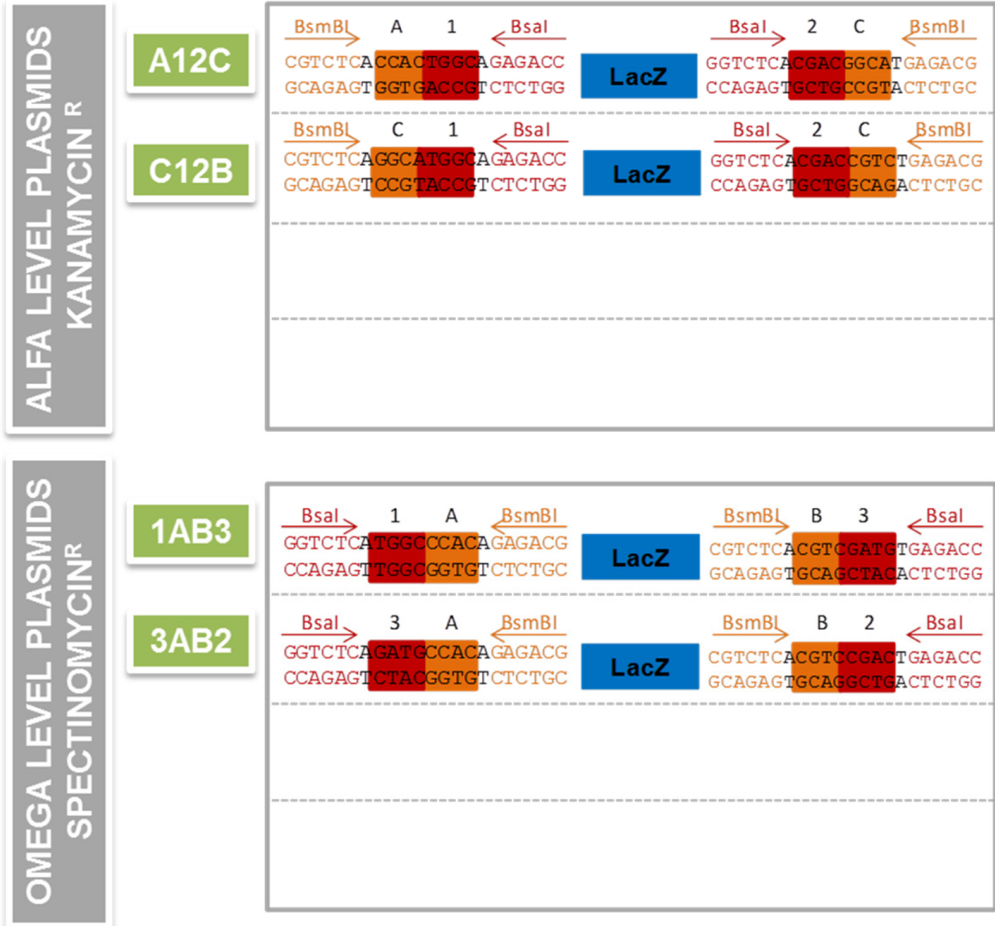
Besides a proposal for a grammar, GB2.0 introduces a number of new elements that modify the original GoldenBraid cloning design to make it simpler and more versatile. Many of the new GB2.0 features rely on the design of the plasmid that harbors GBparts and GBSparts, the Universal Domesticator (pUPD). The pUPD cassette is designed to serve as a polyvalent entry vector for all the different GBparts and GBSparts, regardless of their category. This is because the four nucleotide barcodes are incorporated into the GBpart by PCR instead of being imprinted in the plasmid itself. Such a universal plasmid enables us to establish a single standard protocol for all the domestication parts based solely on its sequence information and category specification.

Another innovative feature of pUPD is the incorporation of both *BtgZ*I and BsaI sites flanking the GB cassette. The enzyme target sites are arranged in such way that both BsaI and *BtgZ*I digestions release exactly the same piece of DNA which contains the same four nucleotide overhangs, regardless of the enzyme used. This opens up the possibility of GBparts being assembled into the α and Ω level vectors indistinctly by using either *Bsa*I-reactions or *BtgZ*I/*BsmB*I-reactions, respectively. To enable this option, the GB cassettes in the pDGBs have also been redesigned and simplified. In the previous version, the sequences of the restriction sites for *Bsa*I (named A, B and C) differed from the restriction sites for *BsmB*I (named 1, 2 and 3). In GB2.0, we made A≡1, B≡2 and C≡3 (see Figure 18A for details). In this way, and by making full use of the dual *Bsa*I/*BtgZ*I release from pUPD, any pDGB can be used as a recipient of a multipartite assembly which, therefore, makes entry in the GB loop fully symmetric. Thus, *Bsa*I reactions are performed to build TUs in α-vectors, and *BsmB*I/*BtgZ*I reactions (*BsmB*I to open pDGB and *BtgZ*I to release the GBpart) are performed to build TUs in Ω-pDGBs. Furthermore by choosing any of the reverse pDGB plasmids as recipients, TU orientation can be inverted. This opens up the possibility of creating new binary assemblies in all the possible relative orientations.

The pUPD design provides yet another interesting new feature to GB2.0 as it enables the use of a non-standard assembly level operating below the standard GBpart level (referred to as the GBpatch level). This feature can be most convenient for a number of applications, including the generation of seamless junctures, introducing combinatorial arrangements into protein engineering, or for promoter tinkering using non-standard positions. The process is similar to the above-described domestication procedure. An example of the use of the GBpatch level for combinatorial antibody engineering is depicted in Supplemental Figure 2

A

GOLDENBRAID



B **RULES FOR MULTIPARTITE ASSEMBLY**

pUPD(01-a)… + pUPD(c-d)… + pUPD(f-21) + $pDGB\alpha_i$ + *Bsa*I + ligase = $pEGB\alpha_i(X)$
pUPD (01-a)…+ pUPD (c-d)…+ pUPD (f-21) + $pDGB\Omega_i$ + *BsmB*I + *BtgZ*I + ligase = $pEGB\Omega_i(X)$

C **RULES FOR BINARY ASSEMBLY**

$pE\alpha1(X_i)$ + $pE\alpha2(X_j)$ + pDGBΩ1 + *BsmB*I + ligase = $pEGB\Omega1(X_i+X_j)$
$pE\alpha1(X_i)$ + $pE\alpha2(X_j)$ + pDGBΩ2 + *BsmB*I + ligase = $pEGB\Omega2(X_i+X_j)$
$pE\Omega1(X_i)$ + $pE\Omega2(X_j)$ + pDGBα1 + *Bsa*I + ligase = $pEGB\alpha1(X_i+X_j)$
$pE\Omega1(X_i)$ + $pE\Omega2(X_j)$ + pDGBα2 + *Bsa*I + ligase = $pEGB\alpha2(X_i+X_j)$

Figure 18. GB2.0 cassettes and assembly rules.

(A) GB2.0 cassettes and their comparison with the previous GoldenBraid version. GBcassettes comprise a LacZ selection cassette flanked by four Type IIS restriction sites (*Bsa*I and *BsmB*I) positioned in inverse orientation. The previous GoldenBraid plasmid kit comprised four destination plasmids, two in each assembly level. GB2.0 incorporates four additional plasmids that permit the assembly of transcriptional units in reverse orientation using the same GBparts. Additionally, the six four-nucleotide barcodes of GoldenBraid (A, B, C, 1, 2, and 3) collapsed in only three GB2.0 barcodes, where A ≡ 1, B ≡ 2, and C ≡ 3. This special design feature permits GBparts to be directly assembled in both level plasmids. Finally, GB2.0 plasmids incorporate distinctive restriction sites flanking the GBcassette as watermarks for plasmid identification. *Bsa*I cleavage sequences are boxed in red, *BsmB*I cleavage sequences are boxed in orange, and sites where both enzymes can digest are boxed in green. The watermark restriction sites are underlined. **(B)** Rules for multipartite assembly. The pUPD elements represent each GBpart and GBSpart that conforms to a grammatically correct TU, pDGBΩi is any level Ω destination vector, pDGBαi is any level α destination vector, and pEGBΩi (X) and pEGBαi (X) are the resulting expression plasmids harboring a well-constructed transcriptional unit X. (C) Rules for binary assembly. (X_i) and (X_j) are composite parts assembled using the multipartite assembly option; (X_i+X_j) is a composite part of (X_i) and (X_j) that follows the same assembly rules as (X_i) and (X_j); pEGBα1(X), pEGBα2(X), pEGBΩ1(X), and pEGBΩ2(X) are expression plasmids hosting a composite part X; and pDGBΩ1, pDGBΩ2, pDGBα1, and pDGBα2 are destination plasmids hosting a LacZ cassette.

2.6.1. Basic expression cassettes for multigene engineering.

Multigene engineering may require the use of different regulatory regions to avoid the silencing associated with the repeated use of a DNA sequence in the same construct. To meet this requirement, we incorporated several regulatory 5' and 3' regions into the GB2.0 collection. Most 5' regulatory regions are (01-12) GBSparts comprising a promoter and 5'-UTR, whereas 3' regulatory regions are (17-21) GBSparts comprising 3'-UTR and terminator regions. According to this basic set up, full (13-16) ORFs can be easily incorporated into tripartite reactions to build a transcriptional unit. In order to undertake Synthetic Biology projects, it is very important to have a range of regulatory regions available, and that the expression strength provided by each promoter/terminator combination is properly characterized so that the multigene expression can be adjusted accordingly. As a first approach toward the characterization of a set of basic expression cassettes, we finely characterized the relative promoter/terminator strength of a number of cassettes using the Luciferase/Renilla system in transiently-transformed *N. benthamiana* leaves. The characterization of (01-12) and (17-21) regions as individual entities is a relatively straightforward procedure using GB2.0 cloning. However as the collection grows, the individual characterization of all the possible combinations becomes an intractable task. We therefore decided to investigate to what extend the transcriptional strength provided by each "promoter/terminator" (i.e., 01-12_17-21) combination can be inferred from the separated contribution of each region. For this purpose, all the (01-12) promoter regions in the collection were tested by the Luciferase/Renilla system in combination with a common (17-21) terminator region (TNos). In parallel, all the (17-21) terminator regions in the collection were tested in combination with a common (01-12) promoter region (PNos). The "Experimental Transcriptional Activity" (ETA) of each region was calculated as being relative to the Luciferase/Renilla values of a (01-12_17-21) reference combination (PNos_TNos), which was arbitrarily set as 1 (see Figure 19A for the construct details). The ETA(01-12) values ranged between 0.47 ± 0.01 and 15.03 ± 1.44 relative luminescence units, whereas the ETA(17-21) values ranged between 0.77 ± 0.18 and 2.61 ± 0.54 (Figure 19B and C). Using these data, "Theoretical Transcriptional Activity" (TTA) was calculated for each cassette combination (Figure 19D) as the product of the individual ETA of the two regulatory regions. Finally, the Luciferase/Renilla ratio of a number of cassette combinations (covering 65% of total possibilities) were also tested experimentally. As we can see in Figure 19E and Supplemental Figure 3, there is a good agreement between the theoretical and experimental activity values. Of the 34 experimental combinations assayed in the evaluation test, 31 showed deviation in relation to the theoretical values below 2-fold (+/- 0.3 in logarithmic values; for detailed information, see Supplemental Figure 3).

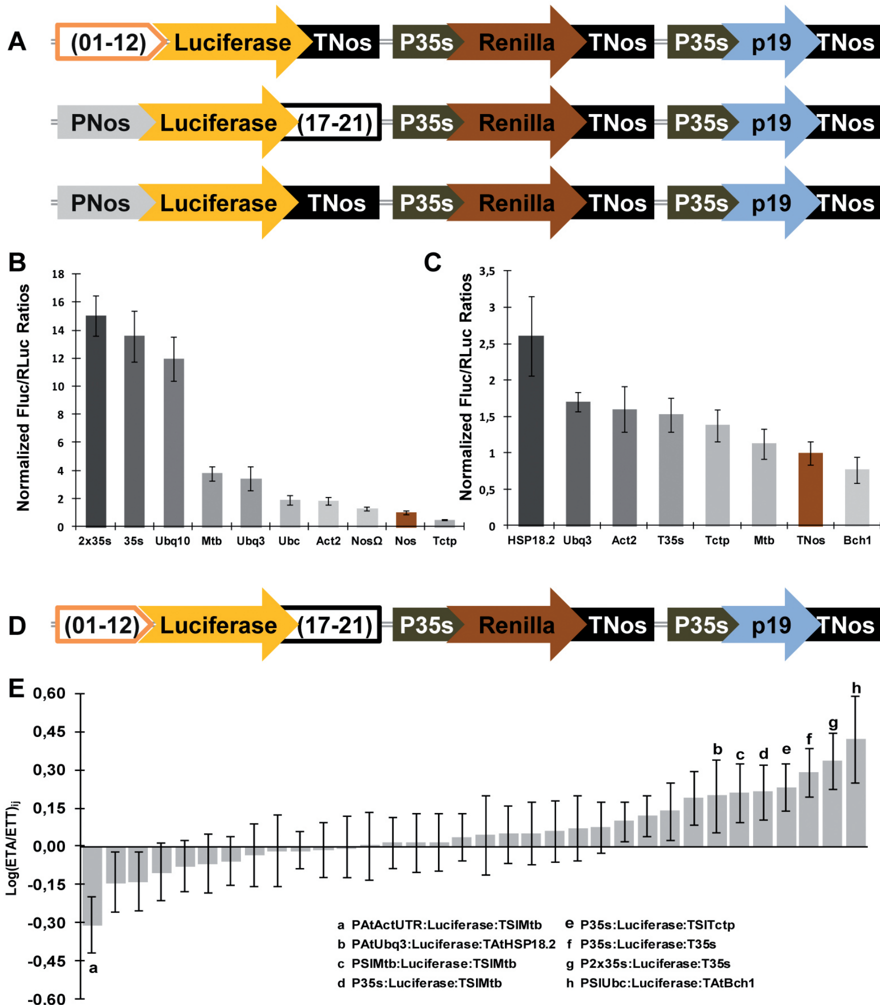


Figure 19. Characterization of regulatory regions for basic expression cassettes.

(A) Constructs for ETA quantification. The promoter (01–12)i_(17–21)TNos constructs comprise a first TU with the (01–12) promoter of interest, the firefly luciferase and the Nopaline synthase terminator, followed by the Renilla reference module (top row). For the (01–12) TNos _(17–21)j terminator constructs, the first TU comprises the (17–21) terminator of interest and the firefly luciferase and the Nopaline synthase promoter (middle row). For activity normalization, the PNos:luciferase:TNos construct combined with the Renilla reference module was used (bottom row). **(B)** The ETA of the promoter regions in (01–12)i_(17–21)TNos constructs was determined as the firefly (FLuc)/Renilla (RLuc) luciferase activity ratio of each construct normalized with the equivalent ratio of the PNos:luciferase:TNos construct. Error bars represent the SD of at least three replicates. **(C)** The ETA of terminator regions in the (01–12)PNos_(17–21) j constructs was estimated as described in B. **(D)** Scheme of the combinatorial promoter/terminator constructs comprising a first TU with a (01–12) promoter, the firefly luciferase, and a (17–21) terminator and combined with the Renilla reference module. **(E)** Correspondence between the ETA and TTA data in the combinatorial constructs. The logarithm of the ratios between the ETA and TTA values for 34 experimental promoter/terminator combinations is plotted. Bars represent average values ± SD. PNos, Nopaline synthase promoter; TNos, Nopaline synthase terminator; P35s, CaMV 35S promoter; P19, Tomato Bushy Stunt Virus silencing suppressor.

2.6.2. Frequently used structures.

There is a limited number of structural types for the majority of synthetic transcriptional units and genetic modules. For instance, many protein-encoding TUs can be constitutively expressed, whereas others are regulated by 5´ (or 3´) operators. The resulting proteins can be preceded by a signal peptide, or may contain C-terminal and N-terminal fusions. Besides, noncoding TUs can be used for silencing purposes. To cope with this functional diversity while simplifying the users´ toolbox, we defined a group of "Frequently Used Structures", for which specific pre-arranged GBparts and GBSparts were developed (depicted in Figure 16B and C). We now go on to describe some of the Frequently Used Structures that are currently included in the GB system and their associated tools.

2.6.3. Regulated expression cassettes.

The GB grammar contains several standard positions for the insertion of regulatory regions. In the 5´ un-transcribed region, we defined three standard GBparts to allow combinatorial promoter tinkering and to facilitate the insertion of synthetic operators. As a functionality proof, we assembled and tested the pre-made cassettes for heat shock and the dexamethasone-regulated expression; the latter is based on the "operated promoter A" scheme shown in Figure 16B. The Luciferase/Renilla/P19 reporter cassettes constructed with promoters pHSP70 and pHSP18.2 showed clear induction after incubation at 37ºC (Supplemental Figure 4). The potential of the GB modular assembly was further demonstrated with the construction of two regulated systems based on the fusion of the glucocorticoid receptor (GR) with the DNA binding domains (BD) of LacI or Gal4 and the activation domain (AD) of Gal4. In this transactivation example, up to 15 pre-made modules comprising coding and noncoding regulatory regions were efficiently assembled de novo to produce two operated luciferase TUs which clearly responded to the presence of dexamethasone (Supplemental Figure 5).

2.6.4. Protein-protein interaction tools.

Reporter fusion partners are powerful analytical tools utilized in the study of protein-protein interactions. However, the use of unlinked co-transformation for the delivery of the interaction partners often compromises the extraction of reliable qualitative data, based on the poorly supported assumption that co-transformation efficiency in each cell is the same for all fusion partners. We reasoned that the linked co-transformation of fusion partners can help improve

the sensitivity and accuracy of the protein-protein interaction analysis. By bearing this use in mind, we designed pre-made modules for the Bimolecular Fluorescence Complementation assays (BiFC). For this purpose, BiFC adaptors with a (01-12) structure were constructed containing the full CaMV 35S promoter and the corresponding YFP or luciferase fusion partners. Based on this set-up, baits and preys with a canonical (13-16) structure can be easily assembled in multipartite reactions to form the required fusion proteins. The prearranged BiFC tools were functionally tested using transcription factors Akin10/Akinβ2 as positive interaction partners, and an spermidine synthase (SPDS) as a negative partner **[135]**. As observed in Supplemental Figure 6, the number of cells showing positive interactions with the GB-assisted linked co-transformation set up outnumbers those of the unliked co-transformation approach.

2.6.5. Silencing tools.

The negative regulation of endogenous genes often proves an engineering requirement. For this reason, special Frequently Used Structures were defined for three RNA silencing strategies: trans-acting small interfering RNAs (tasiRNA); artificial micro RNA (amiRNA); hairpin RNA (hpRNA) (Supplemental Figure 7). Details of all the elements used in the RNAi designs are provided in Supplemental Table 1.

For the generation of tasiRNA constructs, special (01-11) GBSparts containing the mir173 trigger sequence are required. A CaM35S-based GBSpart for the constitutive tasiRNA expression is currently available in the GB collection. A regulated or tissue-specific tasiRNA expression can be designed using the GBpatch special feature of GB2.0. For the functional characterization of the tasiRNA structure, a 410-bp fragment of *A. thaliana* phyotoene desaturase (PDS) **[136]** was incorporated as a (12-16) GBSpart and was transformed into *A. thaliana* to yield approximately 0.1% seedlings with the albino phenotype (Supplemental Figure 7C). TasiRNA constructs require the co-expression of miR173 for effective silencing in plant species other than *Arabidopsis* **[136]**. To extend the species range of the tasiRNA tool, a new TU with a constitutively expressed miR173 was constructed and incorporated into the collection. The functionality of the dual construct was tested transiently in *N. benthamiana* using PDS as the silencing target, which resulted in the bleaching of the infiltrated area (see Supplemental Figure 7D).

An amiRNA silencing tool was also enabled with the creation of two special GBSparts, namely 5'FS and 3'FS, respectively. These GBSparts require noncanonical barcodes to allow the seamless assembly of 5'FS and 3'FS in the amiRNA precursor. The special categories are

denoted as (12-13B) and (16B), respectively, where B indicates the four noncanonical flanking nucleotides (GTGA and TCTC, respectively). The standard (01-11) promoters without ATG and the (17-21) terminators were used in the amiRNA design. The central region (14B-15B), containing a fragment of the gene target sequence, was constructed using gene-specific oligonucleotides, as described in Supplemental Figure 7B. In order to validate the proposed structure, *A. thaliana* PDS silencing was assayed using a gene target fragment which was formerly described by Yan et al. **[137]** (Supplemental Figure 7E). The resulting amiRNA construct was transformed into *A. thaliana* yielding seedlings with the albino phenotype.

Finally in the hairpin RNA (hpRNA) structure, the regulatory regions lacking ATG are inserted as (01-11) parts. An intron from *S.lycopersicum* (SGN-U324070) was incorporated into the collection to serve as an (14-15) Intron GBpart. The inverted fragments of the target gene-of-interest can be cloned at positions (12-13) and at position (16).

2.6.6. GW-GB adapter tool.

A GW-GB adapter tool was incorporated into the GB2.0 collection in order to facilitate the transition between the Gateway (GW) and GB2.0 assembly methods (Supplemental Figure 8). GW-GB adapters are GBparts or GBSparts (e.g., a (12-16) GBSpart to adapt coding regions) made of a GW cassette flanked by *att*R1-*att*R2 sites and embedded inside the pUPD plasmid. As such, adapter vectors can be used directly as destination plasmids for GW entry clones flanked by *att*L1-*att*L2 sites. In this way, GW entry clones can be transferred individually or in bulk to the pUPD plasmid, and become ready-to-use GBSparts. Alternatively, the GW-GB adapter can be employed as an ordinary GBSpart to create a new multigene construct in a binary vector. Consequently, the resulting multigene construct becomes a GW destination vector containing an *att*R1-*att*R2 GW cassette, where GW entry clones can be inserted individually or in bulk. It should be noted that direct GW to GB2.0 adaptation does not remove internal enzyme target sites, therefore the efficiency of subsequent assembly reactions can lower.

2.7 GB collection and software tools.

When this manuscript was being written, our in-house GB collection contained more than 400 entries. As the collection grows, engineering is becoming increasingly easy and fast because, on occasion, the required GBparts, GBSparts and TUs are already domesticated and/or constructed. To efficiently handle this collection, we developed a web framework which hosts a GBdatabase and offers software tools to facilitate the assembly process.

The GB2.0 website was implemented using Django, a Python web framework that supports rapid design and the development of web-based applications **[138]**. Object-relational database management system PostgreSQL was chosen to host our schema, which allowed the incorporation of the sequences of all the elements included in the collection. Additional relevant information on part identity, functionality and indexing is also provided.

Given the simplicity of the GB assembly rules, it was relatively straightforward to develop software tools that assist in GB2.0 assembly. We therefore developed a software package comprising three programs, each program corresponding to one of the three basic processes in GB2.0 assembly. The first program, named GBDomesticator, assists the part adaptation process to the GB standard. It takes an input DNA sequence provided by the user, and it offers the best PCR strategy to remove internal enzyme target sites and to add flanking nucleotides to it according to the specified category. A second program, known as the TUassembler, takes GBparts and GBSparts from the database and simulates a multipartite assembly *in silico*. The TUAssembler includes shortcuts to Frequently Used Structures assembly, as well as a free-hand option. Finally, a third program, namely BinaryAssembler, performs *in silico* binary assemblies between the composite parts stored in the GB database. BinaryAssembler offers the possibility of choosing the relative orientation of each member of the assembly. All three programs generate a detailed lab protocol to perform the domestication/assembly and to return a GenBank formatted file containing the final domesticated/assembled sequence. The GB database and software tools are available at www.gbcloning.org

3. Discussion

The aim of this work is to provide a standard framework for DNA assembly in Plant Synthetic Biology. We, and others, realized that the modularity of the multipartite assembly based on Type IIS enzymes offers a great opportunity for standardization by following a positional information scheme that resembles the grammar of a sentence in many natural languages. Indeed it is illustrative to conceive the transcriptional unit as a similar structure to a sentence, which is made up of hierarchically assembled elements like morphemes, words and phrases. It is also interesting to envision the whole engineering process as a way to imprint instructions using DNA strings. Therefore we, in concert with MoClo developers, propose a common grammar where the four nucleotide overhangs are pre-defined for each position within the transcriptional unit. Overhangs assignation is mainly arbitrary, but some decisions are made to make them scar-benign. For instance, the 12-13 boundary defining the beginning of CDS was designed to include the start codon, whereas the 13-14 boundary was made compatible with signal-peptide cleavage sites.

In our view, this new GB2.0 cloning scheme has a number of features which makes it a good candidate for a plant assembly standard. Many of those features are consubstantial to the Golden Gate system: very high efficiency, modularity and the ability to produce scare-benign assemblies. GB2.0 also incorporates the reusability and modularity of the GoldenBraid and MoClo systems and goes beyond them in that it provides a standardized framework, goes deep into the versatility and the minimalist design of the GoldenBraid loop, and incorporates new tools to assist cloning.

A major drawback of defining a standard is loss in versatility since no standard can cope with all custom design requirements. To deal with this problem, we incorporated an underlying non-standard assembly level which makes full use of the newly designed pUPD vector. At this level, non-standard GBpatches can be custom-designed for, e.g., scar-less assembly, by choosing the appropriate four nucleotide overhangs. GBpatches are assembled together into standard GBparts or GBSparts. We made full use of the GBpatch level for BiFC, amiRNA and antibody engineering. Other possible uses include promoter tinkering or non-standard combinatorial assemblies within the CDS, as exemplified in the construction of customized TAL effectors **[139-140]**. Additionally, the GBpatch level is used for GBpart domestication; that is, for the removal of internal enzyme recognition sites. This feature is also enabled by the special design of the new entry vector pUPD, which introduces inversely oriented *BsmB*I sites into the GB cassette. This new design turns pUPD plasmid into a universal entry vector as the four nucleotides conferring part identity are not located in the entry vector as they are in previous designs **[89, 139]**. Instead in the present setup, the four-nucleotide “barcode” is incorporated into the primers used during initial part/patch isolation. As a toll, this strategy involves the requirement of longer PCR primers during initial part isolation. This minor drawback is by far compensated by the simplicity introduced by the universal domesticator: in the absence of this solution, a minimum of eleven different entry vectors would be required to harbor the different categories in the GB grammar, along with an unaffordable amount of additional vectors to allow the formation of all the possible “phrasal” combinations.

The underlying GoldenBraid cloning pipeline has been substantially simplified in the GB2.0 version to reduce redundancy and to achieve a minimalist design. Figure 20 depicts the comparison of GB2.0 made with the previous GoldenBraid structure. Once again, most of the improvement achieved stems from the specific design of the new entry vector pUPD. First, the asymmetry of the cloning loop is corrected in GB2.0 with the introduction of a *BtgZ*I site into the entry vector. *BtgZ*I is a special enzyme that cuts 10 nucleotides away from its recognition site. This feature enables a dual release option for each GBpart: *Bsa*I release allows cloning in α destination vectors, whereas *BtgZ*I release allows cloning in Ω destination vectors. We noted that *BtgZ*I/*BsmB*I assemblies are less efficient than *Bsa*I ones. Despite

this drawback, the ability to create new TUs in both destination vectors can save one cloning step, which therefore speeds up the construction of new multigene assemblies and opens up new possibilities for automation.

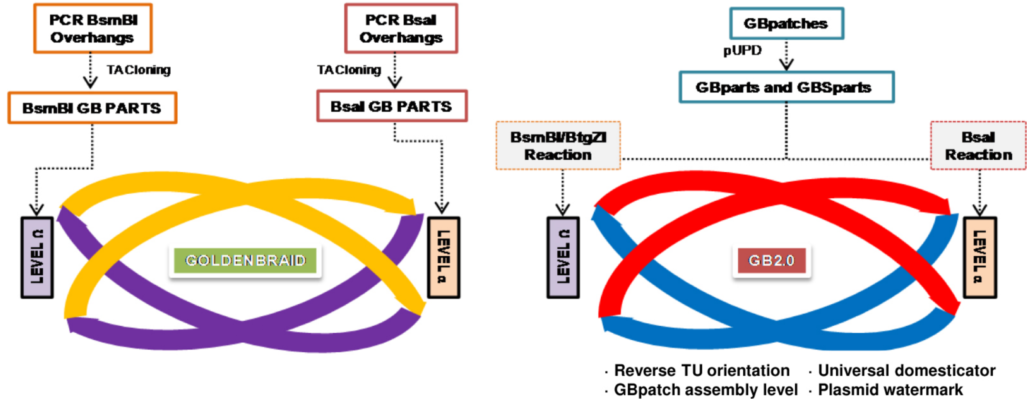


Figure 20 . GoldenBraid versus GB2.0.

(A) The previous GoldenBraid version had an asymmetric assembly flow (left). GBparts incorporated either the *Bsa*I or the *BsmB*I releasable overhangs. *Bsa*I-released GBparts were incorporated into the GoldenBraid cloning loop through level α vectors, whereas the *BsmB*I GBparts were used to build composite parts through the level Ω entry point. In the new GB2.0 symmetric design, the same GBparts can be incorporated into level α plasmids by a *Bsa*I restriction/ligation reaction or into level Ω vectors by a mixed *BsmB*I/*BtgZ*I reaction. Other differences between the previous GoldenBraid version and GB2.0 are also listed. TA Cloning, PCR bands cloned using adenosine/thymine overhangs.

We also developed a number of tools to assist users in their engineering projects. First, we anticipated genetic designers' needs by pre-arranging a number of FUS. Then, we populated our in-house collection with all the elements (GBparts, GBSparts and software tools) required to enable the Frequently Used Structures use. Finally, we assayed the functionality of newly developed elements using *in planta* assays. In certain cases, this implied an initial step toward part characterization. One of the hallmarks of Synthetic Biology is its ability to predict the behavior of a system based on the characteristics of its constitutive parts. We show herein that it is possible to infer the activity provided by a "promoter + terminator" pair from the activities that each individual element displays when separately assayed. The differences observed between the theoretical and experimental activity values fall within a narrow range which comes close to 0, with very few combinations showing deviations that are slightly above 2-fold (+/- 0.3 in log values). This finding is important for engineering attempts which, as in complex metabolic engineering, require the combination of many different non coding parts to create large metabolic pathways, while avoiding the introduction of unstable repetitive regions into the genetic design. The promoter parts assayed herein reveal a wider range

of activities than terminators. Nevertheless, we confirm that the use of strong terminators like TAtHSP18.2 can promote the promoter's transcriptional activity, as previously described **[141]**. It is interesting to note that most of the observed positive deviations result from the combinations involving CaMV 35S-derived parts, suggesting a nonlinear behavior of the CaMV 35S regulatory elements. We employed *N. benthamiana* transient expression and Luciferase/Renilla reporter system **[142]** as a first step towards characterization of regulatory elements. This transient methodology is simple and accurate and therefore facilitates the analysis. A more detailed characterization may need to include the developmental and tissue-specificity information obtained through stable plant transformation.

Both GB2.0 and the GB collections come into being with to an open-source vocation. We reinforced this point by developing a new set of GB-destination vectors based on open-source pCAMBIA binary vectors **[143-144]**. As we see it, the intellectual commons IP model is that which best suits the requirements for the free exchange of parts and modules in Plant Synthetic Biology **[145]**. Nevertheless, a number of issues, such as the IP of individual parts and the ability to freely distribute them, need to be addressed in a concerted manner. Undoubtedly, community effort made to create publicly available collections of synthetic parts will have an impact on the progress of this discipline.

Plant Synthetic Biology has the potential of bringing about a significant impact on crop production. Engineering enhanced abiotic stress tolerance for growth in marginal lands, turning C3 plants into C4 **[146]**, constructing whole-organism biosensors or sentinels **[8]**, engineering highly challenging metabolic routes **[5]**, and combinations of these, are just some examples of high-impact goals with biotechnologists' reach. Also, it has not escaped our notice that the proposed grammar can be easily adopted by other non-plant systems as well. We believe that technologies like GB2.0, which enable the standardization and facilitate the characterization and exchange of genetic parts and modules, are important contributions for the achievement of the challenging biotechnology goals ahead.

4. Material and Methods

4.1. Strains and growth conditions.

E. coli DH5α was used for cloning. *A. tumefaciens* strain GV3101 was used for transient expression and transformation experiments. Both strains were grown in LB medium under agitation (200 rpm) at 37°C and 28°C, respectively. Ampicillin (50 μg ml^{-1}), kanamycin (50 μg ml^{-1}) and spectinomycin (100 μg ml^{-1}) were used for *E. coli* selection. Rifampicin, tetracycline

and gentamicin were also used for *A. tumefaciens* selection at 50, 12.5 and 30 μg ml^{-1}, respectively. XGal (0.5 mM) and IPTG (40 μg ml^{-1}) were used in LB agar plates for the white/blue selection of clones.

4.2. Restriction-Ligation assembly reactions.

Restriction-Ligation reactions were set up as described elsewhere **[90]** using *Bsa*I, *BsmB*I, *BtgZ*I or BbsI as restriction enzymes (New England Biolabs, Ipswich, MA, USA) and T4 Ligase (Promega, Madison, WI, USA). Reactions were set up in 25 or 50 cycles digestion/ligation reactions (2' at 37º C, 5' at 16º C), depending on assembly complexity. One μl of the reaction was transformed into *E. coli DH5*α electrocompetent cells and positive clones were selected in solid media. Plasmid DNA was extracted using the E.Z.N.A. Plasmid Mini Kit I (Omega Bio-Tek, Norcross, GA, USA). Assemblies were confirmed by restriction analysis and sequencing.

4.3. GBpart Domestication.

GBparts and GBpatches were obtained by PCR amplification using suitable templates. The Phusion® High-Fidelity DNA Polymerase (ThermoScientific, Waltham, MA, USA) was used for amplification following the manufacturer's protocols. Primers smaller than 60 mers were purchased from Sigma-Aldrich (St. Louis, MO, USA). 60-mer or longer oligonucleotides were synthesized by IDTDNA (Coralville, IO, USA) by the Ultramer™ technology. Amplified bands were purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and were quantified in a Nano Drop Spectrophotometer 2000. Then 40 ng of each amplicon and 75 ng of the domestication vector (pUPD) were mixed and incubated in a *BsmB*I restriction-ligation reaction. The pUPD sequence is deposited in the GBdatabase. Positive clones were selected in the ampicillin-, XGal- and IPTG -containing plates, and the correct assembly was confirmed by restriction analyses and sequencing. A description of the GBparts and GBSparts employed in this work is provided in Supplementary Table 1. The nucleotide sequence of all the GB parts in the collection is deposited in the GB database.

4.4. pDGB Construction.

Two pDGB series, pDGB1 and pDGB2, were constructed. pDGB1 is based on the pGreenII backbone **[131]** and pDGB2 is based on pCAMBIA **[143]**. For pDGB construction, the backbone of each binary vector was divided into fragments (vector modules). The pDGB1

backbone comprised two fragments, whereas the pDGB2 backbone was divided into four modules given the presence of internal sites. To build vector modules, each fragment was amplified by PCR in a similar procedure to that described for GBparts and was cloned into a vector domestication plasmid (pVD) using a *Bsa*I digestion-ligation reaction. The pVD vector was derived from pUPD; its sequence is deposited in GBdatabase. In addition to the backbone modules, a number of common modules were built: eight GB-cassettes (α1, α1R, α2, α2R, Ω1, Ω1R, Ω2 and Ω2R) and two fragments encoding spectinomycin and kanamycin resistance. To assemble each pDGB, a BbsI restriction-ligation reaction was performed by combining the modules of the vector backbone, the desired GB-cassette and appropriate antibiotic resistance.

4.5. *N. bentamiana* transient transformation.

For the transient expression experiments, plasmids were transferred to *A. tumefaciens* strain GV3101 by electroporation. Agroinfiltration was performed, as previously described **[132]**. Overnight-grown bacterial cultures were pelleted and resuspended in agroinfiltration medium (10 mM MES pH 5.6, 10 mM $MgCl_2$, 200 μM acetosyringone) to an optical density at 600 nm =0.5. Infiltrations were carried out using a needle-free syringe in leaves 2, 3 and 4 of 4–5 weeks old *Nicotiana benthamiana* plants (growing conditions: 24°C day/20°C night in a 16 h light/8 h dark cycle). Depending on the purpose of the experiments, leaves were harvested 3-5 days post-infiltration (d.p.i.) and examined for transgene expression.

4.6. *A. thaliana* stable transformation.

A. thaliana Col-0 accession plants were transformed by the floral-dip method **[147]**. Seeds were sterilized-plated in plates of MS medium with 0.8% (w/v) agar and 1% (w/v) sucrose (growing conditions: 24ºC day/20ºC night in a 16h light/8h dark cycle). Transgenic lines were selected without antibiotic resistance as PDS silencing transformed lines showed the albino phenotype.

4.7. Luciferase/Renilla expression assays.

In order to measure the activity of Luciferase/Renilla reporters **[142]**, 3 or 4 *N. bentamiana* leaves were agroinfiltrated following the above-described procedure. Leaves were harvested 3 d.p.i. *Firefly Luciferase* and *Renilla Luciferase* were assayed from 100-mg leaf extracts

following the Dual-Glo Luciferase Assay System (Promega, Madison, WI, USA) standard protocol and were quantified with a GloMax 96 Microplate Luminometer (Promega, Madison, USA). The "ETA" of each region (ETA) was calculated in relation to the Luciferase/Renilla values of a (01-12_17-21) reference combination (PNos_TNos), which was arbitrarily set as 1, according to the formulae:

$$\mathrm{ETA(01\text{-}02)_i} = \frac{\mathrm{FLuc}/_{\mathrm{RLuc}}\,[(01\text{-}12)_{i}_(17\text{-}21)_{TNos}]}{\mathrm{FLuc}/_{\mathrm{RLuc}}\,[(01\text{-}12)_{PNos}_(17\text{-}21)_{TNos}]}\,;\ \mathrm{ETA(17\text{-}21)_j} = \frac{\mathrm{FLuc}/_{\mathrm{RLuc}}\,[(01\text{-}12)_{PNos}_(17\text{-}21)_{j}]}{\mathrm{FLuc}/_{\mathrm{RLuc}}\,[(01\text{-}12)_{PNos}_(17\text{-}21)_{TNos}]}$$

where FLuc/RLuc $[(01\text{-}12)_{i}_(17\text{-}21)_{j}]$ refers to the ratio between the Firefly luciferase activity (FLuc) of a $(01\text{-}12)_i$:Luciferase:$(17\text{-}21)_j$ construct and the Renilla luciferase activity (RLuc) of a 35S:Renilla:TNos internal standard construct. TTA was calculated for each cassette combination as the product of the individual ETA of the two regulatory regions, as follows:

$$\mathrm{TTA_{ij}} = \mathrm{ETA(01\text{-}02)_i} \times \mathrm{ETA(17\text{-}21)_j}$$

Finally, the FLuc/RLuc of a number of cassette combinations was tested experimentally, and the ETA of each combination (ETAij) was calculated with the formula:

$$\mathrm{ETA_{ij}} = \frac{\mathrm{FLuc}/_{\mathrm{RLuc}}\,[(01\text{-}12)_{i}_(17\text{-}21)_{j}]}{\mathrm{FLuc}/_{\mathrm{RLuc}}\,[(01\text{-}12)_{PNos}_(17\text{-}21)_{TNos}]}$$

4.8. Glucocorticoid Receptor induction and Heat shock treatments.

One-cm^2 disks from agroinfiltrated leaves were harvested at 3 d.p.i., placed in a 350 µl solution containing 5 to 20 µM Dexamethasone (Sigma Aldrich St. Louis, MO, USA) in 0.02% Tween-80 and incubated overnight in a growth chamber. *Firefly Luciferase* and *Renilla Luciferase* activities were measured after 24-hour treatment. For the heat shock treatments, 1 cm^2 of 3 d.p.i. leaves were placed in 350 µl water at 37ºC for 2h. Samples were collected at 3 h and 14 h after treatment.

/C3

CHAPTER 3

ENGINEERING MODULAR AND ORTHOGONAL LOGIC GATES FOR PLANT SYNTHETIC BIOLOGY

Chapter 3: Engineering modular and orthogonal logic gates for Plant Synthetic Biology.

1. Introduction

Synthetic Biology introduces engineering principles to the design of new artificial functions in living organisms, associating genetic expression units to functional devices and genetic regulatory networks to modules or circuits, all of which can be simulated and predicted with mathematical models **[101, 148]**. The first attempts to describe biological regulation mechanisms with mathematical models date back to the 70s **[149-150]** but it is only 20 years ago when researchers started applying electrical circuit analogies to biological pathways. One of the first examples of these recreations was the description of the bacteriophage λ lisogeny vs. lisis life cycles as a decision circuit network using basic Boolean logic gates and simple differential equation models **[151]**. The year 2000 marked an inflexion point for Synthetic Biology. Two regulatory genetic circuits consistent with its mathematical predictions were described, including an repressillator **[94]** and a toggle-switch circuit that permit the flipping between two stable states using external induction **[93]**. Since then, many developments have been reported, including the reprogramming of cells to perform sophisticated electronic-like computations **[152]**, counters **[96]** and cells showing long-term maintenance of memory and therefore able to serve as data storage units **[95-96, 104]**. Furthermore, complex functional circuits have been engineered in different organisms using GFP or Luciferase as reporters **[153-154]**. Practical applications on the biomedical field include *E. coli* sentinels that specifically detect and inhibit the growth of the infectious *Pseudomonas aeruginosa* **[155]** or tumor growth control in mice using engineerly attenuated *Salmonella enterica* carrying an expression circuit triggered by acetyl salicylic acid **[156]**.

An initial mandatory step in the construction of genetic circuits is the development of logic gates with orthogonal elements, which are essential for the robust assembly of more complex circuits. The Plant branch of Synthetic Biology is a nascent research area and therefore the construction and testing of most basic engineering tools as logic gates is still nonexistent or rudimentary. The most common logical gates are binary operations, which involve two inputs and produce a single output. There are sixteen different binary operations, also called

Boolean functions and all of them can be assembled by the combination of simple one-input one-output operations. Boolean logic gates could be used directly for the generation of gene expression profiles that lay beyond the spectrum of natural promoters. As an example, an AND gate can be used to limit the response of a chemically inducible promoter to a particular cell type, whereas an OR gate can connect two different environmental signals to produce a single response. Indirectly, logic gates will probably be on high demand, together with synthetic promoters or artificial transcription factors, to build higher order components such as artificial toggle switches, counters and oscillators, that can be used for fine-control the gene expression in plants **[35]**.

In this chapter, we describe the initial steps in the development of orthogonal logic gates in plants. We focus first on testing the set of one-input one-output operations. Later, we combine some of these basic operations to implement the Boolean OR gate. We also undertake the initial steps for the construction of an AND gate. All the composition we describe here have been constructed using the GB2.0 Framework **[90-91]**, demonstrating how a multigene assembly standard can be used in the creation of *plug-and-play* devices.

2. Results

All the constructs in this chapter (see Table 1) were obtained by combination of GBparts and GBSparts, which are listed in Supplemental Table 2 and described in more detail in the GBDataBase. All operators were tested in transient expression agroinfiltration experiments on *N. bentamiana* leaves using a Luciferase/Renilla reporter system integrated in the same T-DNA.

2.1. Simple operations.

Simple operations having one input and one output are the basis for the development of the Boolean Logic gates. We have established and characterized the full set of basic operations. Furthermore the operations have been tested as double-input logic gates, which can be used later as parts of more complex circuits.

INPUT	TRUE	FALSE	PROPOSITION	NOT
0	1	0	0	1
1	1	0	1	0

Figure 21. Truth tables for the one-input one-output models.

Truth Tables are composed of one column for each input variable and one column for each of the possible results of the logical operations included in the table (1 is for *true* and 0 is for *false*).

	OPERATION NAME	SYMBOL	BIOLOGICAL EXAMPLE NAME	CONSTRUCT
(A) SIMPLE OPERATION ONE IMPUT-ONE OUTPUT	TAUTOLOGY	⊤	⊤LACI::GAL4AD	pEGB2Ω1 35S:LacIBD::Gal4AD:TNos-6xOpLacI:mini35S:Luciferase:TNos-PAtUbq10:Renilla:TMtb
	CONTRADICTION	⊥	⊥UAS	pEGB2Ω1 UAS:mini35S:Luciferase:TNos-PAtUbq10:Renilla:TMtb
	PROPOSITION	*P*	*P*GR	pEGB2Ω1 TNos:KanR:PNos-35S:GR::LacIBD::Gal4AD:TNos-6xOpLacI:mini35S:Luciferase:TNos-PAtUbq10:Renilla:TMtb
			*P*ER	pEGB2Ω1 TNos:KanR:PNos-35S:ER::LacIBD::Gal4AD:TNos-6xOpLacI:mini35S:Luciferase:TNos-PAtUbq10:Renilla:TMtb
			*P*HS	pEGB2Ω1 TNos:KanR:PNos-PHSP70B::LacIBD::Gal4AD:TNos-6xOpLacI:mini35S:Luciferase:TNos-PAtUbq10:Renilla:TMtb
	NEGATION (NOT)	¬	¬ GR.LexA::SRDX	pEGB2α1 TNos:KanR:PNos-PAtUbq10:Renilla:TMtb-THSP:SRDX:LexABD::GR:35S-TNos:Luciferasa:mini35S:4xOpLexA:35S(CORE)
			¬ GR.LexA::BDR	pEGB2α1 TNos:KanR:PNos-PAtUbq10:Renilla:TMtb-THSP:BDR:LexABD::GR:35S-TNos:Luciferasa:mini35S:4xOpLexA:35S(CORE)
(B) BINARY OPERATIONS TWO IMPUT-ONE OUTPUT	DISJUNCTION	∨	GR∨HS	pEGB2Ω1 TNos:KanR:PNos-35S:GR::LacIBD::Gal4AD:TNos-HSP70B::LacIBD::Gal4AD:TNos-6xOpLacI:mini35S:Luciferase:TNos-PAtUbq10:Renilla:TMtb
			GR∨ER	pEGB2Ω1 TNos:KanR:PNos-35S:GR::LacIBD::Gal4AD:TNos-35S:ER::LacIBD::Gal4AD:TNos-6xOpLacI:mini35S:Luciferase:TNos-PAtUbq10:Renilla:TMtb
	CONJUNCTION	∧	INTERACTION PDZ-VKESLV	pEGB 2Ω1 35S:PDZ::Gal4AD::TNos-35S::VKESLV::Gal4AD::TNos

Table 1. Logic operations for Plant Synthetic Biology.

2.1.1. Tautology and Contradiction.

The most trivial gates are those corresponding to the TRUE and FALSE states. The TRUE gate (tautology, designed with the symbol ⊤) implies that any input produces a positive value. The FALSE gate (contradiction, ⊥) has the opposite effect that is any input produces unconditionally a negative output (Figure 21). Among all the possibilities that could be engineered with the basic units available as GBparts, two multigenic constructs have been built to illustrate the TRUE and FALSE operations.

The TRUE gate named ⊤LACI::GAL4AD was built as a three TUs construct. The first TU, expressed the LacIBD::Gal4AD module under the control of the constitutive CaMV 35S promoter (LacIBD is the LacI DNA-binding domain and Gal4AD is the transcriptional activation domain of Gal4). The second unit corresponded to the reporter and was composed of six copies of the LacI operator (LacIOp) upstream of a minimal CaMV 35S promoter and the firefly luciferase. The third TU consists of the renilla luciferase driven by the AtUbq10 promoter and serves as internal standard, hereinafter referred as Renilla Module (Figure 22A. See Supplemental Figure 9 for the complete GoldenBraid building pathway). Although the mechanism is trivial, this gate was formally tested to discard unexpected interference and to test the reproducibility with different trigger molecules. This gate was tested by agroinfiltration in N. benthamiana leaves in presence of two different inducers (10 µM dexamethasone and a heat shock treatment). The output levels were similar in all the cases (Figure 22B).

The FALSE operation, named ⊥UAS, incorporated a first TU with an UAS-operated minimal CaMV 35S promoter and the firefly luciferase, and in a second TU the Renilla Module (Figure 22C). The ⊥UAS gate was tested with the same two inputs as done earlier with the TRUE gate, producing a negligible luciferase expression in all the cases (Figure 22D).

2.1.2. Proposition (*P*).

A very simple gate, named proposition, corresponds to the logic operation that only produces a value of *true* when the input is *true*. In genetic terms, it corresponds to an inducible system in which the reporter is only expressed on the presence of the trigger molecule.

We have engineered three versions of this gate, each responding to a different inducer which is dexamethasone for gate *P.GR* (Figure 23A), β-estradiol for gate *P.ER* (Figure 23C), and a 2 hours 37ºC heat shock for *P.HS* (Figure 23E). See Supplemental Figure 10 for details on the building pathways. The genetic structure of the gates is equivalent for all of them: (i) a

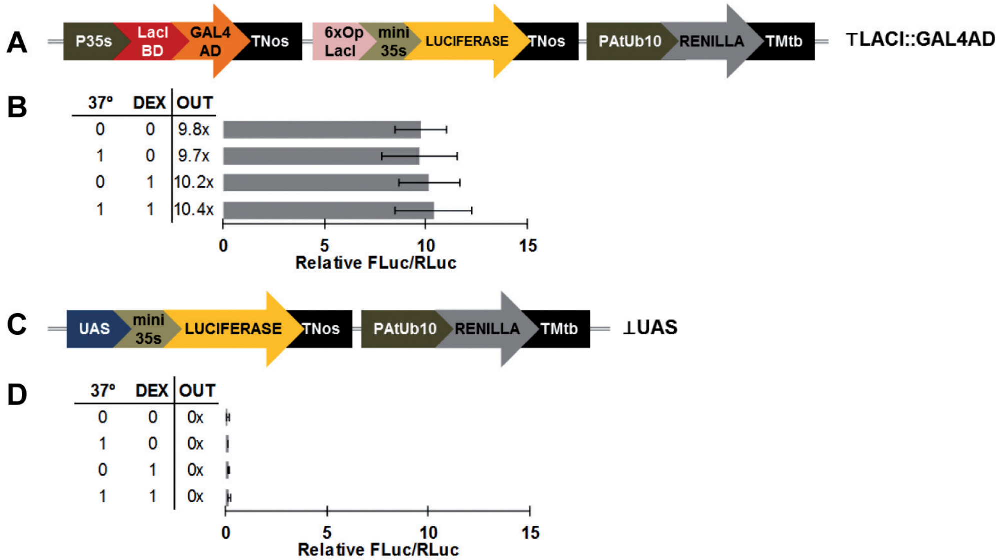


Figure 22. Tautology and contradiction.

(A) Multigenic construct for the tautology ⊤LACI::GAL4AD gate. **(B)** Digital combination of two input inducers results in similar levels of analog gene expression outputs. **(C)** Construct for the contradiction ⊥UAS gate (D) Digital combination of two input inducers result in a minimal background output. The ratios between Firefly (FLuc) and Renilla (RLuc) luciferase activities were normalized using a Nopaline synthase construct. Error bars represent the SD of at least three replicates. 1x corresponds to the reference PNos:Luciferase:TNos construct. ⊤ and ⊥ denote the tautology and contradiction logic gates. P35S is CaMV 35S promote; LacI BD indicates residues 1-330 of the Y17H mutant of lac repressor; Gal4AD indicates transcription activation domain II (residues 768-881) of *Saccharomyces cerevisiae* Gal4; TNos is the Nopaline synthase terminator; 6xOpLacI are six closely spaced lac operators; mini35S is the minimal CaMV 35S promoter; Luciferase and Renilla are the *Photinus pyralis* and *Renilla reniformis* Luciferase genes; PAtUbq10 is the Promoter of the *A. thaliana* Ubiquitin 10 gene; TMtb is the *S.lycopersicum* Metalothionein-like protein terminator, 6xUAS is the upstream activation sequence where Gal4 binds to.

TU encoding the activator module, (ii) a reporter unit consisting of the LacIOp upstream of a minimal CaMV 35S promoter and the firefly luciferase, and (iii) the Renilla Module. The first two propositions were based on the constitutive expression of a fusion protein with three domains, which are the glucocorticoid or estradiol receptor, the LacIBD and the Gal4AD proteins. This fusion protein is primarily located in the cytoplasm forming a heteroduplex in combination with Hsp90 in the absence of ligand. It moves to the nucleus after binding the trigger molecule where it will bind the DNA and activate the transcription [157-158]. The third proposition was driven by the *Arabidopsis* 70B heat shock promoter to result in the expression of the modular protein LacIBD::Gal4AD.

The three gates were responsive to the corresponding inducer in transient expression experiments (10 μM dexamethasone, 5 μM β-estradiol and a 2h 37ºC heat shock treatment).

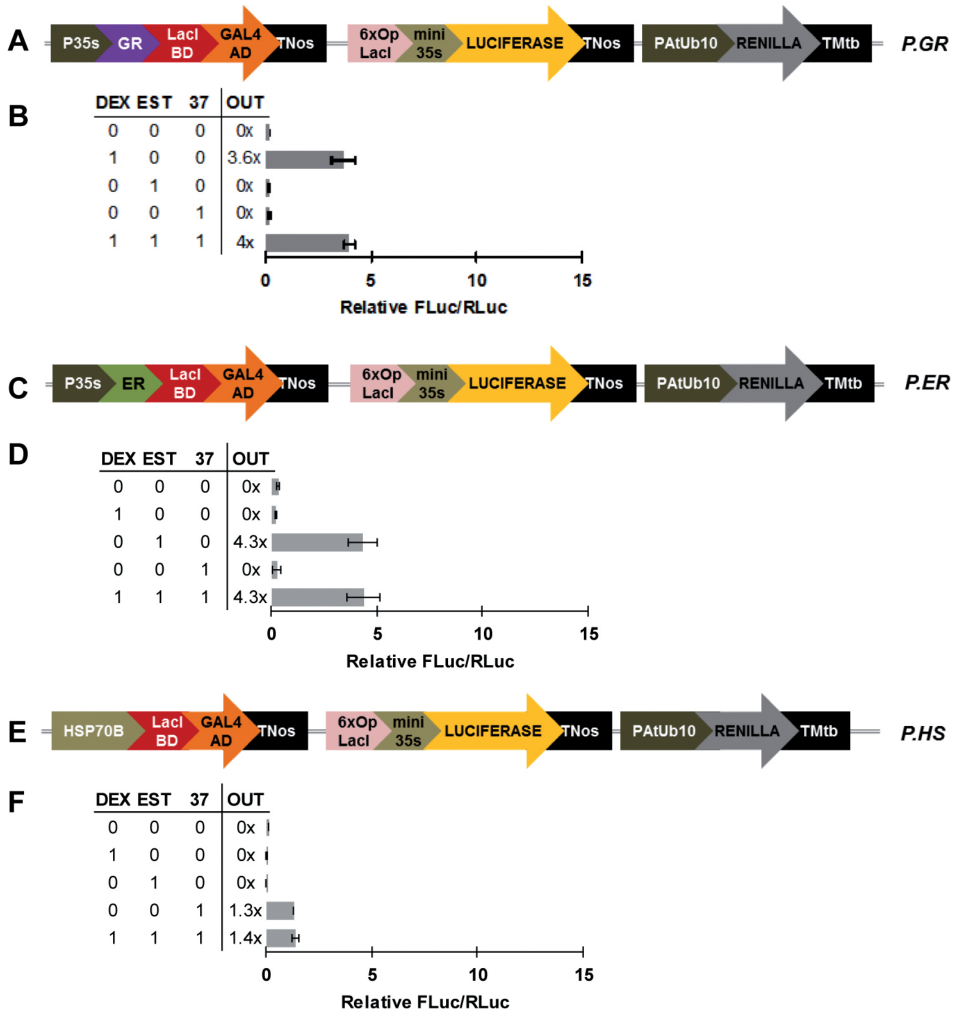


Figure 23. Propositions.

(A) Genetic structure for the *GR* proposition. **(B)** Digital combinations of three input inducers. Luciferase was only expressed when dexamethasone was introduced. **(C)** Genetic structure for the *ER* proposition. **(D)** Digital combinations of three input inducers. The reporter was only expressed in presence of β-estradiol. **(E)** Genetic structure for the *HS* proposition. **(F)** Digital combinations of three input inducers resulted on the expression of luciferase only after a 2 hours treatment at 37ºC. The ratios between Firefly (FLuc) and Renilla (RLuc) luciferase activities were normalized using a Nopaline synthase construct. Error bars represent the SD of at least three replicates. 1x corresponds to the reference PNos:Luciferase:TNos construct. *P* denotes a proposition. P35S is CaMV 35S promoter; GR is the rat glucocorticoid receptor; LacI BD indicates residues 1-330 of the Y17H mutant of lac repressor; Gal4AD indicates transcription activation domain II (residues 768-881) of *Saccharomyces cerevisiae* Gal4; TNos is the Nopaline synthase terminator; 6xOpLacI are six closely spaced lac operators; mini35S is the minimal CaMV 35S promoter; Luciferase and Renilla are the *Photinus pyralis* and *Renilla reniformis* Luciferase genes; PAtUbq10 is the Promoter of the *A. thaliana* Ubiquitin 10 gene; TMtb is the *S.lycopersicum* Metalothionein-like protein terminator; ER is the human estradiol receptor; HSP70B is the *A. thaliana* Heat Shock Promoter 70.

The propositions are theoretically considered as single-entry gates but we could also verify the lack of cross-interference between all three inducers as indicated by the negligible background levels of firefly luciferase detected in all cases (Figure 23B, D and F).

2.1.3. Logical Negation (NOT).

The NOT gate (¬A) is an operation that has a single input which produces a value of false when the input is present and a true value when the input is absent. In genetic terms, it can be configured as a transcriptional repressor down-regulating an otherwise constitutive promoter.

The development of this logic operation required the introduction of transcriptional repressors elements in our system. We built a set of GBparts containing the minimal sequences LDLERLGFA and GNSKTLRLFGVNMEC, corresponding to the A. thaliana transcriptional repressors SRDX **[159]** and BDR **[160]** fused to three different DNA binding domains (LacIBD, LexABD and Gal4BD). GBparts were assembled with the CaMV 35S promoter and the Ubiquitin3 terminator to test the performance of the repressor domains by transient expression in *N. bentamiana* leaves. The DNA binding domains without repressor domain were also set as negative controls (Figure 24A).

A second element in the NOT gate setup is the reporter module. In our approach, the expression of the reporter is driven by operated 35S-based promoters. Following Ikeda and Ohme-Takagi **[160]**, a DNA operator was inserted between the enhancer elements of the CaMV 35S promoter and the minimal CaMV 35S promoter. We tested three operator conformations: (i) 6 tandem copies of the LacI operator (LacIOp), (ii) 4 tandem copies of the LexA operator (LexAOp), and (iii) 4 tandem copies of the UAS sequence (Figure 24B).

The reporter was first combined with the Renilla Module in order to test the transcriptional activity of the engineered versions with the CaMV 35S promoter, as described previously **[91]**. The transcriptional strength of the different versions of the synthetic promoter was found to be inversely proportional to the length of the DNA operator. Values ranged from 1.67±0.29 to 7.98±0.91 relative luminescence units. This represented a significant reduction of the activity of the modified promoters in comparison to the 10.83±0.74 units of the standard CaMV 35S promoter (Figure 24C).

Next, all combinations between the reporter and the corresponding repressor constructs were assayed by transient co-expression in *N. bentamiana* leaves. The strongest repression

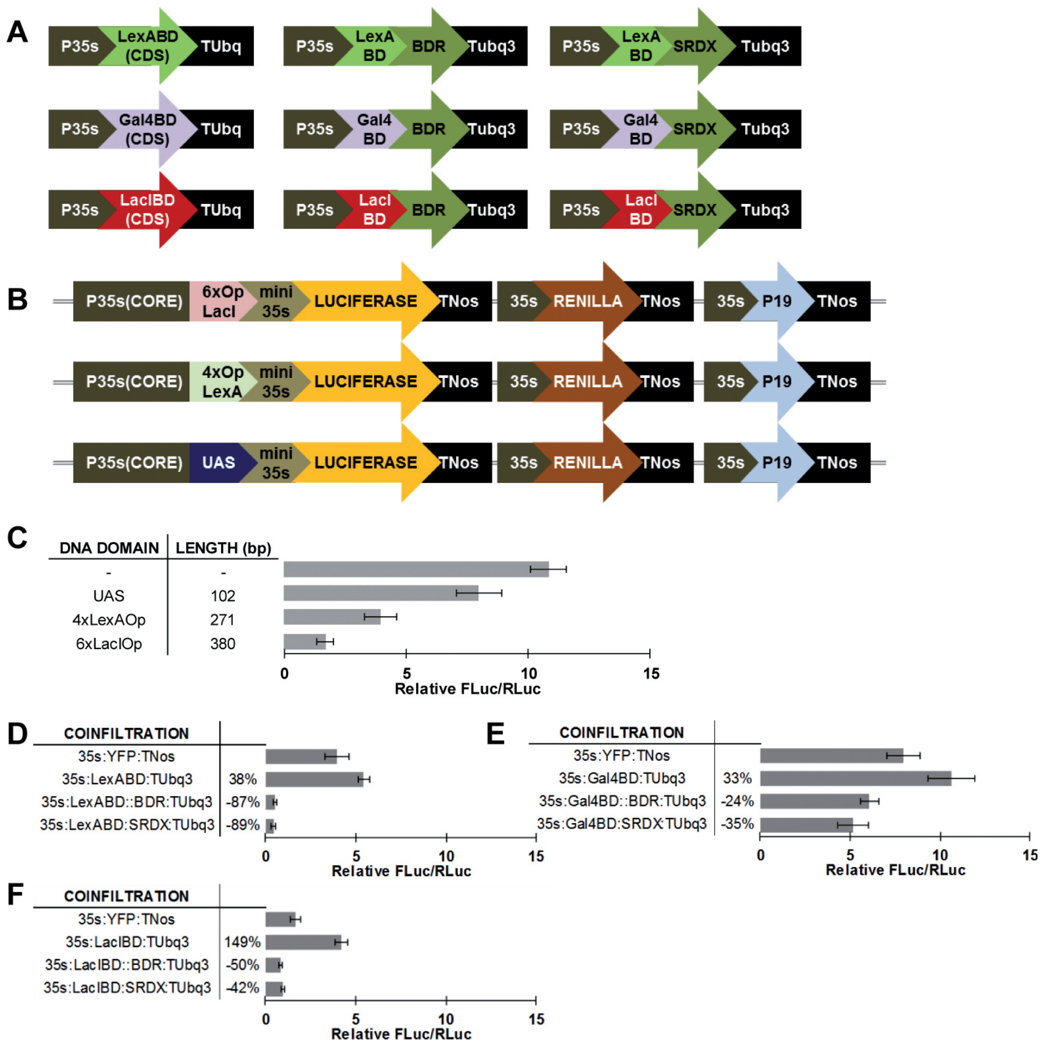


Figure 24. Transcriptional strength of the engineered versions of the CaMV 35S promoter and effect of the synthetic transcriptional repressors.

(A) Generated constructs to test the effect of the transcription repressor domains by agroinfiltration of the different combinations in *N. bentamiana* leaves. **(B)** Constructs incorporating the engineered 35S-promoters. **(C)** Transcriptional strength of the engineered 35S-promoters, compared to the original CaMV 35S promoter. **(D)** Effect of the transcriptional repressor domains on the 35S:OpLexA:mini35S engineered promoter. **(E)** Effect of the transcriptional repressor domains on the 35S:UAS:mini35S engineered promoter. **(F)** Effect of the transcriptional repressor domains on the 35S:OpLacI:mini35S engineered promoter. YFP was coinfiltrated with the reporter constructs as a control to each of the experiments. The ratios between Firefly (FLuc) and Renilla

(RLuc) luciferase activities were normalized using a Nopaline synthase construct. A promoter-less construct was also set as negative control. Error bars represent the SD of at least three replicates. P35S is CaMV 35S promoter; mini35S is the minimal CaMV 35S promoter residues (-60, +1); 35S(CORE) are the enhancer elements of the CaMV 35S promoter; PNos is the Nopaline synthase Promoter; TNos is the is the Nopaline terminator; Luciferase and Renilla are the *Photinus pyralis* and *Renilla reniformis* Luciferase genes; P19 is the Tomato Bushy Stunt Virus silencing suppressor; 6xOpLacI and 4xLexAOp are six and four closely spaced Lac and LexA operators; UAS indicates upstream activation sequence where Gal4 binds to; Gal4 BD indicates residues 1-74 of Gal4; LacI BD indicates residues 1-330 of the Y17H mutant of lac repressor; LexA indicates residues 1-87 of LexA; BDR is the GNSKTLRLFGVNMEC transcriptional repressor domain; SRDX is the LDLERLGFA transcriptional repressor domain; TUbq3 is the *A. thaliana* Ubiquitin3 terminator.

effect was obtained with the LexABD::BDR and LexABD::SRDX fusions, which reduced in an 87% and 89% the transient expression levels of the reporter construct (Figure 24D). In contrast, LacIBD and Gal4BD operated repressors showed weaker repression values (Figure 24E and F). Negative controls were tested in parallel and surprisingly, all three DNA binding proteins showed a certain enhancer effect on the expression of the reporter construct. This transcriptional activation had no statistical significance in the case of LexABD, but LacIBD significantly increased the transcriptional strength of the reporter module 2.5 times.

Once the transient evaluation of the LexA operators worked as expected, the next step was to build a chemically operable NOT gate for stable transformation based on the LexA-operated promoter and the LexABD::BDR and LexABD::SRDX transcriptional repressors. Two dexamethasone-controlled NOT gates were assembled. An additional construct including only the LexABD sequence without a repressor domain was also built, to serve as negative control (named ¬GR.LexABD::BDR, ¬GR.LexABD::SRDX and ∅GR.LexABD respectively. See Figure 25A and Supplemental Figure 11 for a description of the building steps of these constructs).

A. thaliana plants stably transformed with the control ∅GR.LexABD construct and the ¬GR. LexA::BDR NOT gate were analyzed. The levels of luciferase expression were determined in leaves from the transformed plants with or without dexamethasone in a single (24h) time-point experiment. Plants with the control constructs responded similarly as earlier observed in transient expression experiments: the expression of the reporter was slightly enhanced in 6/8 plants after dexamethasone treatment (Figure 25B), indicating a possible activation effect. In contrast 5/6 plants transformed with ¬GR.LexABD::BDR had a reduced luciferase expression after a 24h treatment with dexamethasone (Figure 25C). Two representative lines were selected for a more detailed analysis and monitored for 30h in the presence of dexamethasone. The luciferase expression of ∅GR.LexABD line #3 consistently showed a slight activation in presence of 10μM dexamethasone, which was maintained during the

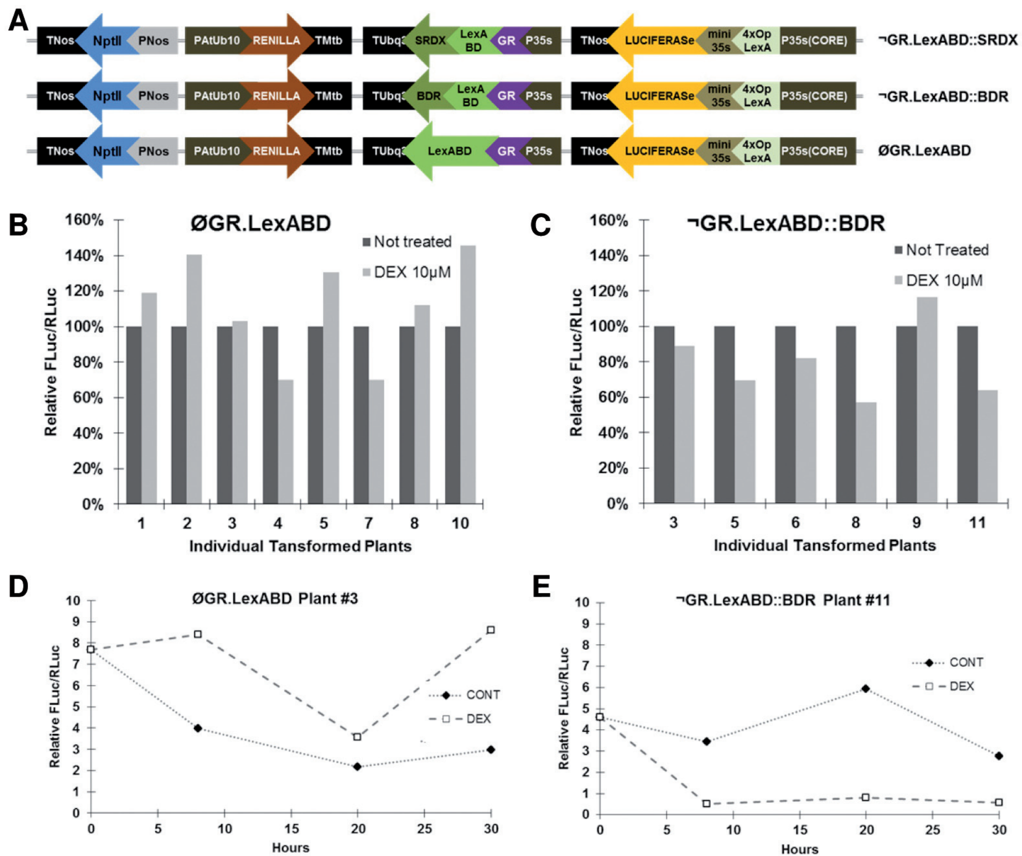


Figure 25. NOT logic gate.

(A) Constructs to for the ¬GR.LexABD::BDR, ¬GR.LexABD::SRDX logic gates, which are based on the use of the LexABD and the LexAOp engineered version of the CaMV 35S promoter. The ØGR.LexABD construct was built as a negative control for the NOT gate. **(B)** Screening of transformed plants with the ØGR.LexABD control construct. **(C)** Screening of transformed plants with the ¬GR.LexABD::BDR construct. **(D)** Performance of the ØGR.LexABD Plant #3 during 30h in presence or absence of dexamethasone. **(E)** Performance of the ¬GR.LexABD Plant #11 during 30h in presence or absence of dexamethasone. ¬ denotes a logical negation. 35S(CORE) are the enhancer elements of the CaMV 35S promoter; mini35S is the minimal CaMV 35S promoter residues (-60, +1); TNos is the is the Nopaline terminator; Luciferase and Renilla are the *Photinus pyralis* and *Renilla reniformis* Luciferase genes; 4xLexAOp are four closely spaced LexAOp; LexA indicates residues 1-87 of LexA; BDR is the GNSKTLRLFGVNMEC transcriptional repressor domain; SRDX is the LDLERLGFA transcriptional repressor domain; TUbq3 is the *A. thaliana* Ubiquitin3 terminator; PAtUbq10 is the Promoter of the *A. thaliana* Ubiquitin 10 gene; TMtb is the *S.lycopersicum* Metalothionein-like protein terminator; GR is the rat glucocorticoid receptor; PNos is the Nopaline synthase promoter; NptII is the Neomycin phosphotransferase II gene, for plant stable transformation.

whole experiment (Figure 25D). Luciferase expression in ¬GR.LexA::BDR line #11 also was monitored for 30h. Dexamethasone treated leaves showed a clear repression to basal levels (less than 0.5 relative luminescence units), in comparison to the non-treated disks, that showed a stronger expression (>3.5 relative luminescence units, Figure 25E).

2.2. Binary operations.

All the two-input Boolean logic gates can be engineered by combining the set of six simple operations described above, provided that an AND gate is also implemented. Two examples of the OR gate have been developed in this work. Furthermore initial experiments to construct an AND gate were also conducted.

2.2.1. Logical Disjunction (OR).

The OR logic gate (A∨B) is an operation that produces a *true* value whenever one or both of its inputs are *true*. The *false* value is only produced when there is no input. The most basic setup comprises two inputs, but additional inputs can be incorporated, resulting in a more complex operation. In genetic terms, a reporter is expressed as a response to two different triggers, regardless of whether one or both are present.

To build the OR logic gate the individual propositions *P.GR*, *P.ER* and *P.HS* described in Supplemental Table 2 were combined, together with the appropriate reporter and the Renilla modules. Two different OR gates were built, configuring a dexamethasone/heat shock inducible OR gate and a dexamethasone/estradiol inducible one (named GR∨HS and GR∨ER respectively. See Figure 26A and Supplemental Figure 12 for the strategy used for building the constructs).

The GR∨HS gate was tested first by transient expression in *N. bentamiana* leaves. This resulted in different levels of reporter gene expression depending on the input(s) applied. The GR∨HS gate produced 3.76±0.176 relative luminescence units when treated with 10 µM dexamethasone and 1.56±0.032 units after a 2h treatment at 37ºC. The additive effect of the two inducers resulted in 4.91±0.48 relative luminescence units (Figure 26B). This construct was later transformed into *A. thaliana* and nine T1 plants were analyzed, resulting in seven of the plants performing as a functional OR gate with different output levels (Figure 26C).

The GRvER gate was also transiently evaluated. Results showed different expression levels when treated with each of the inducers (4.26±0.17 relative luminescence units to 10 μM dexamethasone and 3.27±0.04 to 5 μM β-estradiol) and a similar semi-additive effect to the GRvHS gate was found when the two inducers were applied at the same time (5.57±0.74 relative luminescence units, Figure 26D).

2.2.2. Logical Conjunction (AND).

The AND logic gate (A∧B) is a two-input logical operator that exclusively results *true* if both of the inputs are also *true*. The value *false* is produced in any other circumstances. In genetic terms, it corresponds to a decoupled transcriptional activation in which two interacting proteins are fused to a DNA binding domain and to a transcription activation domain, respectively. As a consequence the reporter is only expressed when both proteins are synthesized and then transported to the nucleus where they physically interact (Figure 27A).

The most crucial part on the design of the AND gate is the choice of the interacting partners since, this will determine the stability and strength of the logic gate. Three pairs of orthologous partners from *E. coli* and mammalian systems were tested. They included (i) two versions of the E. coli DnaA (version I comprising amino acids 1-86 and version II a fusion between the amino acids 1-86 and 135-467) and DnaB proteins **[161]**, and (ii) the PDZ domain from the mammalian protein α1-syntrophin (amino acids 87-170) and its interacting partner, the small peptide VKESLV **[162]**. The strong interacting *A. thaliana* FRUTITFULL (FUL) and SUPRESSOR OF OVEREXPRESSION OF CO (SOC1) **[163-164]** transcriptional factors were also included in the experiment (Figure 27B). The interaction of all these partners was first verified by Bimolecular Fluorescence Complementation assay in *N. bentamiana* leaves (data not shown).

Next, the PDZ domain, FUL and the two versions of DnaA were fused to the N-terminus of the mentioned DNA binding domains and assembled to the CaMV 35S promoter and the TNos terminator into pDGB1α1. The small peptide VKESLV, DnaB and SOC1 were fused to the N-terminus of Gal4AD and assembled to the CaMV 35S promoter and the TNos terminator into pDGB1α2. Both TUs were later combined into pDGB2Ω1 (Figure 27C, only the assembly of the PDZ/VKESLV set of constructs is depicted. Supplemental Figures 13 and 14 for the rest). The reporter construct consisted of three TUs, one containing the DNA operator (UAS, LacIOp, LexAOp) upstream of the minimal CaMV 35S promoter and the firefly luciferase, a second TU which is the Renilla Module and a third one containing the Tomato Bushy Stunt Virus P19 silencing suppressor **[121]** (Figure 27D).

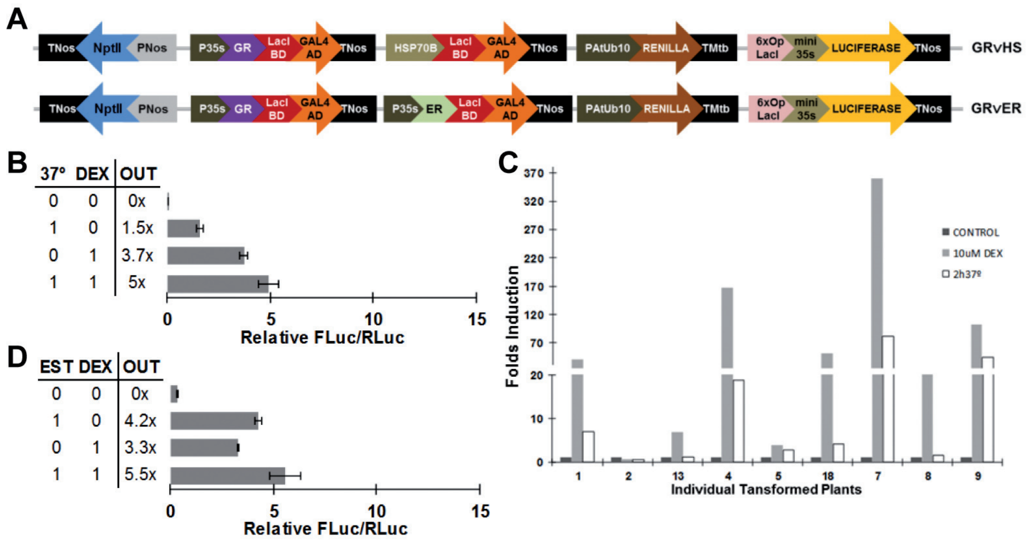
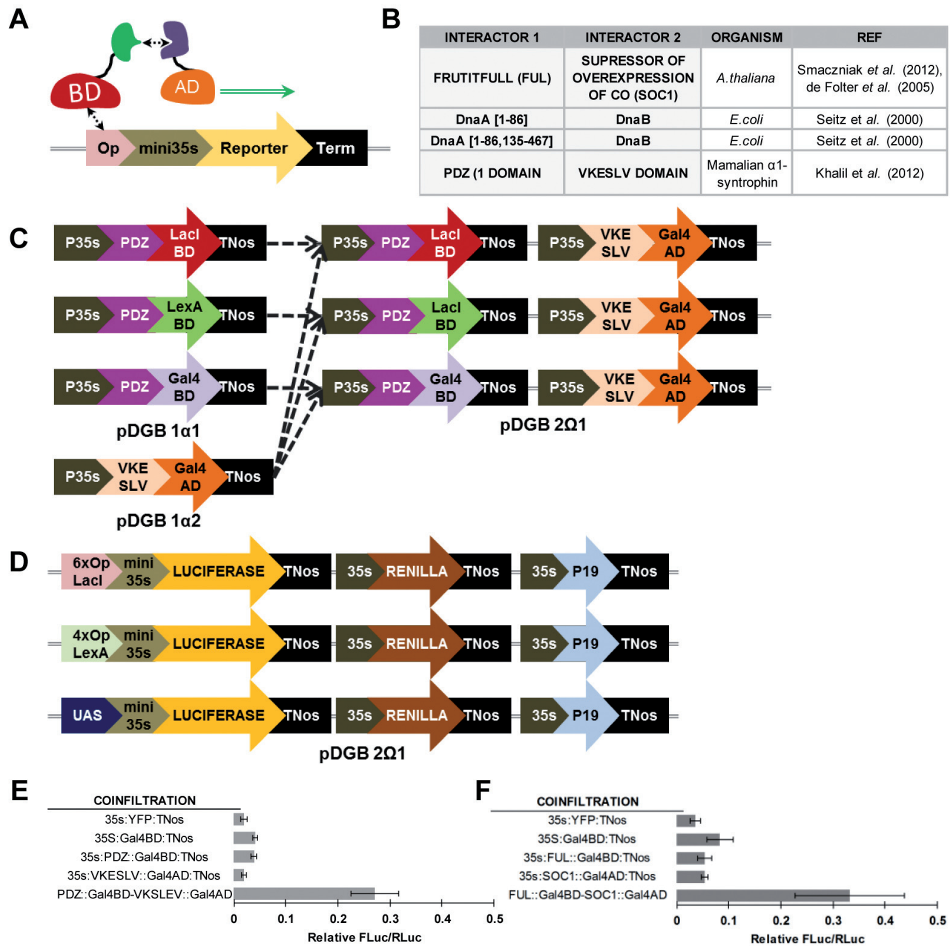


Figure 26. Logical Disjunction .

(A) Genetic structure for the GRvHS and GRvER OR gates. **(B)** Digital combinations of two input inducers (10μM dexamethasone and a 2h treatment at 37ºC) to the GRvHS, tested by transient expression in N. bentamiana leaves. Luciferase was expressed as a response to any of the inducers. **(C)** Screening of 9 lines transformed with the GRvHS operation. **(D)** Digital combinations of two input inducers (10μM dexamethasone and 5μM β-estradiol) to the GRvER, tested by transient expression in N. bentamiana leaves. The ratios between Firefly (FLuc) and Renilla (RLuc) luciferase activities were normalized using a Nopaline Synthase construct (except on the stable plants experiments, which was referred to the control without inducers). Error bars represent the SD of at least three replicates. 1x corresponds to the reference PNos:Luciferase:TNos construct. **v** denotes a logical disjunction. TNos is the Nopaline synthase terminator; NptII is the neomycin phosphotransferase II gene, for plant stable transformation; PNos is the Nopaline synthase promoter; P35S is CaMV 35S promoter; GR is the rat glucocorticoid receptor; ER is the human estradiol receptor; HSP70B is the *A. thaliana* Heat Shock Promoter 70; LacIBD indicates residues 1-330 of the Y17H mutant of lac repressor; Gal4AD indicates transcription activation domain II (residues 768-881) of *Saccharomyces cerevisiae* Gal4; 6xOpLacI are six closely spaced lac operators; mini35S is the minimal CaMV 35S promoter; Luciferase and Renilla are the *Photinus pyralis* and *Renilla reniformis* Luciferase genes; PAtUbq10 is the Promoter of the *A. thaliana* Ubiquitin 10 gene; TMtb is the *S.lycopersicum* Metalothionein-like protein terminator.

Of all the combinations transiently tested in *N. bentamiana* leaves, only the PDZ::Gal4BD-VKESLV::Gal4AD, performed as an AND operator (see Figure 27E and Supplemental Figures 13 and 14 for the remaining combinations). In the PDZ::Gal4BD-VKESLV::Gal4AD combination, the activation of the reporter unit was weak (0.3±0.04 relative luminescence units) but at least 15 times stronger than the background levels. In addition, this was the only combination whose individual TUs had no auto-activation effect on the reporter construct. Similar results were also obtained with the non-orthogonal interactors FUL::Gal4BD-SOC1::Ga4AD (Figure 27F).



INTERACTOR 1	INTERACTOR 2	ORGANISM	REF
FRUTITFULL (FUL)	SUPRESSOR OF OVEREXPRESSION OF CO (SOC1)	*A.thaliana*	Smaczniak *et al.* (2012), de Folter *et al.* (2005)
DnaA [1-86]	DnaB	*E.coli*	Seitz et *al.* (2000)
DnaA [1-86,135-467]	DnaB	*E.coli*	Seitz et *al.* (2000)
PDZ (1 DOMAIN	VKESLV DOMAIN	Mamalian α1-syntrophin	Khalil et *al.* (2012)

Figure 27. Logical Conjunction.

(A) Model for the AND logic gate. It is based on the reconstitution of an artificial functional transcription factor when two proteins or polypeptides interact. **(B)** List of interactors tested in this chapter. **(C)** GoldenBraid pathway for the construct of the PDZ-VKESLV interaction tests. **(D)** Reporter constructs, including three different DNA operators (6xLacI, 4xLexA and UAS). **(E)** PDZ::Gal4BD-VKESLV::Gal4AD co-infiltration experiments in *N. bentamiana* leaves. Two constructs, which are 35S:YFP:TNos and 35S:Gal4BD:TNos were included as negative controls of the experiment. **(F)** FUL::Gal4BD-SOC1::Gal4AD co-infiltration experiments in *N. bentamiana* leaves. The ratios between Firefly (FLuc) and Renilla (RLuc) luciferase activities were normalized using a Nopaline synthase construct (except on the stable plants experiments, which was referred to the control without inducers). A promoter-less construct was also set as negative control. Error bars represent the SD of at least three replicates.

LacI BD indicates residues 1-330 of the Y17H mutant of lac repressor; LexA indicates residues 1-87 of LexA; Gal4 BD indicates residues 1-74 of Gal4; Gal4AD indicates transcription activation domain II (residues 768-881) of *Saccharomyces cerevisiae* Gal4; 6xOpLacI and 4xLexAOp are six and four closely spaced Lac and LexA operators; UAS indicates upstream activation sequence where Gal4 binds to; mini35S is the minimal CaMV 35S promoter; Luciferase and Renilla are the *Photinus pyralis* and *Renilla reniformis* Luciferase genes; TNos is the Nopaline synthase terminator; 35S is the CaMV 35S promoter; P19 is the Tomato Bushy Stunt Virus silencing suppressor.

3. Discussion

Logic gates are important tools for the future development of Plant Synthetic Biology as they are required for fine-control of genetic expression and to build genetic circuits. Recently, June Medford's group produced the first examples of synthetic programmed circuits in Plant Biotechnology **[165-166]**. They took advantage of the modular fashion of the histidine kinase signaling systems conserved across species to build a synthetic circuit responding to the presence of TNT in the soil that resulted in a degreening phenotype **[8]**. Although the design of this circuit was not based on any logic operation, it represents an interesting case that could integrate plant-adapted logic gates to define the flux of information in the circuit.

In this work, we have taken the first steps toward the development of Boolean logic gates adapted to Plant Biotechnology. We have generated a set of standardized modular building blocks following the GB2.0 *grammar* to implement the described logic gates. GB2.0 facilitated the assembly of the genetic constructs and the combination of pre-tested operations to create double-input gates. Moreover, the assembly of the components of the gates plus the reporter in a single T-DNA undoubtedly contributed to the reproducibility and accuracy of the experimental setup.

The construct of the three simple one-input operations (tautology, contradiction and proposition) was trivial as most of the GBparts used had been previously characterized **[91, 167]**. We have built the following operations: ⊤LACI::GAL4AD, ⊥UAS and the propositions *P.ER*, *P.GR* and *P.HS*. All of them were tested in different situations to confirm the lack of interference, a requirement for the assembly of double-input logic gates.

The NOT gate has an outstanding interest for plant biotechnologists. There are many examples of inducible systems in plants (extensively reviewed by Moore *et al.* **[167]**) and the constitutive repression of a housekeeping gene has been previously demonstrated **[168-170]**, but to our knowledge, the inducible repression of housekeeping gene was not described before. In order to develop an inducible transcriptional repressor, some preliminary experiments were conducted. First, the transcriptional strength of the operated 35S-based promoters were verified. Next, all synthetic transcriptional repressors were tested, but only LexABD::SRDX and LexABD::BDR were able to repress the activity of the operated 35S promoter to basal levels. We included a set of constructs to verify the effect of the DNA binding domains in the absence of a repression domain. Remarkably, LacIBD alone activated the LacI-operated 35S promoter. A possible explanation to this is that the canonical LacIBD interacts with the transcription machinery in plants in a different way that it interacts in bacteria. Stably transformed *A. thaliana* plants were produced for the dexamethasone inducible versions

of the NOT gates. Results with T1 plants of the ¬GR.LexABD::BDR gate showed a strong inducible repression of the transcriptional activity of the LexA-operated 35S promoter. More experiments are needed to explore the remaining versions of the operations and to confirm the results in T2 and T3 plants.

Double-input gates need the combination of two simpler operations. Therefore several TUs had to be assembled in the same composite part. Two examples of the OR gate were developed by combining previous described propositions. GR**v**HS and GR**v**ER resulted in two functional OR gates that responded as it was expected from the data obtained on the evaluation of the individual propositions. This indicates that the combination of different operations produce the predicted results. Stable transformation experiments with the GR**v**HS and GR**v**ER constructs are also initiated.

Various approaches have been employed on the design of AND gates in different organisms, such as the use of RNAs, RNA binding proteins, synthetic promoters, transcription factors and co-activating proteins, among others **[105, 171-173]**. We have followed an interacting-protein design similar to the yeast-two-hybrid, which is based on the reconstitution of an artificial functional transcription factor when two proteins or polypeptides interact. We performed a full set of experiments in order to find a suitable pair of orthogonal interacting molecules that would follow the model for the AND operation (Figure 27A). The experiments performed so far indicate that the pair PDZ::Gal4BD-VKESLV::Gal4AD is a good candidate for the construction of a AND gate in plants. The next step will consist on the construction of an inducible version of the gate and the stable transformation of *A. thaliana*.

In this chapter, we have described the tautology and contradiction operations, three different propositions and two alternative ORs. The results obtained from the NOT gate indicated that the inducible operation is functional and we got interesting initial results for the AND gate. In summary, we have given the first steps towards the development of a full set of orthogonal and modular logic gates adapted to Plant Biotechnology.

4. Materials and Methods.

4.1. Bacterial strains and growth conditions.

E. coli DH5α and TOP10 were used for cloning. *A. tumefaciens* GV3101 was used for transient expression and stable transformation experiments. Both strains were grown in LB medium under agitation (200 rpm) at 37°C and 28°C, respectively. Ampicillin (100 μg mL^{-1}), kanamycin

(50 µg mL^{-1}), and spectinomycin (100 µg mL^{-1}) were used for *E. coli* selection. 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside acid (X-gal, 40 µg mL^{-1}) and isopropylthio-β-galactoside (IPTG, 0.5 mM) were used on LB agar plates for the white/blue selection of clones. Rifampicin (50 µg mL^{-1}), tetracycline (12.5 µg mL^{-1}) and gentamicin (30 µg mL^{-1}) were also used for selection in *A tumefaciens*.

4.2. Restriction-Ligation assembly and Domestication reactions.

Assembly and domestication reactions were set as described before **[90-91]** using *Bsa*I, *BsmB*I and *BtgZ*I as restriction enzymes (New England Biolabs) and T4 Ligase (Promega). One microliter of the reaction was transformed into *E. coli* DH5α or Top10 electrocompetent cells, and positive clones were selected in solid medium including X-gal and IPTG. Plasmid DNA was extracted using the E.Z.N.A. Plasmid Mini Kit I (Omega Bio-Tek) and correct assemblies were confirmed by restriction analysis and/or sequencing.

4.3. *N. bentamiana* transient transformation.

Overnight-grown *A. tumefaciens* cultures were pelleted and resuspended in agroinfiltration medium (10 mM MES, pH 5.6, 10 mM MgCl2, and 200 µM acetosyringone) to an optical density at 600 nm of 0.5 as described previously **[132]**. Transient transformation *in cis* (all the assayed TUs were inserted on a single binary vector) or *in trans* (TUs were transformed into different *Agrobacterium*, therefore more than one has to be co-transformed at the same time). Cultures were mixed 1:1 for in trans experiments. 4 to 5-week-old *N. bentamiana* plants were infiltrated using a needle-free syringe and cultivated at 24°C day/20°C night in a 16-h-light/8-h-dark cycle. Leaves were harvested 3 to 5 days post infiltration (depending on the purpose of the experiment) and assayed for transgene/reporter expression.

4.4. *A. thaliana* stable transformation and growth conditions.

A. thaliana Col-0 plants were transformed by the floral dip method **[147]**. Seeds were sterilized and sown on plates of Murashige and Skoog medium with 0.8% (w/v) agar and 1% (w/v) sucrose (growing conditions: 24°C day/20°C night in a 16-h-light/8-h-dark cycle).

Transgenic lines were selected in 100 μg ml kanamycin. Plants were grown in the glass house in a 24°C day/20°C night in a 16-h-light/8-h-dark cycle.

4.5. Glucocorticoid and estradiol receptor Induction.

One cm^2 discs were punched off leaves and placed in a 350-μL solution containing 5 to 20 μM dexamethasone (Sigma-Aldrich) or 1 to 10 μM β-estradiol (Sigma-Aldrich) in 0.02% Tween 80. Discs were incubated in a growth chamber (6-h-light/8-h-dark, 24°C) and harvested after the indicated time and expression of the reporter gene determined.

4.6. Heat-shock treatments.

One cm2 leaf discs were placed in 350 μl of water at 37°C for 2 h. Samples were collected at 3 h after treatment and assayed for transgene reporter expression.

4.7. Luciferase/Renilla expression assays.

Firefly luciferase and Renilla luciferase were assayed from 100-mg leaf extracts following the Dual-Glo Luciferase Assay System (Promega) standard protocol and were quantified with a GloMax 96 Microplate Luminometer (Promega). Results were analyzed as described previously **[91]**.

/GD

GENERAL
DISCUSSION

General Discussion

In this thesis, a new technological framework for multigene engineering in Plant Synthetic Biology has been established, that includes a powerful and modular DNA assembly tool and a collection of standardized parts and modules. The latest addition to the collection consisted in a number of primary logic gates for the future implementation of digitalized gene networks. Detailed discussions around these genetic tools and parts have already been presented in each chapter. Furthermore, the discussion in Chapter 2 complements and expands that of the first version of GoldenBraid in Chapter 1. For this reason, this general discussion chapter will be directed to critically examine the GB2.0 Framework, and to delve into the future perspectives for this technology.

1. The GoldenBraid Framework: an insight look.

The GoldenBraid Framework has a number of specific features that account for its success but also for its limitations. Some decisions were taken during the design of the platform, as the definition of the GBgrammar, the selection and characterization of basic GBparts and Frequently Used Structures, the requirements for domestication or the creation of the software tools. These decisions have shaped the platform to its current form and need to be discussed and compared with similar arbitrary decisions taken by the developers of other DNA assembly technologies.

1.1. A binary loop fed by a multipartite standard.

A standardized assembly requires first the definition of a set of basic rules that governs how the DNA parts have to be created and assembled. Usually, DNA assembly standards categorize their basic parts using flanking prefix and a suffix sequences. The simplest

assembly standard, defined by BioBricks ten years ago, has the property of idempotency **[37]**. This means that all parts have identical prefix and suffix, and that the assembly of two DNA parts produces a new composite part which is flanked by the same prefix and suffix that their original constituents. The toll for idempotency is that only binary assemblies are allowed, thus sacrificing speed for simplicity.

The minimalist design of the GoldenBraid loop is guided by the ideal of idempotency, ending up with four assembly rules instead of a single one as in BioBricks. This is far from impotency, yet the chances that two randomly selected TUs from the collection are compatible one another are 25%, and in the worst scenario they can be made compatible with a single additional reaction (using the twister plasmids). The simplicity of the GB design translates in the small size of the assembly toolkit. The basic GB toolbox comprises only 8 destination plasmids plus the universal domesticator plasmid (as defined in Chapter 2). MoClo requires at least 28 plasmids whereas others like pSAT **[73]** need 19 plasmids to permit the assembly of TUs in both directions. In contrast, simpler idempotent approaches require only three plasmids (i.e. two for the construction of TUs in both orientations and a third functioning as destination vector **[62]**).

GB2.0 completely deviates from idempotency in the multipartite assembly strategy used for the construction of TUs, and it does so deliberately to accommodate a standard *grammar*. In concert with MoClo developers, we defined a positional notation system to pre-define the sequence of the four nucleotide overhangs for each connecting position within the transcriptional unit. This *grammar*, described in Chapter 2, defines the minimum elements that compose a functional TU and aims at facilitating the exchange of DNA parts among laboratories. This permits a common definition of the DNA parts but still does not guarantee the full exchangeability of the parts between users, which is partially limited by the different bacterial resistance genes used on the different assembly levels and by the restriction enzymes used for the restriction-ligation reactions (see Table 2). MoClo Level 0 parts can be used for GB Level α multipartite assemblies, as both systems share the same *Bsa*I sites and compatible antibiotic selection at this level. Conversely, the incorporation of GBparts into MoClo Level 1 is not possible because both share ampicillin as bacterial resistance so counterselection of correctly assembled clones is not possible. This could be solved with (i) a change in the resistance selection markers used in any of the levels, or (ii) the introduction of a highly efficient digestion step after the restriction-ligation reaction, to guarantee the cleavage of all the unassembled GBparts.

	MoClo		GoldenBraid		
	Selection	Type IIS Enzyme	Selection	Type IIS Enzyme	
Level 0 Parts	Spectinomycin	*Bpi*I	Ampicillin	*BsmB*I	GBparts
Level 1	Ampicillin	*Bsa*I	Kanamycin	*Bsa*I	Level α
Level 2	Kanamycin	*Bpi*I *Bpi*I+*Bsa*I* *Bpi*I+*BsmB*I*	Spectinomycin	*BsmB*I *BsmB*I+*BtgZ*I**	Level Ω

Table 2. **Bacterial Resistances and Type IIS Enzymes used in MoClo and GoldenBraid for the different restriction-ligation reactions.** * For the incorporation of TUs after intermediate levels. ** For the multipartite assembly of the GBparts into Ω-level vectors.

1.2. The generation, selection and characterization of the GBparts.

A standardized procedure for the domestication of the GBparts was described in Chapter 2 that substitutes the previous procedure described in Chapter 1. The GB2.0 domestication process involves the removal of *Bsa*I, *BsmB*I and *BtgZ*I internal sites, which are the enzymes further used for multipartite and binary assemblies. In the new strategy, prefix and suffix are incorporated into the primers used for amplification, instead of being located in the domestication vector as previously conceived by Weber *et al.* **[89]**. As a consequence, all categories in the *grammar* can be generated using a single domesticator vector, the universal domesticator plasmid (pUPD). GBparts can then be released by *Bsa*I or *BtgZ*I and used for multipartite assembly. It should be noticed that, once a part is domesticated, no additional PCR amplifications are required, this representing an important advantage compared to other powerful overlap-assisted methods that imply serial PCR amplification and sequence verification **[50, 170]**.

At this stage, all DNA parts deposited in the publically available GB2.0 collection have been functionally-tested. The collection of promoters and terminators is of special interest for the creation of complex circuits and pathways. The use of different promoters and/or terminators in multigenic constructs may avoid spontaneous recombination and/or the unintended transcriptional gene silencing through promoter methylation and inactivation **[174]**. The experimental transcriptional activity (ETA) of all promoters and terminators in the collection was determined by transient expression in *N. bentamiana* leaves using the luciferase/renilla

reporter assay, in a similar approach to other characterization attempts made in bacteria **[169, 171, 175]**. This is an initial step towards promoter characterization, as it only includes a snapshot of the relative strength of the different promoters. A more comprehensive description including spatio-temporal activity profiles will be required for a more accurate modeling and engineering in plants.

1.3. Size and speed considerations.

The topology of GoldenBraid was designed to allow for the indefinite exponential growth of the size and complexity of the construct by making binary assemblies alternating between α and Ω levels. The obvious limitation is imposed by the insert size that can be incorporated into the destination vector. Two set of destination plasmids were created, one based on the pGreenII backbone **[131]** and another based on the pCAMBIA backbone **[143]**. We assembled up to 5 TUs (14.3 Kb) in one pGreenII-based pDGB and up to 7 TUs (23.7 Kb) into the pCAMBIA. In our hands, big constructs were more stable in pCAMBIA when transferred to *A. tumefaciens*. Also, the transformation efficiency in tomato of the latter vector was higher. No obvious difference was found on the transformation of *A. thaliana* though. In order to efficiently transfer large assemblies to the plants, it would be of interest the GB-adaptation of a BiBAC or TAC vector. A good candidate is the pHUGE vector **[176]**, which is a pYLTAC7 derived vector **[177]**. It was recently adapted to the Gateway technology and used to transfer an 8-gene 74kb T-DNA into tobacco, tomato, strawberry and poplar.

The GoldenBraid binary technology simplified the design of the cloning system but imposed a limitation in speed. A 5-gene pathway can be engineered using GoldenBraid in 11 working days (see Figure 28). If compared to other assembly systems that permits the incorporation of just one TU at a time (MISSA **[68]**, pSAT **[73]** or MRG **[62]**), GB is a faster option. On the other hand, using MoClo **[89]**, which permits the multipartite assembly of up to 6 TUs at level two, the generation of a 5-gene pathway will be completed in only 5 working days. This speed limitation in GoldenBraid can be partially solved with the fast-track assembly step described in Chapter 1, which would reduce in 3 days the total time to accomplish the construct. The speed disadvantage of GB against MoClo is partially compensated by the simplicity of the system.

1.4. The software tools.

In Chapter 2 we described the implementation of a set of software tools, which are available in our website www.gbcloning.org. These tools serve to (i) handle the GBCollection, (ii) assist

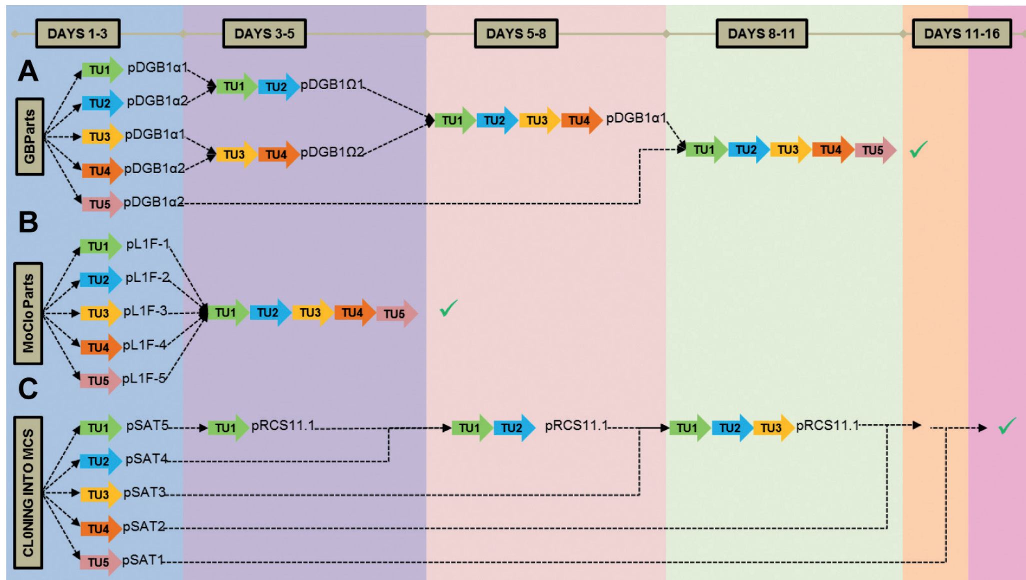


Figure 28. Pathways the generation of a 5-gene construct using three different assembly systems.

It includes (A) MoClo (B) GoldenBraid and (c) pSAT. The minimum number of required working days to achieve the constructs through the different pathways is indicated on the diagram.

the user through the domestication of their sequences of interest, and to (iii) facilitate the *in silico* multipartite and binary assemblies, including the users-developed GBparts. Other platforms such as SEVA **[178]** or the Ghent Collection of Gateway Vectors **[52]** include a well-organized online database but do not assist users on the combination of the elements nor provide detailed lab protocols. In our opinion, the GB2.0 tools are one of the strengths of the technology and set a qualitative difference with other assembly standards.

The ultimate goal of dedicated assembly software is the automated design of optimized assemblies, something not yet implemented in the GB2.0 software package, where the assembly design is defined by the user. The GB2.0 Framework will benefit from the incorporation of computational algorithms that minimize the time and cost required for making the genetic assemblies thus guiding the user through the most effective assembly paths. This has already been implemented in several software web-based tools **[15, 179-180]**. Other platforms also include algorithms that help predicting the behavior of an engineered biological network **[129]**. This last characteristic is still far from the scope of GoldenBraid, but surely the development of increasingly sophisticated gene combinations as logic gates and gene circuits will, in the future, require from the implementation of predictive tools.

2. GoldenBraid perspectives: An outlook.

Many ingenious solutions for the *in vitro* and *in vivo* assembly of multigene constructs have been proposed in the last years, and were described at the introduction of this thesis. However, a simple analysis of the literature shows that few methods have been widely adopted (see Table 3). On the contrary, many of the solutions have been adopted only by developers or have run into a stone wall. There are probably different reasons for that. One of them is that many of these methods are hard to be understood and even more complex to implement since they usually do not include a complete ready-to-use toolbox. GoldenBraid aims to be a good candidate for an assembly standard, specially dedicated to Plant Biotechnology. For this reason, during this thesis the objective was not restricted to the development of the DNA assembly system itself but also included the implementation of a characterized collection of parts. To this date, the GBCollection kept *in house* at the IBMCP is composed by more than 350 characterized GBparts and over 500 assembled devices. This is to our knowledge the most extensive collection of standardized DNA parts publically available for Plant Biotechnology. We aim to keep the collection growing with the contribution of the actual GB users in other participating labs.

2.1. Compatibility between different multigene assembly methods.

The success of Plant Synthetic Biology will be surely require the combination of different multigene assembly strategies. Some effort should be dedicated to the compatibility between the different methods. As a first example of this, MoClo and GB2.0 converged into a common definition of the assembly standard. GB2.0 has also incorporated a Gateway-GB compatibility module, facilitating the transit from one technology to the other.

Each of the assembly systems described in the introduction section have got interesting features that could be added to the design of the others (see Table 3 for a list of the advantages and disadvantages of each assembly systems). An example of the possible interconnections would be the creation of a GB-MISSA/RMDAP adapter tool. GoldenBraid would benefit from the efficiency of the different *in vivo* recombination methods for the creation of the large constructs **[68, 70]**. Conversely MISSA would benefit from the introduction of a multipartite entry point for the assembly of basic TUs. Another interesting feature that GoldenBraid could implement is the ability to remove assembled units from a multigenic structure using ZFNs or HENs as described by Zeevi et al. **[73]**. This would require a global redefinition of the pDGB set of vectors and would possibly be at the cost of the simplicity of the basic toolbox.

METHOD	REF	CLONING SYSTEM	MAX NUMBER OF STACKED GENES	MAIN ADVANTAGE	MAIN DISADVANTAGE	ADDITIONAL REFERENCES
MultiRound Gateway	**[60, 62]**	Gateway	9	Many labs are used to Gateway	Subsequent reactions can be difficult to understand	**[61]**
pGATE Vectors	**[63]**	Gateway	3	Easy to understand for not experienced users	In principle permits only the assembly of up to 3 TUs	**[181]**
ISSI	**[65]**	Cre Recombinase + ΦC31 integrase	2*	Up to 150kb have been assembled	ΦC31 homologous recombination is not a widely-used technique	**[66-67]**
MISSA	**[68]**	Cre recombinase. Phage λ recombination. Conjugational transfer.	9+4 MARs	Simplifies the *in vitro* manipulations to a simple mixing of bacterial strains	Labs not used to *in vivo* work may find this method too complicated	**[69]**
RMDAP	**[70]**	Gateway™. Cre recombinase. Recombineering	5	Toolbox with promoters, terminators and special vectors	Red-Recombination requires the use of special *E. coli* strains.	NR
pSAT	**[75-76]**	Homing Endonucleases	**3**	Permits gene replacement in already built constructs	Small number of HENs commercially available	NR
pSAT + ZFNs	**[73]**	Homing Endonucleases. Zing Finger Nucleases.	9	Modular. Includes a ready-to-use set of vectors and DNA parts	ZFNs overcome the problematic of the few available HENs, but have to be expressed and purified	NR
pZK3BCSPS	**[80]**	Homing Endonucleases. Type II Restriction Enzymes	9	Two TUs can be assembled in one step using *SfiI*	Limited number if TUs to be incorporated as there are only 4 HEN entry points	NR
Type II	**[28, 40]**	Type II Restriction Enzymes	5	Cheap and simple technology	Unique restriction enzymes dramatically reduced with the growth of the constructs.	**[38]**
MoClo	**[89, 92]**	Golden Gate	17	Assembly up to 7 TUs in a one-pot-one-step reaction	Complex full-set of basic vectors	NR
GoldenBraid	**[90-91]**	Golden Gate	7	Simplified standardized assembly rules. Software assisted.	The assembly of the TUs is made by binary assemblies what may slow the process	**[182]**

Table 3 Multigene Assembly methods for Plant Biotechnology.

NR = Not Reported. * Not exploited for multigene transformation in plants.

2.2. The creation of a community of users.

To this date, we have delivered the GB Starter Kit to more than 35 labs around the world. This included important research centers in Europe and USA and also two non-profit initiatives such as the *do-it-yourself* (DIY) *BioCurious* lab in California and the KickStarter funded *Glowing Plant Project* **[130]**. This warm welcome demonstrates an interest of other labs for a technology that facilitates multigene engineering but so far only one publication that came out from our lab has used GoldenBraid for genetic engineering. Juárez *et al.* **[182]** reported the combinatorial assembly and transient expression in *N. benthamiana* of 16 combinatorial versions of a human secretory immunoglobulin A against the VP8* rotavirus. Additional efforts should be directed towards the institution of a centralized repository of standardized and characterized building blocks and assembled units, promoting the interaction between researchers, similarly as has been done with the BioBricks Catalogue, the widely used Ghent Collection of Gateway Vectors **[52]** or other collections that had been deposited in non-profit plasmid repositories like Addgene.

3. The development of Logic Gates adapted to Plant Biotechnology.

The generation of operations and circuits is important in Synthetic Biology to the control of gene expression for the most diverse applications (e.g. producing a therapeutic output or generating a sensor device). A first approach to confront this complex programming is the development of simple and reliable logic gates. In chapter 3 we have described a set of modular logic gates based on the control of the transcription of a group of engineered promoters. This included the tautology and contradiction operations, three different propositions, two alternative ORs and the first steps towards the development of a functional AND gate.

Single-input operations are the base of further developments as the double-input gates. Outputs from two propositions can be wired to engineer a double-input gate, similarly as we did in Chapter 3, where two independent propositions were combined to perform an OR gate. The modularity and reusability of the GoldenBraid multigene constructs will simplify the development of other gates such as the NOR gate, that could be now engineered by connecting two NOT gates (see Figure 29). The connections described above were very easy to implement since the same elements that were used for the development of the operations served as connectors themselves. Further developments may include more complex wires such as those based on hormones or metabolites, permitting a whole-organism connectivity

as was engineered in bacteria where different chemical compounds were used as wires for establishing cross-talk between bacterial populations **[183-184]**.

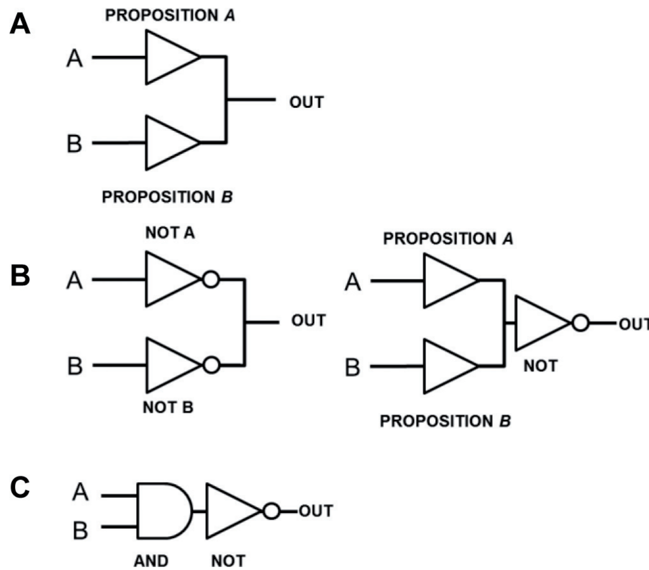


Figure 29. Wiring options for the development of Boolean Logic gates.

(A) Engineering of an OR gate by combining two independent propositions. (B) Two alternative options for the development of a NOR gate, including the combination of two independent NOTs or the connection of an OR and a NOT gate. (C) NAND gate, by combination of an AND and a NOT gate.

Perhaps the most important goal achieved in chapter 3 is the implementation of the NOT gate. This is, to our knowledge, the first example in which the constitutive CaMV 35S promoter is turned into a regulated promoter with the introduction of a repressible operator. Two transcriptional repressor domains SRDX **[159]** and BDR **[160]** were fused to the LexABD to implement a dexamethasone inducible NOT gate. Experiments in T1 generation of *A. thaliana* plants showed a strong inducible repression of the activity of the LexA-operated CaMV 35S promoter. Further refinements on the NOT operations may include the BTB/POZ domain of the *A. thaliana* disease resistance protein NPR1, that has been described to interact with the repression domain of TGA2 to negate its function **[174]**. This domain could also be interesting as a part of two-input Boolean gate such as NOR or A NIMPLY B. Another element that should be taken into account, not only for the NOT gate but also for other operations, is the inclusion of Transcriptional Activator-Like Effectors (TALEs) instead of the LexA, LacI or Gal4 binding domains. TALEs can be customized to bind any sequence of interest with

different strengths and this can be an alternative to the DNA operators used in Chapter 3 to modify the CaMV 35S promoter **[185]**. Similarly, the Cas9 protein of the CRISP-CAS9 system can be retargeted to bind and cleave any DNA sequence of interest and could be therefore considered as an alternative to the used DNA binding domains **[186-187]**.

The AND gate is still under development. In this thesis, different orthogonal interactors were tested and at least a promising pair of interactors was found to produce the expected AND outcome. A general observation in all experiments was that, although clearly rising above background levels (> 15 times in best interaction tested) the response of the reporter unit in the ON state was very weak when compared with strong promoters as CaMV 35S. This difference between the ON and OFF state might be sufficient to trigger a response when coupled in a circuit. Other examples of AND gates reported a 20-fold **[188]**, 50-fold **[111]** and 180-fold **[107]** change between the non-induced and induced state; therefore our result is promising. The inducible version of the operation is currently being developed. Additional interacting molecules should also be tested, including examples that have been used to develop AND gates in other organisms such as the HrpR and HrpS proteins from *Pseudomonas syringae* described in Wang *et al.* **[105]**. Moreover, the final steps will require connecting the interacting partners with inducible modules (GB individual propositions) and to bring the whole network to stable transformed plants where the logical operations can be tested. The achievement of a functional AND gate will be the last piece of this set of operations that would enable the engineering of the rest of logic gates.

We believe that these first examples of simple and reliable logic gates will be used as elementary control elements in more sophisticated networks in the near future.

/CN

CONCLUSIONS

Conclusions

C1. A new DNA assembly system named GoldenBraid was designed for Plant Biotechnology, which allows the indefinite growth of reusable genetic modules made of standard DNA parts.

C2. The GoldenBraid system adopts the principles of modularity and standardization of Synthetic Biology, and proposes a modular cloning schema with positional notation that resembles the grammar of natural languages.

C3. A collection of functionally characterized DNA parts conforming to the GBgrammar was established and made available in a public repository. This collection includes promoter regions, terminators and a group of pre-made structures that facilitate the construction of new transcriptional units.

C4. As a first step towards the engineering of programmable genetic networks in plants, a set of logic operations in GoldenBraid format was developed, including a NOT gate, two OR gates and the rudiments of a transcriptional AND gate.

/RF

REFERENCES

References

1. Zurbriggen, M.D., A. Moor, and W. Weber, Plant and bacterial systems biology as platform for plant synthetic bio(techno)logy. J Biotechnol, 2012. 160(1-2): p. 80-90.
2. Farre, G., S. Maiam Rivera, R. Alves, E. Vilaprinyo, A. Sorribas, R. Canela, S. Naqvi, G. Sandmann, T. Capell, C. Zhu, and P. Christou, Targeted transcriptomic and metabolic profiling reveals temporal bottlenecks in the maize carotenoid pathway that may be addressed by multigene engineering. Plant J, 2013. 75(3): p. 441-55.
3. Jiang, L., X. Yu, X. Qi, Q. Yu, S. Deng, B. Bai, N. Li, A. Zhang, C. Zhu, B. Liu, and J. Pang, Multigene engineering of starch biosynthesis in maize endosperm increases the total starch content and the proportion of amylose. Transgenic Res, 2013.
4. Ruiz-Lopez, N., R.P. Haslam, M. Venegas-Caleron, T. Li, J. Bauer, J.A. Napier, and O. Sayanova, Enhancing the accumulation of omega-3 long chain polyunsaturated fatty acids in transgenic *Arabidopsis thaliana* via iterative metabolic engineering and genetic crossing. Transgenic Res, 2012. 21(6): p. 1233-43.
5. Farre, G., S. Naqvi, G. Sanahuja, C. Bai, U. Zorrilla-Lopez, S.M. Rivera, R. Canela, G. Sandman, R.M. Twyman, T. Capell, C. Zhu, and P. Christou, Combinatorial genetic transformation of cereals and the creation of metabolic libraries for the carotenoid pathway. Methods Mol Biol, 2012. 847: p. 419-35.
6. von Caemmerer, S., W.P. Quick, and R.T. Furbank, The Development of C4 Rice: Current Progress and Future Challenges. Science, 2012. 336(6089): p. 1671-1672.
7. Gowik, U. and P. Westhoff, The path from C3 to C4 photosynthesis. Plant Physiol, 2011. 155(1): p. 56-63.
8. Antunes, M.S., K.J. Morey, J.J. Smith, K.D. Albrecht, T.A. Bowen, J.K. Zdunek, J.F. Troupe, M.J. Cuneo, C.T. Webb, H.W. Hellinga, and J.I. Medford, Programmable ligand detection system in plants through a synthetic signal transduction pathway. PLoS One, 2011. 6(1): p. e16292.
9. Juarez, P., S. Presa, J. Espi, B. Pineda, M.T. Anton, V. Moreno, J. Buesa, A. Granell, and D. Orzaez, Neutralizing antibodies against rotavirus produced in transgenically labelled purple tomatoes. Plant Biotechnol J, 2012. 10(3): p. 341-52.
10. Yi, D., L. Cui, L. Wang, Y. Liu, M. Zhuang, Y. Zhang, J. Zhang, Z. Lang, Z. Zhang, Z. Fang,

and L. Yang, Pyramiding of Bt cry1Ia8 and cry1Ba3 genes into cabbage (*Brassica oleracea L. var. capitata*) confers effective control against diamondback moth. Plant Cell, Tissue and Organ Culture (PCTOC), 2013: p. 1-10.
11. Ramana Rao, M.V., C. Parameswari, R. Sripriya, and K. Veluthambi, Transgene stacking and marker elimination in transgenic rice by sequential *Agrobacterium*-mediated co-transformation with the same selectable marker gene. Plant Cell Rep, 2011. 30(7): p. 1241-52.
12. Li, Z., B.P. Moon, A. Xing, Z.B. Liu, R.P. McCardell, H.G. Damude, and S.C. Falco, Stacking multiple transgenes at a selected genomic site via repeated recombinase-mediated DNA cassette exchanges. Plant Physiol, 2010. 154(2): p. 622-31.
13. Chen L, Marmey P, Taylor NJ, Brizard JP, Espinoza C, D'Cruz P, Huet H, Zhang S, de Kochko A, Beachy RN, Fauquet CM. (2002) Expression and inheritance of multiple transgenes in rice plants. Transgenic Research 11, 533-541.
14. Zorrilla-Lopez, U., G. Masip, G. Arjo, C. Bai, R. Banakar, L. Bassie, J. Berman, G. Farre, B. Miralpeix, E. Perez-Massot, M. Sabalza, G. Sanahuja, E. Vamvaka, R.M. Twyman, P. Christou, C. Zhu, and T. Capell, Engineering metabolic pathways in plants by multigene transformation. The International journal of developmental biology, 2013. 57(6-7-8): p. 565-576.
15. Leguia, M., J. Brophy, D. Densmore, and J.C. Anderson, Automated assembly of standard biological parts. Methods in enzymology, 2011. 498: p. 363-97.
16. Naqvi, S., C. Zhu, G. Farre, K. Ramessar, L. Bassie, J. Breitenbach, D. Perez Conesa, G. Ros, G. Sandmann, T. Capell, and P. Christou, Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. Proc Natl Acad Sci U S A, 2009. 106(19): p. 7762-7.
17. Zhu, C., S. Naqvi, J. Breitenbach, G. Sandmann, P. Christou, and T. Capell, Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. Proc Natl Acad Sci U S A, 2008. 105(47): p. 18232-7.
18. Ali, Z., H.M. Schumacher, E. Heine-Dobbernack, A. El-Banna, F.Y. Hafeez, H.J. Jacobsen, and H. Kiesecker, Dicistronic binary vector system-A versatile tool for gene expression studies in cell cultures and plants. J Biotechnol, 2010. 145(1): p. 9-16.
19. Ha, S.H., Y.S. Liang, H. Jung, M.J. Ahn, S.C. Suh, S.J. Kweon, D.H. Kim, Y.M. Kim, and J.K. Kim, Application of two bicistronic systems involving 2A and IRES sequences to the biosynthesis of carotenoids in rice endosperm. Plant Biotechnol J, 2010. 8(8): p. 928-38.
20. Moldrup, M.E., F. Geu-Flores, C.E. Olsen, and B.A. Halkier, Modulation of sulfur

metabolism enables efficient glucosinolate engineering. BMC Biotechnol, 2011. 11: p. 12.

21. Moldrup, M.E., F. Geu-Flores, M. de Vos, C.E. Olsen, J. Sun, G. Jander, and B.A. Halkier, Engineering of benzylglucosinolate in tobacco provides proof-of-concept for dead-end trap crops genetically modified to attract *Plutella xylostella* (diamondback moth). Plant Biotechnol J, 2012. 10(4): p. 435-42.
22. Krichevsky, A., B. Meyers, A. Vainstein, P. Maliga, and V. Citovsky, Autoluminescent plants. PLoS One, 2010. 5(11): p. e15461.
23. Bohmert-Tatarev, K., S. McAvoy, S. Daughtry, O.P. Peoples, and K.D. Snell, High levels of bioplastic are produced in fertile transplastomic tobacco plants engineered with a synthetic operon for the production of polyhydroxybutyrate. Plant Physiol, 2011. 155(4): p. 1690-708.
24. Kumar, S., F.M. Hahn, E. Baidoo, T.S. Kahlon, D.F. Wood, C.M. McMahan, K. Cornish, J.D. Keasling, H. Daniell, and M.C. Whalen, Remodeling the isoprenoid pathway in tobacco by expressing the cytoplasmic mevalonate pathway in chloroplasts. Metab Eng, 2012. 14(1): p. 19-28.
25. Lu, Y., H. Rijzaani, D. Karcher, S. Ruf, and R. Bock, Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. Proc Natl Acad Sci U S A, 2013. 110(8): p. E623-32.
26. Bock, R., Genetic engineering of the chloroplast: novel tools and new applications. Current Opinion in Biotechnology, 2014. 26: p. 7.
27. Dafny-Yelin, M. and T. Tzfira, Delivery of multiple transgenes to plant cells. Plant Physiol, 2007. 145(4): p. 1118-28.
28. Sun, Q., J. Liu, Y. Li, Q. Zhang, S. Shan, X. Li, and B. Qi, Creation and validation of a widely applicable multiple gene transfer vector system for stable transformation in plant. Plant Mol Biol, 2013.
29. Regulation (EC) No 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed, 2003: Official Journal of the European Union.
30. Ellis, T., T. Adie, and G.S. Baldwin, DNA assembly for synthetic biology: from parts to pathways and beyond. Integr Biol, 2011. 3(2): p. 109-118.
31. Bock, R., Strategies for metabolic pathway engineering with multiple transgenes. Plant Mol Biol, 2013. 83(1-2): p. 21-31.
32. Wang, Y., Y.Y. Yau, D. Perkins-Balding, and J.G. Thomson, Recombinase technology: applications and possibilities. Plant Cell Rep, 2011. 30(3): p. 267-85.
33. Naqvi, S., G. Farre, G. Sanahuja, T. Capell, C. Zhu, and P. Christou, When more is better: multigene engineering in plants. Trends Plant Sci, 2009.
34. Sainsbury, F., M. Benchabane, M.C. Goulet, and D. Michaud, Multimodal protein

constructs for herbivore insect control. Toxins (Basel), 2012. 4(6): p. 455-75.

35. Liu, W., J.S. Yuan, and C.N. Stewart Jr, Advanced genetic tools for plant biotechnology. Nature reviews. Genetics, 2013.

36. Anderson, J.C., J.E. Dueber, M. Leguia, G.C. Wu, J.A. Goler, A.P. Arkin, and J.D. Keasling, BglBricks: A flexible standard for biological part assembly. J Biol Eng, 2010. 4(1): p. 1.

37. Knight, T.F., Idempotent Vector Design for Standard Assembly of BioBricks. Tech. rep., MIT Synthetic Biology Working Group Technical Reports, 2003.

38. Kiani, S., B.B. Mohamed, K. Shehzad, A. Jamal, M.N. Shahid, A.A. Shahid, and T. Husnain, Chloroplast-targeted expression of recombinant crystal-protein gene in cotton: an unconventional combat with resistant pests. J Biotechnol, 2013. 166(3): p. 88-96.

39. Cheng, B., G. Wu, P. Vrinten, K. Falk, J. Bauer, and X. Qiu, Towards the production of high levels of eicosapentaenoic acid in transgenic plants: the effects of different host species, genes and promoters. Transgenic Res, 2010. 19(2): p. 221-9.

40. Kuroda, M., M. Kimizu, and C. Mikami, A Simple Set of Plasmids for the Production of Transgenic Plants. Bioscience, Biotechnology, and Biochemistry, 2010. 74(11): p. 2348-2351.

41. Hanafy, M.S., S.M. Rahman, Y. Nakamoto, T. Fujiwara, S. Naito, K. Wakasa, and M. Ishimoto, Differential response of methionine metabolism in two grain legumes, soybean and azuki bean, expressing a mutated form of *Arabidopsis* cystathionine gamma-synthase. J Plant Physiol, 2013. 170(3): p. 338-45.

42. Hatakeyama, K., K. Suwabe, R.N. Tomita, T. Kato, T. Nunome, H. Fukuoka, and S. Matsumoto, Identification and Characterization of *Crr1a*, a Gene for Resistance to Clubroot Disease (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. PLoS ONE, 2013. 8(1): p. e54745.

43. Saika, H., S. Nonaka, K. Osakabe, and S. Toki, Sequential Monitoring of Transgene Expression Following *Agrobacterium*-Mediated Transformation of Rice. Plant and Cell Physiology, 2012. 53(11): p. 1974-1983.

44. Shi, Z., A.G. Wedd, and S.L. Gras, Parallel *in vivo* DNA assembly by recombination: experimental demonstration and theoretical approaches. PLoS One, 2013. 8(2): p. e56854.

45. Chen, W.H., Z.J. Qin, J. Wang, and G.P. Zhao, The MASTER (methylation-assisted tailorable ends rational) ligation method for seamless DNA assembly. Nucleic Acids Res, 2013. 41(8): p. e93.

46. Wang, R.Y., Z.Y. Shi, Y.Y. Guo, J.C. Chen, and G.Q. Chen, DNA fragments assembly based on nicking enzyme system. PLoS One, 2013. 8(3): p. e57943.

47. Moriarity, B.S., E.P. Rahrmann, V.W. Keng, L.S. Manlove, D.A. Beckmann, N.K. Wolf,

T. Khurshid, J.B. Bell, and D.A. Largaespada, Modular assembly of transposon integratable multigene vectors using RecWay assembly. Nucleic Acids Res, 2013. 41(8): p. e92.

48. Nour-Eldin, H.H., F. Geu-Flores, and B.A. Halkier, USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories. Methods Mol Biol, 2010. 643: p. 185-200.
49. Sleight, S.C., B.A. Bartley, J.A. Lieviant, and H.M. Sauro, In-Fusion BioBrick assembly and re-engineering. Nucleic Acids Res, 2010. 38(8): p. 2624-36.
50. Gibson, D.G., L. Young, R.-Y. Chuang, J.C. Venter, C.A. Hutchison, and H.O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Meth, 2009. 6(5): p. 343-345.
51. Gibson, D.G., Enzymatic assembly of overlapping DNA fragments. Methods in enzymology, 2011. 498: p. 349-61.
52. Karimi, M., A. Depicker, and P. Hilson, Recombinational cloning with plant gateway vectors. Plant Physiology, 2007. 145: p. 1144-1154.
53. Estornell, L.H., D. Orzaez, L. Lopez-Pena, B. Pineda, M.T. Anton, V. Moreno, and A. Granell, A multisite gateway-based toolkit for targeted gene expression and hairpin RNA silencing in tomato fruits. Plant Biotechnol J, 2009. 7(3): p. 298-309.
54. Fernandez, A.I., N. Viron, M. Alhagdow, M. Karimi, M. Jones, Z. Amsellem, A. Sicard, A. Czerednik, G. Angenent, D. Grierson, S. May, G. Seymour, Y. Eshed, M. Lemaire-Chamley, C. Rothan, and P. Hilson, Flexible tools for gene expression and silencing in tomato. Plant Physiol, 2009. 151(4): p. 1729-40.
55. Tanaka, Y., S. Nakamura, M. Kawamukai, N. Koizumi, and T. Nakagawa, Development of a series of gateway binary vectors possessing a tunicamycin resistance gene as a marker for the transformation of *Arabidopsis thaliana*. Biosci Biotechnol Biochem, 2011. 75(4): p. 804-7.
56. Mann, D.G., P.R. Lafayette, L.L. Abercrombie, Z.R. King, M. Mazarei, M.C. Halter, C.R. Poovaiah, H. Baxter, H. Shen, R.A. Dixon, W.A. Parrott, and C. Neal Stewart, Jr., Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum L.*) and other monocot species. Plant Biotechnol J, 2012. 10(2): p. 226-36.
57. Karimi, M., D. Inze, M. Van Lijsebettens, and P. Hilson, Gateway vectors for transformation of cereals. Trends Plant Sci, 2013. 18(1): p. 1-4.
58. Karimi, M., A. Bleys, R. Vanderhaeghen, and P. Hilson, Building blocks for plant gene assembly. Plant Physiology, 2007. 145: p. 1183-1191.
59. Estornell, L.H., A. Granell, and D. Orzaez, Exploiting Multisite Gateway and pENFRUIT plasmid collection for fruit genetic engineering. Methods Mol Biol, 2012. 847: p. 351-68.

60. Chen, Q.J., H.M. Zhou, J. Chen, and X.C. Wang, A Gateway-based platform for multigene plant transformation. Plant Mol Biol, 2006. 62(6): p. 927 - 36.
61. Ren, F., Q.J. Chen, M. Xie, L.J. Li, W.H. Wu, J. Chen, and X.C. Wang, Engineering the K+ uptake regulatory pathway by MultiRound Gateway. J Plant Physiol, 2010. 167(16): p. 1412-7.
62. Buntru, M., S. Gartner, L. Staib, F. Kreuzaler, and N. Schlaich, Delivery of multiple transgenes to plant cells by an improved version of MultiRound Gateway technology. Transgenic Res, 2013. 22(1): p. 153-67.
63. Vemanna, R.S., B.K. Chandrashekar, H.M. Hanumantha Rao, S.K. Sathyanarayanagupta, K.S. Sarangi, K.N. Nataraja, and M. Udayakumar, A modified MultiSite gateway cloning strategy for consolidation of genes in plants. Mol Biotechnol, 2013. 53(2): p. 129-38.
64. Lin, L., Y.-G. Liu, X. Xu, and B. Li, Efficient linking and transfer of multiple genes by a multigene assembly and transformation vector system. Proceedings of the National Academy of Sciences of the United States of America, 2003. 100(10): p. 5962-5967.
65. Dafhnis-Calas, F., Z. Xu, S. Haines, S.K. Malla, M.C.M. Smith, and W.R.A. Brown, Iterative in vivo assembly of large and complex transgenes by combining the activities of φC31 integrase and Cre recombinase. Nucleic Acids Research, 2005. 33(22): p. e189.
66. De Paepe, A., S. De Buck, K. Hoorelbeke, J. Nolf, I. Peck, and A. Depicker, High frequency of single-copy T-DNA transformants produced by floral dip in CRE-expressing *Arabidopsis* plants. Plant J, 2009. 59(4): p. 517-27.
67. De Paepe, A., S. De Buck, J. Nolf, E. Van Lerberge, and A. Depicker, Site-specific T-DNA integration in *Arabidopsis thaliana* mediated by the combined action of CRE recombinase and varphiC31 integrase. Plant J, 2013. 75(1): p. 172-84.
68. Chen, Q.J., M. Xie, X.X. Ma, L. Dong, J. Chen, and X.C. Wang, MISSA is a highly efficient in vivo DNA assembly method for plant multiple-gene transformation. Plant Physiol, 2010. 153(1): p. 41-51.
69. Ma, D.M., W.R. Xu, H.W. Li, F.X. Jin, L.N. Guo, J. Wang, H.J. Dai, and X. Xu, Co-expression of the *Arabidopsis* SOS genes enhances salt tolerance in transgenic tall fescue (*Festuca arundinacea Schreb.*). Protoplasma, 2013.
70. Ma, L., J. Dong, Y. Jin, M. Chen, X. Shen, and T. Wang, RMDAP: a versatile, ready-to-use toolbox for multigene genetic transformation. PLoS One, 2011. 6(5): p. e19883.
71. Mosberg, J.A., M.J. Lajoie, and G.M. Church, Lambda red recombineering in *Escherichia coli* occurs through a fully single-stranded intermediate. Genetics, 2010. 186(3): p. 791-9.
72. Zeevi, V., A. Tovkach, and T. Tzfira, Artificial zinc finger nucleases for DNA cloning.

Methods Mol Biol, 2010. 649: p. 209-25.
73. Zeevi, V., Z. Liang, U. Arieli, and T. Tzfira, Zinc finger nuclease and homing endonuclease-mediated assembly of multigene plant transformation vectors. Plant Physiol, 2012. 158(1): p. 132-44.
74. Goderis, I.J.W.M., M.F.C. De Bolle, I.E.J.A. François, P.F.J. Wouters, W.F. Broekaert, and B.P.A. Cammue, A set of modular plant transformation vectors allowing flexible insertion of up to six expression units. Plant Mol Biol, 2002. 50(1): p. 17-27.
75. Tzfira, T., G.W. Tian, B. Lacroix, S. Vyas, J. Li, Y. Leitner-Dagan, A. Krichevsky, T. Taylor, A. Vainstein, and V. Citovsky, pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. Plant Mol Biol, 2005. 57(4): p. 503-16.
76. Dafny-Yelin, M., S.M. Chung, E.L. Frankman, and T. Tzfira, pSAT RNA interference vectors: a modular series for multiple gene down-regulation in plants. Plant Physiol, 2007. 145(4): p. 1272-81.
77. Chen, M.H., G.W. Tian, Y. Gafni, and V. Citovsky, Effects of calreticulin on viral cell-to-cell movement. Plant Physiol, 2005. 138(4): p. 1866-76.
78. Zrachya, A., E. Glick, Y. Levy, T. Arazi, V. Citovsky, and Y. Gafni, Suppressor of RNA silencing encoded by Tomato yellow leaf curl virus-Israel. Virology, 2007. 358(1): p. 159-65.
79. Glick, E., A. Zrachya, Y. Levy, A. Mett, D. Gidoni, E. Belausov, V. Citovsky, and Y. Gafni, Interaction with host SGS3 is required for suppression of RNA silencing by tomato yellow leaf curl virus V2 protein. Proc Natl Acad Sci U S A, 2008. 105(1): p. 157-61.
80. Fujisawa, M., E. Takita, H. Harada, N. Sakurai, H. Suzuki, K. Ohyama, D. Shibata, and N. Misawa, Pathway engineering of *Brassica napus* seeds using multiple key enzyme genes involved in ketocarotenoid formation. J Exp Bot, 2009. 60(4): p. 1319-32.
81. Engler, C., R. Kandzia, and S. Marillonnet, A one pot, one step, precision cloning method with high throughput capability. PLoS One, 2008. 3(11): p. e3647.
82. Engler, C., R. Gruetzner, R. Kandzia, and S. Marillonnet, Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS One, 2009. 4(5): p. e5553.
83. Bendandi, M., S. Marillonnet, R. Kandzia, F. Thieme, A. Nickstadt, S. Herz, R. Frode, S. Inoges, A. Lopez-Diaz de Cerio, E. Soria, H. Villanueva, G. Vancanneyt, A. McCormick, D. Tuse, J. Lenz, J.E. Butler-Ransohoff, V. Klimyuk, and Y. Gleba, Rapid, high-yield production in plants of individualized idiotype vaccines for non-Hodgkin's lymphoma. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO, 2010. 21(12): p. 2420-7.
84. Canonne, J., D. Marino, L.D. Noel, I. Arechaga, C. Pichereaux, M. Rossignol, D. Roby, and S. Rivas, Detection and functional characterization of a 215 amino acid

N-terminal extension in the *Xanthomonas* type III effector XopD. PLoS One, 2010. 5(12): p. e15773.

85. Cermak, T., E.L. Doyle, M. Christian, L. Wang, Y. Zhang, C. Schmidt, J.A. Baller, N.V. Somia, A.J. Bogdanove, and D.F. Voytas, Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res, 2011. 39(12): p. e82.
86. Zhang, F., L. Cong, S. Lodato, S. Kosuri, G.M. Church, and P. Arlotta, Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nature biotechnology, 2011. 29(2): p. 149-53.
87. Liang, J., J.C. Ning, and H. Zhao, Coordinated induction of multi-gene pathways in *Saccharomyces cerevisiae*. Nucleic Acids Res, 2013. 41(4): p. e54.
88. Emami, S., M.C. Yee, and J.R. Dinneny, A robust family of Golden Gate Agrobacterium vectors for plant synthetic biology. Frontiers in plant science, 2013. 4: p. 339.
89. Weber, E., C. Engler, R. Gruetzner, S. Werner, and S. Marillonnet, A modular cloning system for standardized assembly of multigene constructs. PLoS One, 2011. 6(2): p. e16765.
90. Sarrion-Perdigones, A., E.E. Falconi, S.I. Zandalinas, P. Juarez, A. Fernandez-del-Carmen, A. Granell, and D. Orzaez, GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. PLoS One, 2011. 6(7): p. e21622.
91. Sarrion-Perdigones, A., M. Vazquez-Vilar, J. Palaci, B. Castelijns, J. Forment, P. Ziarsolo, J. Blanca, A. Granell, and D. Orzaez, GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. Plant Physiol, 2013. 162(3): p. 1618-31.
92. Werner, S., C. Engler, E. Weber, R. Gruetzner, and S. Marillonnet, Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. Bioengineered, 2012. 3(1): p. 38-43.
93. Gardner, T.S., C.R. Cantor, and J.J. Collins, Construction of a genetic toggle switch in *Escherichia coli*. Nature, 2000. 403(6767): p. 339-42.
94. Elowitz, M.B. and S. Leibler, A synthetic oscillatory network of transcriptional regulators. Nature, 2000. 403(6767): p. 335-8.
95. Ajo-Franklin, C.M., D.A. Drubin, J.A. Eskin, E.P. Gee, D. Landgraf, I. Phillips, and P.A. Silver, Rational design of memory in eukaryotic cells. Genes & development, 2007. 21(18): p. 2271-6.
96. Bonnet, J., P. Subsoontorn, and D. Endy, Rewritable digital data storage in live cells via engineered control of recombination directionality. Proc Natl Acad Sci U S A, 2012. 109(23): p. 8884-9.
97. Weber, W., M. Daoud-El Baba, and M. Fussenegger, Synthetic ecosystems based on

airborne inter- and intrakingdom communication. Proc Natl Acad Sci U S A, 2007. 104(25): p. 10435-40.

98. Tabor, J.J., H.M. Salis, Z.B. Simpson, A.A. Chevalier, A. Levskaya, E.M. Marcotte, C.A. Voigt, and A.D. Ellington, A synthetic genetic edge detection program. Cell, 2009. 137(7): p. 1272-81.

99. Ro, D.K., E.M. Paradise, M. Ouellet, K.J. Fisher, K.L. Newman, J.M. Ndungu, K.A. Ho, R.A. Eachus, T.S. Ham, J. Kirby, M.C. Chang, S.T. Withers, Y. Shiba, R. Sarpong, and J.D. Keasling, Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature, 2006. 440(7086): p. 940-3.

100. Martin, V.J., D.J. Pitera, S.T. Withers, J.D. Newman, and J.D. Keasling, Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. Nature biotechnology, 2003. 21(7): p. 796-802.

101. Khalil, A.S. and J.J. Collins, Synthetic biology: applications come of age. Nature reviews. Genetics, 2010. 11(5): p. 367-79.

102. Kramer, B.P., C. Fischer, and M. Fussenegger, BioLogic gates enable logical transcription control in mammalian cells. Biotechnology and bioengineering, 2004. 87(4): p. 478-84.

103. Buchler, N.E., U. Gerland, and T. Hwa, On schemes of combinatorial transcription logic. Proc Natl Acad Sci U S A, 2003. 100(9): p. 5136-41.

104. Siuti, P., J. Yazbek, and T.K. Lu, Synthetic circuits integrating logic and memory in living cells. Nature biotechnology, 2013. 31(5): p. 448-52.

105. Wang, B., R.I. Kitney, N. Joly, and M. Buck, Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. Nature communications, 2011. 2: p. 508.

106. Muller, K., R. Engesser, S. Schulz, T. Steinberg, P. Tomakidi, C.C. Weber, R. Ulm, J. Timmer, M.D. Zurbriggen, and W. Weber, Multi-chromatic control of mammalian gene expression and signaling. Nucleic Acids Res, 2013. 41(12): p. e124.

107. Shis, D.L. and M.R. Bennett, Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. Proc Natl Acad Sci U S A, 2013. 110(13): p. 5028-33.

108. Xie, Z., L. Wroblewska, L. Prochazka, R. Weiss, and Y. Benenson, Multi-input RNAi-based logic circuit for identification of specific cancer cells. Science, 2011. 333(6047): p. 1307-11.

109. Auslander, S., D. Auslander, M. Muller, M. Wieland, and M. Fussenegger, Programmable single-cell mammalian biocomputers. Nature, 2012. 487(7405): p. 123-7.

110. Callura, J.M., C.R. Cantor, and J.J. Collins, Genetic switchboard for synthetic biology applications. Proc Natl Acad Sci U S A, 2012. 109(15): p. 5850-5.

111. Bonnet, J., P. Yin, M.E. Ortiz, P. Subsoontorn, and D. Endy, Amplifying genetic logic gates. Science, 2013. 340(6132): p. 599-603.
112. Benenson, Y., Engineering. Recombinatorial logic. Science, 2013. 340(6132): p. 554-5.
113. Silva-Rocha, R. and V. de Lorenzo, Mining logic gates in prokaryotic transcriptional regulation networks. FEBS letters, 2008. 582(8): p. 1237-44.
114. Check, E., Synthetic biology: Designs on life. Nature, 2005. 438(7067): p. 417 - 418.
115. Haseloff, J. and J. Ajioka, Synthetic biology: history, challenges and prospects. J R Soc Interface, 2009. 6 Suppl 4: p. S389-91.
116. Carter, G.W., C.G. Rush, F. Uygun, N.A. Sakhanenko, D.J. Galas, and T. Galitski, A systems-biology approach to modular genetic complexity. Chaos, 2010. 20(2): p. 026102.
117. Matzas, M., P.F. Stahler, N. Kefer, N. Siebelt, V. Boisguerin, J.T. Leonard, A. Keller, C.F. Stahler, P. Haberle, B. Gharizadeh, F. Babrzadeh, and G.M. Church, High-fidelity gene synthesis by retrieval of sequence-verified DNA identified using high-throughput pyrosequencing. Nat Biotech, 2010. 28(12): p. 1291-1294.
118. Kosuri, S., N. Eroshenko, E.M. LeProust, M. Super, J. Way, J.B. Li, and G.M. Church, Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. Nat Biotech, 2010. 28(12): p. 1295-1299.
119. Gibson, D.G., J.I. Glass, C. Lartigue, V.N. Noskov, R.-Y. Chuang, M.A. Algire, G.A. Benders, M.G. Montague, L. Ma, M.M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E.A. Denisova, L. Young, Z.-Q. Qi, T.H. Segall-Shapiro, C.H. Calvey, P.P. Parmar, C.A. Hutchison, III, H.O. Smith, and J.C. Venter, Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. Science, 2010. 329(5987): p. 52-56.
120. Knight, T.F., Idempotent vector design for standard assembly of biobricks. MIT Synthetic Biology Working Group Technical Reports, 2003.
121. Voinnet, O., S. Rivas, P. Mestre, and D. Baulcombe, An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant Journal, 2003. 33(5): p. 949-956.
122. Butelli, E., L. Titta, M. Giorgio, H.P. Mock, A. Matros, S. Peterek, E.G. Schijlen, R.D. Hall, A.G. Bovy, J. Luo, and C. Martin, Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nature biotechnology, 2008. 26(11): p. 1301-8.
123. Koltunow, A.M., J. Truettner, K.H. Cox, M. Wallroth, and R.B. Goldberg, Different Temporal and Spatial Gene Expression Patterns Occur during Anther Development. Plant Cell, 1990. 2(12): p. 1201-1224.
124. Kobayashi, K., I. Munemura, K. Hinata, and S. Yamamura, Bisexual sterility conferred

by the differential expression of Barnase and Barstar: a simple and efficient method of transgene containment. Plant Cell Rep, 2006. 25(12): p. 1347-1354.

125. Karimi, M., D. Inze, and A. Depicker, GATEWAY vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci, 2002. 7(5): p. 193-5.

126. Lin, L., Y.-G. Liu, X. Xu, and B. Li, Efficient linking and transfer of multiple genes by a multigene assembly and transformation vector system. Proc Natl Acad Sci, 2003. 100(10): p. 5962-5967.

127. Castilho, A., R. Strasser, J. Stadlmann, J. Grass, J. Jez, P. Gattinger, R. Kunert, H. Quendler, M. Pabst, R. Leonard, F. Altmann, and H. Steinkellner, *In planta* protein sialylation through overexpression of the respective mammalian pathway. J Biol Chem, 2010. 285(21): p. 15923-30.

128. Hamilton, C.M., A binary-BAC system for plant transformation with high-molecular-weight DNA. Gene, 1997. 200(1-2): p. 107-16.

129. Beal, J., T. Lu, and R. Weiss, Automatic compilation from high-level biologically-oriented programming language to genetic regulatory networks. PLoS One, 2011. 6(8): p. e22490.

130. Callaway, E., Glowing plants spark debate. Nature, 2013. 498(7452): p. 15-6.

131. Hellens, R.P., E.A. Edwards, N.R. Leyland, S. Bean, and P.M. Mullineaux, pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. Plant Mol Biol, 2000. 42(6): p. 819-32.

132. Wieland, W.H., A. Lammers, A. Schots, and D.V. Orzaez, Plant expression of chicken secretory antibodies derived from combinatorial libraries. J Biotechnol, 2006. 122(3): p. 382-91.

133. Aslanidis, C. and P.J. de Jong, Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res, 1990. 18(20): p. 6069-74.

134. Geu-Flores, F., H.H. Nour-Eldin, M.T. Nielsen, and B.A. Halkier, USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res, 2007. 35(7): p. e55.

135. Belda-Palazon, B., L. Ruiz, E. Marti, S. Tarraga, A.F. Tiburcio, F. Culianez, R. Farras, P. Carrasco, and A. Ferrando, Aminopropyltransferases involved in polyamine biosynthesis localize preferentially in the nucleus of plant cells. PLoS One, 2012. 7(10): p. e46907.

136. Felippes, F.F., J.W. Wang, and D. Weigel, MIGS: miRNA-induced gene silencing. Plant J, 2012. 70(3): p. 541-7.

137. Yan, H., X. Deng, Y. Cao, J. Huang, L. Ma, and B. Zhao, A novel approach for the construction of plant amiRNA expression vectors. J Biotechnol, 2011. 151(1): p. 9-14.

138. Django, Django (Version 1.5) Retrieved from http://djangoproject.com 2013, Django Software Foundation.

139. Weber, E., R. Gruetzner, S. Werner, C. Engler, and S. Marillonnet, Assembly of designer TAL effectors by Golden Gate cloning. PLoS One, 2011. 6(5): p. e19722.
140. Li, L., M.J. Piatek, A. Atef, A. Piatek, A. Wibowo, X. Fang, J.S. Sabir, J.K. Zhu, and M.M. Mahfouz, Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. Plant Mol Biol, 2012. 78(4-5): p. 407-16.
141. Nagaya, S., K. Kawamura, A. Shinmyo, and K. Kato, The HSP terminator of *Arabidopsis thaliana* increases gene expression in plant cells. Plant Cell Physiol, 2009. 51(2): p. 328-32.
142. Grentzmann, G., J.A. Ingram, P.J. Kelly, R.F. Gesteland, and J.F. Atkins, A dual-luciferase reporter system for studying recoding signals. RNA, 1998. 4(4): p. 479-86.
143. Roberts, C., Rajagopal,S., Smith,L.M., Nguyen,T.A., Yang,W.,, S. Nugrohu, Ravi,K.S., Vijayachandra,K., Harcourt,R.L.,, L. Dransfield, Desamero,N., Slamet,I., Hadjukiewicz,P., Svab,Z.,, and P. Maliga, Mayer,J.E., Keese,P.K., Kilian,A. and Jefferson,R.A., A comprehensive set of modular vectors for advanced manipulations and efficient transformation of plants. pCAMBIA Vector Release Manual, 1997.
144. Chi-Ham, C.L., S. Boettiger, R. Figueroa-Balderas, S. Bird, J.N. Geoola, P. Zamora, M. Alandete-Saez, and A.B. Bennett, An intellectual property sharing initiative in agricultural biotechnology: development of broadly accessible technologies for plant transformation. Plant Biotechnol J, 2012. 10(5): p. 501-10.
145. Oye, K.A. and R. Wellhausen, The Intellectual Commons and Property in Synthetic Biology, in Synthetic Biology: The technoscience and its societal consequences, M. Schmidt, Editor 2010, Springer Netherlands. p. 121-140.
146. Caemmerer, S.v., W. Quick, and R. Furbank, The Development of C4 Rice: Current Progress and Future Challenges. Science (New York, N.Y, 2012. 336: p. 2.
147. Clough, S.J. and A.F. Bent, Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J, 1998. 16(6): p. 735-43.
148. McAdams, H.H. and A. Arkin, Towards a circuit engineering discipline. Current biology : CB, 2000. 10(8): p. R318-20.
149. Glass, L. and S.A. Kauffman, The logical analysis of continuous, non-linear biochemical control networks. Journal of theoretical biology, 1973. 39(1): p. 103-29.
150. Kauffman, S., The large scale structure and dynamics of gene control circuits: an ensemble approach. Journal of theoretical biology, 1974. 44(1): p. 167-90.
151. McAdams, H.H. and L. Shapiro, Circuit simulation of genetic networks. Science, 1995. 269(5224): p. 650-6.
152. Purnick, P.E.M. and R. Weiss, The second wave of synthetic biology: from modules to

systems. Nat Rev Mol Cell Biol, 2009. 10(6): p. 410-422.
153. Weber, W., N. Link, and M. Fussenegger, A genetic redox sensor for mammalian cells. Metab Eng, 2006. 8(3): p. 273-80.
154. Blount, B.A., T. Weenink, and T. Ellis, Construction of synthetic regulatory networks in yeast. FEBS letters, 2012. 586(15): p. 2112-21.
155. Gupta, S., E.E. Bram, and R. Weiss, Genetically Programmable Pathogen Sense and Destroy. ACS synthetic biology, 2013.
156. Royo, J.L., P.D. Becker, E.M. Camacho, A. Cebolla, C. Link, E. Santero, and C.A. Guzman, *In vivo* gene regulation in *Salmonella spp.* by a salicylate-dependent control circuit. Nat Methods, 2007. 4(11): p. 937-42.
157. Devin-Leclerc, J., X. Meng, F. Delahaye, P. Leclerc, E.E. Baulieu, and M.G. Catelli, Interaction and dissociation by ligands of estrogen receptor and Hsp90: the antiestrogen RU 58668 induces a protein synthesis-dependent clustering of the receptor in the cytoplasm. Molecular endocrinology, 1998. 12(6): p. 842-54.
158. Echeverria, P.C., G. Mazaira, A. Erlejman, C. Gomez-Sanchez, G. Piwien Pilipuk, and M.D. Galigniana, Nuclear import of the glucocorticoid receptor-hsp90 complex through the nuclear pore complex is mediated by its interaction with Nup62 and importin beta. Molecular and cellular biology, 2009. 29(17): p. 4788-97.
159. Hiratsu, K., N. Mitsuda, K. Matsui, and M. Ohme-Takagi, Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in *Arabidopsis*. Biochemical and biophysical research communications, 2004. 321(1): p. 172-178.
160. Ikeda, M. and M. Ohme-Takagi, A novel group of transcriptional repressors in *Arabidopsis*. Plant Cell Physiol, 2009. 50(5): p. 970-5.
161. Li, M.Z. and S.J. Elledge, SLIC: a method for sequence- and ligation-independent cloning. Methods Mol Biol, 2012. 852: p. 51-9.
162. Seitz, H., C. Weigel, and W. Messer, The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*. Molecular microbiology, 2000. 37(5): p. 1270-9.
163. de Folter, S., R.G. Immink, M. Kieffer, L. Parenicova, S.R. Henz, D. Weigel, M. Busscher, M. Kooiker, L. Colombo, M.M. Kater, B. Davies, and G.C. Angenent, Comprehensive interaction map of the *Arabidopsis* MADS Box transcription factors. Plant Cell, 2005. 17(5): p. 1424-33.
164. Smaczniak, C., R.G. Immink, J.M. Muino, R. Blanvillain, M. Busscher, J. Busscher-Lange, Q.D. Dinh, S. Liu, A.H. Westphal, S. Boeren, F. Parcy, L. Xu, C.C. Carles, G.C. Angenent, and K. Kaufmann, Characterization of MADS-domain transcription factor complexes in *Arabidopsis* flower development. Proc Natl Acad Sci U S A, 2012. 109(5): p. 1560-5.

165. Antunes, M.S., S.B. Ha, N. Tewari-Singh, K.J. Morey, A.M. Trofka, P. Kugrens, M. Deyholos, and J.I. Medford, A synthetic de-greening gene circuit provides a reporting system that is remotely detectable and has a re-set capacity. Plant Biotechnol J, 2006. 4(6): p. 605-22.
166. Antunes, M.S., K.J. Morey, N. Tewari-Singh, T.A. Bowen, J.J. Smith, C.T. Webb, H.W. Hellinga, and J.I. Medford, Engineering key components in a synthetic eukaryotic signal transduction pathway. Molecular systems biology, 2009. 5: p. 270.
167. Moore, I., M. Samalova, and S. Kurup, Transactivated and chemically inducible gene expression in plants. Plant J, 2006. 45(4): p. 651-83.
168. Heyl, A., E. Ramireddy, W.G. Brenner, M. Riefler, J. Allemeersch, and T. Schmulling, The transcriptional repressor ARR1-SRDX suppresses pleiotropic cytokinin activities in *Arabidopsis*. Plant Physiol, 2008. 147(3): p. 1380-95.
169. Kelly, J., A. Rubin, J. Davis, C. Ajo-Franklin, J. Cumbers, M. Czar, K. de Mora, A. Glieberman, D. Monie, and D. Endy, Measuring the activity of BioBrick promoters using an *in vivo* reference standard. Journal of Biological Engineering, 2009. 3(1): p. 4.
170. Casini, A., J.T. Macdonald, J.D. Jonghe, G. Christodoulou, P.S. Freemont, G.S. Baldwin, and T. Ellis, One-pot DNA construction for synthetic biology: the Modular Overlap-Directed Assembly with Linkers (MODAL) strategy. Nucleic Acids Res, 2013.
171. Chappell, J., K. Jensen, and P.S. Freemont, Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. Nucleic Acids Res, 2013. 41(5): p. 3471-81.
172. Keren, L., O. Zackay, M. Lotan-Pompan, U. Barenholz, E. Dekel, V. Sasson, G. Aidelberg, A. Bren, D. Zeevi, A. Weinberger, U. Alon, R. Milo, and E. Segal, Promoters maintain their relative activity levels under different growth conditions. Molecular systems biology, 2013. 9: p. 701.
173. Teo, W.S. and M.W. Chang, Development and characterization of AND-gate dynamic controllers with a modular synthetic GAL1 core promoter in *Saccharomyces cerevisiae*. Biotechnology and bioengineering, 2013.
174. Peremarti, A., R.M. Twyman, S. Gomez-Galera, S. Naqvi, G. Farre, M. Sabalza, B. Miralpeix, S. Dashevskaya, D. Yuan, K. Ramessar, P. Christou, C. Zhu, L. Bassie, and T. Capell, Promoter diversity in multigene transformation. Plant Mol Biol, 2010. 73(4-5): p. 363-78.
175. Silva-Rocha, R. and V. de Lorenzo, A GFP-lacZ bicistronic reporter system for promoter analysis in environmental gram-negative bacteria. PLoS One, 2012. 7(4): p. e34675.
176. Untergasser, A., G.J. Bijl, W. Liu, T. Bisseling, J.G. Schaart, and R. Geurts, One-step *Agrobacterium* mediated transformation of eight genes essential for *rhizobium*

symbiotic signaling using the novel binary vector system pHUGE. PLoS One, 2012. 7(10): p. e47885.
177. Liu, Y.G., Y. Shirano, H. Fukaki, Y. Yanai, M. Tasaka, S. Tabata, and D. Shibata, Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. Proc Natl Acad Sci U S A, 1999. 96(11): p. 6535-40.
178. Silva-Rocha, R., E. Martinez-Garcia, B. Calles, M. Chavarria, A. Arce-Rodriguez, A. de Las Heras, A.D. Paez-Espino, G. Durante-Rodriguez, J. Kim, P.I. Nikel, R. Platero, and V. de Lorenzo, The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. Nucleic Acids Res, 2013. 41(Database issue): p. D666-75.
179. Densmore, D., T.H. Hsiau, J.T. Kittleson, W. DeLoache, C. Batten, and J.C. Anderson, Algorithms for automated DNA assembly. Nucleic Acids Res, 2010. 38(8): p. 2607-16.
180. Hillson, N.J., R.D. Rosengarten, and J.D. Keasling, j5 DNA assembly design automation software. ACS synthetic biology, 2012. 1(1): p. 14-21.
181. Babitha, K.C., S.V. Ramu, V. Pruthvi, P. Mahesh, K.N. Nataraja, and M. Udayakumar, Co-expression of AtbHLH17 and AtWRKY28 confers resistance to abiotic stress in *Arabidopsis*. Transgenic Res, 2013. 22(2): p. 327-41.
182. Juarez, P., E. Huet-Trujillo, A. Sarrion-Perdigones, E.E. Falconi, A. Granell, and D. Orzaez, Combinatorial Analysis of Secretory Immunoglobulin A (sIgA) Expression in Plants. International journal of molecular sciences, 2013. 14(3): p. 6205-22.
183. Tamsir A Fau - Tabor, J.J., C.A. Tabor Jj Fau - Voigt, and C.A. Voigt, Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. Nature, 2011. 469(7329): p. 212-5.
184. Silva-Rocha, R. and V. de Lorenzo, Engineering Multicellular Logic in Bacteria with Metabolic Wires. ACS synthetic biology, 2013.
185. Lienert, F., J.P. Torella, J.H. Chen, M. Norsworthy, R.R. Richardson, and P.A. Silver, Two- and three-input TALE-based AND logic computation in embryonic stem cells. Nucleic Acids Res, 2013.
186. Cong, L., F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P.D. Hsu, X. Wu, W. Jiang, L.A. Marraffini, and F. Zhang, Multiplex Genome Engineering Using CRISPR/Cas Systems. Science, 2013. 339(6121): p. 819-823.
187. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in biotechnology, 2013. 31(7): p. 397-405.
188. Sayut, D.J., Y. Niu, and L. Sun, Construction and enhancement of a minimal genetic and logic gate. Applied and environmental microbiology, 2009. 75(3): p. 637-42.
189. De Loose, D., Van Bockstaele, Van Montagu, Depicker, Different 5' leader sequences

modulate beta-glucuronidase accumulation levels in transgenic *Nicotiana tabacum* plants. Euphytica, 1995. 85: p. 209-216.
190. Mitsuhara, I., M. Ugaki, H. Hirochika, M. Ohshima, T. Murakami, Y. Gotoh, Y. Katayose, S. Nakamura, R. Honkura, S. Nishimiya, K. Ueno, A. Mochizuki, H. Tanimoto, H. Tsugawa, Y. Otsuki, and Y. Ohashi, Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants. Plant Cell Physiol, 1996. 37(1): p. 49-59.
191. Aoyama, T. and N.H. Chua, A glucocorticoid-mediated transcriptional induction system in transgenic plants. Plant J, 1997. 11(3): p. 605-12.
192. Hartley, J.L., G.F. Temple, and M.A. Brasch, DNA cloning using in vitro site-specific recombination. Genome Research, 2000. 10(11): p. 1788-1795.
193. Hawkins, E., M. Beck, B. Butler, and K. Wood, Increased Renilla luciferase sensitivity in the Dual-Luciferase® Reporter Assay System. Promega Notes Magazine, 2003. 85: p. 31-33.
194. Moore, I., L. Galweiler, D. Grosskopf, J. Schell, and K. Palme, A transcription activation system for regulated gene expression in transgenic plants. Proc Natl Acad Sci U S A, 1998. 95(1): p. 376-81.
195. Schwab, R., S. Ossowski, M. Riester, N. Warthmann, and D. Weigel, Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. Plant Cell, 2006. 18(5): p. 1121-33.
196. Qiu, W.P., J.W. Park, and H.B. Scholthof, Tombusvirus p19-mediated suppression of virus-induced gene silencing is controlled by genetic and dosage features that influence pathogenicity. Molecular Plant-Microbe Interactions, 2002. 15(3): p. 269-280.
197. Bevan, M., W.M. Barnes, and M.D. Chilton, Structure and transcription of the nopaline synthase gene region of T-DNA. Nucleic Acids Res, 1983. 11(2): p. 369-85.
198. Wielopolska, A., H. Townley, I. Moore, P. Waterhouse, and C. Helliwell, A high-throughput inducible RNAi vector for plants. Plant Biotechnology Journal, 2005. 3(6): p. 583-590.
199. Yoshida, K., T. Kasai, M.R. Garcia, S. Sawada, T. Shoji, S. Shimizu, K. Yamazaki, Y. Komeda, and A. Shinmyo, Heat-inducible expression system for a foreign gene in cultured tobacco cells using the HSP18.2 promoter of *Arabidopsis thaliana*. Appl Microbiol Biotechnol, 1995. 44(3-4): p. 466-72.
200. Sung, D.Y., E. Vierling, and C.L. Guy, Comprehensive expression profile analysis of the Arabidopsis hsp70 gene family. Plant Physiology, 2001. 126(2): p. 789-800.
201. Koo, J.C., S. Asurmendi, J. Bick, T. Woodford-Thomas, and R.N. Beachy, Ecdysone agonist-inducible expression of a coat protein gene from tobacco mosaic virus confers viral resistance in transgenic *Arabidopsis*. Plant J, 2004. 37(3): p. 439-48.

202. Engineer, C.B., K.C. Fitzsimmons, J.J. Schmuke, S.B. Dotson, and R.G. Kranz, Development and evaluation of a Gal4-mediated LUC/GFP/GUS enhancer trap system in *Arabidopsis*. BMC Plant Biol, 2005. 5: p. 9.

203. Odell, J.T., F. Nagy, and N.H. Chua, Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature, 1985. 313(6005): p. 810-2.

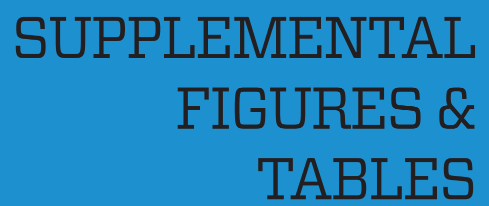
204. Grefen, C., N. Donald, K. Hashimoto, J. Kudla, K. Schumacher, and M.R. Blatt, A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. Plant J, 2010. 64(2): p. 355-65.

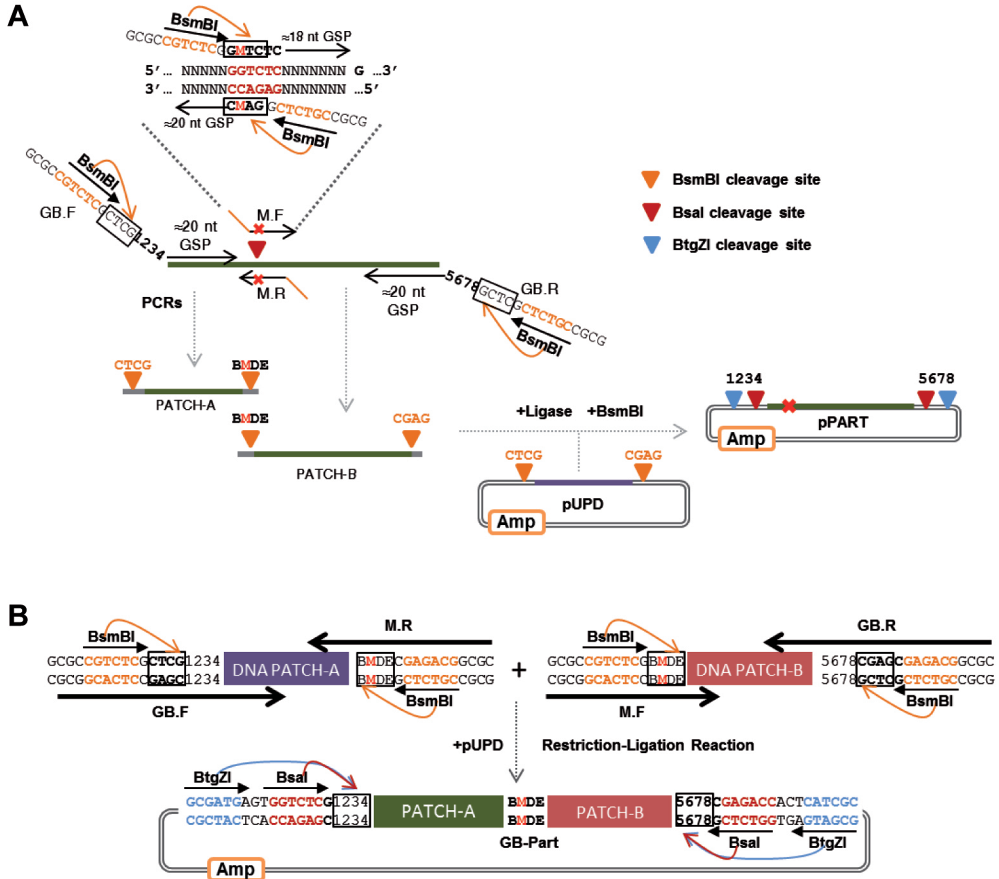
205. Beck, E., G. Ludwig, E.A. Auerswald, B. Reiss, and H. Schaller, Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. Gene, 1982. 19(3): p. 327-36.

206. Ishige, F., M. Takaichi, R. Foster, N.-H. Chua, and K. Oeda, A G-box motif (GCCACGTGCC) tetramer confers high-level constitutive expression in dicot and monocot plants. The Plant Journal, 1999. 18(4): p. 443-448.

207. Lehming, N., J. Sartorius, M. Niemoller, G. Genenger, B. v Wilcken-Bergmann, and B. Muller-Hill, The interaction of the recognition helix of lac repressor with lac operator. EMBO J, 1987. 6(10): p. 3145-53.

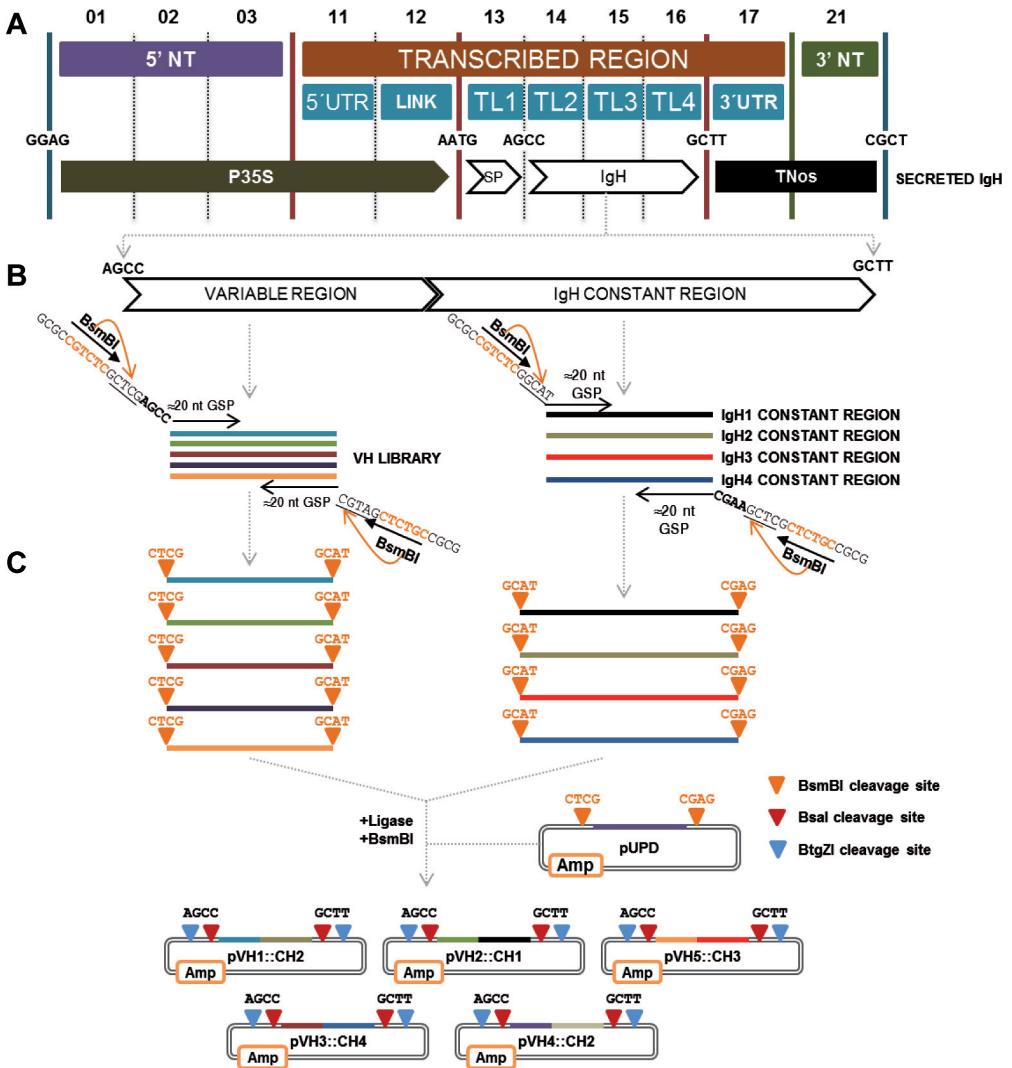
/SF


SUPPLEMENTAL
FIGURES &
TABLES



Supplemental Figure 1. GB Domestication with the removal of internal restriction sites.

(A) Internal Type IIS recognition sites (exemplified here with the GGTCTC *Bsa*I recognition site) are mutagenized during domestication following a standard procedure. In addition to the GB.F and GB.R primers described in Figure 3, two other primers (M.F and M.R) are required here, which incorporate the flanking *BsmB*I overhangs and the single nucleotide change (C>M). Each oligo pair is used to amplify a GBpatch by PCR, and the resulting fragments are assembled together in a BsmBI restriction-ligation reaction into pUPD. The resulting GBpart is free of internal recognition sites and can be released from pUPD using *Bsa*I or *BtgZ*I. **(B)** A detailed view of the *BsmB*I restriction-ligation reaction that follows the amplification of both GBpatches. The *BsmB*I recognition sequences are shown in orange in the DNA sequence and are marked with orange triangles in the schemes; *Bsa*I and *BtgZ*I are labeled in red and blue, respectively; the enzymes cutting sites are boxed. Single nucleotide mismatches are depicted in red in the DNA sequence or are labeled with a cross in the oligo scheme.



Supplemental Figure 2. Combinatorial antibody engineering using the GBpatch assembly level.

A combinatorial fusion between a library of antibody variable regions and a collection of constant immunoglobulin domains can be obtained using the GBpatch functionality. A non-standard boundary (GCAT) is required here to ensure seamless assemblies within the VH-IgH linker region, therefore avoiding the presence of undesired amino-acids in the final antibody sequence. **(A)** Diagram showing the position of the non-standard junction. **(B)** Generation of GBpatches by PCR. **(C)** Combinatorial assembly of GBpatches to generate a new collection of the (14-16) standard GBparts ready for multipartite assembly.

A

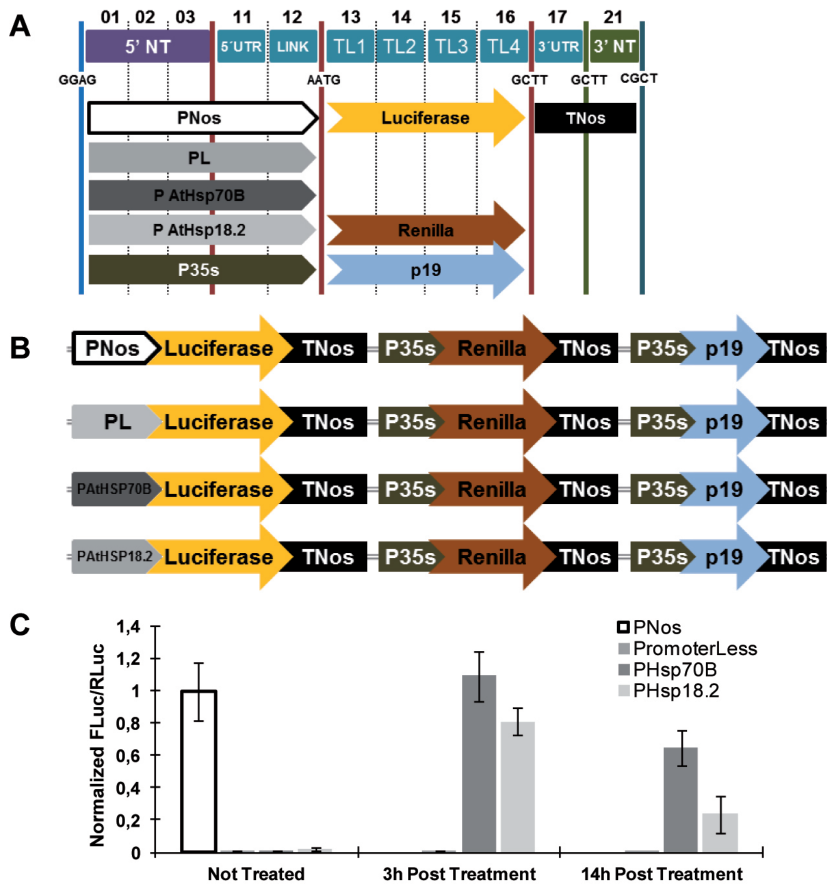
	TAtHSP18.2	TAtUBQ3	TAtAct2	T35s	TSlTctp	TSlMtb	TNos	TAtBch1
P2x35s	40.38 ± 2.08		23.52 ± 3.01	49.86 ± 9.07			15.03 ± 1.44	
P35s	39.68 ± 3.66	22.36 ± 1.64	18.66 ± 2.50	40.41 ± 3.43	32.11 ± 2.03	24.98 ± 2.62	13.57 ± 1.80	11.96 ± 0.87
PAtUbq10	32.27 ± 2.91			18.27 ± 4.28		15.79 ± 3.04	11.92 ± 1.58	
PSlMTb	10.23 ± 1.09	8.11 ± 0.72	4.83 ± 0.54	6.90 ± 0.83	5.71 ± 0.30	6.94 ± 0.95	3.79 ± 0.52	2.82 ± 0.50
PAtUbq3	14.06 ± 1.25	8.00 ± 1.54		5.76 ± 1.27			3.41 ± 0.85	
PSlUbc	4.92 ± 0.49			2.07 ± 0.39			1.89 ± 0.31	*3.82 ± 1.05
PAtAct2	4.36 ± 0.55			2.41 ± 0.10		1.00 ± 0.09	1.80 ± 0.27	
PNosΩ	2.39 ± 0.33		2.25 ± 0.41	1.60 ± 0.23			1.25 ± 0.14	
PNos	2.61 ± 0.54	1.70 ± 0.13	1.60 ± 0.31	1.52 ± 0.23	1.38 ± 0.22	1.12 ± 0.21	1.00 ± 0.16	0.77 ± 0.18
PSlTctp	0.89 ± 0.16			0.94 ± 0.10			0.47 ± 0.01	

B

	TAtHSP18.2	TAtUBQ3	TAtAct2	T35s	TSlTctp	TSlMtb	TNos	TAtBch1
P2x35s	39.20 ± 8.94	25.62 ± 3.14	24.08 ± 5.20	22.92 ± 4.10	20.77 ± 3.86	16.89 ± 3.55	15.03 ± 1.44	11.56 ± 2.92
P35s	35.40 ± 8.70	23.13 ± 3.53	21.75 ± 5.10	20.70 ± 4.15	18.76 ± 3.88	15.26 ± 3.49	13.57 ± 1.80	10.44 ± 2.81
PAtUbq10	31.10 ± 7.64	20.32 ± 3.11	19.11 ± 4.48	18.18 ± 3.65	16.48 ± 3.41	13.40 ± 3.07	11.92 ± 1.58	9.17 ± 2.47
PSlMTb	9.90 ± 2.46	6.47 ± 1.01	6.08 ± 1.44	5.79 ± 1.18	5.24 ± 1.10	4.27 ± 0.99	3.79 ± 0.52	2.92 ± 0.79
PAtUbq3	8.90 ± 2.88	5.82 ± 1.52	5.47 ± 1.72	5.20 ± 1.52	4.72 ± 1.39	3.84 ± 1.19	3.41 ± 0.85	2.62 ± 0.90
PSlUbc	4.92 ± 1.31	3.22 ± 0.59	3.03 ± 0.77	2.88 ± 0.65	2.61 ± 0.60	2.12 ± 0.53	1.89 ± 0.31	1.45 ± 0.42
PAtActUTR	4.69 ± 1.20	3.06 ± 0.52	2.88 ± 0.71	2.74 ± 0.59	2.48 ± 0.55	2.02 ± 0.49	1.80 ± 0.27	1.38 ± 0.39
PNosΩ	3.26 ± 0.76	2.13 ± 0.29	2.00 ± 0.45	1.91 ± 0.36	1.73 ± 0.33	1.41 ± 0.30	1.25 ± 0.14	0.96 ± 0.25
PNos	2.61 ± 0.54	1.70 ± 0.13	1.60 ± 0.31	1.52 ± 0.23	1.38 ± 0.22	1.12 ± 0.21	1.00 ± 0.16	0.77 ± 0.18
PSlTctp	1.23 ± 0.26	0.80 ± 0.07	0.75 ± 0.15	0.72 ± 0.11	0.65 ± 0.11	0.53 ± 0.10	0.47 ± 0.01	0.36 ± 0.09

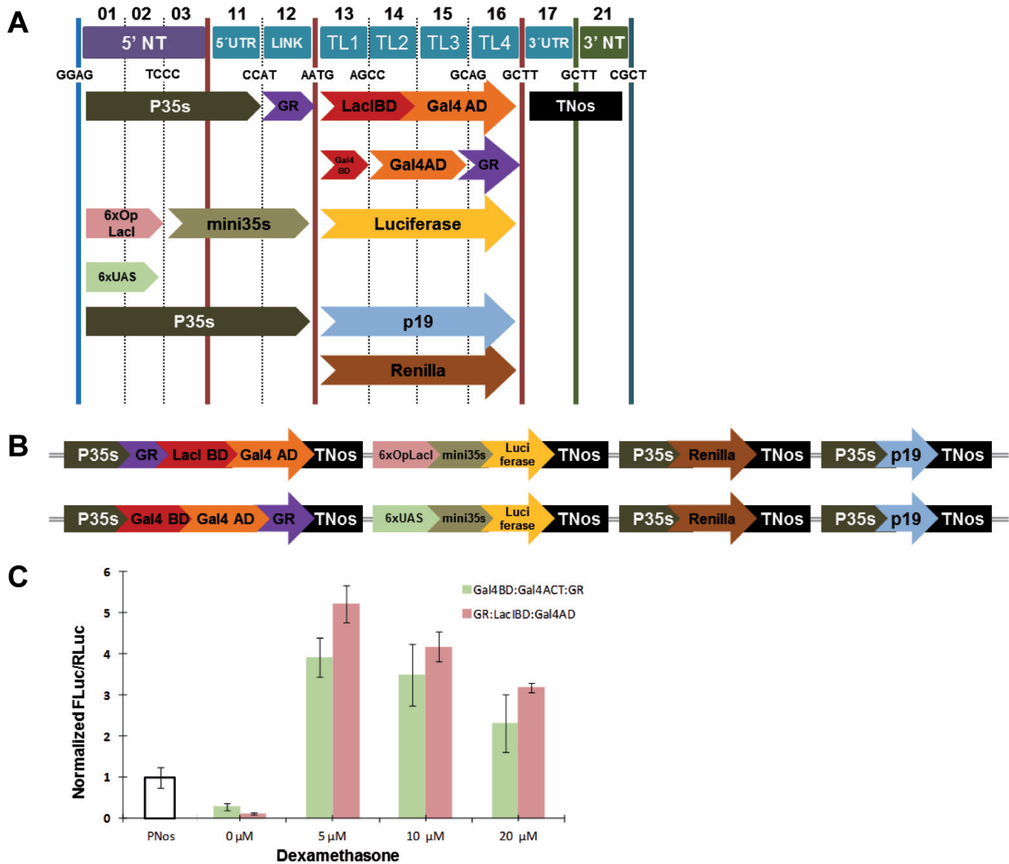
Supplemental Figure 3. Theoretical and experimental transcriptional activity of different promoter/ terminator combinations.

(A) ETA of 62 experimental (01-12)_(17-21) combinations. The SD of at least three replicates is indicated. **(B)** The TTA of all the possible combinations in the grid. The SD propagated from the ETA measures is also indicated. Colors indicate activity ranges from high (purple) to low (yellow).



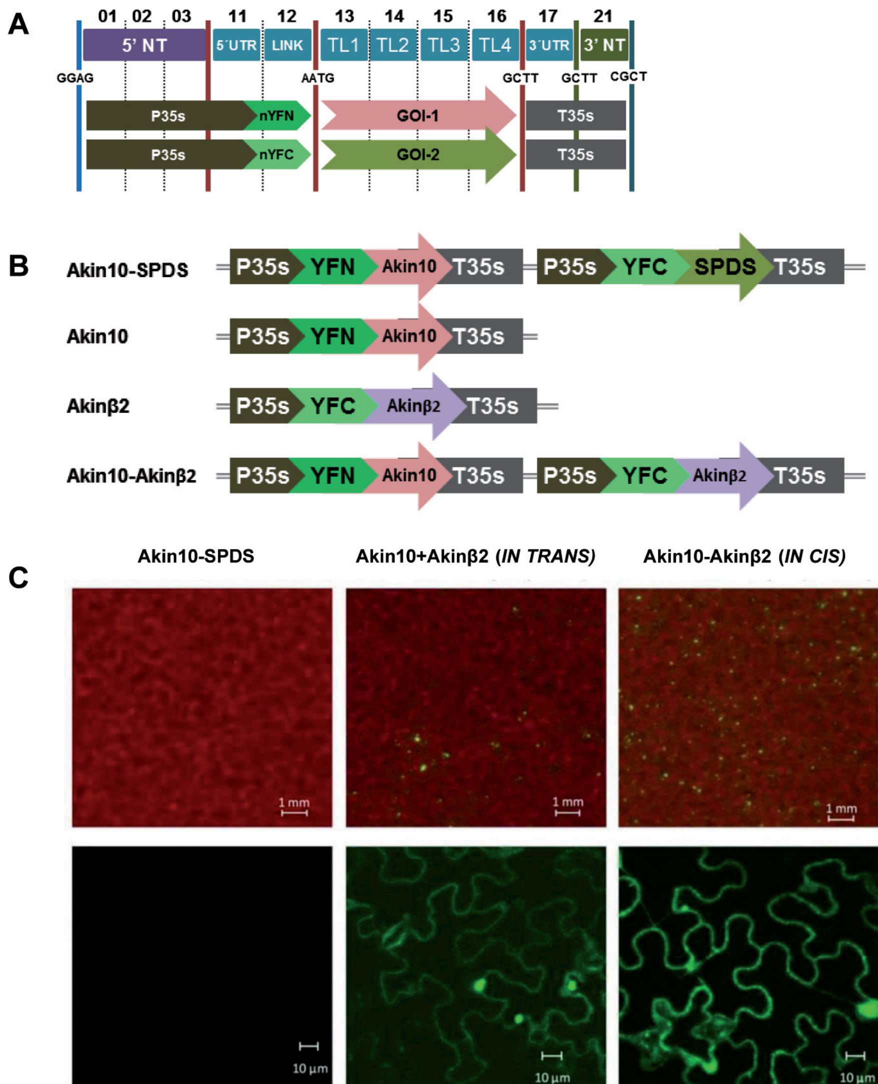
Supplemental Figure 4. Functional characterization of heat-shock promoters.

(A) GBparts used to build the heat shock transcriptional units (TUs). PNos indicates the Nopaline synthase promoter; TNos is the Nopaline synthase terminator; PL indicates a promoter-less construct; PHSP70B and PHSP18.2 are the promoter regions of the *Arabidopsis* AtHSP70B and AtHSP18.2 genes; P35S is the CaMV 35S promoter; P19 is the TBSV silencing suppressor. **(B)** Structure of the heat shock-regulated constructs. Multigene constructs were built by assembling the promoter of interest with the reporter luciferase and the TNos terminator in pDGBα1; the resulting unit was then combined in pDGBΩ1 with the previously assembled units 35S:Renilla:TNos-35S:P19:TNos. **(C)** Effect of heat shock treatment (2 h x 37ºC) on the heat shock promoters (HSPs) Activity. The ratios between HSP-driven Firefly (FLuc) and 35S-driven Renilla (RLuc) luciferase activities were normalized using a Nopaline Synthase construct. A promoter-less construct was also set as a negative control. Error bars represent the SD of at least three replicates.



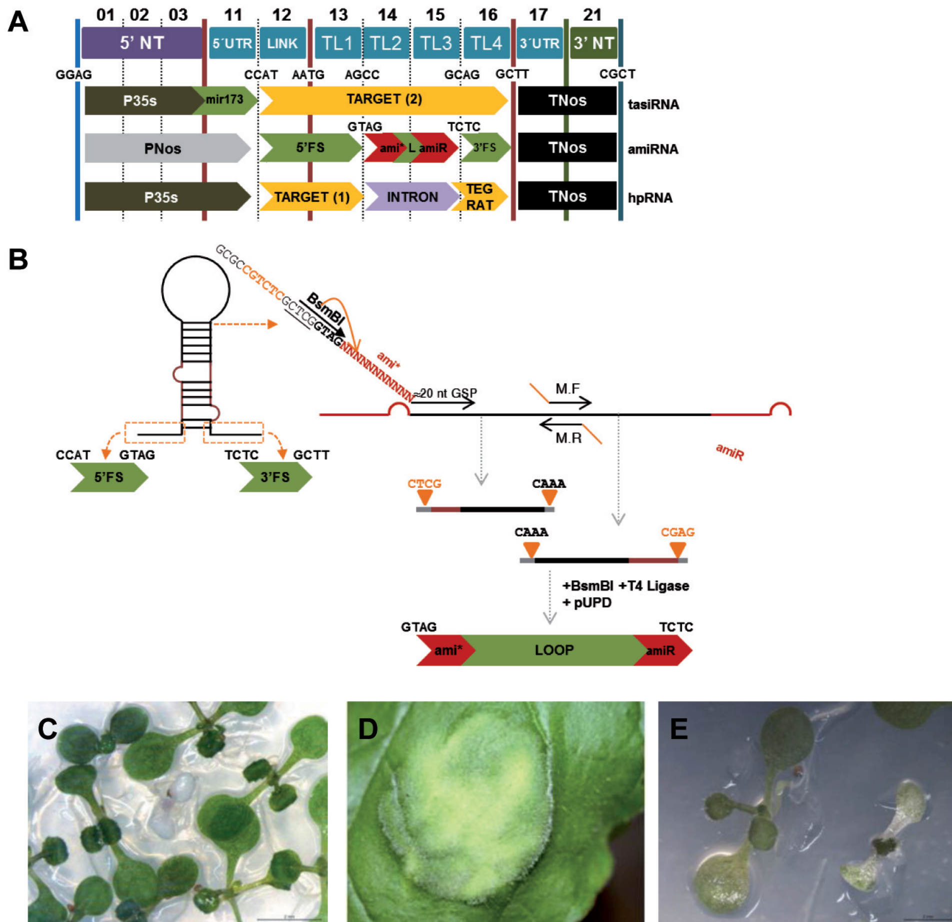
Supplemental Figure 5. Functional characterization of transactivation constructs.

(A) GBparts used to build the dexamethasone-regulated constructs. P35S is the CaMV 35S promoter; GR is the rat glucocorticoid receptor; BD is the binding domain; AD is the activation domain; TNos is the Nopaline synthase terminator; 6xOpLacI are six closely spaced *lac* operators; mini35S is the minimal CaMV 35S promoter; 6xUAS is an upstream activation sequence to which Gal4 binds; P19 is the TBSV silencing suppressor. **(B)** Structure of the dexamethasone regulated constructs. Multigene constructs were built by assembling the chimeric transcription factors into pDGBα1 and the operated promoter with the luciferase gene into pDGBα2. Both TUs were combined into pDGBΩ1. The resulting construct was later combined with the Renilla/P19 reference module into pDGBα1. **(C)** Effect of glucocorticoid treatment on the activity of the operated promoters. Firefly (FLuc) and Renilla (RLuc) Luciferase activities were measured after 24 h of treatment with different concentrations of dexamethasone. Ratios were normalized to the reference construct driven by the Nopaline Synthase Promotor. Error bars represent the SD of at least three replicates.



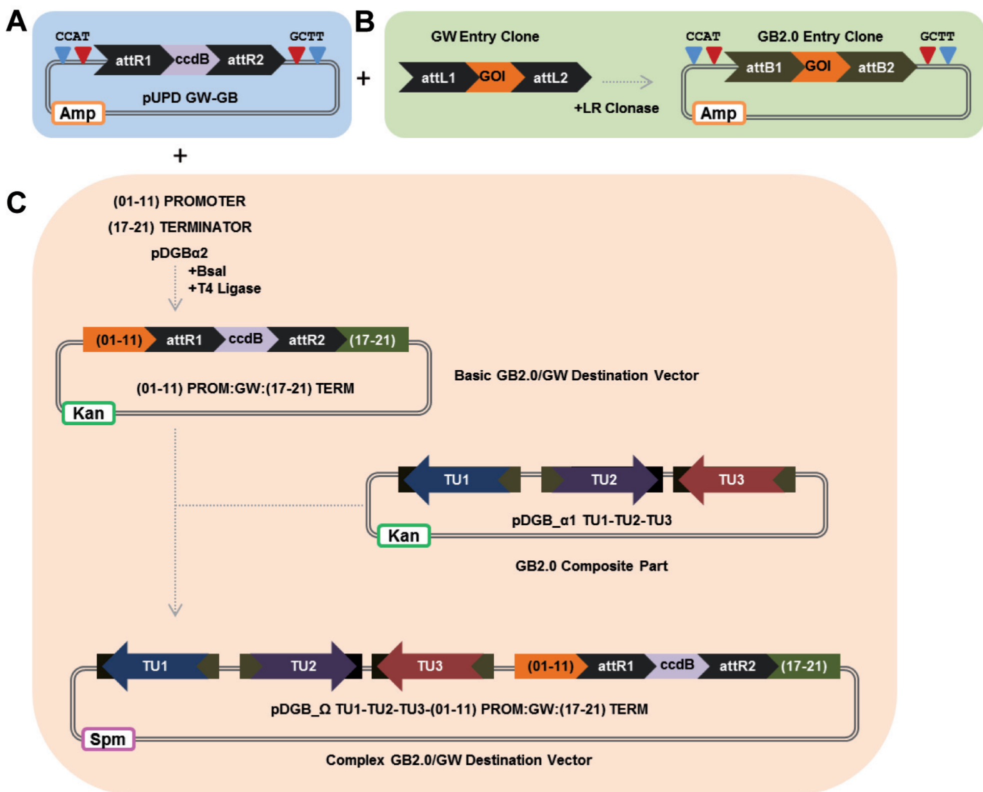
Supplemental Figure 6. Frequently Used Structures (FUS) for the protein-protein interaction analysis.

(A) *Grammar* of the FUS used in Bifluorescent Complementation. **(B)** The constructs used in the evaluation of the linked and unlinked co-transformation analysis. **(C)** The BIFC analysis of two negative interaction partners in a linked co-transformation (Akin10-SPDS), two positive interaction partners co-transformed in trans (Akin10 + Akinβ2) and the same two interaction partners assayed in a linked co-transformation (Akin10-Akinβ2). Upper images were taken with a fluorescence-coupled binocular lens. Lower panels are the confocal micrographies.



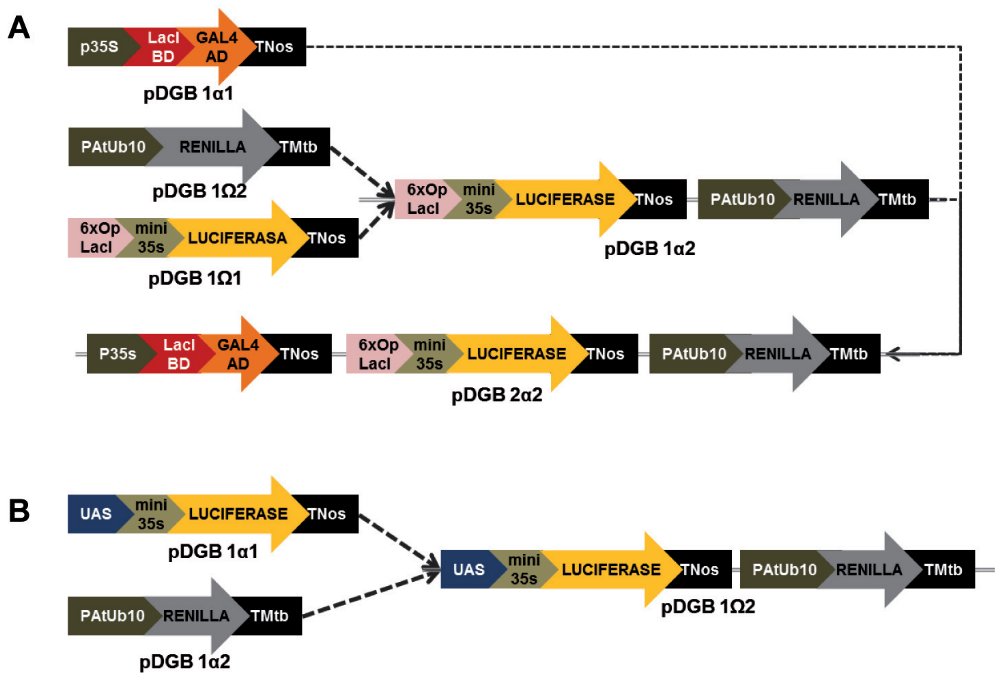
Supplemental Figure 7. The Frequently Used Structures (FUS) for endogenous gene silencing.

(A) *Grammar* of the FUS used in the construction of transcriptional units (TUs) for gene silencing: hpRNA, amiRNA, tasiRNA. **(B)** Construction of the non-standard GB parts for amiRNA. Flanking elements 5'FS and 3'FS are cloned as non-standard (12-13) and (16) GBparts, respectively. The ami* and amiR sequences are included in the primers used to build the custom central GBpart (a dedicated 14-15 GBpart). **(C)** Functional assays of the phytoene deaturase (PDS) tasiRNA constructs in *A. thaliana*. The seeds transformed with the 35S:mir173:PDS:TNos construct were plated in MS; transformed plants exhibited the albino phenotype (indicated with a red arrow). **(D)** The dual construct 35S:mir173:PDS*:TNos-PNos:mir173:TNos transiently transformed into *N. bentamiana* yielded a bleaching effect on the infiltrated area recorded at 7 d.p.i. **(E)** *A. thaliana* seeds transformed with 35S:PDSamiRNA:TNos were plated in MS; transformed plants exhibited a slight albino phenotype (indicated with a red arrow), whereas negative seedlings showed normal growth. P35S-mir173 is the CaMV promoter fused to the mir173 target site; TARGET(1) is a fragment of the gene to be silenced; PNos is the Nopaline synthase promoter and P35S is the CaMV promoter. TARGET(2) and TEGRAT are representative fragments of a target gene in the inverted orientation. 5'FS and 3'FS are the 5' and 3' flanking sequences of the miRNA precursor, respectively; L refers to the loop of the amiRNA structure; amiR and ami* are the complementary target gene sequences.

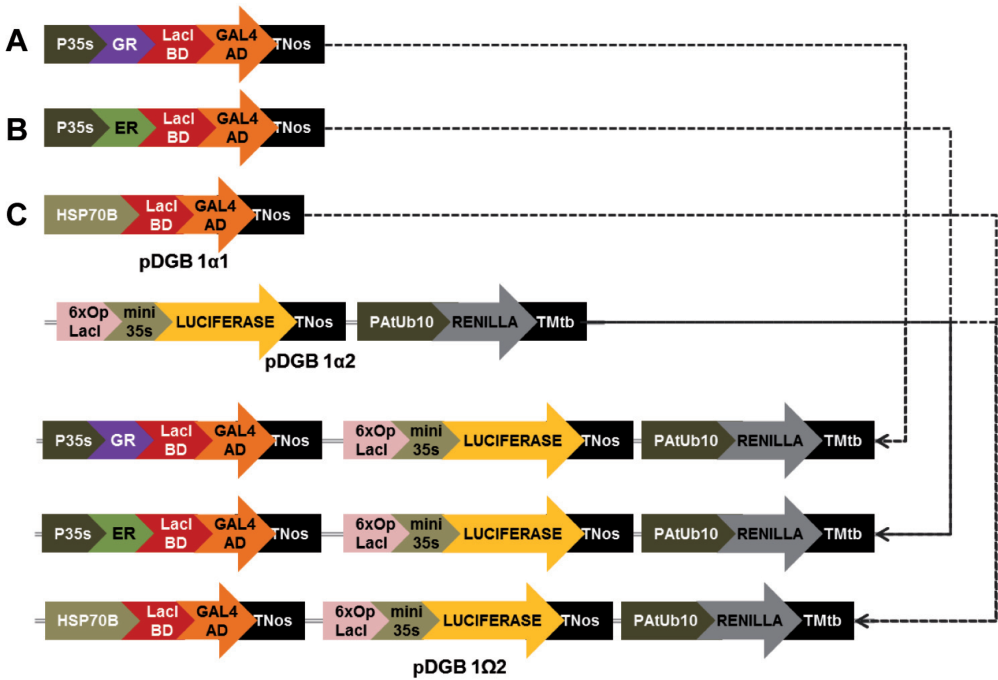


Supplemental Figure 8. Adapting Gateway (GW) technology to GB2.0.

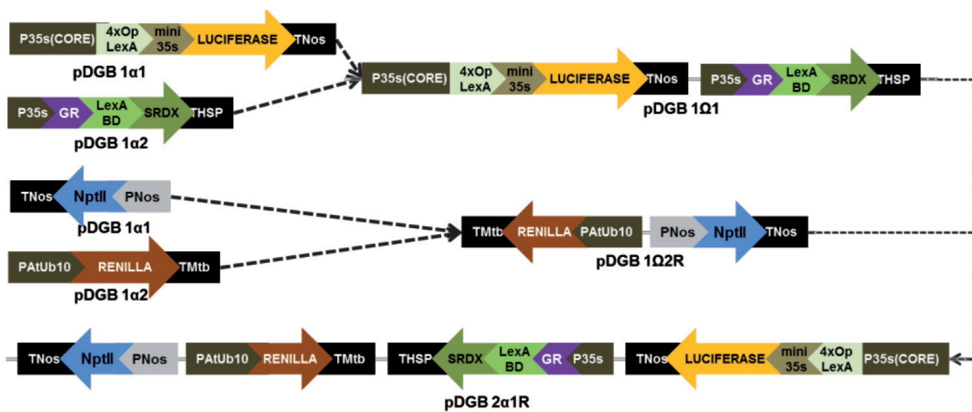
(A) The GW-GB adaptor is a GBSpart (e.g., a (12-16) GBSpart) with a GW cassette flanked by the *artt*R1 and *att*R2 sites. **(B)** The adapter vector can be used as destination plasmids for the GW entry clones, therefore translating GW entry clones into GB2.0 entry clones. **(C)** Alternatively, the GW-GB adapter can be used as an ordinary GBSpart to build increasingly complex GW destination vectors, where GW entry clones can be LR-cloned.



Supplemental Figure 9. GoldenBraid Building pathways for the (A) ⊤LACI::GAL4AD and (B) ⊥UAS gates.

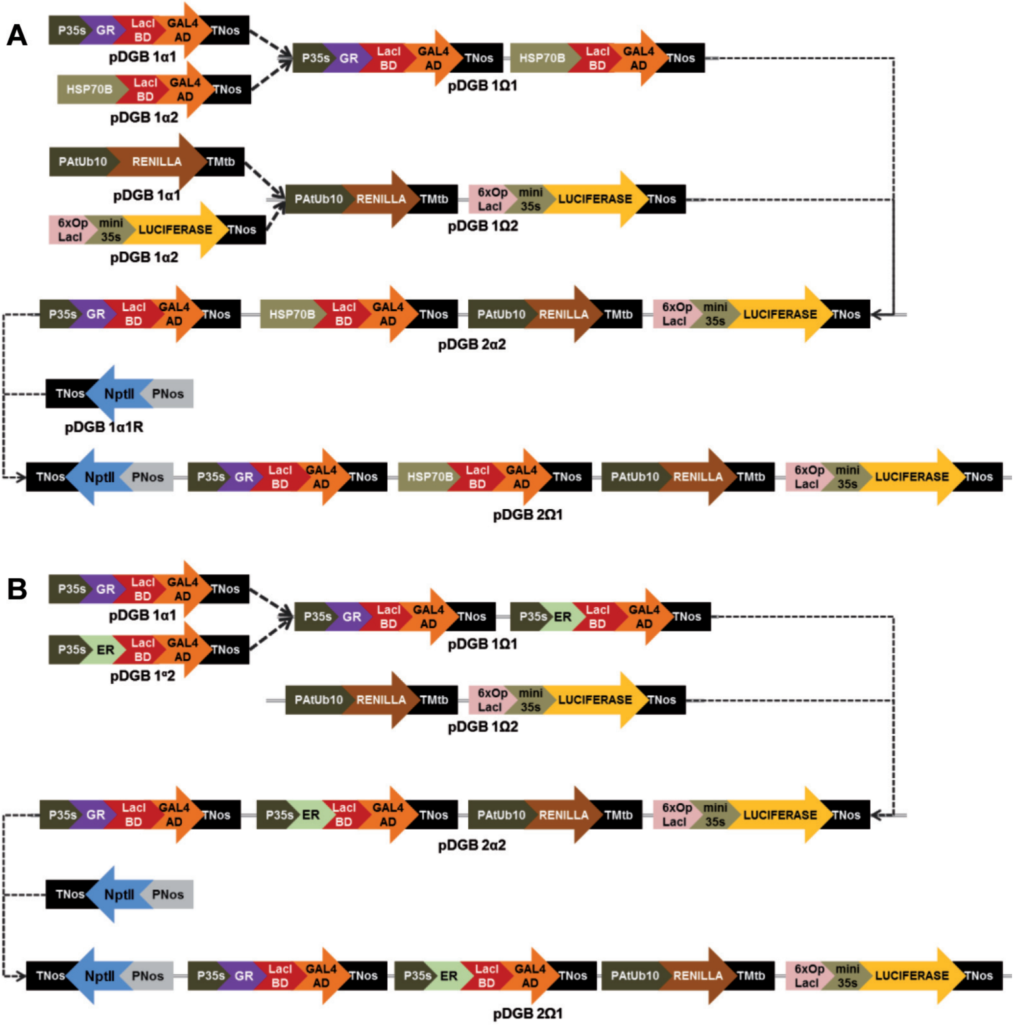


Supplemental Figure 10. GoldenBraid Building pathways for the (A) *P.GR* (B) *P.ER* and (C) *P.HS* logic gates.

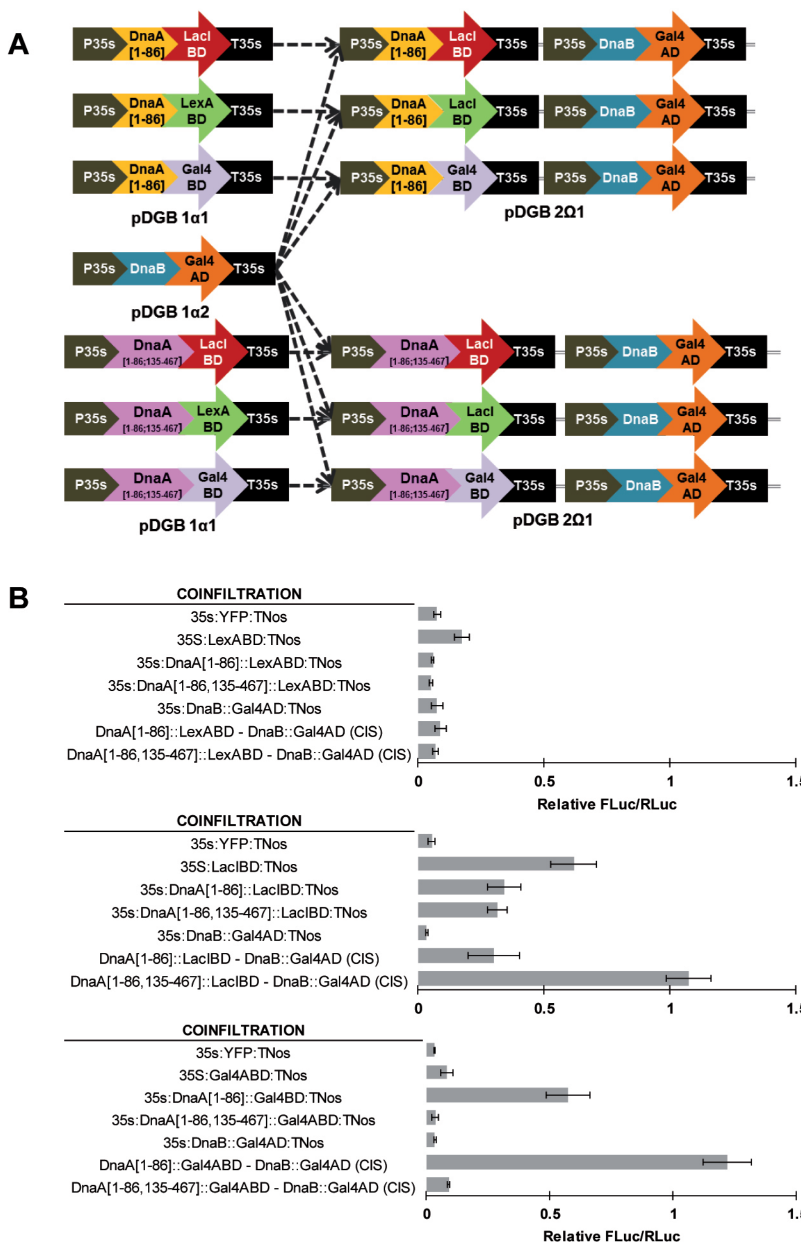


Supplemental Figure 11. GoldenBraid Building pathway for the ¬GR.LexA::SRDX version of the logic gate.

The rest of operations were built following the same steps.

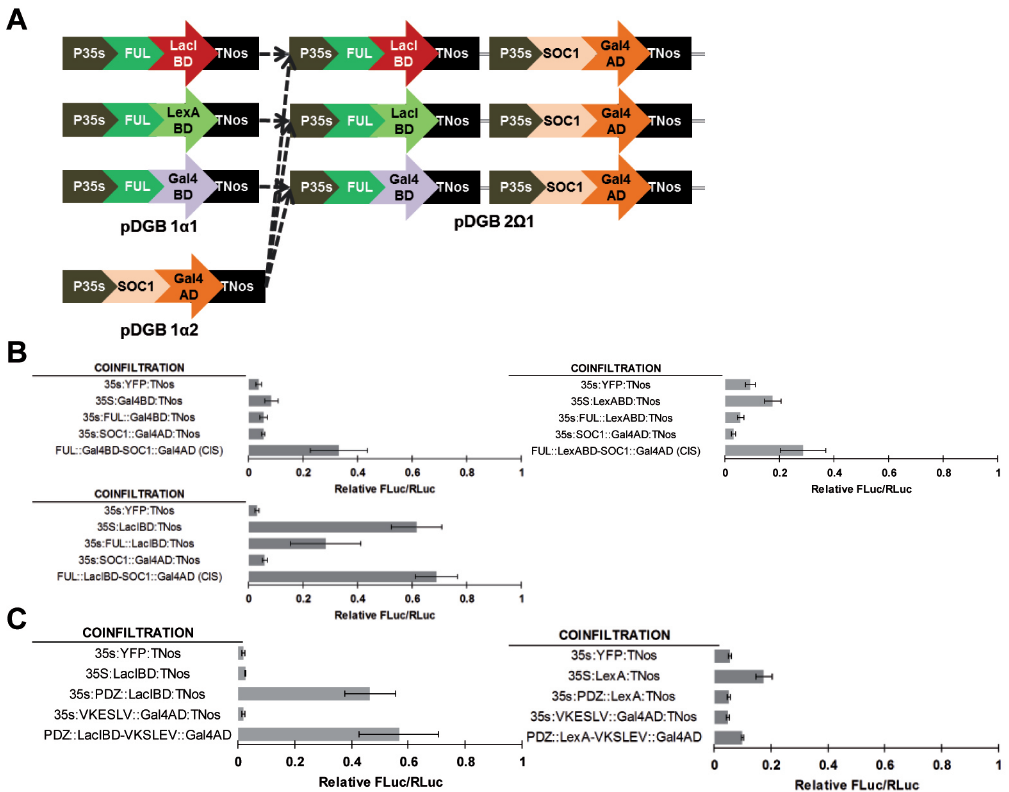


Supplemental Figure 12. GoldenBraid Building pathways for the (A) GRvHS (B) GRvER gates



Supplemental Figure 13. GoldenBraid Building pathways and results for the AND tests constructs (I).

(A) Building steps for the Interactors DnaA [1-86] and DnaA [1-86, 135-467] vs DnaB **(B)** Transient expression experiments for the DnaA/DnaB constructs, coinfiltrated with the reporter constructs. The ratios between Firefly (FLuc) and Renilla (RLuc) luciferase activities were normalized using a Nopaline synthase construct (except on the stable plants experiments, which was referred to the control without inducers). A promoter-less construct was also set as negative control. Error bars represent the SD of at least three replicates.



Supplemental Figure 14. GoldenBraid Building pathways and results for the AND tests constructs (II).

(A) Building steps for the FUL/SOC pair of interactors **(B)** Transient expression experiments for the FUL/SOC set of constructs, coinfiltrated with the corresponding reporter constructs. **(C)** Transient expression experiments for the PDZ/VKESLV constructs, fused to LexABD and LacIBD and coinfiltrated with the correspondent reporter constructs. The ratios between Firefly (FLuc) and Renilla (RLuc) luciferase activities were normalized using a Nopaline synthase construct (except on the stable plants experiments, which was referred to the control without inducers). A promoter-less construct was also set as negative control. Error bars represent the SD of at least three replicates.

Supplemental Table 1. GBparts used in Chapter 2.

GB ID	GBPart	Cathegory	Description	Reference
30	pP35s	01-02-03-11-12	CaMV 35s promoter	**[189]**
32	pmini35s	03-11-12	Minimal CaMV 35s promoter (residues -60,+1)	**[189]**
35	TAtHSP	17-21	*Arabidopsis thaliana* Heat Shock Terminator	**[141]**
36	T35s	17-21	CaMV 35s Terminator	**[190]**
39	pGR	16	Rat glucocorticoid receptor	**[191]**
55	pGW-GB Adaptor	12-13-14-15-16	Gateway *att*R1-ccdB-*att*R2 cassette	**[192]**
56	pRenilla	13-14-15-16	*Renilla reniformis* Luciferase gene	**[193]**
57	pGal4AD	14-15	Transcription activation domain II (residues 768-881) of *Saccharomyces Cerevisiae* Gal4	**[194]**
66	p3'FS	16B	3' Flanking sequence of miRNA precursor, from pRS300	**[195]**
67	pP19	13-14-15-16	TBSV Silcencing Suppressor	**[196]**
68	p35s:YFN	01-02-03-11-12B	CaMV35s Fused to the NT portion of the Yellow fluorescent Protein	**[135]**
69	p35s:YFC	01-02-03-11-12B	CaMV35s Fused to the CT portion of the Yellow fluorescent Protein	**[135]**
72	pPNos	01-02-03-11-12	Nopaline synthase promoter	**[197]**
73	pPNos	01-02-03-11	Nopaline synthase promoter	**[197]**
74	p5'FS	12-13B	5' Flanking sequence of miRNA precursor, from pRS300	**[195]**
80	pPSlMTb	01-02-03-11-12	*Solanum lycopersicum* Metalothionein-like protein promoter	This work
81	pTSlTctp	17-21	*Solanum lycopersicum* Translationally-controlled tumor protein homolog terminator	This work
83	pPSlTctp	01-02-03-11-12	*Solanum lycopersicum* Translationally-controlled tumor protein homolog promoter	This work
94	p6xLacIOp	01-02	Six closely spaced *lac* operators	**[198]**
96	pLuciferase	13-14-15-16	*Photinus pyralis* Luciferase gene	**[193]**
99	pINTRONhp	14-15	*Solanum lycopersicum* Intron SGN-U324070	**[53]**
103	p35s:mir173	01-02-03-11	CaMV 35s promoter fused to the mir173 precursor	**[136]**
142	pTSlMtb	17-21	*Solanum lycopersicum* Metalothionein-like protein terminator	This work
145	pPSlUbc	01-02-03-11-12	*Solanum lycopersicum* Ubiquitin Carrier protein promoter	This work
153	pPAtHsp18.2	01-02-03-11-12	*Arabidopsis thaliana* Heat Shock Promoter 18.2	**[199]**

155	pPAtHsp70b	01-02-03-11-12	*Arabidopsis thaliana* Heat Shock Promoter 70	**[200]**
177	pGal4BD	13	DNA-binding domain (1–147) of *Saccharomyces Cerevisiae* Gal4	**[201]**
179	p6xUAS	01-02	Upstream activation sequence where Gal4 binds to.	**[202]**
180	pami*(PDS):Loop:ami:(PDS)	14B-15B	PDS target sequences	**[137]**
187	pRosea.GOI-TARGET(1)	12-13	*Antirrhinum majus* Rosea1 transcription factor (nucleotides 319-448)	**[59]**
188	pRosea.IOG-TEGRAT	16	*Antirrhinum majus* Rosea1 transcription factor (nucleotides 448-319)	**[59]**
189	pPDS(2)	12-13-14-15-16	*Arabidopsis thaliana* phyotoene desaturase gene cDNA (nucleotides 799-1208)	**[136]**
192	pPAct2	01-02-03	*Arabidopsis thaliana* Actin 2 Promoter	**[89]**
193	pUTRAct2	11-12	*Arabidopsis thaliana* Actin 2 5'UTR	**[89]**
210	pTAtAct2	17-21	*Arabidopsis thaliana* Actin 2 Terminator	This work
218	pTAtBch1	17-21	*Arabidopsis thaliana* Terminator	This work
219	pTAtUbq3	17-21	*Arabidopsis thaliana* AtBch1 Terminator	This work
222	pP2x35s	01-02-03-11-12	Double CaMV 35s promoter	**[203]**
250	pPromoterLess	01-02-03-11-12	*Solanum lycopersicum* Intron SGN-U324070	**[53]**
253	pΩ	11-12	Omega Enhancer Element	**[189]**
272	pAtUbq3	17-21	*Arabidopsis thaliana* Ubiquitin 3 Promoter	This work
273	pAtUbq10	17-21	*Arabidopsis thaliana* Ubiquitin 10 Promoter	**[204]**
465	pLacIBD::Gal4AD	13-14-15-16	Residues 1-330 of the Y17H mutant of *lac* repressor fused to the transcription activation domain II (residues 768-881) of *Saccharomyces cerevisiae*	This work
531	pGR	12	Rat glucocorticoid receptor	**[191]**
552	p35s	01-02-03-11	CaMV 35s promoter	**[189]**
560	pAkin10	13-14-15-16	*Arabidopsis thaliana* transcription factor	**[135]**
561	pAkinβ2	13-14-15-16	*Arabidopsis thaliana* transcription factor	**[135]**
562	pSPDS	13-14-15-16	*Arabidopsis thaliana* spermidine synthase	**[135]**

Supplemental Table 2. GBparts used in Chapter 3.

GB ID	GBPart	Cathegory	Description	Reference
30	pP35s	01-02-03-11-12	CaMV 35s promoter	**[189]**
32	pmini35s	03-11-12	minimal CaMV 35s promoter (residues -60,+1)	**[189]**
34	pPNos:NptII:TNos	01-[…]-21	promoter:Neomycin phosphotransferase II gene:Nopaline synthase terminador	**[205]**
37	pTNos	17-21	Nopaline synthase terminator	**[197]**
56	pRenilla	13-14-15-16	*Renilla reniformis* Luciferase gene	**[193]**
94	p6xLacIOp	01-02	Six closely spaced *lac* operators	**[198]**
96	pLuciferase	13-14-15-16	*Photinus pyralis* Luciferase gene	**[193]**
118	p4xLexAOp	01-02	Four copies of the lexA operator	**[206]**
142	pTSlMtb	17-21	*Solanum lycopersicum* Metalothionein-like protein terminator	**[91]**
155	pPAtHsp70b	01-02-03-11-12	*Arabidopsis thaliana* Heat Skock Promoter 70	**[200]**
179	p6xUAS	01-02	Upstream activation sequence where Gal4 binds to	**[202]**
212	pSOC1	13-14-15	*A.thaliana* SUPRESSOR OF OVEREXPRESSION OF CO transcription factor	**[164]**
224	pFUL	13-14-15	*A.thaliana* FRUITFUL transcription factor	**[163]**
272	pAtUbq3	17-21	*Arabidopsis thaliana* Ubiquitin 3 Promoter	This work
273	pAtUbq10	17-21	*Arabidopsis thaliana* Ubiquitin 10 Promoter	**[204]**
424	pGR::LacIBD::SRDX	13-14-15-16	LacI Binding Domain fused to the LDLERLGFA SRDX domain	This work
531	pGR	12	Rat glucocorticoid receptor	**[191]**
534	p6xLacIOp::mini35s	02-03	Six closely spaced lac operators	**[198]**
535	pP35s(CORE)	01	Enhancer elements of the CaMV 35S promoter	**[160]**
552	p35s	01-02-03-11	CaMV 35s promoter	**[189]**
633	pGal4AD	16	Transcription activation domain II (residues 768-881) of *Saccharomyces Cerevisiae* Gal4	**[194]**
634	pPDZ	13-14-15	Amino acids 87-170 of the mammalian protein α-sintropin	**[162]**
638	pVKESLV::Gal4AD	13-14-15-16	Domain VKESLV fused to Gal4AD	This work
672	p6xUAS::mini35s	02-03	Upstream activation sequence where Gal4 binds to	**[202]**
674	pLacIBD(CDS)	13-14-15-16	Residues 1-330 of the Y17H mutatnt of *lac* repressor	**[207]**
674	pLacIBD	16	Residues 1-330 of the Y17H mutatnt of *lac* repressor	**[207]**
675	pDnaA [1-86,35-467]	13-14-15	Fusion between amino acids 1-86 and 135-467 from the *E. coli* DnaA protein	**[161]**
676	pDnaA [1-86]	13-14-15	Amino acids 1-86 from the *E. coli* DnaA protein	**[161]**

677	pDnaB	13-14-15-16	*E. coli* DnaB protein	**[161]**
678	pER	12	Human Esteroid Receptor	**[157]**
703	pLacIBD::BDR	13-14-15-16	LacI Binding Domain fused to the GNSKTLRLFGVNMEC BDR domain	This work
732	pLexABD::SRDX	13-14-15-16	LexA Binding domain fused to the LDLERLGFA SRDX domain	This work
733	p4xLexAOp::mini35s	02-03	Four copies of the lexA operator sequence	**[206]**
734	pGal4BD::BDR	13-14-15-16	Gal4 Binding domain fused to the GNSKTLRLFGVNMEC BDR domain	This work
735	pLexABD::BDR	13-14-15-16	LexA Binding domain fused to the GNSKTLRLFGVNMEC BDR domain	This work
857	pGal4BD::SRDX	13-14-15-16	Gal4 Binding domain fused to the LDLERLGFA SRDX domain	This work
858	pLacIBD	13-14-15	Residues 1-330 of the Y17H mutatnt of *lac* repressor	**[207]**
859	pGal4BD	16	DNA-binding domain (1–147) of *Saccharomyces Cerevisiae* Gal4	**[201]**
863	pLexABD	16	DNA-binding domain of lexA (residues 1–87)	**[206]**
938	pLexABD(CDS)	13-14-15-16	DNA-binding domain of lexA (residues 1–87)	**[206]**
939	pGal4BD(CDS)	13-14-15-16	DNA-binding domain (1–147) of *Saccharomyces Cerevisiae* Gal4	**[201]**