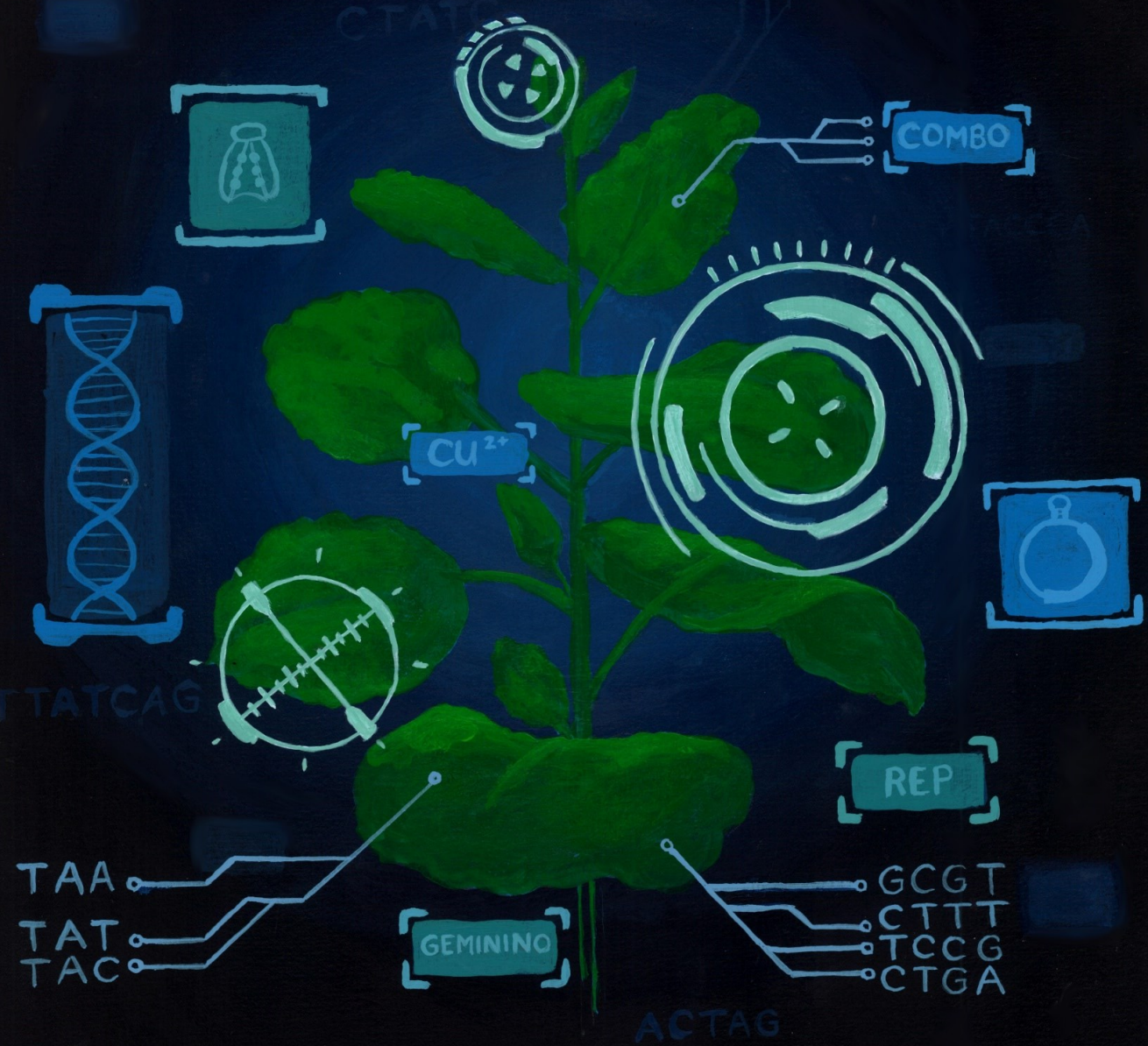
PhD in Biotechnology

December 2024

# Development of a Copper Sensor and Geminivirus-Based Processors for Engineering Synthetic Gene Circuits in Plants

Elena García Pérez




UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Advisors:
Diego Orzáez Calatayud
Marta Vázquez Vilar

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

PhD in Biotechnology

# Development of a Copper Sensor and Geminivirus-Based Processors for Engineering Synthetic Gene Circuits in Plants

Elena García Pérez

Advisors:
Dr. Diego Orzáez Calatayud
Dra. Marta Vázquez Vilar

Valencia, December 2024

El Dr. Diego Orzáez Calatayud, Científico Titular del Consejo Superior de Investigaciones Científicas, perteneciente al Instituto de Biología Molecular y Celular de Plantas (IBMCP, UPV-CSIC) de Valencia, y la Dra. Marta Vázquez Vilar, investigadora del Instituto de Biología Molecular y Celular de Plantas (IBMCP, UPV-CSIC) de Valencia, CERTIFICAN que Elena García Pérez, ha realizado bajo su dirección en el Instituto de Biología Molecular y Celular de Plantas, el trabajo titulado "Development of a Copper Sensor and Geminivirus-Based Processors for Engineering Synthetic Gene Circuits in Plants", y que autoriza su presentación para optar al grado de Doctor en Biotecnología de la Universitat Politècnica de València.

Y para que así conste, firman el presente certificado en Valencia a 17 de Diciembre de 2024.

Dr. Diego Orzáez

Dra. Marta Vázquez Vilar

Esta tesis va dedicada con todo mi corazón a la memoria de mi amigo Ángel Bógalo Gallego.

Ángel y yo vivíamos y soñábamos en paralelo. Compartíamos muchas cosas, pero lo que más ilusión me hacía tener en común con él era la pasión por la ciencia. Éramos los únicos niños que soñaban con un laboratorio cuando aún no sabían ni lo que era una pipeta. Recuerdo las tardes hablando sobre experimentos o sobre cómo hacer un combustible con huesos de aceituna para los coches y los debates sobre si estudiar Química o Biología. ¡Hubiéramos flipado si por aquella época hubiésemos sabido lo que era la Biotecnología!

A los catorce años, Ángel se marchó, pero siempre ha estado presente en mi mente. Su sonrisa, sus ganas y sus chistes siguen en mí con la misma inocencia y ternura de la niñez. Viene a mi cabeza en algunas ocasiones por el laboratorio, cuando necesito fuerza o cuando celebro mis pequeñas victorias, que son suyas también. Pues sé que, si estuviera aquí, estaríamos plantándole cara a la ciencia juntos. Quizás nos llamaríamos para ver qué se le ocurría hacer al otro con ese experimento que no sale. Quizás estaría en la sala sonriéndome mientras defiendo mi tesis. Quizás se habría leído este librito que escribo hoy aquí.

En el instituto, me prometí a mí misma que cumpliría el sueño de ser científica no solo por mí, sino también por él. Y me prometí que, si algún día escribía una tesis, sería por los dos.

Ángel, ya está, amigo; lo hemos conseguido. Esta tesis es tuya también.

# Acknowledgements

***«Aguantamos porque nos ayudan. Sola no se puede ni con la felicidad».***

Quien me conoce sabe que me apasiona mi trabajo. Pero siempre he mantenido una máxima en mi cabeza: *los papers no abrazan*. Las que sí que abrazan, literal o metafóricamente, de una manera que aglutina mis piezas hasta recomponerlas, son las personas que caminan a mi lado. ¡Vaya suerte la mía! Aquí van mis más sinceros y profundos agradecimientos a todas y todos los que me han acompañado en este viaje tan bonito y enriquecedor que ha sido el de la tesis.

En primer lugar, a mis supervisores. Diego, nunca podré agradecerte lo suficiente que confiaras en aquella estudiante de máster desconocida que te llamaba desde Wageningen con la esperanza de que la acogieras como doctoranda en tu laboratorio. Supe que quería hacer la tesis contigo desde el momento que leí el primer artículo de tu laboratorio y jamás me he arrepentido de esa decisión. Gracias por disipar mis *crisis existenciales* con tus razonamientos, tu perspectiva práctica y ese ingenio y grado de detalle que lo mejora todo. Por trasmitirme tu pasión por la investigación y por el apoyo cuando más lo he necesitado, incluso para conseguir que me operaran una rodilla. No he visto persona más resolutiva e imparable que tú; no hay seguros médicos universitarios ni líneas de tranvía que se te resistan. Eres un referente. Marta, gracias por guiarme y por enseñarme con paciencia y cariño a desenvolverme en el laboratorio. Por estar siempre ahí, con tu escucha atenta y tu disposición incondicional para ayudarme a comprender y resolver mis dudas y preocupaciones. Me llevo grandes lecciones de tu manera de trabajar: tu enfoque realista, entusiasta y analítico ha sido un ejemplo constante para mí.

Continuamos con el resto de (ex)tripulantes del 2.09/2.10. Asier, gracias por ser la primera persona en integrarme en el grupo con ese don social tan encantador que tienes. Gracias por cuidarme como si fueras un hermano mayor y haberme escuchado y apoyado siempre. Es un bálsamo hablar contigo y te adoro. Sara ¡SELMA!, fuiste pilar fundamental de mi alegría en el laboratorio. Hay gente con la que simplemente tienes química y todo fluye, y eso me pasó contigo desde el minuto uno. Gracias por rebautizarme, por tus abrazos recomponedores y el millón de risas que hemos compartido. Has dejado un buen legado en mi persona, amiga. Marta Martita, la chica de mis ojos y de mi corazón. Verte todas las mañanas en el labo ha sido siempre motivo suficiente para sonreír. Gracias por ser incondicional, por todas las charlas, reflexiones, ánimos, risas, canciones... Tu capacidad de hacerme sentir tan a gusto, comprendida y querida ha sido un regalo en este camino. No puedo estar más agradecida de llevarme de aquí una amiga como tú. Asun, gracias por estar siempre dispuesta a darme consejo sobre mis tímidos anticuerpos y por darme paz mental con tus cálculos, experiencias y sugerencias en momentos nebulosos. Ha sido un verdadero placer sentarme a tu lado en el 2.09; eres una institución y una fuente de calma y sensatez. Bea, gracias por aportarme perspectiva y ánimo en los inicios de la tesis. Tener compañeros con un corazón tan grande como el tuyo es una suerte. Silvia Gianoglio, gracias por los ratitos de filosofía, libros, conciertos y desahogos que tanto bien me han hecho. Dani, gracias por llenar de color la etapa final de mi tesis. Eres una de las personas más divertidas, buenas y encantadoras del planeta, le alegras el día a cualquiera y el tiempo contigo —ya sea pipeteando, corriendo bajo la lluvia o haciendo planes de intensitos— siempre es mucho mejor. Y si la vida fuera realmente una lenteja, no dudaría en compartirla contigo, porque, ya sabes, *la vida está hecha para nosotros*. Víctor, gracias por confiar en mí para supervisar tu TFM y hacer el camino del *geminino* mucho más entretenido y emocionante. Hemos crecido y aprendido un montón juntos y ha sido un auténtico placer trabajar con alguien tan entusiasta, apasionado y buen compañero como tú. Xenja, thank you so much for being such a lovely student and for helping me throughout the challenging journey of regulating rebel Reps. Supervising your internship's work was an absolute pleasure; your dedication and excitement for science made every moment rewarding. On top of that, your incredible talent in design and illustration has left a special mark on this thesis with the beautiful cover you created. I will always treasure it. Jose, the goat,

gracias por devolverme la fe en los *fifes*, un mundo mejor es posible con chicos como tú. Ha sido un placer verte crecer con las qPCRs, seguir de cerca tu entusiasmo por la ciencia, oírte hablar de Murcia y compartir música, fútbol y charlas con un sol de persona como tú. Silvia Presa, gracias por ser la capitana firme entre las aguas turbulentas de nuestras poyatas. Sin ti, nos hundiríamos todos; tu ayuda siempre es imprescindible. Sophie, gracias por los intentos de clases de francés, los consejos y por alegrarme las mañanas con chocolatitos o con las visitas de mi futura científica favorita: la pequeña Alys. Gracias al resto de compañeros del laboratorio: Rubén, Nuccio, Camilo, Elena Moreno, Antonio Granell, Maria, Jordi, Joan, Clara, Víctor, Pietro, Dorotea, Ave, Ángel y Mavi, con los que he compartido laboratorio, cafés, comidas, ideas y, por supuesto, muchas risas. Mención especial para la última incorporación, Javi; estoy segura de que serás un gran biotecnólogo y harás unos tomates chulísimos. Mis pipetas se quedan en buenas manos contigo. En definitiva, millones de gracias a unos grandes compañeros de trabajo, a quienes solo puedo desear lo mejor en sus vidas y carreras. Ha sido un privilegio compartir esta etapa con vosotros. Fuera de nuestro laboratorio, en el IBMCP también agradezco a gente estupenda como Marisol, por su ayuda con microscopía y a Aure por enseñarme a remar entre las entretenidas mareas de comandos en Linux.

I would also like to extend a great gratitude to the entire GeminiTeam for welcoming me so warmly during my doctoral stay in Tübingen and for everything I learnt of geminiviruses from them. To Rosa Lozano, a brilliant and inspiring scientist and a wonderful person who made me feel like part of her team even before I arrived in Germany. To Laura Medina, for teaching me so much about microscopy, for the interesting and funny discussions, and for always including me in the plans both inside and outside the lab. To Delphine, for the endless brainstormings; sharing ideas with such enthusiasm was truly delightful. To Wei Hua, for the Chinese lessons, the delicious dishes, and all the laughter we shared. To Tan, Chaonan, Pan, Yu, Zhihao, Julio, Betina and, in general, all the people in the lab for the support and for insightful feedback. My time in Tübingen will always hold a very special place in my heart.

Fuera del ámbito investigador, durante estos años de tesis he disfrutado mi vida junto a personas maravillosas, a las que admiro, me han hecho sentir inmensamente feliz y se han convertido en mi familia valenciana. Ángela, gracias por ser mi mujer de verde, ¡mi superheroína! Por estar siempre preparada para escuchar, acompañar y cuidarme. Por ser chispa y cortafuegos a la vez. Qué pasada y qué divertida es la vida a tu lado, amiga. Iría contigo al fin del mundo. María López, mi pepito grillo. Gracias por estar siempre ahí, con tu apoyo, tu risa, cariño, cervezas y disposición a decirme las palabras que necesito. Eres hogar. Tere, gracias por llenarme de paz y alegría en cada abrazo, por la infinidad de (des)conexiones y llenar mi vida en Valencia de luz. Espero que nunca dejes de alumbrarme. Merc, gracias por hacerme sentir tan acogida y valorada desde el primer momento. Tienes un corazón enorme. Además, tú fuiste quien marcó el inicio del cronómetro para la escritura de esta tesis, ¡así que tienes un papel importante en este libro! Cris, gracias por abrirme la mente con nuestras conversaciones y por las recomendaciones literarias, cinematográficas y teatrales; eres cultura pura y me siento muy afortunada de haber aprendido tanto de una persona tan guay como tú. Sabela, gracias por romper esquemas, por la galleguización, por todo el cariño y por convertirte en experta en cultivo hidropónico conmigo. Siempre llevaré algo de tinta azul y blanca en el corazoncito. Maca, gracias por levantar torbellinos de emociones y pensamientos, por las reflexiones acompañadas de mandarinas y por abrir tantas ventanas. Aida, Rocío, Mario, Pablo, Alba y Óscar, gracias por ser una gran vía de escape durante el proceso de escritura, por hacerme sentir *más joven* y por enseñarme el lado más alternativo de Valencia. Estoy segura de que formaréis una comuna increíble. A mis compañeras del club de lectura de la Rossa por hacerme ver más allá, conectar con otras ideas y no dejar caer mi pasión por la lectura. A mi equipo de fútbol sala Samarucs, gracias por devolverme la motivación tras mi lesión y por traer de vuelta una de las cosas que más feliz me hace en la vida: el fútbol. Ha sido súper divertido ser parte de este equipo y ha merecido la pena todos y cada uno de los entrenamientos y terceros tiempos con vosotres. Sois geniales.

Por otro lado, quiero agradecer al pueblo valenciano en general y a mi querido barrio Benimaclet en particular por ser ejemplo de solidaridad y humanidad. Defender la tierra y cuidar de los vecinos son valores imprescindibles que me llevo de mi paso por Valencia. Ocurran incendios, inundaciones o procesos de gentrificación, todos nos levantamos a una, arrimamos el hombro y luchamos por salir adelante. Gràcies por hacerme sentir una más, siempre llevaré conmigo este orgullo de barrio obrero.

No quiero dejar sin homenajear a los grandes amigos que llevo en la mochila y que no me abandonan por mucho tiempo que pase. Mis amigas de Infantes. Sois refugio desde que era una niña y me seguís comprendiendo y aconsejando a pesar de lo mucho que se dispersan nuestros caminos. Me ancláis a la realidad y me recordáis quién soy. Os llevo enganchadas en el corazón siempre. En especial, quiero agradecer a Sara por sus llamadas sorpresa, que tanta alegría me traen siempre, y por esa fuerza que me transmite con cada palabra y abrazo. Siempre conmigo, querida Monchi. Mi clan de biotecs: Eva, Zule, Carmen, Nico y Patri. Con vosotros empecé la aventura científica y me llena de alegría saber que, por muy lejos que estemos, seguimos estando juntos. Eva, gracias por tus visitas a Valencia que tanto me han cargado las pilas y por llenarme la cocina de alegría con tus plantitas; ya sabes que eres de mis personas favoritas de este mundo. Y, como no, no me puedo olvidar de mis sestras: Lara y Patri. Si hoy soy biotecnóloga es en buena parte gracias a vuestro apoyo. Habéis jugado un papel esencial en mi vida cuando más lo he necesitado y me habéis empujado a ser libre; con vosotras empezó todo.

Y, por último, mi familia. Mamá, papá, sois mi bandera. Os he dicho millones de veces que os quiero muchísimo y que me siento orgullosísima de vosotros y agradecida por todo lo que habéis hecho por mí. Sois honradez, lealtad, amor y un sinfín de valores que os hacen personas admirables. Desde pequeñita, me habéis transmitido la importancia del trabajo duro, desde la más profunda humildad y conciencia social. Con pocos recursos, me lo habéis dado todo. Gracias por todos los sacrificios que habéis hecho para que yo pudiese tener el futuro que soñaba. Mamá, gracias por el cariño constante y tener el oído siempre preparado para escucharme. Eres mi mayor confidente y, de verdad, que admiro la capacidad que tienes para soportar todas las chapas que te suelto cuando me pongo a reflexionar sobre la vida en voz alta contigo. ¡Qué santa paciencia la tuya! Gracias también por cuidarme tantísimo y hacerme morir de risa cuando quería llorar. Gracias por enseñarme el significado más puro y bonito del querer. Papá, si he llegado hasta aquí, es gracias a que no dejaste que me rindiera nunca, porque tú me has enseñado que el único camino que tenemos en la vida es hacia delante. Una sola palabra tuya es suficiente para recomponerme y hacerme sonreír. Gracias también por ser el mejor modelo de dispensador de cobre del mundo y mi mejor compañero en la poda de olivos y de tabacos. Eres sencillamente genial. Gracias a mi hermano David, quien en mi primera crisis de frustración científica me llamó para decirme que ‘es normal que las cosas no me salieran a la primera; si fuese tan fácil, todo el mundo sería científico y todo estaría descubierto’. Gracias por restarle hierro a mis pequeños dramas. A mi abuela Dolores, por llenarme de amor solo con una mirada y un montón de besos. A mi Abuela María, porque un solo apretón de manos tuyo es suficiente para querer comerse el mundo. Eres la voz de la experiencia, la mejor poetisa improvisada que conozco y mi referente de fuerza. La primera vez que salí del pueblo, me dijiste: 'que no se te olvide nunca que una se llama María aquí y al otro lao' y, desde entonces, avanzo con la conciencia tranquila, porque sé que quien me quiere, lo hace por lo que soy y no por lo que aparento.

Lo mejor que me ha dado la vida son las personas con las que la he compartido. Qué plena y feliz me hace saberme rodeada de personas tan fantásticas. Gracias y gracias a todos los que estáis ahí y me habéis ayudado tanto a crecer profesional y personalmente durante estos años. Sin vosotros, no sería quien soy hoy.

---

Esta tesis doctoral ha sido financiada mediante la Subvención para la contratación de personal investigador de carácter predoctoral (ACIF/2020/309) de la Generalitat Valenciana.

*It is your ability to discern facts that makes you an individual,*
*and our collective trust in common knowledge that makes us a society.*
*The individual who investigates is also the citizen who builds.*

Timothy Snyder

# Table of contents

# Abstract

# ABSTRACT

Plant Synthetic Biology is a rapidly advancing field with significant potential to enhance plant-based biomanufacturing. Plants can serve as sustainable and cost-effective biofactories for producing valuable compounds, yet sustained production often leads to issues like toxicity for the plant itself and low yield. A major goal in plant biomanufacturing is to enable precise and inducible control over gene expression through the use of synthetic gene circuits containing modular sensor and processor elements. This thesis focuses on the development of a copper sensor and geminivirus-based processors, designed as modular components for synthetic gene circuits in *Nicotiana benthamiana*, a model plant with strong potential for transient and stable gene transformation.

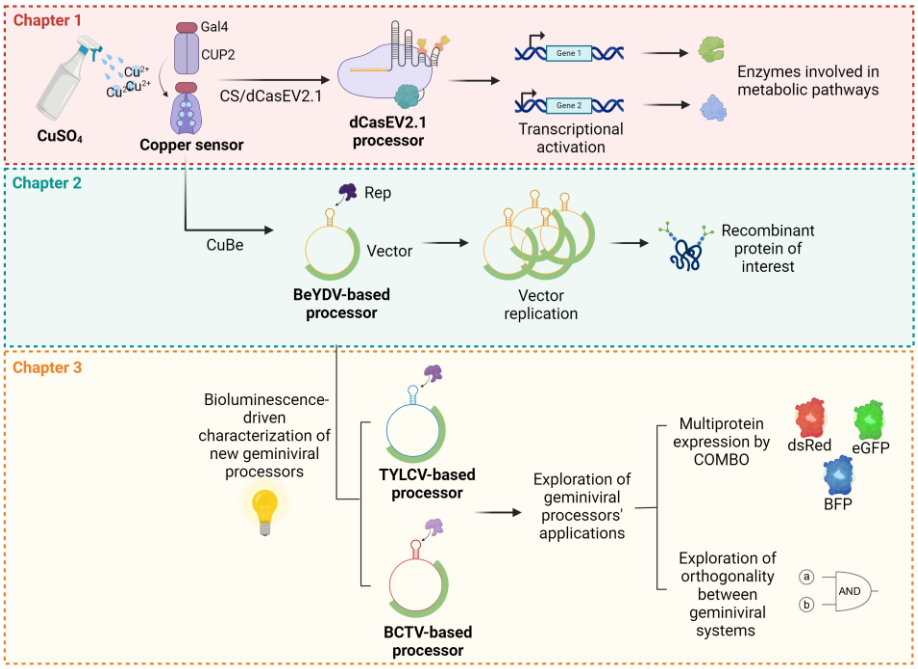
The first chapter of this thesis presents a copper-inducible gene expression system designed to achieve precise, copper sulfate ($CuSO_4$)-dependent regulation of transcriptional activation. This system comprises a sensor element consisting of the copper-binding transcription factor CUP2 fused to the Gal4 activation domain, and a synthetic promoter containing copper-binding sites (CBS) positioned upstream of a minimal promoter. The copper sensor was successfully applied to control the activation of a CRISPR/dCas9-based programmable transcriptional activator (dCasEV2.1) in *N. benthamiana*. The copper sensor, combined with the dCasEV2.1 processor (CS/dCasEV2.1), conferred robust control of gene expression, with minimal basal activity and strong activation in response to copper sulfate, providing a reliable tool for fine-tuned induction of endogenous genes. This system offers significant potential for targeted gene regulation in Plant Synthetic Biology, especially for applications requiring on-demand activation with minimal background expression.

The second chapter focuses on the development of a copper-regulated gene amplification circuit called CuBe, which takes its name from the chemical symbol for copper ($Cu^{2+}$) and the initials of the bean yellow dwarf virus (BeYDV). In this system, the copper sensor is combined with a replicative vector and the associated replication proteins (Rep/RepA), both derived from BeYDV. This integration enables copper-inducible activation of the replicon, allowing for controlled expression of recombinant proteins in *N. benthamiana*. Designed to achieve low basal expression and high inducibility, CuBe facilitates controlled expression of recombinant proteins with pharmaceutical relevance, such as antibodies. Stable transgenic lines carrying CuBe were generated, confirming its long-term functionality and robustness with alternative copper application methods, including post-harvest and hydroponic. The CuBe system highlights the potential of copper-inducible circuits for scalable Molecular Farming applications in controlled conditions.

The third chapter expands on geminivirus-based systems by investigating and optimizing various geminiviral vectors (derived from BeYDV, TYLCV, and BCTV) for plant biotechnology applications. Using a self-sustained bioluminescence-based reporter system, this study characterizes the dynamics and basal expression levels of different vector configuration, comparing their efficacy for the transient expression of a recombinant enzyme linked to

bioluminescence. These analyses provided insights into the strengths and limitations of each vector for specific applications, such as multi-gene expression, thereby enhancing the geminiviral toolbox available for Plant Synthetic Biology.

In conclusion, this thesis contributes novel tools for Plant Synthetic Biology, with significant implications for Molecular Farming and plant-based biocomputing. By integrating inducible systems with optimized vector design, this research provides a foundation for future advancements in synthetic gene circuits tailored for plant biotechnology applications that demand tightly regulated and customizable gene expression.



# RESUMEN

La Biología Sintética de Plantas es un campo en rápida expansión, con un gran potencial para potenciar la producción de compuestos en plantas. Las plantas pueden actuar como biofactorías sostenibles y económicas para la producción de moléculas de alto valor; sin embargo, la producción continua a menudo genera problemas de toxicidad para la propia planta y bajos rendimientos. Un objetivo clave en este ámbito es lograr un control preciso e inducible sobre la expresión génica mediante circuitos genéticos sintéticos que incluyen sensores y procesadores modulares. Esta tesis se centra en el desarrollo de un sensor de cobre y procesadores basados en geminivirus, diseñados como componentes modulares para circuitos sintéticos en *Nicotiana benthamiana*, una planta modelo con gran potencial para transformaciones génicas tanto transitorias como estables.

El primer capítulo de esta tesis presenta un sistema de expresión génica inducible por cobre, diseñado para permitir una regulación precisa dependiente de sulfato de cobre ($CuSO_4$). Este sistema incluye un elemento sensor compuesto por un factor de transcripción que se une al cobre, CUP2, fusionado al dominio de activación Gal4 y un promotor sintético con sitios de unión al cobre (CBS) ubicado antes de un promotor mínimo. El sensor de cobre se utilizó exitosamente para controlar un activador transcripcional programable basado en CRISPR/dCas9 (dCasEV2.1) en *N. benthamiana*. Este sensor, combinado con el procesador dCasEV2.1 (CS/dCasEV2.1), confirió un control robusto de la expresión génica, con mínima actividad basal y fuerte activación en respuesta a cobre, ofreciendo una herramienta robusta para la activación específica de genes endógenos. Este sistema tiene un gran potencial para la regulación dirigida de genes en Biología Sintética de Plantas, especialmente en aplicaciones que requieren activación bajo demanda y baja expresión de fondo.

El segundo capítulo se centra en el desarrollo de un circuito de amplificación génica regulado por cobre, denominado CuBe, en referencia al ion cobre ($Cu^{2+}$) y al virus del enanismo amarillo del frijol (BeYDV). CuBe combina el sensor de cobre con un vector replicativo derivado específicamente del geminivirus BeYDV y sus correspondientes proteínas asociadas a replicación (Rep/RepA). Esta combinación permite la activación inducida por cobre del replicón y la posterior expresión de proteínas recombinantes en *N. benthamiana*. Diseñado para lograr una baja expresión basal y alta inducibilidad, CuBe facilita la producción controlada de proteínas de interés farmacéutico, como anticuerpos. Se generaron líneas transgénicas estables con CuBe, confirmando su funcionalidad a largo plazo y robustez con métodos alternativos de aplicación de cobre, como en post-cosecha e hidroponía. Este sistema destaca el potencial de los circuitos regulados por cobre para aplicaciones escalables de Agricultura Molecular en condiciones controladas.

El tercer capítulo amplía el uso de sistemas basados en geminivirus investigando y optimizando varios vectores geminivirales (derivados de BeYDV, TYLCV y BCTV) para aplicaciones en biotecnología vegetal. Utilizando un sistema reportero de bioluminiscencia autosostenida, este estudio caracteriza la dinámica y los niveles de expresión basal de diferentes configuraciones de vectores, evaluando su eficacia para la expresión transitoria de una enzima recombinante ligada a la bioluminiscencia. Estos análisis proporcionaron

información sobre los puntos fuertes y las limitaciones de cada vector para aplicaciones específicas, como la expresión de múltiples genes, lo que enriquece el conjunto de herramientas geminivirales disponibles para la Biología Sintética de Plantas.

En conclusión, esta tesis ofrece herramientas innovadoras para la Biología Sintética de Plantas, con importantes aplicaciones en agricultura molecular y biocomputación basada en plantas. Al integrar sistemas inducibles con diseños de vectores optimizados, este trabajo establece una base sólida para futuros avances en circuitos genéticos sintéticos ajustados a aplicaciones biotecnológicas vegetales que requieren una expresión génica controlada y adaptable.

# RESUM

La Biologia Sintètica de Plantes és un camp en ràpida expansió, amb un gran potencial per a potenciar la producció de compostos en plantes. Les plantes poden actuar com biofactories sostenibles i econòmiques per a la producció de molècules d'alt valor; no obstant això, la producció contínua sovint genera problemes de toxicitat per a la pròpia planta i baixos rendiments. Un objectiu clau en aquest àmbit és aconseguir un control precís i induïble sobre l'expressió gènica mitjançant circuits genètics sintètics que inclouen sensors i processadors modulars. Aquesta tesi se centra en el desenvolupament d'un sensor de coure i processadors basats en geminivirus, dissenyats com a components modulars per a circuits sintètics en Nicotiana benthamiana, una planta model amb gran potencial per a transformacions gèniques tant transitòries com estables.

El primer capítol d'aquesta tesi presenta un sistema d'expressió gènica induïble per coure, dissenyat per a permetre una regulació precisa dependent de sulfat de coure (CuSO4). Aquest sistema inclou un element sensor compost per un factor de transcripció que s'uneix al coure, CUP2, fusionat al domini d'activació Gal4 i un promotor sintètic amb llocs d'unió al coure (CBS) situat abans d'un promotor mínim. El sensor de coure es va utilitzar amb èxit per a controlar un activador transcripcional programable basat en CRISPR/dCas9 (dCasEV2.1) en N. benthamiana. Aquest sensor, combinat amb el processador dCasEV2.1 (CS/dCasEV2.1), va conferir un control robust de l'expressió gènica, amb mínima activitat basal i forta activació en resposta a coure, oferint una eina robusta per a l'activació específica de gens endògens. Aquest sistema té un gran potencial per a la regulació dirigida de gens en Biologia Sintètica de Plantes, especialment en aplicacions que requereixen activació sota demanda i baixa expressió de fons.

El segon capítol se centra en el desenvolupament d'un circuit d'amplificació gènica regulat per coure, denominat CuBe, en referència a l'ió coure ($Cu^{2+}$) i al virus del nanisme groc del fesol (BeYDV). CuBe combina el sensor de coure amb un vector replicatiu derivat específicament del geminivirus BeYDV i les seues corresponents proteïnes associades a replicació (Rep/RepA). Aquesta combinació permet l'activació induïda per coure del replicon i la posterior expressió de proteïnes recombinants en N. benthamiana. Dissenyat per a aconseguir una baixa expressió basal i alta inducibilitat, CuBe facilita la producció controlada de proteïnes d'interés farmacèutic, com a anticossos. Es van generar línies transgèniques estables amb CuBe, confirmant la seua funcionalitat a llarg termini i robustesa amb mètodes alternatius d'aplicació de coure, com en post-collita i hidroponia. Aquest sistema destaca el potencial dels circuits regulats per coure per a aplicacions escalables d'Agricultura Molecular en condicions controlades.

El tercer capítol amplia l'ús de sistemes basats en geminivirus investigant i optimitzant diversos vectors geminivirals (derivats de BeYDV, TYLCV i BCTV) per a aplicacions en biotecnologia vegetal. Utilitzant un sistema reporter de bioluminescència autosostinguda, aquest estudi caracteritza la dinàmica i els nivells d'expressió basal de diferents configuracions de vectors, avaluant la seua eficàcia per a l'expressió transitòria d'un enzim recombinant lligat a la bioluminescència. Aquestes anàlisis van proporcionar informació

sobre els punts forts i les limitacions de cada vector per a aplicacions específiques, com l'expressió de múltiples gens, la qual cosa enriqueix el conjunt d´eines geminivirals disponibles per a la Biologia Sintètica de Plantes.

En conclusió, aquesta tesi ofereix eines innovadores per a la Biologia Sintètica de Plantes, amb importants aplicacions en Agricultura Molecular i biocomputació basada en plantes. En integrar sistemes inducibles amb dissenys de vectors optimitzats, aquest treball estableix una base sòlida per a futurs avanços en circuits genètics sintètics ajustats a aplicacions biotecnològiques vegetals que requereixen una expressió gènica controlada i adaptable.

# Abbreviations

# ABBREVIATIONS

a.u.: arbitrary units

AD: activation domains

ASCL: *Physcomitrella patens* anther-specific chalcone synthase

attB/attP: recognition sites for PhiC31 recombinase

AUC: area under the curve

BCIR: BCTV intergenic region

BCTV: beet curly top virus

BeLIR: BeYDV long intergenic region

BeYDV: bean yellow dwarf virus

BFP: blue fluorescent protein

bp: DNA base pair

Cas: CRISPR-associated protein

CBS: copper binding site (CUP2 binding site)

cDNA: complementary DNA

CDS: coding sequence

CI-switch: copper-inducible switch

CPH: *Neonothopanus nambi* caffeylpyruvate hydrolase

CRISPR: clustered regularly interspaced short palindromic repeats

CRISPRa: CRISPR activation

crRNA: CRISPR RNA

CS/dCasEV2.1: copper sensor coupled to dcasEV2.1

$Cu^{2+}$: copper ion

CuBe: copper-regulated BeYDV-based synthetic circuit

CUP2: *Saccharomyces cerevisiae* copper-binding transcription factor

$CuSO_4$: copper sulfate

DBTL: design, build, test and learn

dCas9: dead Cas9, nuclease-deactivated Cas9 protein

dCasEV2.1: dCas9:EDLL + MS2:VPR + gRNA2.1

DFR: dihydroflavonol 4-reductase

DNA: deoxyribonucleic acid

dpca: days post-copper application

dpci: days post-copper incubation

dpi: days post-infiltration

DSB: double-strand break

dsDNA: double-stranded DNA

eGFP: enhanced green fluorescent protein

ELISA: enzyme-linked immunosorbent assay

FBP: fungal bioluminescence pathway

FE: fully expanded

Flp: flippase

Fluc: Firefly luciferase

FRT: recognition site for flippase recombinase

FW: fresh weight

GB: GoldenBraid

GOI: gene of interest

gRNA: guide RNA

h: hours

H3H: *Neonothopanus nambi* hispidin-3 hydroxylase

HBP: hybrid bioluminescence pathway

HispS: *Neonothopanus nambi* hispidin synthase

IR: intergenic region

Kb: Kilobase

KDa: KiloDalton

LB[1]: left border

LB[2]: Luria Bertani medium

LIR: long intergenic region

LSL vector: gene of interest flanked by an upstream LIR and a downstream element comprising a SIR and a LIR

Luz: *Neonothopanus nambi* luciferase

MAR: matrix attachment region

MES: 2-(N-morpholino)ethanesulfonic acid

min: minutes

minDFR: minimal DFR promoter

mM: milimolar

MRE: metal responsive element

MS: Murashige and Skoog medium

n.s.: no significant

n72_hγHC: single-chain anti-SARS-CoV-2 antibody

Nb: *Nicotiana benthamiana*

NE: non-fully expanded

NPGA: *Aspergillus nidulans* 4'-phosphopantetheinyl transferase acid synthase

OCSt: octopine synthase gene transcriptional terminator

OD: optical density

ORF: open reading frame

p35S: promoter of Cauliflower Mosaic Virus 35 S (CaMV35S)

PAL: phenylalanine ammonia-lyase

PAM: protospacer adjacent motif

PCR: polymerase chain reaction

pDGB3: GoldenBraid destination vector

PKS2: *Plumbago zeylanica* polyketide synthase

pNOS: promoter of *Agrobacterium tumefacients* nopaline synthase gene

Pol-II: RNA polymerase II

PTA: programmable transcriptional activator

pUPD2: GoldenBraid universal domesticator plasmid

qPCR: quantitative PCR

RB: right border

RBD: receptor binding domain

RCR: rolling-circle-replication

RDR: recombination-dependent-replication

Rep: replication-associated protein

RepA: replication-associated protein A

Rluc: renilla luciferase

RNA: ribonucleic acid

ROS: reactive oxygen species

rpm: revolutions per minute

RPUs: relative promoter units

RS: recombination site

RT: room temperature

RTA: relative transcriptional activities

RT-qPCR: real time-quantitative PCR

s: seconds

SAM: synergistic activation mediator

ScFv: single-chain variable fragment

scRNAs: CRISPR scaffold RNAs

SD: standard deviation

SDS-PAGE: sodium dodecyl-sulfate polyacrylamide gel electrophoresis

SGC: synthetic gene circuit

sgRNA: single guide RNA

SIR: short intergenic region

*Sl: solanum lycopersicum*

ssDNA: single-stranded DNA

SynBio: Synthetic Biology

T0: transgenic plant generation 0

t35S: transcriptional terminator of Cauliflower Mosaic Virus 35 S (CaMV35S)

TAD: transcriptional activator domain

T-DNA: Ti-plasmid from *Agrobacterium tumefaciens*

TF: transcription factor

TMV: Tobacco Mosaic Virus

tNOS: transcriptional terminator of *Agrobacterium tumefacients* nopaline synthase gene

tracrRNA: trans-activating CRISPR RNA

tRNA: transfer RNA

TSP: total soluble protein

TSS: transcriptional start site

TU: transcriptional unit

TV: 6×TAL + 4×VP64 activation domain

TYIR: TYLCV intergenic region

TYLCV: tomato yellow leaf curl virus

UTR: untranslated region

VPR: VP64- p65-Rta transcription factor

WT: wild type

# Introduction

# INTRODUCTION

## 1. Synthetic Biology

In 1988, Francis Crick stated '*what we see was not designed, but rather evolved*.' Undoubtedly, he would be impressed if he could see the current potential of humans in designing and manipulating biological systems. Genetically modified species are no longer limited to the realm of films like Jurassic Park or Splice. Science fiction became a reality. The past few decades have witnessed a remarkable convergence of advancements in genetic engineering, computational biology, and technology. This convergence has fuelled the rise of Synthetic Biology (SynBio), a rapidly evolving field with the potential to harness the power of life.

### 1.1. A brief history of Synthetic Biology origin and evolution

The term "Synthetic Biology" has a long history, with the first recorded use in a book written by Stéphane Leduc in 1912, *La Biologie Synthétique*. The work of Leduc deviated from the prevailing theories of that time by viewing life not as a mystic but as a purely physical phenomenon driven by physicochemical forces. He proposed that understanding these forces could enable the reconstruction of life-like phenomena outside living organisms. Leduc pioneered the concept of biomimicry, as evidenced by his experiments to create non-living systems that mimicked the characteristics of living beings (Leduc, 1912; Peretó, 2016). While Leduc introduced the term Synthetic Biology, the contemporary definition of the field is more closely linked to the work of Wacław Szybalski in 1974. Unfamiliar with Leduc's earlier usage, Szybalski employed the term in the context of advancements in DNA recombination. He envisioned Synthetic Biology as the engineering of novel genetic control systems, proposing their integration into existing genomes and even the creation of entirely new ones (Davies, 2018; Szybalski, 1974).

The emergence of genetics and molecular biology during the 20$^{th}$ century enabled the incorporation of new attributes into existing living organisms. The birth of genetic engineering was marked by Stanley Cohen and collaborators, who introduced recombinant DNA into bacteria, leading to the production of a functional protein from the transferred gene (Berg & Mertz, 2010; Cohen et al., 1972). While traditional genetic engineering and conventional biotechnology primarily focused on transferring a single foreign gene into a recipient organism to introduce or enhance specific traits (Liu & Stewart, 2015), the field quickly evolved towards more sophisticated engineering strategies that involved assembling multi-component systems within the recipient organism. A pivotal social-impacting example of this evolution was the generation of Golden Rice. Developed in the late 1990s (Ye et al., 2000) and optimized in 2005 (Paine et al., 2005), Golden Rice incorporates genes from different organisms to engineer a metabolic pathway for enhanced β-carotene production within the rice grain. This pathway was not naturally present in rice, and the combination of these transgenes significantly increased the provitamin A content in the grain. Golden Rice

exemplifies how genetic engineering can be used to manipulate complex biological processes within an organism, introducing some of the core principles of Synthetic Biology.

The shift from transferring genes or modifying existing pathways to constructing entirely new biological 'devices' with distinct functionalities marked the dawn of Synthetic Biology as a distinct discipline. While the concept of engineering genetic circuits existed before, the year 2000 marked a turning point with the creation of two synthetic biological devices in *Escherichia coli*: the genetic toggle switch and the repressilator (Figure 1). These pioneer devices demonstrated the foundations of modern Synthetic Biology: creating synthetic biological devices that mimic the functionalities of electronic circuits (Andrianantoandro et al., 2006; De Lorenzo & Danchin, 2008).

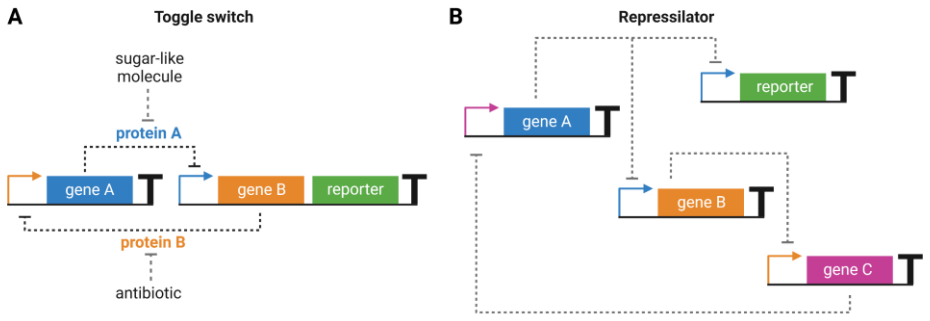



**Figure 1. Synthetic gene circuits from the early years of Synthetic Biology. A)** The toggle switch consists of two mutually repressing genes (genes A and B), creating a bistable system where only one gene is active at a time. Environmental signals, such as a sugar-like molecule to inhibit protein A or an antibiotic to inhibit protein B, can switch the circuit between transcriptional states. **B)** The repressilator is composed of three repressor genes (genes A, B, and C) arranged in a negative feedback loop. This circuit generates periodic oscillations in reporter expression.

The toggle switch operates similarly to a classic computer set-reset memory circuit, while the repressilator behaves like a ring oscillator in electronics. The toggle switch, developed by Gardner et al. (2000), consists of a synthetic genetic regulatory network composed of two genes that mutually repress each other's expression. Thus, this system can exist in one of two stable states: "gene A on, gene B off" or "gene A off, gene B on." Additionally, the system incorporated an external control mechanism mediated by small molecules. Specifically, the activity of the repressor protein produced by gene A can be modulated by an exogenous sugar-like molecule, while the repressor protein of gene B can be inhibited by an antibiotic. This inducible feature allows the system to be flipped between the stable states by adding or removing these specific molecules. The ability to manipulate the state of the system and achieve bistability with external control underscores the potential for synthetic genetic circuits to function as cellular memory elements. Building upon the success of the toggle switch, another landmark achievement in 2000 was the repressilator, a genetic oscillator developed by Elowitz & Leibler, (2000). This device represents a more complex synthetic circuit, exhibiting periodic oscillations in gene expression through a cascade of repression.

Gene A's protein product represses the activity of gene B. In turn, gene B's protein product represses gene C. Finally, gene C's protein product feeds back to repress gene A, completing the regulatory loop. The oscillatory behaviour arises from inherent time delays within the system. There is a time lag between gene activation and the production of functional protein molecules. Parallelly, a delay is associated with the degradation and disappearance of existing proteins. These time delays create a natural imbalance within the circuit, leading to the observed oscillations.

The groundbreaking achievements in 2000, like the toggle switch and the repressilator, not only provided the first examples of functional synthetic biological devices but also paved the way for the development of increasingly complex and sophisticated biological systems. The first meeting dedicated to Synthetic Biology was held at MIT in 2004, marking a pivotal moment in consolidating this discipline (Patil & Dhar, 2015). This meeting brought together researchers from diverse backgrounds –like genetics, engineering, and computer science– all united by the goal of reprogramming biological systems to have novel applications. This multidisciplinary effort has led to the modern definition of Synthetic Biology, a discipline based on the rational design and engineering of circuits that confer novel functionalities in living systems (Kitano et al., 2023). The most common synthetic circuits are focused on gene regulation since proteins interacting with nucleic acids, like transcriptional factors, are easily configurable to regulate gene activity. Nevertheless, protein-level circuits based on a modification of activity, stability, or localization of proteins can also be implemented to offer a new level of control (Gao et al., 2018).

Nowadays, Synthetic Biology stands out as one of the most rapidly advancing research domains in the 21$^{st}$ century (Munnelly, 2013). Its global market has seen explosive growth, reaching an estimated value of USD 11.4 billion in 2022, and is projected to maintain a compound annual growth rate (CAGR) of 25.6% until 2027 (Markets & Markets, 2023). Today, synthetic biologists can engineer programmable microbial biosensors to detect pollutants and pathogens in real-time, with applications in environmental monitoring and bioremediation (Liu et al., 2022; Nnachi et al., 2022). Further current applications of SynBio include living therapeutics such as engineered bacteria cells designed to produce therapeutic molecules directly inside the human body. These engineered bacteria offer targeted and sustained drug delivery for cancer treatment and other diseases, acting exclusively in the desired microenvironment to minimize side effects on healthy tissue (Li et al., 2021; Patil et al., 2023). In the plant sector, rewiring plant metabolism through Synthetic Biology, for instance, offers a path toward sustainable biofuel production (Yang et al., 2022). The future of Synthetic Biology is promising, with the potential to revolutionize industry, medicine, agriculture, and environmental protection. As an additional note, it is also remarkable that the significance of Synthetic Biology extends beyond technological advancements and direct applications. It also offers potent tools to answer fundamental scientific questions and understand biological processes like multicellular networks, and fitness trade-offs (Budin & Keasling, 2019; Serrano, 2007).

## 1.2. The Engineering Influence on Synthetic Biology

Synthetic Biology aims to be an established engineering discipline by applying core engineering principles to create predictable and controllable biological functions (Benner & Sismour, 2005). The engineering principles can be used in the design of either novel or modified biological systems. One approach within SynBio, known as the bottom-up approach, proposes to build entirely new biological systems from fundamental chemical elements. This involves redesigning a biological system from its foundation (Smith et al., 2003). This approach faces significant challenges due to the complexity of living systems (Koide et al., 2009). By contrast, a more prevalent approach, top-down Synthetic Biology, takes advantage of existing cellular organisms as a foundation or chassis. This second approach focuses on engineering new functionalities on top of existing biological systems (Hirschi et al., 2022). Thus, top-down SynBio aims to construct novel genetic devices within an existing organism or cell. Nowadays, this approach offers a more practical and controllable strategy for near-term applications than the bottom-up approach. The novel devices generated through top-down SynBio and composed of gene circuits are designed to process specific inputs and generate predictable and optimized outputs (Benner & Sismour, 2005).

Mutations, evolution, environment interaction, and signalling pathways subjected to complex connections and inherent randomness of molecular diffusion (stochasticity) in living organisms make it challenging to approach biology as a predictable science (De Lorenzo & Danchin, 2008). Indeed, while gene expression descriptions may utilize digital terms like "on" and "off," the underlying biological systems operate in a more continuous, analogical manner (Moon, 2023). Nevertheless, Synthetic Biology pursues the goal of reducing the intrinsic complexity of biological systems by applying stochastic mathematical models, computational design of experiments, and engineering principles (Chen et al., 2009; Skerker et al., 2009).

The three engineering principles governing SynBio are standardization, modularity, and abstraction. These principles are applied to the rational design and construction of biological information processing systems in a chassis cell (Garner, 2021). In these systems, DNA serves as the code for building a biological part or 'biopart'. These bioparts are combined to create biological devices, and finally, the devices are structurally assembled into biological systems. The principle of standardization involves the creation of bioparts with well-defined and uniform characteristics and functionalities (Cameron et al., 2014). Robust characterization of the bioparts allows for a device made of them to have highly predictable and reproducible outputs. Bioparts are designed to be modularly assembled, like electronic components in a circuit. The modularity principle is based on the idea that large (composite) bioparts are created by assembling two (or more) smaller bioparts that share common structural features and following standard assembly rules. This principle facilitates the design and construction of complex biological systems from pre-existing and interchangeable bioparts. Finally, the principle of abstraction refers to describing biological systems at different levels of

complexity. Abstraction implies decomposing complex systems into essential components, allowing one to focus on details at specific levels. This facilitates the understanding, design, and analysis of systems without requiring one to be an expert in all aspects of a complex biological system (Davies, 2018; Garner, 2021; Iram et al., 2024; Liu & Stewart, 2015).

Apart from principles, Synthetic Biology also borrows the work pipeline from engineering. SynBio employs an iterative Design-Build-Test-Learn (DBTL) cycle to develop new biological systems (Kitano et al., 2023) (Figure 2). This methodology assumes that the assembly and organization of a biological system may need optimization through several rounds of the DBTL process to achieve the desired output (Opgenorth et al., 2019). Each round starts with a Design phase consisting of a simulation of the performance system, preferably employing *in-silico* mathematical modelling. Then, the simulated system is assembled in the Build phase and tested in the Test stage for validation. Finally, in the Learn stage, the simulation is compared with the performance of the built system, and differences between predictions and observations provide new insights into the system. These insights feed and refine the model in a subsequent Design phase. The DBTL iteration can continue until the desired biological output is achieved (Carbonell et al., 2018).

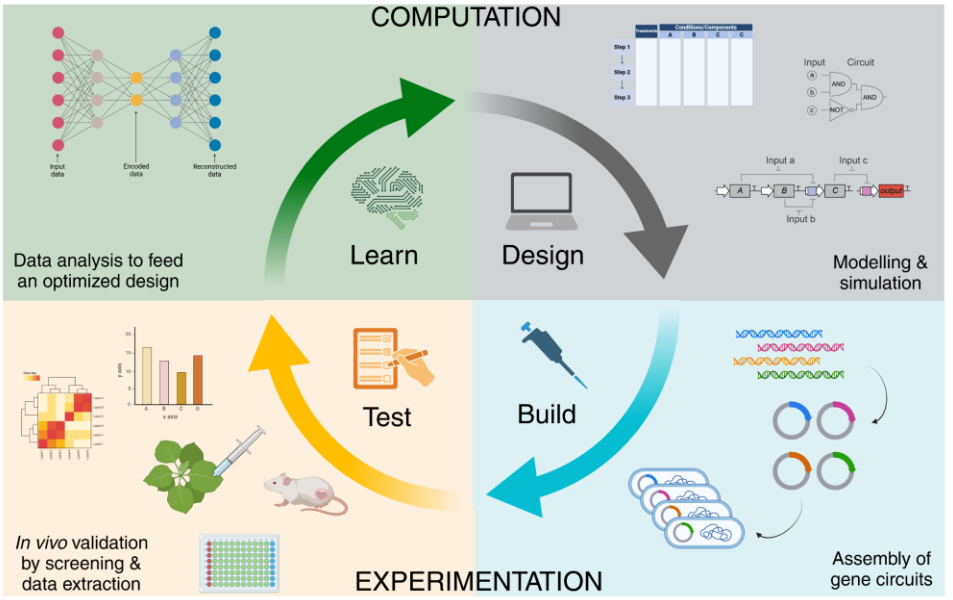



**Figure 2. Schematic representation of the Design-Build-Test-Learn (DBTL) cycle in Synthetic Biology**. The iterative DBTL cycle serves as the foundational framework for the systematic development and optimization of engineered biological systems. The cycle begins with the Design phase, where computational tools and models are used to plan and draft genetic constructs and predict their behaviour. In the Build phase, the designs are physically constructed using molecular biology techniques. The Test phase involves experimental evaluation of the engineered constructs to assess their functionality and performance. The Learn phase incorporates data analysis and machine learning algorithms to refine models and new design iterations. This cyclical process allows for continuous improvement, leading to the development of organisms with highly optimized and predictable

behaviours. This figure contains details modified from https://sbl.postech.ac.kr/. Figure includes images from Biorender (biorender.com).

### 1.3. Engineering challenges in multicellular systems

While implementing the DBTL cycle represents a solid approach to reprogramming gene expression in prokaryotic systems like bacteria precisely (Gurdo et al., 2023), its direct application in multicellular organisms presents a significant challenge. Unlike bacteria, which are unicellular and operate with a more readily digitalized gene expression landscape, multicellular organisms present complex tissue differentiation and intricate regulatory networks. In these complex organisms, Synthetic Biology currently faces challenges in the Design phase due to the difficulty of computational prediction of bioengineering's impact on the entire host. The limited capacity to predict outcomes of bioengineering causes the aims of SynBio to often be accomplished by tedious Trial-and-Error cycles. This way, the *modus operandi* in multicellular systems looks more like a T-E (Trial-Error) cycle than a DBTL one. However, the flourishing of artificial intelligence and machine learning are expected to complement the limited statistics and current data mining techniques like clustering and pattern identification, which need to be revised in the Design phase in multicellular Synthetic Biology (Eslami et al., 2022). Continuing with the cycle flow, while advancements in modular cloning systems have streamlined the Build phase over the past decade, testing remains a significant bottleneck, mainly because of the heterogeneity of cell types and the long time to produce transgenic or mutant organisms. Nevertheless, in multicellular systems like plants, protoplast-based systems (Pouvreau et al., 2020) and transient expression systems (Brophy et al., 2022) partially alleviate this issue, allowing for rapid and high-throughput evaluation of designs. Each cycle offers valuable insights into crucial considerations, which can be incorporated into subsequent iterations for improved efficiency (Vazquez-Vilar et al., 2023). In recent years, machine learning tools like BioAutomata and ART have boosted the Learn step. BioAutomata utilizes a probabilistic model to select the most promising designs from an initial dataset. This model prioritizes designs with a higher likelihood of improving the overall biosystem. The targeted approach significantly reduces the total number of experiments required to identify the optimal design within the search space (HamediRad et al., 2019). ART also utilizes experimental data, including both input variables (e.g., proteomics data, transcriptomics, and promoter combinations) and corresponding responses (e.g., production levels), to provide recommendations for the next experiments (Radivojević et al., 2020).

## 2. Plant Synthetic Biology

Plants are the foundation of life on Earth, providing oxygen and essential resources for humans and ecosystems (Knapp, 2019). Global biomass is dominated by plants, which represent approximately 80% of the biomass across all taxa (Bar-On et al., 2018). The abundance of plants and their vast capacity to supply goods make them a highly promising platform for biomanufacturing. Plants can be engineered for diverse purposes like food

production, biofuels and valuable chemicals and proteins for health and industry (Fesenko & Edwards, 2014). Synthetic Biology emerges as a powerful tool to engineer plants with novel, desirable traits that maximize plant potential by designing new genetic and metabolic circuits, physiological traits, and regulation strategies (Khalil & Collins, 2010).

The main approach within Synthetic Biology for engineering plants lies in the design of synthetic gene circuits (SGCs). These circuits function like biological computers in the living cell, allowing gene expression control in response to specific internal or external inputs. A critical aspect of Plant Synthetic Biology is the development of a robust library of standardized and interchangeable bioparts, whose combination results in functional genetic modules of a SGC. The bioparts that are essential for an operative SGC are genetic components like coding sequences and transcriptional promoters and terminators. Standardized libraries of bioparts enable the efficient assembly and modification of synthetic gene networks within plants, similarly to how electronic components allow for the building of complex circuits. Collections of standardized parts, coupled with user-friendly online design tools and modular cloning, has significantly enriched the implementation of SynBio across plant systems (Sarrion-Perdigones et al., 2011; Weber et al., 2011).

Efforts to standardize plant genetic bioparts received a major boost with the development of the Phytobrick syntax (Patron et al., 2015). Phytobricks are standardized DNA parts (bioparts) designed for plants, which can be systematically assembled into transcriptional units (TUs) and then into complex multigene constructs. Each Phytobrick includes sequences of one or more functional elements found in eukaryotic genetic units for transcription, such as promoter, coding sequence, and untranslated regions, and harbours specific border sequences that allow for interchangeable use in one-step cloning reactions, allowing a combinatorial assembly process. The Phytobrick approach defines a set of DNA cloning rules based on Type IIS restriction enzymes compatible with existing modular cloning strategies like MoClo, GoldenBraid, Mobius, and Loop assembly (Cai et al., 2020). Adopting the Phytobrick standard facilitates the exchange of bioparts and ensure fair functional comparisons.

## 2.1. Modular cloning for plant synthetic gene circuit assembly

Modular cloning techniques enable the assembly of multiple bioparts into a single vector within a single step through a simultaneous restriction-ligation reaction (Marillonnet & Grützner, 2020; Wang et al., 2021). A milestone within modular cloning was the creation of the Golden Gate cloning strategy. This strategy is based on Type IIS restriction enzymes, which cleave outside their recognition sequence, enabling the design of specific flanking sequences for each biopart to assemble in a directional positioning in a complex module (Engler et al., 2009).

One of the most consolidated DNA assembly strategies for plant expression based on Golden Gate principles is GoldenBraid (GB), whose online software enables convenient design assistance (https://goldenbraidpro.com/). GB enables the gradual construction of

complex genetic modules in different sets of binary vectors from individual bioparts. These vectors carry T-DNA left and right borders (LB and RB, respectively) and are compatible with *Agrobacterium*-mediated plant transformation (Sarrion-Perdigones et al., 2011; Sarrion-Perdigones et al., 2017). Each biopart has user-defined four-nucleotide overhangs that ensure precise assembly, since each part only binds to its designated partner with complementary overhangs. These overhangs, acting as barcodes according to Phytobrick syntax, determine the specific order and orientation in which elements are assembled to create a transcriptional unit.

The GB workflow is a three-level process consisting of the domestication of bioparts (level 0), assembly of these bioparts in TUs (level 1), and finally, the binary assembly of the TUs (level > 1) (Figure 3). GB makes use of two Type IIS restriction enzymes, BsaI and BsmBI, which, next to the T4 DNA ligase, mediate the restriction-ligation cloning of the DNA elements in a destination vector. Domestication (level 0) involves adapting bioparts to the GoldenBraid grammar. This entails removing internal BsaI and BsmBI recognition sites that could interfere with assembly and incorporating the flanking BsmBI-target sites and the overhang sequences that give an identity to each part. These modified parts are cloned into a pUPD2 plasmid through BsmBI-mediated digestion. Next, in level 1, GB parts are extracted from the pUPD2 backbone by a BsaI-digestion and assembled according to the compatibility of their overhangs to form a TU in a pDGBα plasmid. Following the generation of TUs in level >1, GB employs a double-loop process (known as 'braid') to facilitate their iterative combination into increasingly complex genetic constructs within the pDGBα and pDGBΩ destination vectors. These vectors harbour strategically positioned restriction sites for the enzymes BsaI and BsmBI. First, two TUs are combined within pDGBΩ by a BsmBI digestion. Then, an illimited number of assemblies can proceed by iterative BsmBI (or BsaI) digestion of two pDGBα1 and α2 (or pDGBΩ1 ad Ω2) donor plasmids and assembly into a pDGBΩ (or pDGBα) destination vector. Therefore, the "braiding" analogy originates from the back-and-forth nature of the assembly process, utilizing the two distinct destination vector types (α and Ω) and their compatible restriction sites. This iterative braiding strategy empowers the construction of complex genetic modules with remarkable efficiency.

## 2.2. Common standardized DNA parts for plant transcriptional units

The precise selection of DNA parts for assembly into transcriptional units is essential for developing finely tuned synthetic circuits in plants. The Phytobrick syntax has facilitated the creation of extensive libraries of standardized parts. While several databases like the iGEM Registry (Vilanova & Porcar, 2014) and BioParts (Plahar et al., 2021) offer collections of engineered bioparts, these primarily focus on prokaryotic and yeast systems, limiting their direct application in plants. Existing plant-specific resources like the GoldenBraid collection, PPDB, PlantCARE, and PlantProm (Kusunoki & Yamamoto, 2017; Lescot, 2002; Sarrion-Perdigones et al., 2011; Shahmuradov, 2003) provide an extensive library of regulatory elements to combine with a coding sequence (CDS) of interest but often lack data on their experimentally validated quantitative strength.

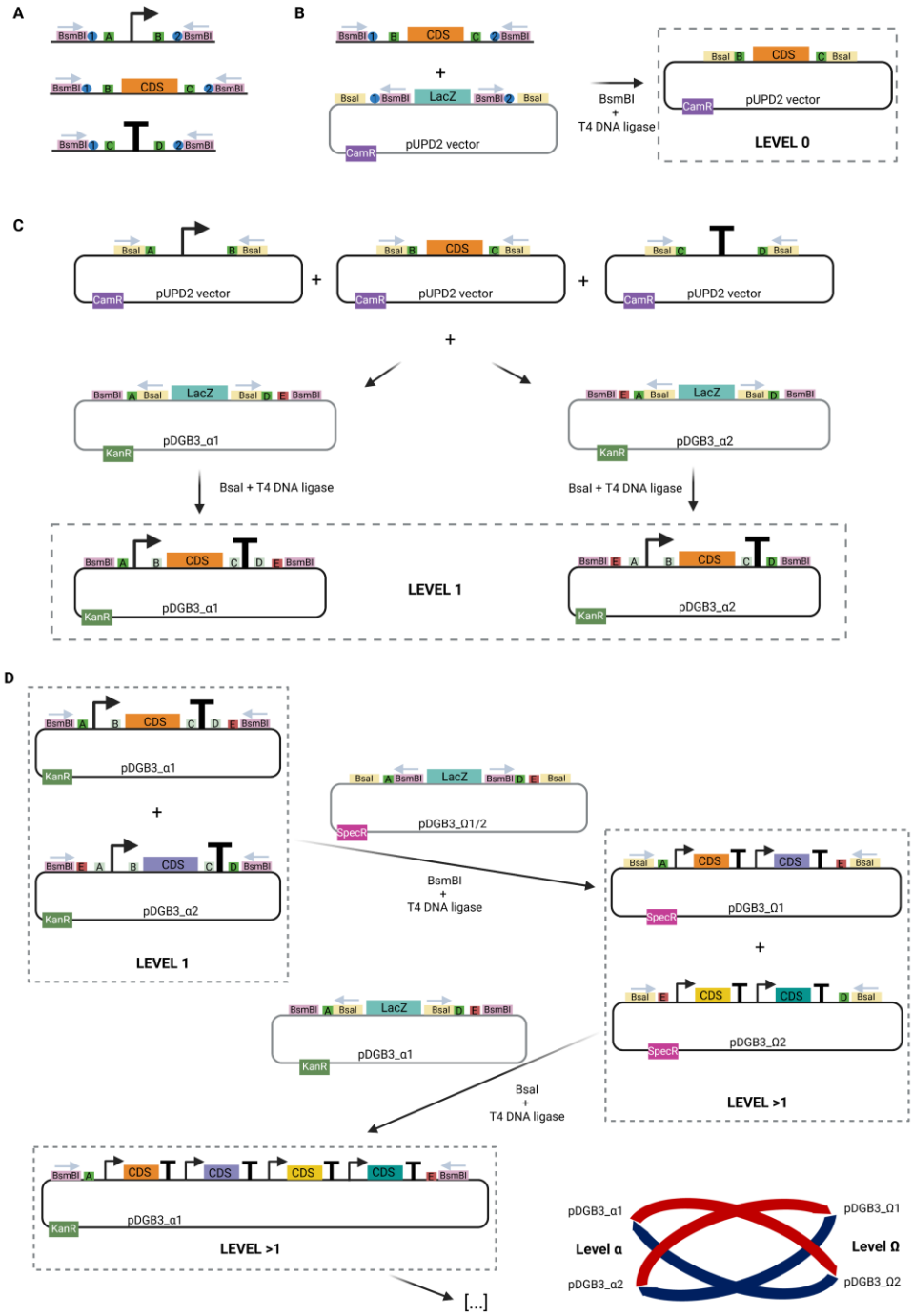



**Figure 3. GoldenBraid assembly strategy. A)** Domestication of basic DNA parts for Level 0 assembly. Each part is flanked by external BsmBI recognition sites and overhang sequences including barcodes for pUPD2 assembly (blue 1 and 2 dots) and part-specific barcodes for transcriptional unit assemblies in Level 1 (green A-D squares). **B)** Cloning of domesticated parts into a pUPD2 entry vector for Level 0 assembly. **C)** Multipartite assembly of promoter (black arrow), CDS (orange box), and terminator (black

T) parts into a pDGBα1 destination vector. **D)** Sequential combination of transcriptional units through iterative binary assemblies, alternating between α and Ω levels, which can lead to an infinite loop of assemblies represented by the blue and red braid symbol of GoldenBraid. BsmBI and BsaI recognition sites are shown in pink and yellow, respectively, with arrows indicating the direction of their cuts. Figure includes images from Biorender (biorender.com).

To fill the quantitative gap in bioparts characterization, Pfotenhauer et al. (2022) tested 91 transgene cassettes combining plant regulatory elements for the transient expression of a green fluorescent protein (GFP) in *Nicotiana benthamiana* leaves. As a result of this study, a well-curated library of 37 plant regulatory parts comprising promoters, untranslated regions (UTRs), and their respective combination was characterized. Here, the combination of the CaMV35S promoter (p35S) with the 5'UTR TMVΩ and the *Agrobacterium tumefaciens OCS* 3'UTR gave the highest expression levels of GFP. Another recent similar characterization study was done by Tian et al., (2022), where they showed that the CsVMV promoter combined with the terminator ATHsp18.2 resulted in the highest expression levels, outperforming CaMV35S. Building upon the value of experimentally validated quantitative data for regulatory bioparts, Tian et al., (2024) developed the Plant Synthetic Biology Database (PSBD). This user-friendly database provides access to a library of 384 well-characterized regulatory elements with quantitative strength data. These elements, alongside 1677 catalytic bioparts and data on species, molecules, and omics assemblies, are curated from PSBD's research and published literature. Additionally, very recently, Kocaoglan et al. (2024) developed a toolkit called Mobius Assembly for Plant Systems (MAPS), which includes optimized binary vectors and a collection of promoters and terminators with different activity levels. They analyzed RNA folding to better understand how these parts interact, aiming to make their behaviour more predictable and reliable for future use. The PSBD network and the MAPS toolkit represent powerful tools for efficiently selecting bioparts for plant synthetic circuits.

After presenting databases for identifying standardized DNA parts, it is important to describe in more detail the different components employed for constructing the transcriptional units comprising plant synthetic circuits. These include promoters, 3'UTR, CDS, tags, and other DNA elements influencing transcriptional unit functionality. Each of these elements plays a vital role in the efficiency of the engineered constructs.

### 2.2.1. Promoters

Traditionally, plant endogenous promoters have been the main source of regulatory sequences for Plant Biotechnology. However, the inherent complexity and incomplete understanding of upstream *cis*-regulatory elements in endogenous promoters present a significant challenge for making accurate predictions in plant genetic engineering. To address this, it is essential to rationally deconstruct endogenous promoters to identify their functional components. Cai et al., (2020) demonstrated that plant constitutive promoters contain multiple functional cis-regulatory elements (CREs) and that the relative position and complexity of these CREs alter the regulatory function of the promoter. This understanding

paved the way for the rational design of synthetic promoters with predictable behaviour in plants. By precisely controlling the expression patterns of these synthetic promoters, it is possible to achieve specific outcomes, such as constitutive, inducible, or organ-specific activity, which enhances the precision and predictability of gene regulation. This is crucial for constructing advanced genetic circuits in plants, offering greater flexibility in Synthetic Biology applications (Pfotenhauer et al., 2022). A synthetic promoter combines a minimal promoter with distal and proximal promoter regions containing one or more *cis*-regulatory elements that respond to one or more transcription factors (TFs). This chimeric promoter interacts with the designated TF and the endogenous transcriptional machinery of the plant. The minimal or core promoter, located immediately upstream of the transcription start site (TSS), is where the RNA-polymerase II and the rest of the pre-initiation complex bind. Therefore, the minimal promoter leads to varying levels of basal transcription. Additionally, the binding of endogenous or non-native regulatory elements in distal and proximal promoter regions upstream of the minimal promoter further controls this expression, offering multiple regulation layers (Belcher et al., 2020; Cai et al., 2020).

For example, Moreno-Giménez et al. (2022) developed a standardized library of synthetic promoter components for fine-tuned recombinant gene expression in plants, mediated by the transcriptional activator dCasEV2.1 (Selma et al., 2019). This collection, named GB_SynP, offers pre-defined building blocks for promoter assembly. GB_SynP includes (i) minimal promoter parts containing the TATA box and 5'UTR region, (ii) proximal parts harbouring the target sequences for gRNA as regulatory boxes, and (iii) randomized distal parts ensuring proper promoter length. These interchangeable parts were used to construct a library of 35 unique promoters, exhibiting minimal background activity and a wide range of tuneable expression levels. All the designed promoters functioned independently of native regulatory plant systems, enabling precise control over gene expression regulation.

Jores et al. (2021) developed another approach to expanding the toolbox of synthetic promoters, focusing on core promoter development. They employed a large-scale analysis (STARR-seq) to identify key features influencing core promoter strength in *Arabidopsis*, maize, and sorghum. Notably, a TATA box in a specific location, positioned around 30-40 bp upstream of the TSS, emerged as the most critical factor. Building upon this knowledge, Jores et al. (2021) designed novel synthetic core promoters by combining essential elements like the TATA box and the *Initiator* elements with randomized sequences. They optimized these sequences using machine learning and in silico evolution, achieving activities comparable to the strong CaMV 35S minimal promoter. While these findings offer valuable resources for synthetic core promoter development, additional regulatory layers beyond the core promoter can influence protein levels (Vollen et al., 2024).

### 2.2.2. 3' untranslated regions (3'UTRs)

Transcriptional termination is an important consideration regarding the influence of genetic elements on protein expression levels. While well-studied 3'UTRs or terminators derived

from viruses (e.g., 35S) and bacteria (e.g., NOS, OCS) or highly expressed plant genes (e.g., RUBISCO, HEAT SHOCK PROTEIN) are generally preferred for their established functionality, specific circumstances might require a more cautious approach. Double terminators offer a potential solution in cases of unintended read-through transcription, particularly in multigene constructs susceptible to silencing effects (Vollen et al., 2024). Several effective double terminators have been identified, including those described by Diamos & Mason (2018), where they showed that the combination of the intron less tobacco extensin 3′ UTR (EU) and the *N. benthamiana* actin 3′ UTR (NbACT3) led to the highest expression of the GFP reporter protein. By rationally selecting both the promoter and terminator, gene expression can be optimized for synthetic constructs and maximization of the desired protein production levels.

### 2.2.3. Coding sequences (CDS)

The choice of a coding sequence (CDS), the segment of DNA that contains the information needed to synthesize a protein, for a synthetic genetic circuit depends on the desired output, resulting in high variability. CDSs can serve various roles within the genetic circuit, such as sensors, processors, or actuators. As sensors, CDSs may detect specific environmental or cellular conditions, triggering downstream genetic responses. As processors, they can integrate multiple signals, performing logical operations to control gene expression. The most critical role of CDSs is as actuators, where they produce a tangible output in response to transcriptional signals. Common actuators include enzymes that catalyse reactions. However, the selection of a CDS as an actuator depends on the type of output desired—whether it is functional (e.g., production of a bioactive compound), qualitative (e.g., visual colour change), or quantitative (e.g., luminescence intensity). These actuators are essential for building synthetic circuits with predictable and tuneable responses, providing a crucial link between the genetic inputs and the observable effects in engineered plant systems.

The CDS not only determines a desired output but also plays a critical role in evaluating the functionality of a circuit. A so-called reporter CDS is required when monitoring the performance of a synthetic circuit. Traditionally, β-glucuronidase (GUS) (Jefferson et al., 1987) or fluorescent proteins like Green Fluorescent Protein (GFP) (Chalfie et al., 1994) and its enhanced variants like eGFP (Heim et al., 1995) and TurboGFP (Evdokimov et al., 2006) have been extensively used, providing a clear visual indication of circuit activity. Another popular choice was luciferase (Luc) (Contag & Bachmann, 2002), offering a quantitative readout through luminescence measurement. In recent years, new alternatives based on metabolic pathways, like RUBY (He et al., 2020; J. Liu et al., 2024) and the fungal bioluminescence pathway (Calvache et al., 2024) have gained prominence. Both systems rely on multiple genes that form complete biosynthetic pathways. For example, RUBY synthesizes betalains, pigments visible to the naked eye, offering a straightforward and reliable way to track circuit activity. Similarly, the fungal bioluminescence pathway produces light through a multi-enzyme cascade, providing a sensitive and continuous readout of gene

expression. These metabolic pathway-based actuators offer distinct advantages in terms of sensitivity, intensity, and simplicity for the assessment of synthetic circuit outputs.

Adjusting the design of the coding sequence is crucial to achieving the desired output of a synthetic circuit. Apart from codon optimization for the host organism, adding functional tags to the CDS is also valuable in refining the design for specific purposes. For instance, tags for subcellular compartmentalization, such as nuclear localization signals (NLS) for targeting the nucleus (Giraud et al., 2014) or a signal peptide (SP) for directing secretion to the apoplast (Diego-Martin et al., 2020), are essential for ensuring that the protein of interest performs its intended function. Affinity tags facilitate protein purification (Pina et al., 2021), while degradation tags, such as degrons (Calvache et al., 2022) or ubiquitination signals (Sun & Chen, 2004), allow for the control of protein levels in the cell. In some cases, splitting the CDS into two fragments and utilizing inteins can offer an additional layer of regulation (Li, 2015). The efficiency of the synthetic circuit's behaviour can be optimized by strategically employing these design elements.

### 2.2.4. Insulators

Transgene insertion in higher eukaryotes often occurs randomly, which can result in integration into transcriptionally repressed genomic regions (heterochromatin) and lead to transgene silencing. Alternatively, transgenes may integrate near endogenous regulatory elements such as repressors or enhancers, leading to abnormal expression (Pérez-González & Caro, 2016). When integrating a synthetic gene circuit in the plant genome, isolating the circuit modules from the positional effect of insertion is important. Chromatin insulators are flanking DNA sequences that define the boundaries between the flanked transgene and genome chromatin domains by blocking close enhancer-promoter interactions and protecting transgenes from position-dependent silencing (Matzat & Lei, 2014).

In plants, scaffold or matrix attachment regions (S/MARs) are among the most studied insulator-like elements. S/MARs are proposed to induce the formation of chromatin loops, delineating discrete chromosomal domain boundaries (Butaye et al., 2004). The impact of each insulator varies. For instance, (Pérez-González & Caro, 2019a) reported that chiMARs can protect transgene insertions from nearby repressive epigenetic states. Additionally, the Rb7 MAR enhances transgene expression, while the AtS/MAR10 reduces the variability among transgenic lines. Therefore, the use of insulators in transgenic constructs should be tailored to the specific requirements of each experiment.

### 2.2.5. Binary T-DNA vectors

*Agrobacterium*-mediated transformation is one of the most common strategies used to deliver transgenes, like the components of a synthetic genetic circuit, to the plant (Thompson et al., 2020). The genes of interest are inserted into the transfer DNA (T-DNA) region of binary vectors. The left and right border repeat sequences (LB and RB, respectively) contain highly conserved 25 bp and define and delimit the T-DNA (Waters et al., 1991). A large collection of binary vector systems has been developed and documented (Lee &

Gelvin, 2008). These binary T-DNA vectors are normally suited for Golden Gate and GoldenBraid cloning as well as for overlap-dependent methods.

The multigenic components of a circuit can be distributed among independent binary T-DNA vectors. This way, the T-DNAs would integrate independently in the plant genome and could be segregated to obtain plants with only one module in case re-transformation with another construct of interest is required. For this distribution of circuit modules, simultaneous use of replication-compatible binary T-DNA vectors is advantageous. Pasin et al., (2017) developed a mini binary vector series of pBBR1-based and RK2-based binary TDNA vectors (so-called pLX) that can be coupled to allow multi-T-DNA delivery from Agrobacterium in a two-plasmid/one-strain strategy. These pLX vectors are also compatible for co-delivery with pDGB3 vectors (Vazquez-Vilar, Quijano-Rubio, Fernandez-del-Carmen, et al., 2017) based on pCAMBIA (Chi-Ham et al., 2012) and containing the replication origin pVS1. The co-existence of replicative binary vectors provides a flexible framework for multigene transfer and DNA parts characterization.

## 2.3. Modularity of plant synthetic gene circuits

The SynBio terminology is often explained using electronic jargon, simplifying complex ideas with familiar technological analogies. Drawing parallels with electronic circuits, Qi et al., (2013) categorized the modules of a transcriptional synthetic gene circuit (SGC) into sensors, processors, and actuators (Figure 4). In electronics, sensors detect environmental inputs and relay them to the processor. The processor then determines the appropriate actions to control the system based on these inputs, and the actuator implements these actions to modify the behaviour of the system. In a transcriptional SGC, the sensor converts non-transcriptional inputs into transcriptional outputs by either activating or repressing downstream transcriptional units. Processors operate the transcriptional input initiated by the sensors, transforming it into a different transcriptional output, which may involve the amplification, distribution, memory storage, or Boolean logical orchestration of the initial transcriptional input. Finally, actuators perform the desired cellular response, converting the transcriptional signal from processors into a non-transcriptional output, such as an enzymatic reaction in a metabolic pathway (Vazquez-Vilar et al., 2023). Integrating these modules within a living 'chassis' like a plant cell enhances its ability to precisely process data, respond to novel stimuli, and generate predictable outputs (Wang et al., 2013).

A major challenge in Plant SynBio is designing SGCs that operate orthogonally, meaning independently from the plant endogenous regulatory systems. The most common approach to minimize interference between the SGCs and the plant chassis is to use bioparts from other organisms, such as viruses, bacteria, or mammals (Vazquez-Vilar et al., 2023). These foreign elements tend to show minimal interaction with the plant machinery, resulting in more predictable and efficient SGCs. The following two subsections explain various sensor and processor modules that can be applied to orchestrate actuators. It also covers tools and strategies for developing new ones for customizable and orthogonal SGCs.

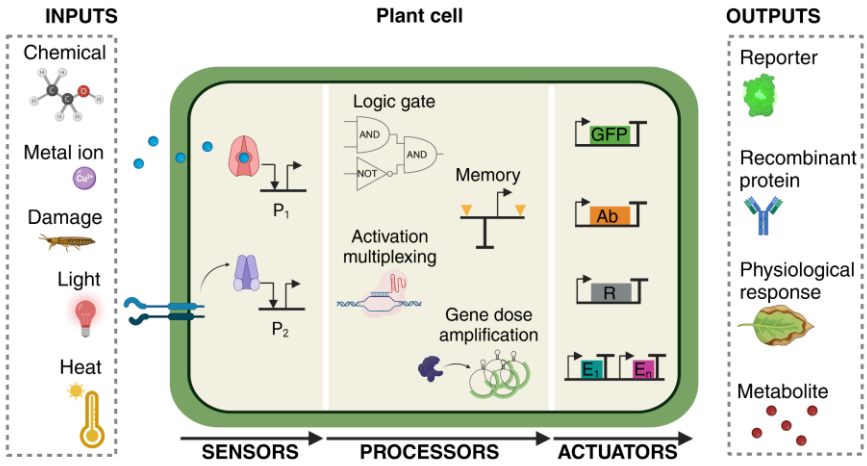



**Figure 4. Schematic representation of synthetic gene circuit components engineered in a plant cell**. The circuit components are organized into three categories: sensors, processors, and actuators. Sensors detect non-transcriptional signals and translate them into transcriptional outputs. Processors modify a transcriptional input, generating a distinct transcriptional output. Actuators then take the transcriptional output and convert it into a non-transcriptional response. Figure includes images from Biorender (biorender.com).

### 2.3.1. Sensor modules

Sensor modules can detect environmental cues or internal signals in the plant. The most basic sensor module includes an activable transcription factor (TF), which, upon detecting a cue or stimulus, can control the expression of a downstream transcriptional unit by binding to the operator sequence within a promoter. The employ of plant native TFs in sensor design can limit its performance to their restricted activation specificities. As an alternative, synthetic TFs are engineered and often combine elements from various sources, creating chimeric TF designs with improved or new capabilities (Del Valle et al., 2021). Synthetic TFs, also known as chimeric, artificial, or orthogonal, can provide greater versatility in terms of activation strength and specificity. These TFs typically consist of two parts: the DNA-binding domain that binds to the operator in the target promoter, and an effector (activator or repressor) domain that regulates the transcription (Koukara & Papadopoulou, 2023; Liu & Stewart, 2016). More sophisticated sensor modules incorporate complex molecular mechanisms, including clusters of receptors and signal transduction pathways. For example, sophisticated sensors might use receptor proteins to detect external stimuli and initiate non-transcriptional signalling cascades that ultimately influence gene expression (Stepanova et al., 2007; Thiruppathi, 2020; Zhao et al., 2021). These cascades can involve multiple steps, including multiplexed activation of gene expression, phosphorylation events and/or the release of second messengers, thus strengthening or fine-tuning the response to the initial signal.

Sensor modules can be classified according to their activation inputs. These inputs include internal signals like hormone levels and developmental cues as well as environmental factors such as chemical, light, biotic, or abiotic stress, temperature, and nutrient availability (recently reviewed in Koukara & Papadopoulou, (2023); Vazquez-Vilar et al., (2023); Wang & Demirer, (2023)). Regarding internal signals, for instance, numerous transcriptional sensors responding to phytohormones like auxins, jasmonic acid (JA), abscisic acid (ABA), gibberellins, cytokinins, and ethylene have been developed to report the levels and distribution of these phytohormones (Vazquez-Vilar et al., 2023). One well-known example involves the binding of the AUXIN RESPONSE FACTOR (ARF) transcription factor to a synthetic auxin response promoter called DR5. The last optimization version of the DR5 promoter contains nine repeats of the AuxRE sequence in inverse orientation upstream of a minimal 35S promoter (Liao et al., 2015). In this case, the ARF sensor drives the expression of reporter genes like fluorescent proteins, GUS, or firefly luciferase (LUC), allowing precise tracking of auxin responses at both cellular and tissue levels (Jedličková et al., 2022). The artificial ARF-responsive DR5 promoter is a great example of how creating synthetic promoters with varying numbers, orientations, or positions of motif repeats in the operator is important for fine-tuning the activation controlled by sensor modules.

Regarding sensors responding to external signals, chemical sensors were among the first ones adapted for use in plants, quickly becoming indispensable components of inducible gene expression systems. One pioneering and model example is the tetracycline sensor system, known as Tet-on, which functions as a de-repressor of transcription (Gatz & Quail, 1988). In the absence of tetracycline, the tetracycline repressor (TetR) binds to the tet operator, thereby blocking the transcription of the downstream gene. This system was successfully adapted to different plant systems like tobacco, tomato, potato, and even BY-2 suspension cells (Bortesi et al., 2012). In 1994, Weinmann et al. developed the Tet-off system. This system uses the fusion of a transactivator VP16 to tetR to activate target gene expression by binding to TRE sequences in the absence of Tet. More recently, Zhou et al. (2017) connected a heat shock transcription factor to TetR and placed it under the control of a heat shock promoter in tobacco plants. This allowed the chimeric TF to be activated by high temperatures and deactivated by tetracycline.

Other traditionally used chemical sensor is based on steroid-regulation. The GVG/UAS sensor is based on the steroid hormone receptor regulation mechanism found in vertebrates (Aoyama & Chua, 1997). This system utilizes a chimeric transcription factor called GVG, which is composed of the DNA-binding domain from the yeast transcription factor Gal4, the transactivation domain from the herpes simplex virus VP16 protein, and the glucocorticoid receptor (GR) from rats. The system is induced using dexamethasone (DEX), a potent synthetic glucocorticoid. When DEX binds to the GR domain of the chimeric GVG protein in cells, it triggers the translocation of GVG to the nucleus. Once in the nucleus, the Gal4 domain specifically binds to upstream activating sequences (UAS) in the promoter region of the target gene, and the VP16 domain initiates the transcription of the gene located downstream of the UAS region.

More recently developed chemical sensors, such as insecticide-, ethanol- or copper-, offer practical and more sustainable applications in field conditions than Tet and steroids (reviewed in Omelina et al., (2022)). The insecticide-induced systems use ecdysone receptors (EcRs) from insects, engineered to recognize synthetic ecdysone agonists. The EcR is combined with the transactivator VP16 and forms a dimer with the retinoid X receptor (RXR). This complex binds to a unique synthetic response operator not targeted by natural nuclear hormone receptors. When ecdysone analogues, like muristerone A and ponasterone, bind to the receptor dimer, it triggers the transcriptional upregulation of genes downstream of the synthetic operator (Koo et al., 2004). The ethanol-induced system uses the fungal transcriptional factor AlcR, which binds to the pAlcA promoter in the presence of ethanol, activating the transcription of the target gene (Caddick et al., 1998; Li et al., 2005). The copper-induced system is based on the copper-responsive factor CUP2, which, upon binding of copper ions, undergoes a conformational change triggering transcription from the copper binding site (CBS) operator (Mett et al., 1993). Saijo & Nagasawa (2014) optimized the original copper-induced system by increasing the number of CBS repeats to recruit more CUP2 molecules and by fusing the transcriptional activator VP16 to CUP2 to enhance its activation capability.

The VP16, previously referred to in the creation of the chimeric TetR, GVG, and copper sensors, is a potent transactivator domain (TAD) derived from the herpes simplex virus (Li et al., 2013). Enhancing its activation potential, various VP16 domains can be fused in tandem to create synthetic activation domains such as VP64 and VP128, expanding the activation range of target genes (Li et al., 2017; Ordiz et al., 2002). Furthermore, VP16-based transactivator can be combined with other transactivation domains, as seen in the tripartite VP64-p65-Rta (VPR) transactivation domain (Chavez et al., 2015) and in TV, which includes six copies of the TALE transactivation domain motif (6TAL) and VP128, corresponding to eight repeats of the VP16 motif (Li et al., 2017).

Aside from viral-based elements, the Gal4 TAD from fungi, as mentioned above, is another effective option. Originally characterized by Carey et al. (1990), Gal4 has been widely employed for transcriptional activation in various eukaryotic systems (Elliott & Brand, 2008). Recently, a library of synthetic transcription factors with varying activation capacities was generated by combining the Gal4 DNA binding domain with truncated putative transcriptional effector domains (Hummel et al., 2023). The Gal4 domain is based on the Nine amino acids Transactivation Domain, 9aaTAD, which is universally recognized by the transcriptional machinery in eukaryotes. The 9aaTAD family comprises over 40 members, highlighting its evolutionary conservation and functional versatility. Additionally, an online 9aaTAD prediction service has been developed, which is relevant in designing new TAD tools (Piskacek et al., 2016). Another fungal-based transcription factor that functions similarly to Gal4 is mlcR. It specifically binds to a 14-base pair operator of mlcA and mlcC to activate expression (Baba et al., 2006). Recently, Pfotenhauer et al. (2024), generated chimeric transcription factors by fusion the DNA-binding domain from mlcR to different

transactivation domains, standing out the QF from the *Neurospora crassa* Q-system (Persad et al., 2020).

Plant endogenous effector domains and receptors can also provide a source for synthetic transcriptional factors. Examples include the ERF2 and EDLL domains from the Ethylene Response Factor (ERF) family (Li et al., 2013; Tiwari et al., 2012). The ERF family also contains repression domains for transcription, such as the EAR motifs. For instance, the SRDX (Hiratsu et al., 2003) and the BRD (Ikeda & Ohme-Takagi, 2009) domains derived from ERF3 (Ohta et al., 2001) are commonly applied for transcriptional repression. Moreover, plant receptors can also be engineered to recognize novel ligands. A very famous example is based on PYR1 (Pyrabactin Resistance 1), an abscisic acid (ABA) receptor characterized by its flexible ligand-binding pocket and the requirement for ligand-induced heterodimerization. Using a structure-based design that relies on ligand docking, Lozano-Juste et al., (2023) created an ABA receptor agonist, iSB09, and engineered a CsPYL1 ABA receptor that binds iSB09 effectively. This optimized receptor-agonist pair activates ABA signalling, significantly enhancing drought tolerance. The engineering of PYL1 receptors serves as a foundation for developing sensors with nanomolar sensitivity to various small molecules, including natural and synthetic cannabinoids and multiple organophosphates (Beltrán et al., 2022).

Recent breakthroughs in machine learning (ML), particularly deep learning algorithms, have opened doors for optimized designs of proteins like sensors. ML-directed evolution (MLDE) efficiently generates improved protein variants compared to mutant screening traditional methods. One successful example is cluster learning–assisted directed evolution (CLADE) by Qiu et al., (2021) to guide protein redesign. This approach addresses the complexity of protein-substrate/ligand interactions and the influence of distant amino acids within the active site, which requires extensive mutant screening. In response to this challenge, Wittmann et al. (2021) proposed a model that improves efficiency by excluding less informative data points during training and identifies optimal protein sequences with the highest fitness. While not yet widely adopted in plant protein engineering, MLDE offers an efficient solution for high-throughput protein engineering (Iram et al., 2024).

Apart from direct evolution, computational prediction of protein structure also can have a critical role in modelling synthetic sensor modules by *de novo* design (Kortemme, 2024). AlphaFold 2 (Jumper et al., 2021) and RoseTTAFold (Baek et al., 2021) were pioneering deep learning-based tools for highly accurate prediction of protein structures. Both methods rely on Multi-Sequence Alignment (MSA), so they are limited to naturally occurring protein sequences. As an alternative, other tools like RGN2 (Chowdhury et al., 2022) make predictions based on language models, making them suitable for predicting the structures of orphan proteins without homologous sequences.

Recently, Watson et al (2023) introduced the RoseTTAFold diffusion (RFdiffusion) method, which enables accurate and versatile protein design by complementing structure prediction networks with a generative diffusion model. Diffusion modelling allows the prediction of raw atomic coordinates, providing precise insights into both structure and molecular interactions. Building on this generative diffusion approach, Abramson et al (2024)

presented AlphaFold 3 for predicting the joint structure of complexes that include proteins, small molecules, nucleic acids, ions, and modified residues. AlphaFold 3 then allows the combination of protein structure prediction with ligand docking, which has the potential to facilitate the creation of highly specific and effective synthetic sensors.

### 2.3.2. Processor modules

The simplest processor module in genetic circuits is a cis-regulatory box located in a proximal promoter. This basic operator, also known as "BUFFER" gate (for an activator) or NOT gate (for a repressor) in Boolean logic (Figure 5), functions as a basic on/off switch, regulating the transcription of a single downstream gene upon the activation mediated by binding of the sensor. However, more sophisticated processors are required to develop complex genetic programs. To master the control of gene expression in plants, a battery of processor modules can manipulate the signal received from a sensor in different ways, allowing a wide range of transcriptional outputs like integration of multiple input signals, storage of the signalling, and signal amplification or distribution (Vazquez-Vilar et al., 2023).

Regarding the integration of multiple input signals, Synthetic Biology aims to mimic computational operating systems. Computer programming relies on a fundamental concept called Boolean logic, which consists of making decisions based on just two values: 0 (false) and 1 (true). By establishing conditional relationships between 0 or 1 input values, Boolean logic allows computers to perform complex calculations and make decisions. Based on this approach, computer circuits use components known as logic gates. These gates take one or multiple inputs (0s or 1s) and perform operations to produce a unique output value of 0 or 1. The two most basic, single-input logic gates are BUFFER and NOT gates mentioned above. In the BUFFER (aka YES) gate, the output value is the same as the input value; in the NOT (INVERTED) gate, the output is the opposite of the input. The next step in complexity is given by two-input gates. In the AND gates, both inputs must be 1 for an output of 1, and in the OR gates, at least one input must be 1 for an output of 1. There are also other more complex gates like NAND (acting like an AND followed by a NOT), NOR (functioning as a NOT after an OR), XOR (giving a true output only when inputs differ), and XNOR (providing a true output only when inputs are the same). Beyond these, there are even more complex gates like IMPLY, which expresses a one-way implication, and NIMPLY (NOT IMPLY), signifying the absence of such an implication. This wide variety of logic gates allows for creating intricate circuits capable of processing complex combinations of inputs and outputs (Ferreira et al., 2023). Like computers, Boolean logic gates have been implemented in Plant Synthetic Biology to introduce new functions. Recently, Brophy et al (2022) showcased the significant potential of logic gates for plant reprogramming. In this work, a library of synthetic transcription factors was created by combining various DNA binding domains with different activation/repression domains, along with a battery of synthetic promoters featuring varied repetition numbers and positioning of binding sites for the transcription factors. By combining these components, different logic gates that can be applied to predictably reprogramme root structure were constructed, generating user-defined spatial patterns of

gene expression. Another remarkable application of logic gates in plants was demonstrated by Lloyd et al (2022). Their work combined logic operations with recombinase enzymes to provide memory to the circuit. A transcription terminator, flanked by specific recognition sequences for recombinase enzymes, was placed between the promoter and the CDS, effectively blocking transcription. The recombinase enzymes acted as molecular switches, rearranging DNA sequences to control gene expression. This approach implies a cellular memory that maintains the output even after the initial signal fades. Using this approach, they were able to combine different recombination systems to create various logic gates, including OR, AND, NOT, NIMPLY, and NAND. However, a drawback is the irreversible nature of these changes, limiting the adaptability of the system for certain applications.

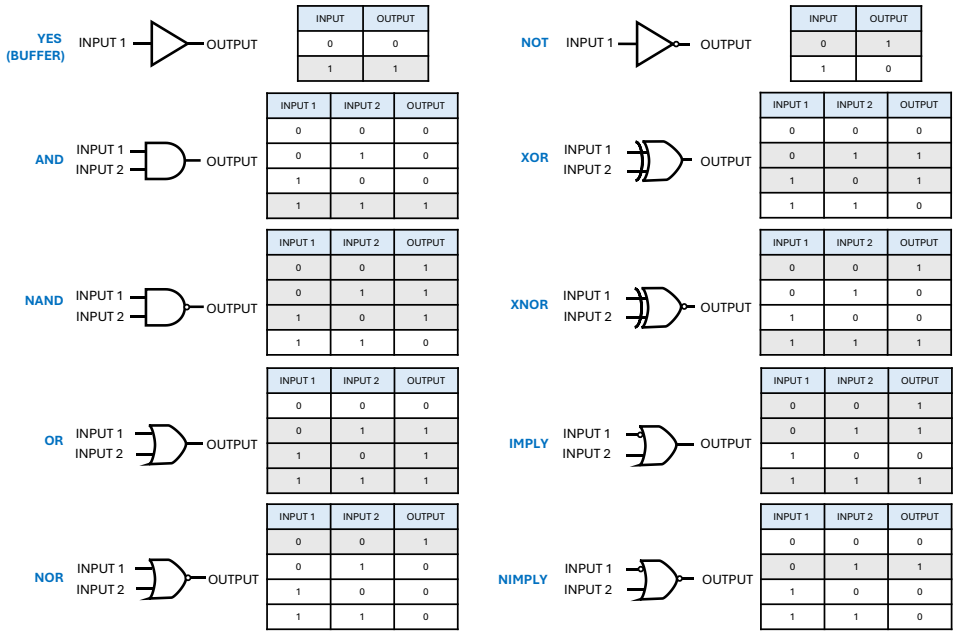



**Figure 5. Boolean logic gates in synthetic gene circuits, including their operation symbols and corresponding truth tables.** A binary notation is used in the truth tables where '1' indicates the presence of a signal, and '0' indicates its absence. The rows in the truth table where the output is '1' are highlighted in grey, signifying an active output response. This binary system allows for the conversion of input signals into a defined output based on the logic gate's function.

Following the mentioned work from Lloyd et al., (2022), another key processor involving memory is the toggle or bistable (ON and OFF) switch (Medford & Prasad, 2016). This system, as mentioned above, eliminates the need for constantly adding external triggers to maintain a desired output, as was the case for sensor modules. Instead, it relies on a single initial detection of the input signal to flip the switch and maintain the output state. Bernabé-Orts et al., (2020) created a bistable memory switch in *Nicotiana benthamiana*, utilizing DNA recombination. This switch employed two reporter genes, and a reversible promoter flanked by specific recombination sites. The orientation of the promoter, and consequently the

active gene, was controlled by alternating between a φC31 integrase enzyme and its partnering recombination directionality factor (RDF).

As depicted above, recombinases are versatile processor modules that can be used for logic gates and memory functions in genetic circuits. Additionally, they can also be used for amplification purposes, as shown in the work of Guiziou et al., (2023). This study focused on using serine integrase recombinases expressed from promoters responding to TFs that are active in specific stages of lateral root development. When integrases were expressed, they mediated an irreversible site-specific DNA recombination event, which was visualized by the switch between fluorescent reporter proteins. Acting as an intermediary between the transcription factor sensor and the reporter proteins, the recombinase amplified the reporter signal and permanently marked all descendant cells. This method enabled easy visualization and recording of gene expression during plant development.

Building on the concepts of signal amplification and distribution, signal amplifiers and signal distributors (multiplexers) have become crucial for routing efficiently the transcriptional output of synthetic genetic circuits towards downstream actuators. While signal amplifiers help boost weak signals, ensuring that they are strong enough to be transmitted, multiplexers distribute these signals to the appropriate actuators. Next, we will focus on two types of processors that operate in different ways. One is a dead Cas9 (dCas9)-based programmable transcription activator, which distributes and directly amplifies transcriptional signals. The other is based on the replicative machinery of geminiviruses, which indirectly amplifies gene expression by driving DNA replication.

#### 2.3.2.1. dCas9-based programmable transcriptional activators

The dead Cas9 (dCas9)-based programmable transcription activators (PTAs) provide a versatile and powerful tool for signal multiplexing in plant synthetic genetic circuits. The amplification signal provided by dCas9 is twofold. First, it can simultaneously modulate the expression of multiple genes (Chen & Qi, 2017). Second, its activating potency is enhanced because multiple activation domains can be coupled to dCas9 (Moradpour & Abdulah, 2020). Moreover, it can be directed to any endogenous genes, representing a 'universal' connection between orthogonal sensors and plant actuators.

The dCas-based PTAs were derived from the CRISPR-Cas system, a naturally occurring immune system in some prokaryotic cells against invading bacteriophages and plasmids by eliminating foreign DNA (Mojica et al., 2005). 'CRISPR' stands for Clustered Regularly Interspaced Short Palindromic Repeats, and 'Cas' for CRISPR-associated proteins. The palindromic repeats flank short spacers coming from past invaders. These spacers are transcribed into non-coding RNA molecules and form complexes with the Cas proteins, which have endonuclease activity. The CRISPR-Cas complex identifies and cleaves complementary target DNA by creating a double-strand break (Barrangou et al., 2007). The most widely used CRISPR-Cas9 system is derived from *Streptococcus pyogenes* (SpyCas9). Cas9 is part of Class 2 of Cas systems, which utilizes a single protein for endonuclease

activity. Within this class, Cas9 falls under Type II, characterized by a multidomain structure. Cas9 has two nuclease domains: the HNH domain cuts the complementary DNA strand, and the RuvC domain cuts the non-complementary strand in the same position, resulting in a blunt double-strand break (Makarova et al., 2020; Takasugi et al., 2022).

The CRISPR-Cas9 system relies on two recognition mechanisms for targeting DNA with high specificity. The first mechanism involves the recognition of a short DNA motif adjacently located on the 3′ end of the protospacer (target) sequence, termed the Protospacer Adjacent Motif (PAM) (Figure 6). The second layer of specificity relies on perfect base pairing between the CRISPR RNA (crRNA) and the target DNA sequence (Mojica & Montoliu, 2016). The crRNA is the molecule that directs the Cas protein to the specific DNA sequence for cleavage. The crRNA pairs with the transactivating CRISPR RNA (tracrRNA) to create a mature guide RNA (gRNA) that keeps stably bound to Cas9 (Chylinski et al., 2014). When re-purposing the CRISPR-Cas system for biotechnological applications, a single guide RNA molecule (sgRNA) containing both the customizable crRNA fused to the tracrRNA was developed to simplify the synthesis and expression processes. The sgRNA, commonly referred as gRNA for simplicity, is made of a 20-nucleotides sequence complementary to the target sequence (spacer) and an RNA scaffold recognized by Cas9, mimicking the function of the tracrRNA (Ran et al., 2013).

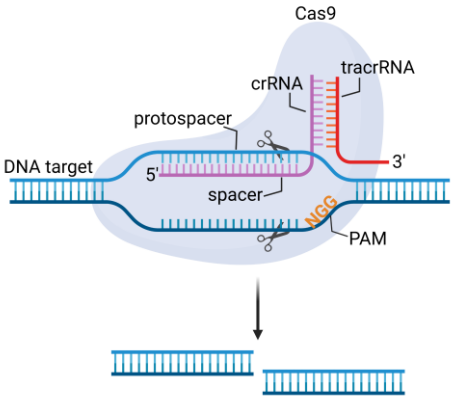



**Figure 6. The CRISPR-Cas9 system**. The CRISPR-Cas9 system consists of a Cas9 nuclease guided by a dual RNA structure composed of a CRISPR RNA (crRNA) and a trans-activating CRISPR RNA (tracrRNA). The crRNA contains the spacer sequence, which is complementary to a target DNA sequence (protospacer), enabling the precise binding of Cas9 to the DNA target. Both the crRNA and tracRNA together are known as guide RNA (gRNA). The tracrRNA connects the crRNA with Cas9 to form the ribonucleoprotein complex. Recognition of a specific protospacer adjacent motif (PAM) on the target by Cas9 is essential for destabilizing the adjacent DNA sequence. The specific PAM recognized by Cas9 comprises three nucleotides NGG. Once the target sequence is recognized, Cas9 induces a blunt double-strand break, which the cell can repair through non-homologous end joining or homology-directed repair.

Unlike the traditional Cas9, which introduces double-strand breaks in DNA, dCas9 is engineered to be catalytically inactive and instead binds to specific DNA sequences without cutting them (Qi et al., 2013). Precise control and amplification or repression of gene expression can be achieved by fusing dCas9 with transcriptional activator or repressor domains and targeting it to gene promoters. These regulation strategies—CRISPRa for activation and CRISPRi for inhibition or repression—allow for robust, easy, and scalable modulation of target genes (Moradpour & Abdulah, 2020). Early efforts to develop Cas9-PTAs in plants utilized transcriptional activator domains (TADs) directly fused to the dCas9 protein, resulting in modest activation rates (Park et al., 2017; Vazquez-Vilar, Quijano-Rubio, Fernandez-del-Carmen, et al., 2017). A next generation of PTAs aimed to enhance activation by incorporating multiple TADs within a single dCas9 protein (Figure 7). PTAs such as the SunTag (Tanenbaum et al., 2014a) employ multi-epitope tags to bind multiple TADs to the dCas9 protein. Other PTAs, like the CRISPR scaffold RNAs (scRNA) systems (Zalatan et al., 2015) and the Synergistic Activation Mediator (SAM) (Konermann et al., 2015a; Y. Zhang, Yin, et al., 2015), use RNA aptamers in the gRNA scaffold as anchoring sites for TADs. In pursuit of improved plant dCas9-PTAs, Selma et al., (2019) evaluated the promoter activation ability of different combinations of SunTag, scRNA, and SAM-based strategies. From this analysis, a novel dCas9-PTA, named dCasEV2.1, was developed and resulted in achieving the highest activation values when targeting both transiently transformed or genome-integrated promoters. This dCasEV2.1 tool contained an EDLL activation domain directly fused to the dCas9 protein and a VPR activation domain bound through the viral MS2 coat protein to the gRNA2.1 aptamer, a mutated version of the gRNA2.0 aptamer used in the scRNA strategy. dCasEV2.1 notably activated the *N. benthamiana* Dihydroflavonol-4-reductase (DFR) gene, which has very low basal expression levels, by up to 10,000-fold. Additionally, analysis by RNAseq reported that dCasEV2.1 exhibited high specificity, with minimal off-target effects along the genome. The combination of high activation strength, remarkable targeting specificity, and orthogonality positions dCasEV2.1 as a powerful transcriptional amplification processor module.

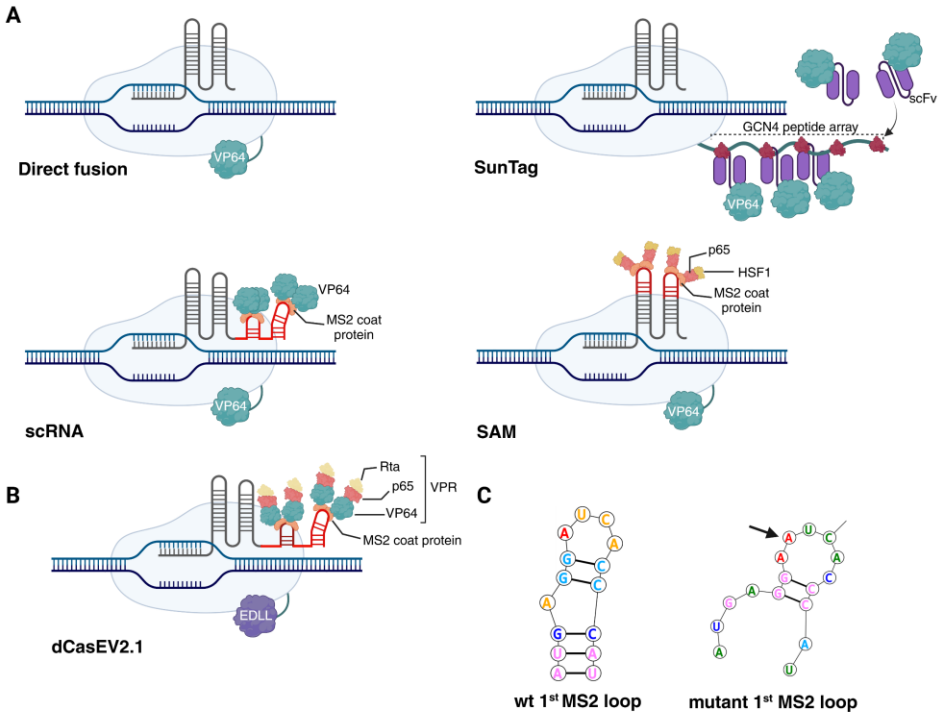



**Figure 7. dCas9-Programmable Transcriptional Activators (PTAs) and MS2 Aptamer Modifications. A)** Different strategies for designing dCas9-programmable transcriptional activators (PTAs). The direct fusion approach involves the gene fusion of a transcriptional activator like VP64 to dCas9. The SunTag system uses a GCN4 polypeptide chain that is bound by multiple scFv antibody fragments-transactivator complexes. The scaffold RNA (scRNA) strategy utilizes single guide RNA (gRNA) in combination with MS2 RNA aptamers to recruit transactivators. Two MS2-binding stem loops are fused at the 3′ end of the sgRNA (sgRNA + 2×MS2). The Synergistic Activation Mediator (SAM) strategy uses a modified gRNA where the minimal hairpin MS2 aptamers are bound to the tetraloop (loop 1) and stem loop 2 of sgRNA. The MS2 coat protein recruits transactivators like p65 and HSF1. **B)** The dCas9-PTA variant known as dCasEV2.1, employed in this thesis. The dCas9 protein is fused to the EDLL transactivator, which is known for its robust transcriptional activation potential. Additionally, the scRNA strategy is employed, where the gRNA binds the VPR transactivator complex (comprising VP64, p65, and Rta), further enhancing transcriptional activation. The first MS2 aptamer in this strategy, shown in dark red, has a point mutation compared to the original scRNA. **C)** Comparison between the original MS2 aptamer loop in scRNA strategy (wt 1st MS2 loop) and the mutant 1st M2S aptamer of dCasEV2.1. The mutated scaffold 2.1 contains an insertion of an adenine nucleotide, as indicated by the arrow, which alters the loop structure. This modification may affect the binding affinity and stability of the MS2 coat protein interaction, enhancing transcriptional activation. Figure includes images from Biorender (biorender.com). Loop structure prediction was built by RNAstructure server at https://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html.

#### 2.3.2.2. Geminiviral-based replicative systems

Plant viruses have long been recognized for their efficiency in hijacking host cellular machinery to amplify their own genetic material. By repurposing their replication strategies, plant viral systems can be employed to boost gene expression within synthetic circuits (Salazar-González et al., 2015). One prominent example is based on the use of genetic elements from geminiviruses, which employ a rolling-circle mechanism to replicate their DNA genome. The replication machinery from geminiviruses can serve as a basis for developing powerful signal amplifiers by means of increasing the copy number of a target gene. The increase in DNA copies subsequently results in enhanced transcriptional activity, as more templates are available for the RNA polymerase to transcribe into messenger RNA (mRNA). The amplified mRNA levels then translate into higher protein production, magnifying the output of the synthetic gene circuit (Baltes et al., 2014).

Geminiviruses (*Geminiviridae* family) are non-enveloped plant viruses with a single-stranded (ss) circular DNA genome. They infect a wide variety of crops, causing significant billionaire economic losses globally, particularly in warmer regions (Hanley-Bowdoin et al., 2013; Kenyon et al., 2014; Varma & Malathi, 2003). Among the fourteen genera most recently classified by the International Committee on Taxonomy of Viruses (ICTV) (Fiallo-Olivé et al., 2021) according to the genomic organization, insects that spread them, host range, and sequence similarity-, the three most prevalent and well-characterized ones are Mastrevirus, Curtovirus, and Begomovirus (Lozano-Durán, 2016; Wu et al., 2022) (Figure 8). Both mastreviruses and curtoviruses have a single genome molecule and are spread by leafhoppers. Mastreviruses can infect both monocotyledonous and dicotyledonous plants, while curtoviruses are limited to dicots. Begomoviruses, the largest genus, can have either monopartite or bipartite genomes (DNA-A and DNA-B) and are transmitted by whiteflies (Wu et al., 2022). The geminivirus genome, one or two circular DNA molecules (monopartite or bipartite, respectively), is enclosed within a twinned incomplete icosahedral particle. The genome ranges from 2.5 to 5.2 kb, making them some of the smallest known autonomously replicating viruses (Zerbini et al., 2017). Geminiviruses produce multifunctional proteins essential for viral replication, transcription, suppression of gene silencing, encapsidation, movement, vector transmission, and pathogenesis (reviewed in Fondong, 2013; Hanley-Bowdoin et al., 2013; Luna & Lozano-Durán, 2020; Wu et al., 2022). Common and basic proteins for all the geminiviruses are the Replication-associated protein (Rep), Capsid Protein (CP), and Movement Protein (MP). Due to their limited coding capacity, the life cycle of geminiviruses extensively relies on the cellular machinery of the plant host.

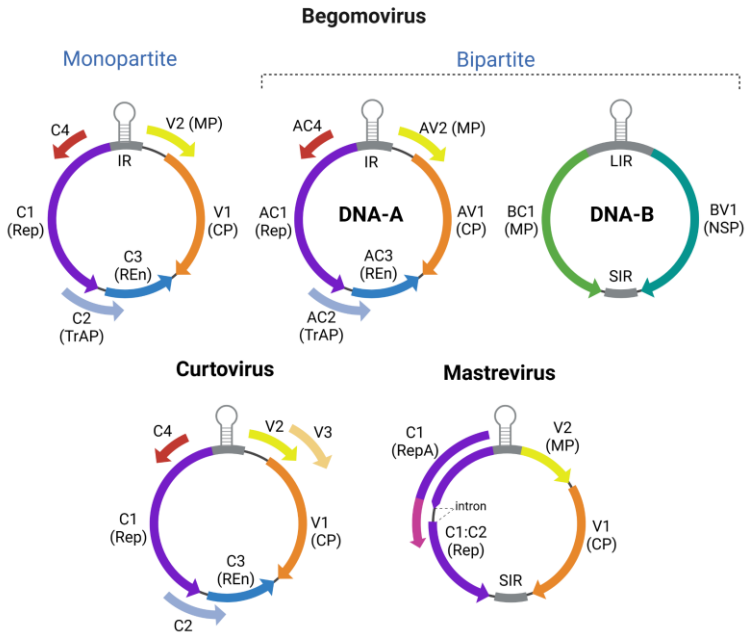



**Figure 8. Geminivirus genome organization of the three best-known genera: Begomovirus, Curtovirus and Mastrevirus.** The Geminivirus genome comprises single-stranded circular DNA. Begomovirus can be monopartite (one DNA molecule) or bipartite (two DNA molecules). Each open reading frame (V1, V2, C1, C2, etc) is color-coded based on the function of the encoded protein: Replication-associated protein (Rep), Replication Enhancer (REn), Transcriptional Activator Protein (TrAP), Capsid Protein (CP), Movement Protein (MP), Nuclear Shuttle Protein (NSP). Key elements involved in DNA replication, including the Intergenic Region (IR), Long Intergenic Region (LIR), and Short Intergenic Region (SIR), are annotated. The IR/LIR comprises a stem-loop structure that harbours the origin of replication. Adapted from Wu et al. (2022). Figure includes images from Biorender (biorender.com).

Every geminivirus genome contains an approximately 300-base pair region known as the intergenic region (IR) or the long intergenic region (LIR) in some genera like Mastrevirus (Figure 9). This intergenic region contains a conserved GC-rich repeat inverted sequence, resulting in a hairpin-loop structure. The loop, usually 10-13 nucleotides long, includes the nonamer sequence TAATATTAC, which acts as the origin for viral dsDNA replication. Mastreviruses possess an additional small intergenic region (SIR) that is essential for the initial stage of viral DNA replication. The binding region for Rep in the IR/LIR sequence contains repeated elements called iterons, which vary in sequence among different geminivirus species, granting specificity for Rep recognition. The IR/LIR sequence also contains promoter elements, like TATA boxes, for RNA polymerase II-driven bidirectional transcription of the virion-sense and complementary-sense strands (Gutierrez, 1999; Lazarowitz & Shepherd, 1992; Wang et al., 2022; Wu et al., 2022).

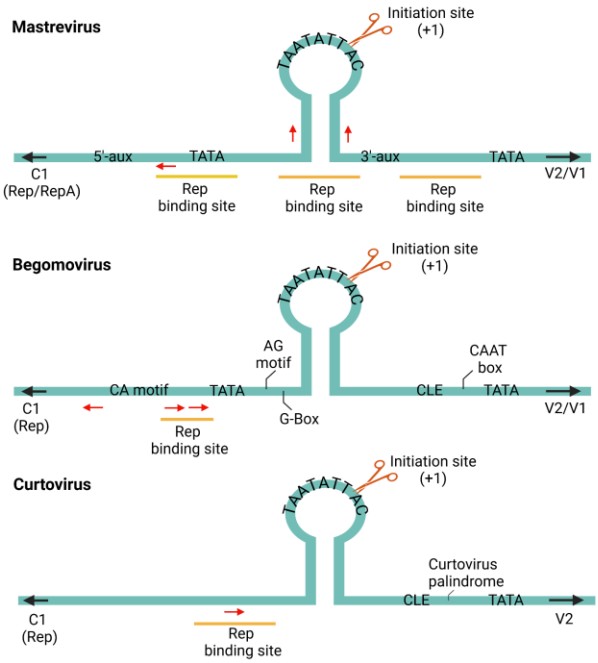



**Figure 9. Geminiviral intergenic regions harbouring the origin of replication in Begomovirus, Mastrevirus, and Curtovirus**. All three genera share a conserved origin of replication sequence, TAATATT↓AC, embedded in the stem-loop DNA structure of the Intergenic Region IR (Begomovirus and Curtovirus) or Long Intergenic Region LIR (Mastrevirus). The replication origin is essential for initiating rolling circle replication (RCR) at the initiation site (+1). Key elements such as iterons (red arrows), Rep-binding sites (underlined), and specific motifs crucial for replication are depicted. Iterons are highly conserved, genus-specific sequences that appear to provide specificity for Rep protein binding. Auxiliary sequences (aux) are flanking regions that stimulate replication. CA motif enhances DNA replication. The AG motif and G-box likely play a role in origin utilization and have been preserved by evolutionary pressure. TATA represent the TATA box required for the initiation of transcription of the downstream open reading frames, as well as the CAAT box. CLE is a Conserved Late Element targeted by the geminiviral transcriptional activator protein (TrAP) to regulate the viral promoter. Begomovirus and Mastrevirus annotations adapted from Wu et al. (2022); Begomovirus annotations complemented by information from Urbino et al. (2022), Vinoth Kumar & Shivaprasad (2020) and Williams et al. (2024); Curtovirus annotations from Cantú-Iris et al. (2019); Hur et al. (2007); Stanley (2008). Figure includes images from Biorender (biorender.com).

The geminivirus infection begins when an insect vector carrying virions feeds on the plant phloem, introducing the virions into the companion cells. Once inside, the viral ssDNA is uncoated and translocated to the nucleus, where the (-) strand (negative strand) is synthesized by the plant DNA polymerase α to form double-stranded DNA (dsDNA) (Wu et al., 2021). In the case of mastreviruses, the synthesis of the (-) strand specifically initiates at the SIR (Gutierrez, 1999; Shafiq et al., 2020). The early genes, such as Rep, are transcribed by RNA polymerase II. Rep plays a critical role in initiating and terminating viral DNA replication by nicking and ligating at specific sites within the intergenic region (IR or LIR). However, the

actual replication of viral dsDNA is carried out by the plant DNA polymerase δ (Wu et al., 2021). This replication occurs through rolling-circle replication (RCR) and recombination-dependent replication (RDR) mechanisms (Jeske, 2001), resulting in the production of dsDNA molecules that form minichromosomes when they associate with histones. These dsDNA minichromosomes act as templates for synthesizing more ssDNA, which the CP then coats to create new virions and move from cell to cell facilitated by the MP. The minichromosome configuration also promotes the transcription of viral open reading frames (ORFs) from both strands by RNA polymerase II (Borah et al., 2016). For more details about the replication process, see Figure 10.

The geminiviral Rep is a multivalent oligomeric protein displaying DNA binding, ATPase, cleavage, ligase, and helicase functions (reviewed in Ruhel & Chakraborty (2019); Shakir et al. (2023); Wu et al. (2022). Rep is encoded in the complementary strand of the geminivirus genome (C1 ORF in Begomovirus and Curtovirus and C1:C2 in Mastrevirus, see Figure 7). The Rep protein in mastreviruses is expressed through alternative splicing when the intron in C1 is processed. When this intron is not processed, RepA (C1) is expressed (Boulton, 2002).

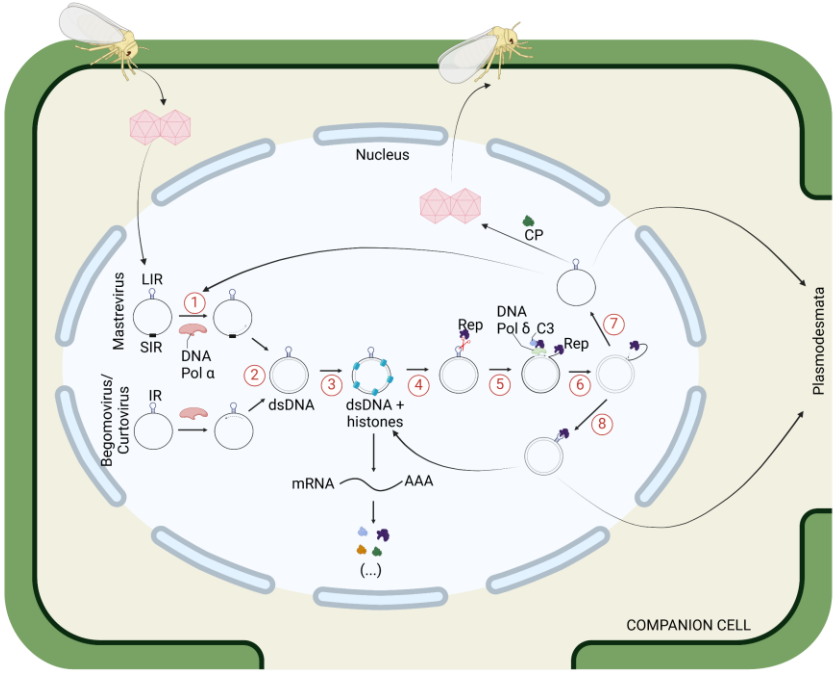



**Figure 10. Geminivirus life cycle and the hypothetical DNA replication process within the plant cell nucleus**. An insect vector introduces geminivirus virions into a companion cell when feeding on the plant. The viral capsid is disassembled, releasing the genomic circular single-stranded DNA (ssDNA), which is transported to the nucleus, where DNA replication takes place. First, the viral ssDNA is converted into a double-stranded DNA (dsDNA) intermediate with the assistance of host DNA polymerase α to synthesize the (-) strand. The dsDNA intermediate then undergoes replication through rolling-circle replication (RCR) and recombination-dependent replication (RDR) assisted by DNA polymerase δ. In mastreviruses, the synthesis of the (-) strand starts in the Short Intergenic Region (SIR).

In begomovirus and curtovirus, the synthesis of (-) strands apparently starts by a region proximate to IR. The DNA pol α mediates the extension of the (-) strand (1) leading to the formation of a full-length covalenty enclosed dsDNA molecule (2). The relaxed closed DNA intermediate is supercoiled by a mechanism dependent on a DNA gyrase where the dsDNA associates with host histones to create a geminiviral minichromosome (3). This supercoiled dsDNA serves as a template for transcription by RNA polymerase II, leading to the synthesis of viral proteins, including Rep. Rep protein binds to the origin of replication of the dsDNA, inducing a nick (4) that initiates rolling-circle replication (RCR) with the help of DNA polymerase δ. In the initiation of RCR, the Rep protein triggers a nucleophilic attack via the hydroxyl (OH) group of a conserved tyrosine residue on the phosphodiester bond between the final T and A of the invariant loop sequence in the viral genome—a process known as transesterification (5). This reaction originates a nick that provides a 3'-OH group, which can then be extended by the host cell's DNA polymerase δ. Meanwhile, Rep remains covalently bound to the 5' end of the DNA strand. Additionally, Rep recruits the host DNA replication machinery to the viral genome and works alongside DNA polymerase δ and the replication enhancer protein C3 in the displacement of the (+) strand during replication. Once the replication of the (+) strand is complete, Rep facilitates the ligation and release of an 'old' circular single-stranded DNA (ssDNA) molecule (7), which can be encapsidated by the capsid protein (CP). Rep also mediates the ligation of the newly synthesized DNA strand (8), leading to the formation of a new dsDNA molecule, ready to re-enter the replication cycle. Newly synthesized viral DNA may move within and between cells, establishing systemic infection, or be encapsidated by CP for transmission by insect vectors to other plants. Adapted from Wu et al. (2022) and complemented by information from Gutierrez. (1999); Wu et al. (2021). Figure includes images from Biorender (biorender.com).

Both Rep and RepA can reprogram the cell cycle to promote re-entry into the S-phase, which is permissive for DNA replication (Wu et al., 2022). For example, Rep has been demonstrated to interact with the retinoblastoma/E2F transcription regulatory network, which enables terminally differentiated plant cells to re-enter the cell cycle and regain the ability to sustain elevated levels of DNA replication (Hanley-Bowdoin et al., 2004). Interestingly, geminiviruses can induce DNA replication competence in cells without causing cell proliferation, indicating that these viruses may exploit the endocycle but not the mitotic division (Bass et al., 2000; Nagar et al., 2002). Moving to the direct implication of Rep in geminiviral genome replication, as previously mentioned, Rep triggers replication of the dsDNA genome by cleaving one strand (nicking) within the IR/LIR TAATATTAC nonamer. Specifically, the nicking happens at the 5′ side of the last adenosine (TAATATT↓AC), resulting in a 5′ phosphoryl group, covalently linked to a tyrosine residue of Rep, and a 3′ hydroxyl group acting as a primer for the synthesis of the viral-strand DNA (C. Wang et al., 2022). Once the viral-strand DNA is fully displaced, it is released through a transesterification reaction catalyzed by Rep. This involves the transfer between the Rep-fixed 5′ end and the 3′ hydroxyl of thymidine 7 within the TAATATTAC nonanucleotide (Bonnamy et al., 2023) (Figure 11).

In addition, studies in *Geminiviridae* provided evidence for an extra replication mode based on recombination-dependent replication (RDR) (Jeske, 2001). This replication consists of homologous recombination between the genome dsDNA and partially replicated molecules of ssDNA (reviewed in Bonnamy et al. (2023)). Apart from prompting the replicative environment in the cell and its role in replication by nicking and ligating the DNA genome,

Rep also participates in the formation of the replisome: the multiprotein complex carrying out the geminiviral DNA replication, by recruiting plant replication factors such as the proliferating cell nuclear antigen (PCNA) (Castillo et al., 2003). Additionally, Rep also acts as a transcription regulator by mediating the repression of its own promoter, which helps to activate the expression of the late viral genes (Eagle et al., 1994). Finally, Rodríguez-Negrete et al. (2013) demonstrated that the geminivirus Rep protein mediates the suppression of transcriptional gene silencing by interfering with the DNA methylation machinery of the plant. Moreover, in mastreviruses, both Rep and RepA have been shown to act as RNA silencing suppressors (Liu et al., 2014).

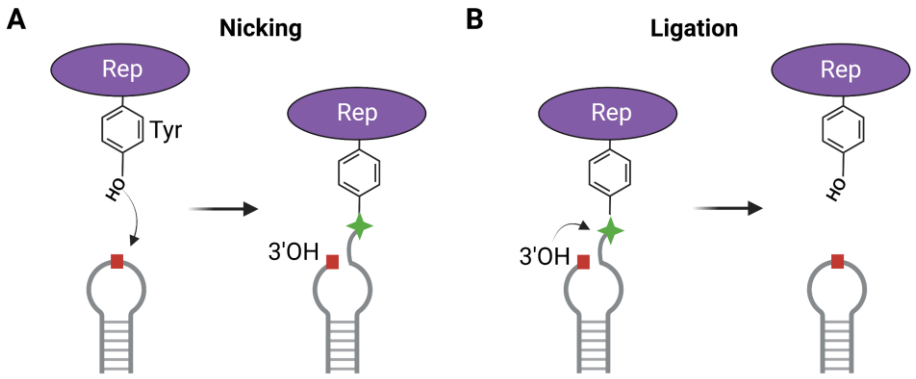



**Figure 11. Mechanisms of A) nicking and B) ligation of DNA catalyzed by geminiviral Rep proteins**. The nicking site, depicted in a hairpin-like structure, is highlighted by a red rectangle. The green star represents the phosphotyrosine bond formed between Rep and the 5′-end of the DNA. This reaction is reversible, as the released 3′-OH end of DNA can function as a nucleophile on the phosphotyrosine bond. Adapted from Wawrzyniak et al. (2017). Figure includes images from Biorender (biorender.com).

After this background on the cell biology and replication process of geminiviruses, we can identify the essential elements for creating a geminiviral amplification processor for plant synthetic circuits. The key components are the Replication-associated proteins (Rep and, in the case of mastreviruses, also RepA) and the intergenic regions (IR/LIR, and SIR in the case of mastreviruses) that act as replication origins. To create a replicative geminiviral vector carrying a gene of interest, the only requirement is to flank the gene by the intergenic regions (Zhang & Mason, 2006). The Rep/RepA bicistronic gene can be trans-delivered to the plant in a different binary vector and will promote the circularization and the replication of the gene of interest after binding the flanking LIR regions (Chen et al., 2011). This versatile system has been successfully used to increase the expression of vaccines, antibodies, enzymes, and virus-like particles in plants (Bhattacharjee & Hallan, 2022; Chen et al., 2011), showcasing the potential of a geminiviral processor in Synthetic Biology applications.

## 3. The need of synthetic gene circuits in Molecular Farming context

The use of plants as "biofactories" to produce recombinant proteins or metabolites for pharmaceutical or industrial applications is referred to as Molecular Farming. By leveraging plants as hosts for these essential products, it becomes possible to reduce reliance on non-renewable resources and create more environmentally friendly production systems (Fischer & Buyel, 2020). Currently, the most popular host for Molecular Farming is *Nicotiana benthamiana*, a close relative of field tobacco (*Nicotiana tabacum*) native to Australia (Ranawaka et al., 2023). *N. benthamiana* offers several advantages for recombinant protein expression. One of its key attributes is its suitability for rapid, transient gene expression via agroinfiltration. This technique involves the infiltration of an *Agrobacterium tumefaciens* suspension carrying the genes of interest into leaf tissue via a syringe or vacuum. The transfer of the genes of interest to the plant cell is mediated by the transfer-DNA (T-DNA) through the type-IV secretion system of the bacterium (Beritza et al., 2024). Additionally, the commonly used 'LAB' strain of *N. benthamiana* harbours a mutation in the RNA-dependent RNA polymerase (Rdr1) gene, compromising its immune response. This mutation facilitates the expression of transgenes, especially those delivered by replicative viral vectors integrated into the T-DNA (Golubova et al., 2024). Moreover, the compact size of the *N. benthamiana* makes it ideal for indoor cultivation, such as in greenhouses, and its short life cycle supports quick protein production. Furthermore, advancements in genomics have resulted in a well-annotated genome, providing valuable insights for optimizing its use in Molecular Farming applications (Ranawaka et al., 2023).

Transient expression in *N. benthamiana* plants has been exploited to produce a large number of recombinant proteins such as antibodies (Paul et al., 2014) and immunomodulatory glycoproteins (Wilbers et al., 2017), viral-like particles (Peyret et al., 2020) and metabolites like precursors of the anti-malarial artemisinin (Van Herpen et al., 2010) or vaccine adjuvants like QS-21 (Martin et al., 2024) in plants. Agroinfiltration also enables temporary modifications to the plant chassis, such as expressing protease inhibitors (Goulet et al., 2012), human chaperones (Margolin et al., 2020) to reduce ER stress caused by protein accumulation, or preventing transcript degradation from gene silencing by co-infiltration with silencing suppressors (Beritza et al., 2024). Moreover, the use of viral elements (Peyret & Lomonossoff, 2013) or 'deconstructed' viruses (reviewed in Hefferon (2017); Peyret & Lomonossoff (2015)) promoting the expression of the genes of interest has also been a boost for transient expression. While agroinfiltration offers a fast and reliable way to obtain transgene expression, factors such as the variability associated with the number of T-DNA copies entering each cell or the high operational costs involving *Agrobacterium* culture and infiltration, limit the viability of this technique for large scale bioproduction.

By contrast to transient agroinfiltration, the stable integration of a transgene into the plant genome enables continuous, scalable and eventually uniform expression across multiple generations, resulting in consistent production yields over time (Figure 12). One prominent

example of the use of stable transformation in Molecular Farming was developed in the frame of the EU Pharma-Planta project, which successfully expressed the HIV-neutralizing human antibody 2G12 in transgenic *Nicotiana tabacum* from a constitutive duplicated 35S promoter (Ma et al., 2015). This project met Good Manufacturing Practice (GMP) standards by characterizing the transgene locus, confirming genetic and phenotypic stability of the plants, and establishing protocols for seed banking, plant cultivation, and antibody purification. The well-defined production process met European Medicines Agency (EMA) guidelines for active pharmaceutical ingredients, enabling authorization for a phase I clinical trial, where P2G12 was shown to be safe and well-tolerated in healthy women. The development of transgenic lines like these offers a consistent and scalable platform for producing biopharmaceuticals.

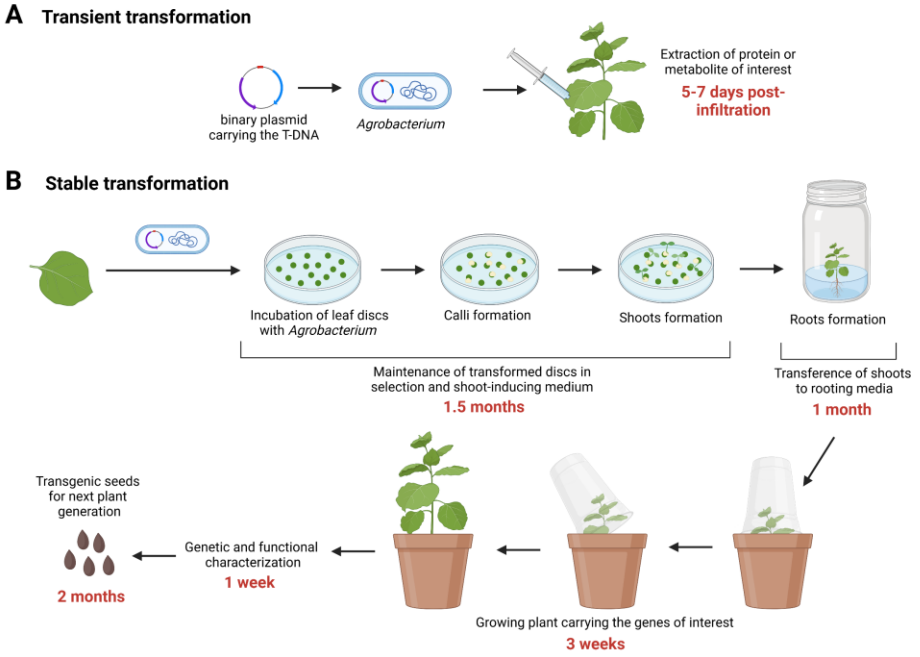



**Figure 12. Schematic representation of (A) transient and (B) stable transformation of transgenes in *Nicotiana benthamiana*.**

Despite the examples of partial success, reaching economically viable yields of the target bioproduct remains one of the main challenges in developing stably transformed biofactories, with several limiting factors. One of them is the metabolic context, which could impede a continuous precursor availability or involve unwanted pathway diversions (Golubova et al., 2024). For example, the first recombinant synthesis of glucoraphanin, a glucosinolate found in broccoli, failed due to the absence of key auxiliary genes in *N. benthamiana*. However, Barnum et al. (2022) boosted the production of glucoraphanin by optimizing the biosynthetic pathway in the plant, successfully achieving high glucoraphanin levels. Similarly, Dudley et al. (2022) enhanced strictosidine production by co-expressing fourteen genes in the synthesis pathway and gene-editing host glycosyltransferases to

eliminate undesirable modifications from host metabolism. These examples underscore the importance of aligning heterologous gene expression with the metabolic processes of the plant to boost efficiency.

Another important issue in yield optimization especially relevant in stable approaches relates to the toxicity of the accumulated products and the resulting effects in plant development and biomass accumulation. As an example, among many others, the recombinant production of the CP933 endolysin in *N. benthamiana* caused the death of the apical portion of the plant (Kovalskaya et al., 2016). To control this toxicity, spatiotemporal control of gene expression can be implemented. For instance, Forestier et al. (2023) produced casbene in *N. benthamiana* by using heat-inducible promoters. Finally, another important factor is transgene dosage. Typically, only single-copy T-DNA insertions are selected in stable transformants strategies due to (unfounded) biosafety considerations, but also as a strategy to minimize the risk of gene silencing. This is in steep contrast with the normal situation in agroinfiltration, where multiple copies of T-DNA are present in each cell maximizing the biosynthesis capacity of the transiently transformed tissue. In this context, viral systems that mediate the replication of the genes of interest, such as those based on the geminivirus replication machinery explained in the previous section, become highly valuable (Zhang & Mason, 2006).

All the above-mentioned limiting factors could be circumvented, in whole or in part, by genetically engineer the plant biofactory chassis. As described, metabolic engineering interventions could provide improved metabolic contexts for the accumulation of the target bioproduct. The optimisation of the spatiotemporal patters of transgene expression by means of synthetic gene circuits could also help to circumvent metabolic burdens and to reduce toxicity effects. Finally, the use of tightly regulated viral replicons imbibed as extra components in such synthetic gene circuits could serve to amplify the effective gene dosage, pushing yields even further. While notable advances have been made in genetically engineering food crops with enhanced agronomic traits, it is in the fields of industrial biofactories and Molecular Farming, both typically involving non-food crops, where the type of Synthetic Biology approaches described here are more likely to translate into real products in the short-midterm. As described in the previous section, the use of modular cloning technologies, and the consequent availability of large collections of easy-to-combine, exchangeable, standardized, and well-document bioparts, constitute an excellent playground for the design, build and test of increasingly productive versions of a given plant biofactory prototype in an iterative manner. Despite this opportunity in place, to this date the practical examples of such biotechnological applications remain limited, highlighting the need for further exploration and development in this area (Wang & Demirer, 2023).

In this thesis, we address the challenge of designing new synthetic gene circuits in plants that integrate modular bioparts, with the final aim to build improved plant biofactory prototypes. A key challenge in this endeavour is the design of new genetic sensors that respond to chemical triggers compatible with common agronomical practices. Equally

important is to connect such sensors with programable synthetic processors that can activate downstream actuators in an easily customizable manner. Finally, we aim to advance in understanding how to “tame” geminivirus-derived replicons so that they can be securely integrated in synthetic circuits as signal amplifiers, hopefully leading to higher product yields. Addressing these challenges is essential for meeting the growing global demand for metabolites and proteins and ensuring that production is both scalable and environmentally sustainable.

# Objectives

# OBJECTIVES

The main objective of this thesis is to advance in the design of new synthetic gene circuits for the improvement of plant biomanufacturing processes. To achieve this, the following specific objectives were proposed:

1. Develop and validate a copper-inducible switch for the control of gene expression in *N. benthamiana* (Chapter 1). This objective will involve the following two sub-objectives:

    a) Design and validate an optimized copper sensor based on the yeast copper responsive factor CUP2.

    b) Create a customizable copper-inducible gene circuit by coupling the newly developed copper sensor to a programmable transcriptional activator based on the CRISPR/Cas9 architecture.

2. Develop and validate a prototype of a *Nicotiana benthamiana* stable biofactory line that synthetises high levels of a recombinant protein in response to copper application (Chapter 2). This objective will involve the following two sub-objectives:

    a) Design and validate a BeYDV (bean yellow dwarf virus) geminiviral-based replicon that functions as a regulable transgene amplifier in *Nicotiana benthamiana*.

    b) Couple the copper sensor optimized in Objective 1a to the BeYDV-based processor developed in Objective 2a, to create a fully functional, customizable gene circuit that amplifies transgene dosage in response to copper.

    c) Customize and integrate the gene circuit developed in the Objective 2b into the *Nicotiana benthamiana* genome to create a stable plant prototype that produces high levels of a recombinant antibody in response to copper treatment.

3. Expand the current modular toolbox of geminiviral-based replicons for new applications in Synthetic Biology (Chapter 3).

    a) Set an autonomous bioluminescence reporter for the characterization of new geminiviral replicons.

    b) Optimize new replicon configurations derived from TYLCV- (tomato yellow leaf curl virus) and BCTV- (beet curly top virus) using the new autonomous bioluminescence reporter.

# Chapter 1

## A copper sensor for inducing CRISPR/Cas9-based transcriptional activation tightly regulates gene expression in *Nicotiana benthamiana*

---

Based on peer-reviewed article: *A copper switch sensor for inducing CRISPR/Cas9-based transcriptional activation tightly regulates gene expression in Nicotiana benthamiana*

Elena Garcia-Perez, Borja Diego-Martin, Alfredo Quijano-Rubio, Elena Moreno-Giménez, Sara Selma, Diego Orzaez and Marta Vazquez-Vilar

BMC Biotechnology (2022), DOI: 10.1186/s12896-022-00741-x

My contribution to this work was essential for its publication. I contributed to the design, performing of the experiments and writing of this manuscript.

## ABSTRACT

CRISPR-based programmable transcriptional activators (PTAs) are used in plants for rewiring gene networks. Better tuning of their activity in a time and dose-dependent manner should allow precise control of gene expression. Here, we report the optimization of a copper sensor for conditional gene activation in *Nicotiana benthamiana*. In the presence of copper, the copper-responsive factor CUP2 undergoes a conformational change and binds a DNA motif named copper-binding site (CBS). In this study, we tested several activation domains fused to CUP2 and found that the non-viral Gal4 domain results in strong activation of a reporter gene equipped with a minimal promoter, offering advantages over previous designs. To connect copper regulation with downstream programmable elements, several copper-dependent configurations of the strong dCasEV2.1 PTA were assayed, aiming at maximizing activation range, while minimizing undesired background expression. The best configuration involved a dual copper regulation of the two protein components of the PTA, namely dCas9:EDLL and MS2:VPR, and a constitutive RNA pol III-driven expression of the third component, a guide RNA with anchoring sites for the MS2 RNA-binding domain. With these optimizations, the CS/dCasEV2.1 system resulted in copper-dependent activation rates of 2,600-fold and 245-fold for the endogenous *N. benthamiana* DFR and PAL2 genes, respectively, with negligible expression in the absence of the trigger. The tight regulation of copper over CS/dCasEV2.1 makes this system ideal for the conditional production of plant-derived metabolites and recombinant proteins in the field.



**Keywords**

CRISPR/Cas9-based programmable transcriptional activator, inducible expression, copper sensor, minimal promoter, Plant Synthetic Biology, *Nicotiana benthamiana*

## INTRODUCTION

Synthetic Biology offers the opportunity to overcome intrinsic limitations of biological systems for biotechnological applications by precisely rewiring genetic circuits. Standardization and modularization of DNA elements lead to the creation of synthetic circuits that can orchestrate the endogenous circuitry in the plant (Liu & Stewart, 2015; Nandagopal & Elowitz, 2011). The implementation of effective synthetic circuits relies on universal devices that incorporate properly characterized genetic elements. Transcriptional regulation based on CRISPR/Cas9 technology has become a popular tool for controlling gene expression in eukaryotes (Konermann et al., 2015; Tanenbaum et al., 2014; Zhang et al., 2015). CRISPR-based genetic circuits provide a precise solution for connecting synthetic devices with endogenous networks and pathways. The binding-ability customization of the nuclease-inactivated Cas9 protein (dCas9) by tailoring guide RNAs allows targeting DNA sequences with high precision (Zhang, 2019). The customizable dCas9 can be linked to transcriptional activation domains leading to the so-called programmable transcriptional activators (PTAs). The use of these synthetic PTAs leads to advanced control over gene expression across diverse plant systems (Pickar-Oliver & Gersbach, 2019).

We previously developed a potent PTA named dCasEV2.1 (Selma et al., 2019). This PTA comprises two transcriptional units (TUs) for expressing two protein components and a third TU for the guide RNAs expression cassette. The first protein component is the dCas9 fused to the EDLL transcriptional activation domain (Tiwari et al., 2012). The second component is a fusion protein made of the MS2 phage coat protein and a VPR activation domain. VPR is a synthetic cluster of viral activators that further recruits the cell machinery for initiating transcription (Chavez et al., 2015). The third component of the dCasEV2.1 system is the guide RNA (gRNA) that targets the dCas9 to a specific position in the DNA and includes two aptamers in its scaffold for the binding of MS2 (Kunii et al., 2018). Together, EDLL and VPR, have a synergic effect on the activation of the targeted gene leading to high expression levels (Selma et al., 2019).

Metabolic engineering and Molecular Farming usually aim at maximizing the expression levels of genes of interest. However, this maximization is often accompanied by serious detrimental effects such as toxicity or developmental defects, which can be mitigated by engineering tight control systems (Bernabé-Orts et al., 2020). The implementation of genetic sensors actuating over PTAs would permit the creation of inducible activation circuits thus resulting in a fine-tuned regulation of genetic networks (Mandegar et al., 2016). Inducers act as stimuli for sensor proteins. Several sensors can additionally act as transmitters and convert a particular stimulus into a transcriptional signal that leads to the activation of a downstream gene. Depending on the type of stimulus, different inducible PTA systems can be established that allow the expression of the dCas9 activator under a specific condition. Traditional inducible systems to control transcription in plants are based on stimuli like chemicals, light, temperature, hormones, stress and wound (Tang et al., 2004). Most of these inducers can trigger deleterious pleiotropic effects in the plant chassis and are complicated

to apply and monitor in large crop fields (Corrado & Karali, 2009). However, among chemical inducers, micronutrients such as copper represent an innocuous and inexpensive group.

Copper is commonly applied to crops as copper sulfate ($CuSO_4$) not only like an essential micronutrient but also like fungicide. Copper sulfate can be easily taken up by plants, it is low-cost, easy to apply, and it is already registered for field use as a fungicide (Kumar et al., 2021). Accordingly, copper represents a potentially advantageous inducer for programmable transcriptional activation systems. Mett et al. (1993) developed a gene expression system inducible by copper based on the copper-metallothionein regulatory system from yeast. This regulatory system stands on a $Cu^{2+}$-regulated transcription factor from *Saccharomyces cerevisiae* (Buchman et al., 1989). In the presence of the trigger, the yeast copper responsive factor CUP2 (also known as ACE1) undergoes a conformational change that enables it to specifically bind a DNA operator sequence known as metal responsive element (MRE) or copper binding site (CBS). The binding of CUP2 to a CBS adjoining a minimal promoter leads to the transcriptional activation of the downstream gene. The copper-inducible system initially described by Mett et al. (1993) was later optimized by Saijo & Nagasawa (2014). This optimization consisted of the concatenation of four repetitions of the CBS for recruiting more molecules of CUP2. Also, a minimum sequence of the 35S promoter was reported to reduce the leaky expression of the targeted gene. Finally, the viral transcription activator VP16 was fused to CUP2 to boost the activation capacity of CUP2. This renewed system worked in planta and overcame the copper-inducible gene expression results reported earlier.

Here we show an optimized copper sensor named copper sensor, which incorporates the Gal4 yeast activator domain fused to CUP2. The copper sensor is free of viral elements and shows considerably reduced background expression in the absence of copper. Moreover, we show the implementation of a small copper-triggered customizable activation cascade that connects the copper sensor to the expression of dCas9EV2.1, with the subsequent activation of either reporters or endogenous genes as downstream elements in *Nicotiana benthamiana*.

## RESULTS

### Optimization of a copper-responsive sensor for efficient reporter gene activation in *N. benthamiana*

To develop an agronomically compatible system for switching on transcription, we explored the optimization of a copper-dependent system in *N. benthamiana*. Four repeats (4x) of the copper binding site (CBS) operator were assembled upstream a minimal promoter of the *Solanum lycopersicum* NADPH-dependent dihydroflavonol reductase (DFR) gene for assessing the transcriptional activation of Firefly luciferase (FLuc) reporter gene (Figure 1A). This minimal promoter spans the 62 bp upstream of the TSS and the 5′ UTR of the DFR promoter, which is characterized by a low and non-variable expression (Selma et al., 2019). This promoter is hereafter referred as CBS:minDFR. The copper responsive transcription

factor CUP2 was expressed under the constitutive Nopaline synthase promoter (pNOS) and fused to different transcriptional activation domains. We compared the performance of the yeast Gal4 activation domain with the activation provided by the viral-based domains VP16, VPR (VP64-p65-Rta), and TV (6x TAL, 4x VP64), commonly used as strong activation elements in plants (Casas-Mollano et al., 2020). Copper sulfate 5 mM was sprayed onto the leaves to induce the transiently expressed system. Luminescence assays showed that the basal expression of the reporter gene in the absence of copper resulted in negligible levels in all tested activation fusions (Figure 1B). The relative transcriptional activities (RTAs) conferred by the CBS:minDFR promoter after copper-mediated activation ranged from 4.5±1.5 to 13.5±6.2, as compared with the expression levels conferred by a pNOS promoter in the same assay conditions (Vazquez-Vilar et al., 2017). VPR was reported as the strongest activator, but also as the most variable one. The RTAs obtained with Gal4, VP16, and TV in the presence of copper were not statistically significant between each other. Therefore, Gal4 resulted as a suitable alternative to viral domains in the copper-responsive sensor. We named this novel combination of the copper-responsive factor CUP2:Gal4 and the 4xCBS operator upstream the minimal DFR promoter as copper sensor.

Next, we optimized the copper concentrations used for the copper sensor. For this, the sprayed $CuSO_4$ volume per leaf was set to 5 ml and different concentrations of $CuSO_4$ were tested. Also, the FLuc activity used as standard reference was analyzed at two time points, one- and two days post-copper-application (dpca), to determine whether the system could result in higher activations rates at times longer than 24 h. Figure 1C shows the dose-dependent response in the system when applying different concentrations of $CuSO_4$. The basal expression of the system without copper was virtually null, thus leakiness of the copper sensor was not observed. The system was not activated with 0.1 mM copper, but higher levels of Fluc expression were induced with increasing $CuSO_4$ concentration. Maximum RTA values of 2.3±0.4 were obtained with 10 mM $CuSO_4$. However, at this concentration necrosis symptoms in the leaf tissue were noticeable (figure 1D). To avoid any potential phytotoxic effect, we selected 5 mM for subsequent experiments. Differences in FLuc activity between one- and two-days post-copper application (dpca) were not statistically significant at any of the assessed concentrations except for the detrimental 10 mM. Therefore, the system does not require more than one day to reach its maximum functionality.

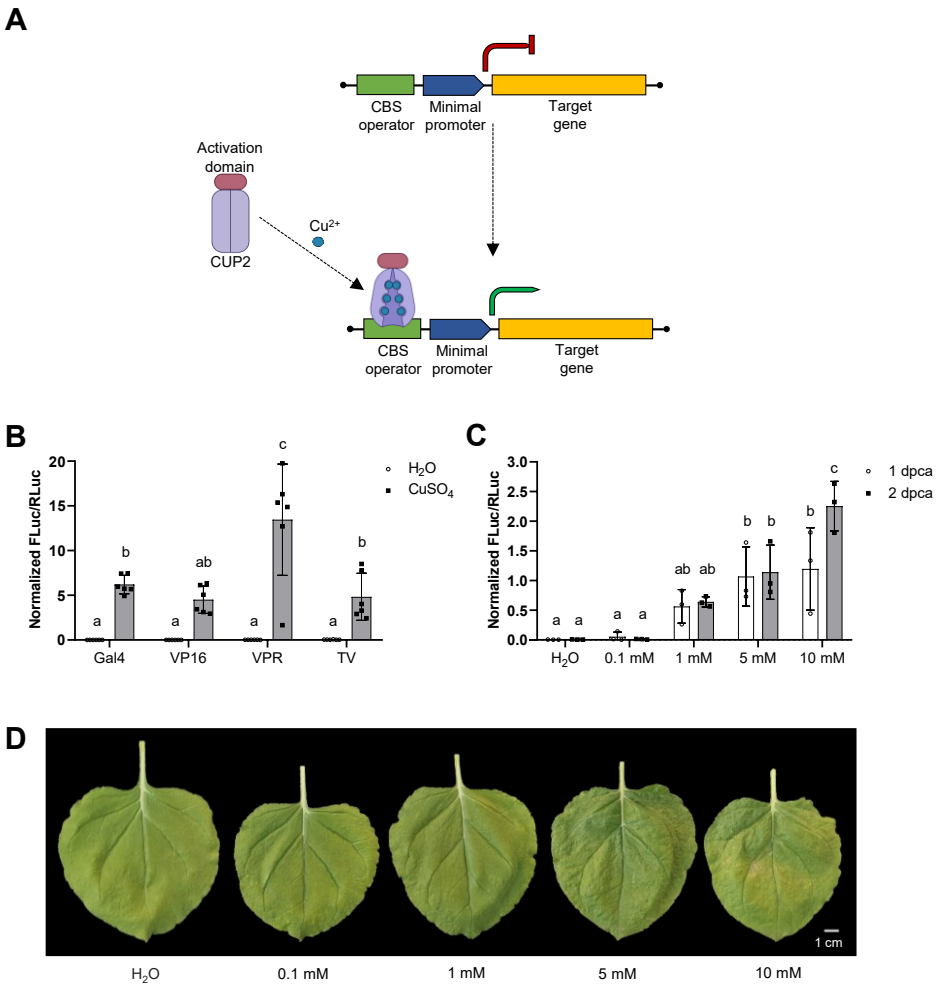



**Figure 1. Copper-inducible reporter gene activation in transient expression *in Nicotiana benthamiana*. A)** Schematic representation of the copper-inducible gene expression system. The CUP2 fused to an activation domain such as Gal4 changes its conformation upon copper binding and binds a copper binding site (CBS) operator allowing transcription of the downstream gene (green arrow). **B)** Copper-mediated activation of Firefly Luciferase (FLuc) reporter gene by different activation domains (Gal4, VP16, VPR, and TV) fused to CUP2 after 5 mM copper sulfate application. **C)** Copper-mediated activation of FLuc by different concentrations of copper sulfate. The construct CUP2:Gal4 was used as the copper-dependent transcriptional factor. **D)** Pictures of leaves 2 days after treatment with different concentrations of copper sulfate. For B) and C), activation is expressed as Normalized FLuc/RLuc ratios of *N. benthamiana* leaves transiently expressing the genetic constructs. All FLuc/RLuc ratios were normalized using the FLuc/RLuc ratios of *N. benthamiana* leaves expressing a constitutive pNOS:Fluc reporter. Error bars indicate SD (n≥3). Statistical analyses were performed using one-way ANOVA (Tukey's multiple comparisons test, P-Value ≤ 0.05). Variables within the same statistical groups are marked with the same letters. The figure includes images from Biorender (biorender.com).

**Optimization of copper-mediated dCasEV2.1 regulation in *N. benthamiana***

To assess the modulation that copper can offer in the regulation of transcriptional gene circuits, we connected the copper sensor to the three-component dCasEV2.1 PTA and tested the performance of the resulting genetic device using a transient luciferase assay in *N. benthamiana*. The CS/dCasEV2.1 tool comprised the CUP2:Gal4 TU, the 3X TUs dCasEV2.1 complex (dCas9:EDLL, MS2:VPR, and the adapted gRNA) (Figure 2A, B), and a FLuc reporter TU that contains the *S. lycopersicum* DFR promoter targeted by the gRNA. We employed two different copper-regulated synthetic promoters, namely CBS:min35S (carrying the 4xCBS box next to a minimal 35S promoter), and the previously tested CBS:minDFR. In our first optimization steps, the CUP2:Gal4 and the gRNA TUs were constitutively expressed (driven by pNOS and Arabidopsis thaliana U6-26 respectively), and the remaining elements were assayed under different copper inducible configurations. In the Cu/Cs configuration, only dCas9:EDLL TU was under copper regulation, whereas MS2:VPR TU was constitutively expressed (Figure 2B). In the Cu/Cu configuration, both TUs were put under copper-regulated promoters. Both Cu/Cs and Cu/Cu configurations were assayed using the two synthetic promoters described above. As can be observed in Figure 2C, when the minDFR element was present, the single Cu configuration (Cu/Cs) showed moderate background levels (0.16±0.04), which were almost completely abolished when dual copper regulation (Cu/Cu) was implemented (0.01±0.05). Furthermore, dual regulation offered higher activation rates of FLuc reporter than single regulation, reaching average RTAs of 1,43±1.00 (Figure 2C).

The same analysis was performed using the CBS:min35S synthetic promoter (Saijo & Nagasawa, 2014). As shown in Figure 2D, the RTAs achieved by using min35S-containing constructs were consistently higher than those obtained with minDFR and even slightly higher than with the constitutive version, Cs/Cs (Figure 2D). However, constructs employing min35S resulted in higher basal expression and there is no statistically significant difference in the activation provided by the constitutive dCasEV2.1 and the Cu/Cu configuration when copper is not present.

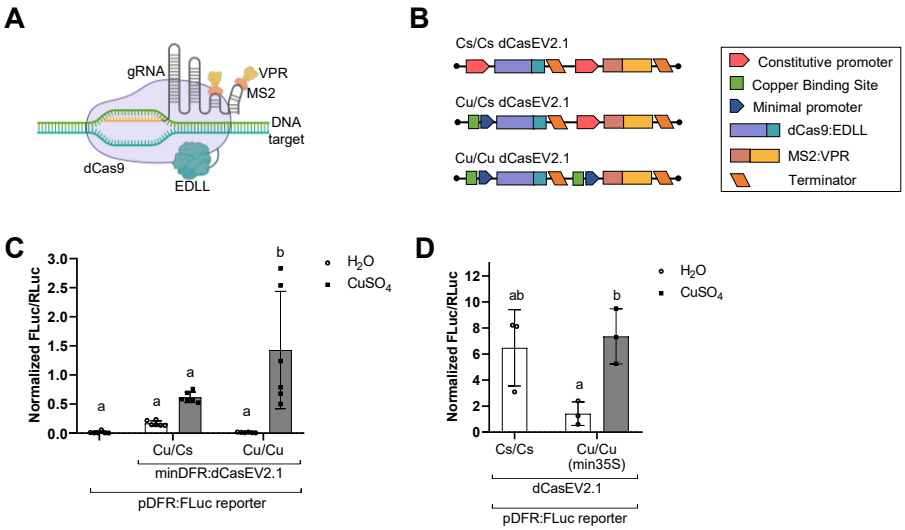



**Figure 2. Copper-inducible reporter gene activation by dCasEV2.1 in *N. benthamiana*. A)** Schematic representation of dCasEV2.1 components. **B)** Genetic constructs for dual constitutive (Cs/Cs) transcription (35S promoter for both dCas9:EDLL and MS2:VPR), single copper-regulated (Cu/Cs) transcription (CBS:minimal DFR promoter for dCas9:EDLL and 35S promoter for MS2:VPR), and dual copper-regulated (Cu/Cu) transcription (CBS:minimal DFR promoter for both dCas9:EDLL and MS2:VPR). **C)** Efficiency of single vs. dual copper-regulated transcription of dCasEV2.1 under a minimal DFR promoter (minDFR) to activate a pDFR:FLuc reporter. **D)** Efficiency of dual copper-regulated transcription of dCasEV2.1 under a minimal 35S (min35S) promoter to activate a pDFR:FLuc reporter. Activation is expressed as Normalized FLuc/RLuc ratios of *N. benthamiana* leaves transiently expressing the genetic constructs. All FLuc/RLuc ratios were normalized using the FLuc/RLuc ratios of *N. benthamiana* leaves expressing a constitutive pNOS:Fluc reporter. Error bars indicate SD (n≥3). Statistical analyses were performed using one-way ANOVA (Tukey's multiple comparisons test, P-Value ≤ 0.05). Variables within the same statistical groups are marked with the same letters. The figure includes images from Biorender (biorender.com).

### Effect of gRNA regulation on copper-inducible dCasEV2.1 activation

Once the regulation of the dCasEV2.1 protein components was optimized, we aimed to further explore the control of the gRNA expression with copper and ethanol. The choice of a second inducer for gRNA expression aims at enabling the activation of different sets of genes under different stimuli. CRISPR/Cas9 guide RNAs are usually transcribed under a RNA polymerase III (pol-III) promoter, like the U6-26 (Cong et al., 2013; Mali et al., 2013). However, most minimal promoters coupled to inducible operators rely on polymerase II (pol-II) transcription (Knapp et al., 2019). As pol-II transcripts are altered with a 5' cap and poly-A tail and exported from the nucleus, the efficiency of a gRNA expressed under pol-II would be highly compromised. Thus, we flanked the sequence of the gRNAs by the A. thaliana tRNAGly sequence (Xie et al., 2015) to ensure the post-transcriptional processing of the 5' and 3' ends generated after pol-II action. To assemble pol-II gRNA expression cassettes using the type IIS DNA assembly system GoldenBraid, we followed the strategy

previously described by Vazquez-Vilar et al. (2016) for pol-III gRNAs. In this strategy, the gRNA unit is assembled in a BsmBI restriction-ligation reaction as a Level 0 part incorporating the spacer sequence as a primer dimer with unique 4-bp overhangs, and the 5' and 3' sequences flanking the spacer are provided from another vector. This vector carries (i) the tRNAGly sequence and (ii) the scaffold-tRNAGly sequence, both flanked by BsmBI restriction sites and unique overhangs for their assembly upstream and downstream the spacer sequence, respectively. In the next step, the gRNA unit is assembled between a Pol-II promoter and a transcriptional terminator in a BsaI restriction-ligation reaction (Figure 3A). A new webtool has been developed during this work to design and assemble tailored Pol-II polycistronic gRNA constructs for gene regulation with (d)Cas9: (https://gbcloning.upv.es/tools/cas9multiplexing_inducible_regulation/).

To assess the feasibility of expressing gRNAs by the pol-II, two constitutive promoters for the pol-II, pNOS and p35S, were tested upstream of the tRNA-protospacer-scaffold-tRNA. To evaluate the regulation capability over the pol-II-synthesized gRNAs, the synthetic CBS:minDFR promoter was used for copper-regulation. In parallel, ethanol regulation was attempted as well, using AlcR as the ethanol-responsive transcriptional factor that binds to AlcA operator in presence of ethanol (Roslan et al., 2001). All the constructs for Pol-II-synthesized gRNAs contained, like in previous experiments, a spacer targeting a DFR promoter upstream of the reporter FLuc gene. The different gRNA constructs were co-infiltrated with the construct for the constitutive expression of dCasEV2.1, the reporter construct, and the copper responsive-transcription factor in the plant. When pNOS and p35S were used for gRNA synthesis, activation of FLuc by dCasEV2.1 reached RTAs of 2.6±0.6 and 3.1±1.1 respectively (Figure 3B). However, this activation is significantly lower (approximately 3.5-fold) than the activation obtained when the gRNA was synthesized under the pol-III promoter U6-26. Unfortunately, when gRNAs were controlled by CBS or AlcA operators, FLuc reporter activity reached similar rates both in the presence and absence of copper/ethanol, probably indicating leaky expression of the gRNA in the absence of the trigger.

Our results indicated that the chemical induction of the gRNA component of dCasEV2.1 alone was not sufficient to regulate downstream gene expression. A possible explanation is that gRNA differential expression, if leaky, may be masked by the high and constitutive expression of the other two components in the system. For this reason, we explored the performance of a triple-controlled system combining the dual copper-inducible dCasEV2.1 with either a CBS:minDFR:gRNA:Tnos or an AlcA:minDFR:gRNA:Tnos. In all combinations assayed, reporter activation in presence of the inducer was observed. However, as expected, transient-expression assays showed that triple regulation (Cu/Cu/EtOH or Cu/Cu/Cu) gave similar results to those observed with the dCasEV2.1 Cu/Cu conformation, as no differences were found when gRNAs were expressed under the constitutive U6-26 promoter or the copper- or ethanol-inducible promoters (Figure 3C). Therefore, we concluded that the most efficient way to regulate the dCasEV2.1 system is via the regulation of its two protein components dCas9:EDLL and MS2:VPR.

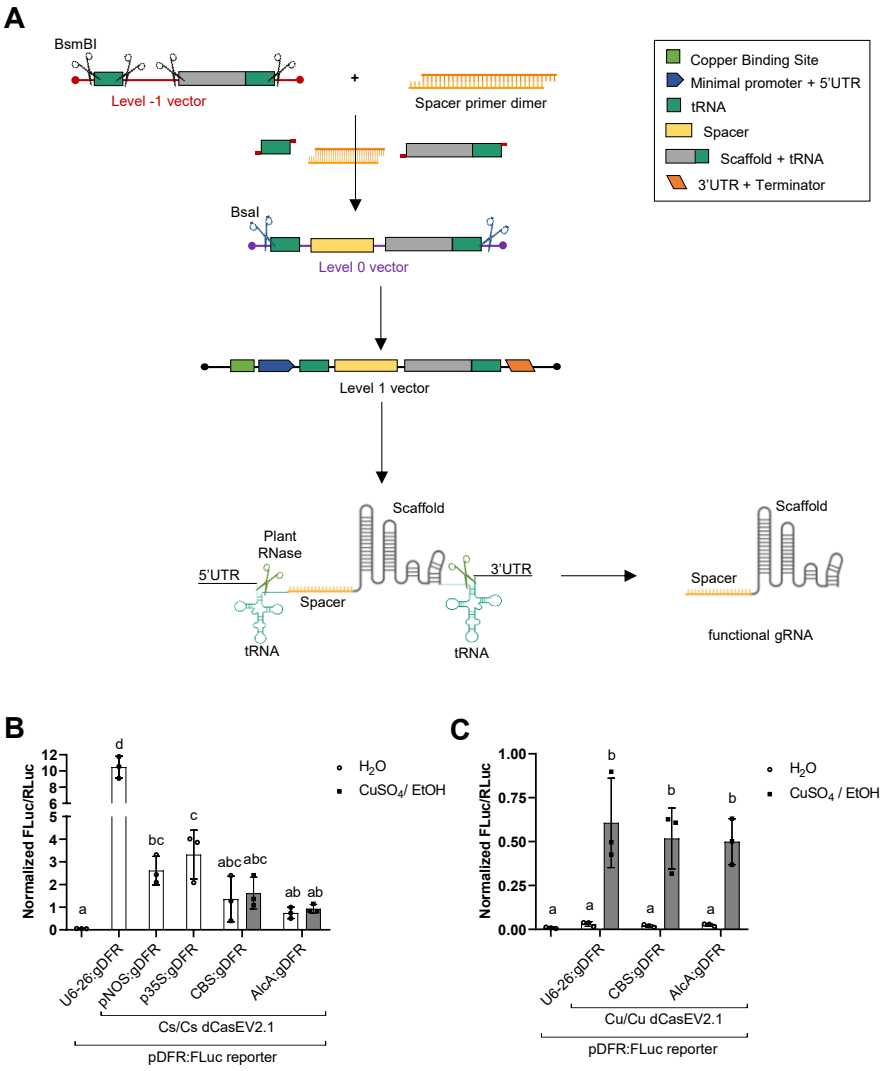



**Figure 3. Reporter gene activation by chemically inducible pol-II expressed guide RNAs in combination with dCasEV2.1. A)** Schematic representation of cloning strategy and post-transcriptional processing of guide RNA (gRNA). The tRNA and scaffold-tRNA sequences are excised from a Level -1 plasmid and assembled together with the spacer in a Level 0 plasmid. The tRNA-spacer-scaffold-tRNA unit is transferred to a Level 1 vector to generate the gRNA expression cassette. **B)** Reporter gene activation with constitutive dCasEV2.1 and differently expressed gRNAs. The pDFR:FLuc reporter was activated by a constitutive dCasEV2.1 under two 35S promoter (Cs/Cs dCasEV2.1) combined with differently expressed gRNAs targeting the DFR promoter of the FLuc reporter. U6-26 is a constitutive promoter for polymerase III, and pNOS and p35S are constitutive promoters for polymerase II. CBS is the operator for copper induction and AlcA is the operator for ethanol induction. Both CBS and AlcA are assembled next to a minimal DFR promoter. **C)** Triple induction of dCasEV2.1 components to activate gene expression. The pDFR:FLuc reporter was activated by a dually induced dCasEV2.1 using a minDFR promoter (Cu/Cu dCasEV2.1) in combination with differently expressed

gRNAs targeting the DFR promoter of the FLuc reporter. Activation is expressed as Normalized FLuc/RLuc ratios obtained for the *N. benthamiana* leaves transiently expressing the genetic constructs. All FLuc/RLuc ratios were normalized using the FLuc/RLuc ratios of *N. benthamiana* leaves expressing a constitutive pNOS:Fluc reporter. Error bars indicate SD (n=3). Statistical analyses were performed using one-way ANOVA (Tukey's multiple comparisons test, P-Value ≤ 0.05). Variables within the same statistical groups are marked with the same letters. The figure includes images from Biorender (biorender.com).

**Copper-regulated dCasEV2.1-mediated activation of endogenous genes**

Using dCasEV2.1 as an intermediate operator to the copper-responsive factor offers the possibility of transferring copper regulation to any endogenous gene of interest in the plant (Figure 4A) only by reprogramming the gRNA component of the dCasEV2.1 complex. To show this, we programmed dCasEV2.1 to target two genes of *N. benthamiana*, the DFR (NbDFR) gene and the phenylalanine ammonia-lyase 2 (NbPAL2) gene. These two genes codify for enzymes involved in the flavonoid biosynthetic pathway in plants. DFR plays an important role in the synthesis of anthocyanins (Tian et al., 2017), while PAL is the key enzyme catalyzing the conversion of aromatic amino acids into secondary metabolites (Olsen et al., 2008). A dCasEV2.1 construct in Cu/Cu conformation (CBS4:minDFR:dCas9 :EDLL:tNOS-CBS4:minDFR:MS2:VPR:tNOS) was targeted with a pol-III expressed multiplexing gRNA to three positions of either the NbDFR promoter (Selma et al., 2019) or the NbPAL2 promoter (Selma et al., 2021). A gRNA targeting the *S. lycopersicum* MTB gene was included as a negative control of the activation. The constitutively expressed dCasEV2.1 was used as a positive control for maximum activation levels. Spraying leaves with 5mM $CuSO_4$ resulted in a remarkable 2600-fold activation of NbDFR mRNA levels and a 245-fold activation of NbPAL2 mRNA levels (Figure 4B). The expression of both NbDFR and NbPAL2 was negligible in absence of copper, an indication of the highly reduced background levels achieved with the optimization of the copper-responsive genetic tool developed in this work.

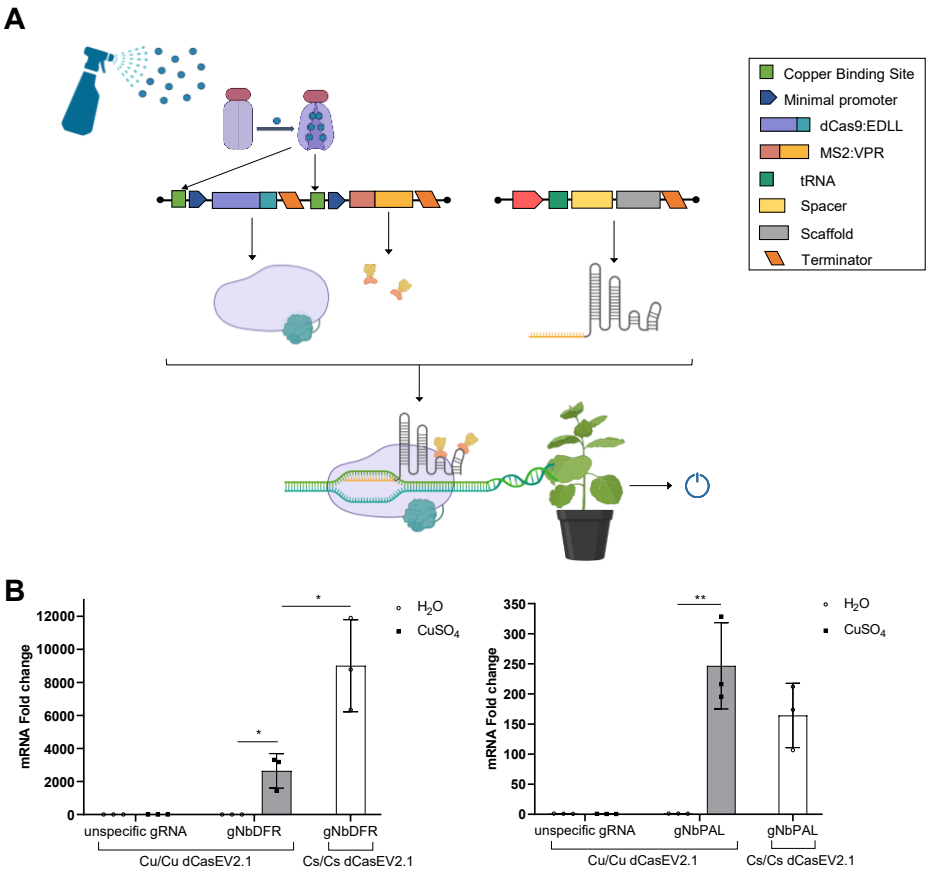



**Figure 4. Application of copper-inducible dCasEV2.1 for activating endogenous genes in *N. benthamiana*. A)** Schematic representation of the strategy of copper-mediated dCasEV2.1 induction for targeted gene activation in planta. **B)** Activation of the *N. benthamiana* DFR (NbDFR) and PAL2 (NbPAL2) genes by the dually copper-regulated dCasEV2.1 under minDFR promoter (Cu/Cu dCasEV2.1) and the constitutive dCasEV2.1 under 35S promoter (Cs/Cs dCasEV2.1). Transcripts were quantified by RT-qPCR at 5 days post-infiltration of dCasEV2.1 in combination with a constitutive guide RNA targeting either NbDFR (gNbDFR), NbPAL2 (gNbPAL2), or the *S. lycopersicum* MTB gene (unspecific gRNA) and a constitutive CUP2:Gal4. Leaves were treated either with 5 mM copper sulfate or water at 3 days post infiltration. Error bars indicate SD (n=3). The asterisks represent a significant difference between treatments. Statistical analysis was performed using an unpaired two-tailed T-test for two samples comparisons (P-value ≤ 0.05). The figure includes images from Biorender (biorender.com).

## DISCUSSION

CRISPR-based PTAs are becoming extensively used for endogenous genes activation in plants due to their high specificity and their effectiveness in boosting transcriptional activation (Konermann et al., 2015; Tanenbaum et al., 2014; Zhang et al., 2015). The exploitation of their full potential as tools for the activation on-demand of metabolic and signalling networks relies on their tight regulation. In plants, several chemical inducible

systems have been used for controlled gene expression. These systems are triggered by different compounds including antibiotics, ethanol, steroids, or copper (Andres et al., 2019). The main advantages of using copper as an inductor for plants are that copper sulfate is cheap and represents a friendly substance in agriculture. Thus, we optimized a copper-inducible system named copper sensor to regulate the programmable transcriptional activator dCasEV2.1 (Selma et al., 2019). Ku et al., (2012) showed that 40 mM and 80 mM $CuSO_4$ applied to *N. benthamiana* by watering was associated with a marked decrease in protein content and water deficit in tobacco shoots. However, they did not significantly report these detrimental effects in the plant with 20 mM $CuSO_4$. The concentrations tested for our system were below 20 mM and $CuSO_4$ was not applied by watering but by spraying intending to diminish any possible systemic copper toxicity. We reported that 10 mM $CuSO_4$ application was associated with necrosis symptoms on the treated leaf. Possibly, 10 mM $CuSO_4$ directly reaching the leaf cells is too high for the homeostasis cell machinery to modulate the redox-active properties of copper in the cytoplasm (DalCorso et al., 2014; Burkhead et al., 2009). An important consideration was that only treated leaves presented necrosis symptoms but not the untreated leaves of the same plant. Possibly, the local toxicity was not declining the global plant growth. Eventually, we selected 5 mM as the $CuSO_4$ concentration to apply to our system because treated leaves did not show necrosis symptoms during the 5 days after $CuSO_4$ application. However, further study is required to understand whether the proposed application of copper can affect plant viability in the long term. Lower applicable concentrations such as the tested 1 mM $CuSO_4$ might be an alternative to mitigate long-term toxic effects at the expense of slightly decreasing gene activation.

Since the first report on the use of the $Cu^{2+}$-regulated transcription factor CUP2 for gene activation in plants (Mett et al., 1993), the system has been adapted for different applications. However, most of these applications were not effective enough because of poor induction (Boetti et al., 1999; Granger & Cyr, 2000; Mohamed et al., 2001), variability between plants, and leakiness (Granger & Cyr, 2001). Saijo & Nagasawa (2014) boosted the initial system by fusing the viral activation domain VP16 to CUP2 and by incorporating 4 repetitions of the copper binding site (CBS) operator next to a minimal 35S promoter upstream the gene of interest. With this improved system, they reported a high transcription activation of GFP in transgenic lines as well as the induction of flowering in Arabidopsis thaliana. In this work, we tested different activation domains in combination with CUP2 to explore a higher activation capacity of the system. Contrary to Saijo & Nagasawa (2014), in our hands, the use of VP16 did not confer a significant improvement to the system. Most interesting, Gal4 fusion resulted in a potent activator and consequently our copper sensor incorporates this domain. According to current regulations (Turnbull et al., 2021), the application of the eukaryotic Gal4 domain at the open field would be more straightforward than the application of the viral VP16. Ultimately, the goal of this work is to provide transcriptional switches which can be realistically deployed in the field, and therefore we decided to continue our characterization with Gal4. We are aware that, despite the choice

of Gal4, the system described in this work is not completely free of viral domains, as dCasEV2.1 contains the VPR domain, a tandem of activation domains containing viral sequences (VP64-p65-Rta). While VPR showed the highest activation rates and therefore was the domain of choice for building the dCasEV2.1 tool, other non-viral domains, namely EDLL, ERF2 and P300, were successfully assayed during dCasEV2.1 optimization which are available in the GoldenBraid collection and the Addgene repository (Selma et al., 2019). Although for the optimization experiments described here we chose the use of dCasEV2.1, the alternative non-viral domains are likely to yield similar responses although probably with smaller activation rates. Recently, Pan et al. reported that the fusion of two copies of the transcription activator-like effector (TALE) TAL Activation Domain (TAD) to MS2 resulted in strong activation (Pan et al., 2021). An alternative configuration for our dCasEV2.1 incorporating TAD may be convenient for developing a free-viral tool to combine with our copper sensor.

When connecting an inducible sensor to a potent transcriptional activator, such as dCasEV2.1, it is important to consider that the signal is highly amplified and a minimal leakiness in the inducible system can result in a non-desired activation of the gene of interest in the absence of inductor. Therefore, for the novel application of copper to regulate a PTA, we explored the single and dual regulation of both transcriptional units that make up dCasEV2.1. We also investigated two different minimal promoters, DFR, characterized for its low basal expression, and 35S, used in previous studies for copper regulation. The results obtained in this work showed that dual regulation presented lower leakiness and resulted in a higher activation than the single regulation. In the case of dual regulation, 110- and 5.18-fold increments of reporter activity were obtained with the minimal DFR and minimal 35S, respectively. To overturn leakiness in the system, minimal DFR resulted as the most suitable promoter to use since the basal expression is practically null. However, if the goal of the activation is reaching a higher expression rate, a minimal 35S promoter would be more promising despite the leakiness. For supplying an extra regulation point to dCasEV2.1 activation, we attempted regulation of the guide RNA. For this regulation, we flanked the gRNA of interest by two $tRNA^{Gly}$ sequences and we tested different pol-II promoters. However, this extra regulation did not lead to efficient control over the activation of the reporter gene. These results are consistent with those reported by Knapp et al. (2019) showing that tRNAs can generate functional gRNAs that are constitutively and promoter-independently expressed. Knapp et al. (2019) designed a $tRNA^{Pro}$ scaffold (Δ, ΔC55A, and ΔC55G) with minimized promoter strength without affecting the processing activity of the tRNA. It would be interesting to explore alternative approaches for gRNA regulation that could include the implementation of these tailored tRNAs or to test the regulation of gRNAs free of tRNA structures.

The strict control of an inducible gene expression system is essential when studying signalling networks or when producing recombinant proteins with toxic effects in plants. Thus, our optimization steps of a copper-inducible PTA aimed to get maximum expression levels while keeping low basal expression. Altogether, our results showed that the best

configuration for the copper sensor consists of a constitutive CUP2:Gal4 transcription factor able to bind and activate the expression of the two protein components of the dCasEV2.1 from a CBS operator and minDFR promoter in the presence of 5 mM copper sulfate. The gRNA targeting the PTA to the gene of interest can be expressed constitutively from a pol III promoter.

We confirmed the robustness of our tool with the activation of two *N. benthamiana* endogenous genes, NbDFR and NbPAL2, which resulted in 2600-fold and 245-fold transcriptional activation rates, respectively, after copper application. Moreover, no leaky activation was reported for our tool in any of the two endogenous activations. The stable transformation of the copper-inducible PTA developed in this work along with user-designed gRNAs would result in a highly specific spatial/temporal expression of the genes of interest, either native or transgenes, in the plant. Furthermore, the application possibilities of the system may be expanded with a regulated expression of the gRNAs. This could be achieved either by the incorporation of tRNAs with minimized promoter strength, as discussed previously, or with tissue-specific promoters. The feasibility of regulating gRNA synthesis would open the possibility of creating AND-logic gates to tightly control endogenous gene activation on-demand.

In this study, we achieved tight control over gene expression in *N. benthamiana* using copper as the inductor. The comparison of different activation domains fused to CUP2 showed that Gal4 gives potent activation, and it is the most promising domain for open field use. The application of 5 mM copper sulfate results in good activation levels without severe phytotoxic effects. Alternative configurations for connecting the copper sensor to the PTA dCasEV2.1 were tested to get maximum expression levels while keeping low basal expression. Our results showed that dual regulation of both transcriptional units that make up dCasEV2.1 driven by four repeats of CBS operator and a minDFR promoter is the most competent configuration. The use of the minDFR promoter is essential for preventing leaky gene expression. When applying the copper-inducible PTA for *N. benthamiana* DFR and PAL2 genes activation, we obtained a 2600-fold and a 245-fold, respectively, transcriptional activation after copper application.

## MATERIALS AND METHODS

### Cloning and assembly of the GoldenBraid constructs

The constructs used in this work were assembled using the GoldenBraid (GB) cloning system (Sarrion-Perdigones et al., 2017). This system is based on subsequent Type IIS-restriction-ligation reactions to clone the desired TUs or combinations of TUs in destination vectors. Shortly, in the GoldenBraid hierarchy, basic DNA parts such as promoters, terminators, and coding sequences of genes of interest are considered as level 0 DNA parts. These level 0 parts are cloned in the pUPD2 plasmid via a BsmBI-mediated restriction-ligation reaction. Transcriptional units represent level 1 and they are assembled via a BsaI-mediated restriction-ligation reaction using a pDGB3 α as destination vector. The transcriptional units

are combined to create complex genetic modules in level >1, by using BsaI- and BsmBI-mediated assembly reactions into pDGB3 α and pDGB3 Ω destination vectors. All the constructs used in this work are listed in Supplementary Table S1, as well as their identification numbers to track their sequence and their assembly history at the GoldenBraid repository (https://gbcloning.upv.es/).

**Agroinfiltration experiments**

Transient expression of the genetic constructs was performed in wild-type *Nicotiana benthamiana* (obtained from Wageningen University & Research seeds collection and grown in-house). Agrobacterium tumefaciens strain C58 harboring the expression vectors were grown in LB medium supplemented with 50 µg/ml kanamycin or spectinomycin and 50 µg/ml rifampicin for 16 hours at 28ºC/250 rpm. Subsequently, the bacteria were pelleted, transferred to infiltration buffer (10 mM MES, 10 mM MgCl2, and 200 µM acetosyringone, pH 5.6), and incubated for 2h at room temperature in a rolling mixer. After this incubation step, the OD600 of the cultures was adjusted to 0.1. For plant co-expression of more than one GB element, the different Agrobacterium cultures were mixed and co-infiltrated in three young fully expanded leaves of five weeks-old *N. benthamiana* plants. The plants were maintained in a growing chamber at 24°C (light)/20°C (darkness) 16h-light/8h dark photoperiod. Infiltrated leaves were treated either with 0.1-10 mM copper sulfate, 20% ethanol, or water in combination with 0.05% fluvius by spray at three days post infiltration (dpi). The spray was applied to both the adaxial and abaxial surfaces of the leaf.

**Quantification of reporter gene expression**

One 8 mm-diameter disc per infiltrated leaf was collected at 4 dpi and stored at –80ºC. For analysis, the samples were homogenized, extracted in 180 µl Passive Lysis Buffer (Promega), and centrifugated for 15 min at 14,000 ×g and 4°C. An aliquot of 10 µl from the resultant supernatant was transferred to a 96-well plate. As the reporter gene that we used to determine gene expression was the Firefly Luciferase (FLuc), quantification of gene expression was done by measuring the luminescence generated by FLuc. For this measurement, the Dual-Glo® Luciferase Assay System (Promega) was used. The 10 µl of crude extract were mixed with 40 µl of LARII and the FLuc activity was measured by a GloMax 96 Microplate Luminometer (Promega), setting a 2-seconds delay and a 10-seconds measurement. After measuring, 40 µl of Stop&Glo reagent were added per sample and Renilla Luciferase (RLuc) activity was quantified with the same protocol. FLuc/RLuc ratios were calculated as the mean value of the three agroinfiltrated leaves from the same plant. Relative transcriptional activities were expressed as the FLuc/RLuc ratios in each sample and normalized to the FLuc/RLuc ratios of a reference reporter (GB1116).

**Real-time quantitative PCR analysis**

Total RNA was isolated from 100 mg of fresh leaf tissue harvested at 5 dpi by using the GeneJET Plant RNA Purification Kit (Thermo Fisher Scientific). The RNA was treated with rDNase (Invitrogen) before cDNA synthesis. The cDNA was synthesized from 1 µg of total RNA using the PrimeScript™ RT-PCR Kit (Takara). Three technical replicates were measured

for gene expression levels in each cDNA sample. The SYBR® fluorescent dye (TB Green® Premix Ex Taq™, Takara) was employed for reporting the amplification of cDNA in the Applied biosystem 7500 Fast Real-Time PCR system with specific primer pairs for either DFR or PAL2 genes (Supplementary Table 2). F-BOX gene was amplified as an internal reference (Liu et al., 2012) for determining DFR/PAL2 gene expression in each sample (ΔCt). The differential DFR/PAL2 gene expression levels were calculated according to the comparative ΔΔCt method as the ratio between the samples agroinfiltrated with the dCasEV2.1 and a gRNA targeting the DFR/PAL2 promoter (gNbDFR/gNbPAL2) and those including the same construct but an unrelated gRNA (gMTB), both for copper and water treatments (Livak & Schmittgen, 2001).

## SUPPLEMENTARY MATERIAL

The dataset generated and analysed in the current study are available in Zenodo, https://doi.org/10.5281/zenodo.6256318

**Supplementary Table 1**. GB level 1 and >1 transcriptional units and modules. Sequences can be found at http://www.gbcloning.upv.es/ using the GB number or GB name.

| GB number | Construct | GB name | Description |
|---|---|---|---|
| GB0107 | Stuffer fragment | pEGB SF | Used as a stuffer construct to get the same infiltration DO for every Agrobacterium mix. |
| GB1116 | pNos:Luciferase:tNos-35S:Renilla:tNos35S:P19: tNos | pEGB PNos:Luciferase:TNos-SF-35S:Renilla:TNos-35S:P19:TNos-SF | Module for the expression of the Firefly Luciferase, the Renilla Luciferase and the P19 silencing suppressor (standard reference in pCambia for luciferase experiments). |
| GB1160 | SlDFR:Luciferase:tNos-35S:Renilla:tNos-35S:P19: tNos | pEGB SlDFR:Luc:TNos-SF-35S:Renilla:TNos-35S:P19:TNos | Module for the expression of the Firefly Luciferase gene driven by the *S. lycopersicum* DFR promoter, the Renilla Luciferase gene and the silencing suppressor P19 driven by the 35s promoter and the Nos terminator. |
| GB1541 | pNos:CUP2:Gal4:tNos | pEGB_3alpha2 PNos:Cup2:Gal4AD:Tnos | Transcriptional unit for constitutive expression of the transcriptional activator protein CUP2 in C-terminal fusion with GAL4 activation domain. Binds to specific DNA operator (CBS) in the presence of copper ions and promotes transcription. |
| GB1838 | U6-26:gRNA-150 SlDFR: scf2.1 | 3alpha1_U6-26-1gRNA-DFR F6x2 MS2scf | Module for the expression by polymerase III of a guide RNA targeting the DFR promoter with two MS2 aptamer copies in the 3' of the scaffold. |
| GB2070 | U6-26:multigRNA SlMTB: scf2.1 | pDGB3_Alpha1_U6-26-5gRNA MTB-F6x2_U6-26-4gRNA MTB-F6x2_U6-26-3gRNA MTB-F6x2_U6-26-2gRNA MTB-F6x2 | Module for the expression of 5,4,3,2 gRNA multiplexing targeting the *S. lycopersicum* Mtb promoter with two Ms2 aptamer copies in the 3' of the scaffold (scaffold 2.1). |
| GB2085 | 35S:MS2:VPR:tNos-35S:dCas9:EDLL:tNos | pDGB3_Omega1_35s-Ms2:VPR-Tnos-35s-dCas9:EDLL-Tnos | Module for the constitutive expression of MS2 fused to VPR and dCas9 fused to EDLL. |

| | | | |
|---|---|---|---|
| GB2170 | U6-26:multigRNA NbDFR: scf2.1 | 3Alpha2_U6-26-sgRNANbDFR -85,-198,-268 F6x2 Multiplex | Multiplex construct with F6x2 aptamer in scaffold to target -85,-198,-268 position of *Nicotiana benthamiana* DFR promoter. |
| GB2396 | U26-26:multigRNA NbPAL2:scf2.1 | alpha2_U2626-NbPAL2432-2.1-gRNA 120-190-268 | Multiplex construct with 2.1 aptamer in scf to target -120,-190,-268 position of NbPAL promoter Niben101Scf02432. |
| GB2725 | pNOS:CUP2:VPR:tNos | 3a2:pNOS:CUP2:VPR:tNOS | Transcriptional unit for the expression of CUP2 (copper responsive transcription factor) fused to VPR under the regulation of a NOS promoter and terminator. |
| GB2727 | pNOS:CUP2:TV:tNos | 3a2:pNOS:CUP2:Tv:tNOS | Transcriptional unit for the expression of CUP2 (copper responsive transcription factor) fused to TV under the regulation of a NOS promoter and terminator. |
| GB2740 | CBS4:miniDFR:Luciferase:t35S-35S:Renilla:tNos-35S:P19:tNos | 3o2:CBS4:DFRmin:Luc:t35S-SF-35S:Ren:tNOS-35S:P19:tNOS | Module for the copper inducible expression of luciferase reporter under the regulation of CBS4 operator and the minimal DFR promoter. |
| GB3754 | 35S:MS2:VPR:tNos-CBS4:miniDFR:dCas9:EDLL:tNos | o1_35S:MS2:VPR:Tnos-CBS:miniDFR:dCas9:EDLL:Tnos | Module for simple copper-regulated dCasEV2.1. Copper-regulated dCas9:EDLL with CBS operator + minimal DFR promoter and constitutive MS2:VPR with 35S. |
| GB3771 | CBS4:mini35S:MS2:VPR:tNos-CBS4:mini35S:dCas9:EDLL:tNos | 3o1:CBS4:35Smin:omegaUTR:MS2:VPR:tNOS-CBS4:35Smin:omegaUTR:dCas9:EDLL:tNOS | Module for double copper-regulated dCasEV2.1. dCas9:EDLL with CBS operator + minimal 35S promoter and MS2:VPR with CBS operator + mini35S promoter. |
| GB3774 | CBS4:miniDFR:MS2:VPR:tNos-CBS4:miniDFR:dCas9:EDLL:tNos | 3o1:CBS4:DFRmin:MS2:VPR:tNOS-CBS4:DFRmin:dCas9:EDLL:tNOS | Module for double copper-regulated dCasEV2.1. dCas9:EDLL with CBS operator + minimal DFR promoter and MS2:VPR with CBS operator + mini35S promoter. |

| | | | |
|---|---|---|---|
| GB3786 | CBS4:miniDFR:SlDFR150:scRNA2.1-tRNA:tNOS | 3a1:CBS4:DFRmin:tRNA-SlDFR-150-scRNA2.1-tRNA:tNOS | Transcriptional unit for copper-inducible expression of a sgRNA targeting SlDFR promoter (position -150) under the regulation of the CBS4 operator and DFRmin promoter. This guide RNA contains a scRNA2.1 useful for transcriptional regulation. |
| GB3787 | AlcA:miniDFR:SlDFR150:scRNA2.1-tRNA:tNOS | 3a1:AlcA:DFRmin:SlDFR-150-scRNA2.1-tRNA:tNOS | Transcriptional unit for ethanol-inducible expression of a sgRNA targeting SlDFR promoter (position -150) under the regulation of the AlcA operator and DFRmin promoter. This guide RNA contains a scRNA2.1 useful for transcriptional regulation. |
| GB3788 | pNos:gRNA-150 SlDFR:scf2.1 | 3a1:pNOS:tRNA-SlDFR-150-scRNA2.1-tRNA:tNOS | Transcriptional unit for the expression of a sgRNA targeting SlDFR promoter (position -150) under the regulation of a NOS promoter. This guide RNA contains a scRNA2.1 useful for transcriptional regulation. |
| GB3789 | p35S: gRNA-150 SlDFR:scf2.1 | 3a1:p35S:tRNA:SlDFR-150-scRNA2.1:tNOS | Transcriptional unit for the expression of a sgRNA targeting SlDFR promoter (position -150) under the regulation of a 35S promoter. This guide RNA contains a scRNA2.1 useful for transcriptional regulation. |
| GB3813 | pNOS:CUP2:VP16:tNos | 3a2:pNOS:CUP2:VP16:tNOS | Transcriptional unit for the expression of CUP2 (copper responsive transcription factor) fused to VP16 under the regulation of the NOS promoter and terminator. |
| GB3820 | pNOS:CUP2:Gal4:tNos-pNOS:AlcR:tNos | 3a1_pNOS:CUP2:Gal4:tNOS-pNOS:AlcR:tNOS | Copper and ethanol responsive transcription factors that binds to CBS operator and AlcA operator respectively. |

**Supplementary Table 2**: Primer pairs for RT-qPCR

| Target gene | Sequence 5'-3' |
|---|---|
| NbDFR | TTCATCTGCGCATCCCATCA |
| | TCCCTACTGAGTTTAAAGGTATCGA |
| NbF-Box | TTGGAAACTCTCTCCCCACTTG |
| | GCTCATTGTTGGATGGGTACCT |
| NbPAL2 | AAGATCGCAAAGCACATTTATTTG |
| | CGAAGTATATGTGATGTTTTGATGGA |

Chapter 1

# Chapter 2

## CuBe: a copper-regulated geminivirus-based processor to express recombinant proteins in *Nicotiana benthamiana*

---

Based on peer-reviewed article: *CuBe: a geminivirus-based copper-regulated expression system suitable for post-harvest activation*

Elena Garcia-Perez, Marta Vazquez-Vilar, Rosa Lozano-Duran, Diego Orzaez

Plant Biotechnology Journal (2024), DOI: 10.1111/pbi.14485

My contribution to this work was essential for its publication. I contributed to the design, performing of the experiments and writing of this manuscript.

## ABSTRACT

The growing demand for sustainable platforms for biomolecule manufacturing has fueled the development of plant-based production systems. Agroinfiltration, the current industry standard, offers several advantages but faces limitations for large-scale production due to high operational costs and batch-to-batch variability. Alternatively, here, we describe the CuBe system, a novel bean yellow dwarf virus (BeYDV)-derived conditional replicative expression platform stably transformed in *Nicotiana benthamiana* and activated by copper sulfate ($CuSO_4$), an inexpensive and widely used agricultural input. The CuBe system utilizes a synthetic circuit of four genetic modules integrated into the plant genome: (i) a replicative vector harbouring the gene of interest (GOI) flanked by cis-acting elements for geminiviral replication and novelly arranged to enable transgene transcription exclusively upon formation of the circular replicon, (ii) copper-inducible Rep/RepA proteins essential for replicon formation, (iii) the yeast-derived CUP2-Gal4 copper-responsive transcriptional activator for Rep/RepA expression, and (iv) a copper-inducible Flp recombinase to minimize basal Rep/RepA expression. Copper sulfate application triggers the activation of the system, leading to the formation of extrachromosomal replicons, expression of the GOI, and accumulation of the desired recombinant protein. We demonstrate the functionality of the CuBe system in *N. benthamiana* plants expressing high levels of eGFP and an anti-SARS-CoV-2 antibody upon copper treatment. Notably, the system is functional in post-harvest applications, a strategy with high potential impact for large-scale biomanufacturing. This work presents the CuBe system as a promising alternative to agroinfiltration for cost-effective and scalable production of recombinant proteins in plants.



### Keywords

Inducible expression, Plant Synthetic Biology, Molecular Farming, *Nicotiana benthamiana*, Copper sensor, Geminivirus, Post-harvest activation

## INTRODUCTION

Manufacturing recombinant biomolecules in plants is becoming a flourishing new bioindustry (Eidenberger et al., 2023). In the last two decades, *Agrobacterium*-mediated transient expression in *Nicotiana benthamiana* by agroinfiltration has become the industry standard for recombinant protein production in plants (Chung et al., 2022; Shanmugaraj et al., 2020; Wani & Aftab, 2022). Product yields obtained with agroinfiltration can be further increased with upstream strategies incorporating viral replicons. These strategies involve the delivery of the gene of interest imbibed in a minimal viral unit capable of self-replicating in the plant cell. The MagnIcon (Marillonnet et al., 2005) and TRBO (Lindbo, 2007) systems are two noteworthy examples of replicative systems for agroinfiltration based on the single-stranded RNA (ssRNA) tobacco mosaic virus (TMV), both producing remarkable yields of recombinant proteins.

While RNA virus-based vectors can offer high titers for transient recombinant protein expression, they have significant limitations. Importantly, they can only integrate inserts of a limited size without disrupting RNA replication, and the absence of proofreading mechanisms in RNA-dependent RNA polymerases reduces transcript fidelity (Ahlquist et al., 2005; Castro et al., 2005; Sainsbury & Lomonossoff, 2008). Furthermore, superinfection exclusion in RNA viruses prevents using the same vector to deliver multiple transgenes (Folimonova et al., 2010; Gleba et al., 2007; Julve et al., 2013).

In contrast to ssRNA viruses, single-stranded circular DNA viruses like geminiviruses handle larger insertions than RNA viral vectors and do not present homologous (or replicative) interference for the co-expression of multiple proteins (Chen et al., 2011). Mor et al. (2003) pioneered the generation of a geminiviral vector for agro-delivery based on bean yellow dwarf virus (BeYDV) (Liu et al., 1998). The minimum viral elements required for BeYDV replication are the long and the short intergenic regions (LIR and SIR, respectively) and the Replication-associated protein (Rep) and RepA. Geminiviral replicons employed in agroinfiltration basically comprise a gene of interest (GOI) flanked by an upstream LIR and a downstream element comprising a SIR and a LIR sequence (LSL vector) (Zhang & Mason, 2006) (Figure 1A). The Rep/RepA bicistronic gene involved in replicating the geminiviral vector can be co-delivered to the plant in a different binary vector or embedded in the replicative vector between SIR and LIR (Chen et al., 2011). Vaccines, antibodies, enzymes, and virus-like particles have been expressed through BeYDV and other geminivirus-based systems (Bhattacharjee & Hallan, 2022; Chen et al., 2011). Recently, the BeYDV technology produced an antibody in *N. benthamiana* that was successfully tested *in vivo*, showing potential as an immunotherapeutic drug (Rattanapisit et al., 2023).

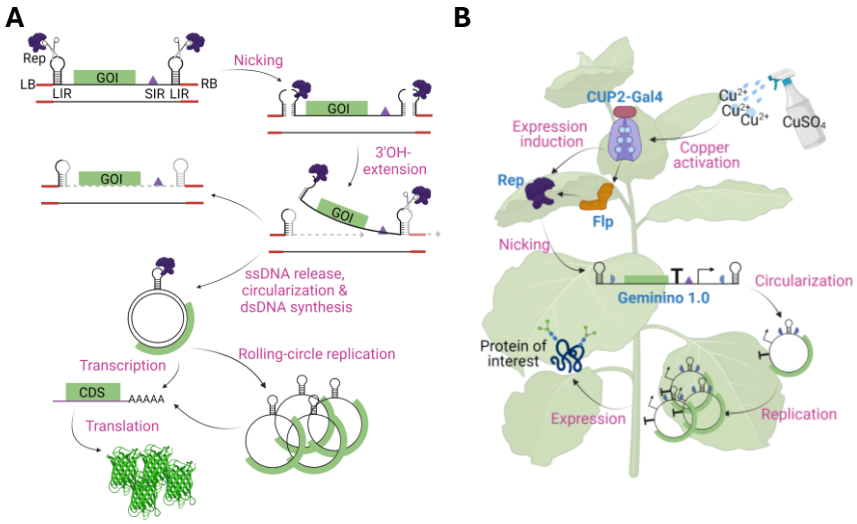



**Figure 1. Functioning models of the BeYDV-derived replicative system and CuBe system for gene expression in plants. A)** Schematic representation of the functioning model of the BeYDV-derived replicative system. The gene of interest (GOI) is flanked by long and short intergenic regions (LIR and SIR, respectively) configuring the LSL vector (pro-replicon) that is embedded between the left and right borders (LB and RB, respectively) of the T-DNA. The BeYDV Replication-associated protein (Rep) nicks the conserved region in LIR and triggers the vector's circularization, enabling plant cell machinery to transcribe and replicate it. **B)** Schematic representation of the functioning of CuBe system stably integrated into the plant. Upon application of $CuSO_4$, the yeast-based copper-responsive factor CUP2-Gal4 activates, binding to the copper binding site (CBS) DNA operator and initiating expression induction. The BeYDV Rep/RepA proteins, under CUP2-Gal4 regulation, have minimized basal expression due to an embedded transcriptional terminator flanked by FRT recombination sites. Copper-induced Flp recombinase expression removes this terminator, allowing Rep expression. The Rep protein then induces nicking and release of the LSL replicative vector (Geminino 1.0), which harbors a transcriptional unit (TU) of interest. In its circular replicative form, the promoter relocates upstream of the gene of interest (GOI), enabling effective expression of the gene of interest and replication of extrachromosomal circular replicons. Figure includes images from Biorender (biorender.com).

Despite its many advantages, agroinfiltration also entails severe limitations for large-scale bioproduction, mainly concerning production costs associated with scaling up (Fischer et al., 2012; Schillberg et al., 2019). Some studies estimate that agroinfiltration represents up to 7.5 % of the operational costs (Buyel & Fischer, 2012), plus the additional capital costs required for bacterial growth, including fermenters, infiltration tanks, and segregated infiltration rooms. Also, transient expression implies a high degree of heterogeneity in the process of vector delivery, leading to batch-to-batch variation that hinders compliance with Good Manufacturing Practices (GMP) (Gleba et al., 2013; Schillberg et al., 2019). Furthermore, the presence of *Agrobacterium* in the resulting biomass makes

downstream processing difficult for specific sensitive applications. Therefore, alternative systems are required to take full advantage of the safety and scalability inherent to agricultural platforms as required by the increasing demand.

One of the possible directions to overcome agroinfiltration shortcomings is the engineering of genome-integrated, conditionally activated replicons. These are silent virus-derived replicative units integrated into the plant genome, which become active for replication only upon activation by a specific trigger. Conditional replicative platforms resemble transient expression systems in that recombinant products only accumulate at the production phase without interfering with plant development and minimizing gene silencing from long-term transgene overexpression (Mortimer et al., 2015). Additionally, an inducible system allows the expression of cytotoxic proteins in the plant without penalty in plant growth (Dugdale et al., 2013). Unfortunately, few examples exist for stable replicons due to the difficulties of engineering conditional systems that efficiently maintain the replicon in a silent state in the absence of the trigger. Werner et al. (2011) developed an ethanol-inducible system based on the MagnIcon TMV strategy. To achieve tight control of replication, the viral vector was deconstructed in a replicon module and a cell-to-cell movement module, with both modules being placed under the control of an inducible promoter.

Compared to ssRNA-based systems, the stable integration of geminivirus-based replicons has proved more amenable. Zhang & Mason (2006) stably transformed potato plants with a BeYDV-based LSL vector in which the BeYDV Rep gene was regulated by an alcohol-inducible promoter (Caddick et al., 1998). In this system, the BeYDV pro-replicon carried a fully functional transcriptional unit (TU) carrying the GOI (GFP), leading to high levels of basal expression. In 2013, Dugdale et al. developed transgenic tobacco plants harbouring the In Plant Activation (INPACT) system based on the tobacco yellow dwarf virus (TYDV). In the INPACT pro-replicon design, the GOI was split into two pieces and only reconstituted upon circularization through the action of ethanol-induced TYDV Rep/RepA proteins. To our knowledge, all plant-made conditional replicons engineered for Molecular Farming processes to this date have employed ethanol as the trigger molecule. Whereas ethanol has low toxicity compared with other triggers employed in plant research, such as glucocorticoids, its use in large-scale installations is problematic due to its relatively high cost, high flammability, volatility, and the requirements of special permissions for its handling (Ma et al., 2021).

In this work, we describe the development of the CuBe system, a conditional BeYDV-derived replicative expression system activated by the $Cu^{++}$ ions of copper sulfate ($CuSO_4$), an inexpensive chemical trigger widely employed in agriculture, including organic agriculture, for crop protection (Lamichhane et al., 2018; Tamm et al., 2022).

The name "CuBe" is derived from the combination of $Cu^{2+}$ and BeYDV. This new system is based on the mild strain of BeYDV (Regnard et al., 2010) and the yeast-based copper responsive factor CUP2-Gal4, which acts on a copper binding site (CBS) DNA operator (Garcia-Perez et al., 2022). Four genetic modules are combined in the plant genome to get a fine-tuned geminiviral control of recombinant protein expression: the LSL replicative vector, the copper-inducible Rep/RepA proteins, the copper-dependent transcriptional activator CUP2-Gal4, and the copper-inducible Flp recombinase. The LSL replicative vector, here called Geminino 1.0, harbours a transcriptional unit (TU) of interest where the promoter is downstream of the GOI and transcriptional terminator. Due to this configuration, the GOI is only expressed from extrachromosomal circular replicons released from the plant chromosome by the Rep protein: upon release, the promoter relocalizes upstream of the GOI in the circular replicative configuration, and the effective transcription can proceed. The expression of the bicistronic Rep/RepA is placed under the transcriptional regulation of the copper-responsive factor CUP2-Gal4. To minimize the basal expression levels of Rep/RepA, a transcriptional terminator flanked by the FRT recombination sites (Lloyd et al., 2022) was embedded into the copper-inducible promoter of Rep/RepA. Upon copper-induced Flp recombinase expression, the terminator flanked by the FRT sites is removed, and Rep expression can occur (Figure 1B). The combination of all four modules forms a synthetic gene circuit that enables refined plant and viral engineering upon induction by an agronomically friendly inducer like copper. Here, we show the development of two *N. benthamiana* plant lines conditionally expressing high levels of eGFP and an anti-SARS-CoV2 antibody following different copper application regimes. Interestingly, we also show that the system is suitable for post-harvest activation regimes, a strategy with interesting implications for large-scale manufacturing.

## RESULTS

### Optimization of the replicative module for the CuBe system

The functioning model for a BeYDV-derived LSL vector (pro-replicon) is depicted in Figure 1A. In the search for an optimal design of the CuBe system with minimal basal expression, the standard sequential order of the transcriptional unit within the pro-replicon (hereafter referred to as standard configuration) was altered, creating alternative configurations where the transcription of the GOI is impeded in the lineal pro-replicon form, but enabled upon circularization. This strategy was earlier proposed by Dugdale et al. (2013) for the so-called INPACT system, where the splitting of the transcriptional unit takes place at the CDS level, leaving a spliceable intron in between, containing the LIR (Figure 2A). Here, we modified the original INPACT configuration to facilitate golden gate-based modular cloning methods (Engler et al., 2014). In the so-called Geminino 1.0 configuration, a leader intron was

placed in the 5´UTR, moving the splitting site and its associated cloning scar outside the CDS, and thus facilitating the cloning of standard DNA elements (Sarrion-Perdigones, et al., 2013). In the Geminino 2.0 design, the intron was located immediately after the start codon (ATG), also avoiding the truncation of the CDS, but reducing chances of spurious CDS expression in the pro-replicon linear state.

The enhanced Green Fluorescence Protein (eGFP) gene was used as the reporter to assay the functionality of each configuration. All four BeYDV-based vectors were agroinfiltrated in *N. benthamiana* leaves in the presence or absence of Rep/RepA, and fluorescence was measured three days post infiltration (dpi). The standard configuration produced the strongest signal, although with high background levels in the absence of Rep/RepA (Figure 2B). In contrast, all three circularization-dependent configurations showed low fluorescence levels in the absence of Rep/RepA. INPACT and Geminino 1.0 were strongly activated upon Rep/RepA-mediated circularization, whereas Geminino 2.0 remained at basal levels. The co-expression of silencing suppressor p19 raised fluorescence in most configurations, including Geminino 2.0, which was then detectable above background levels (Figure 2C). Based on these results, we selected Geminino 1.0 to develop a modular and tightly controllable BeYDV-based replicative expression system.

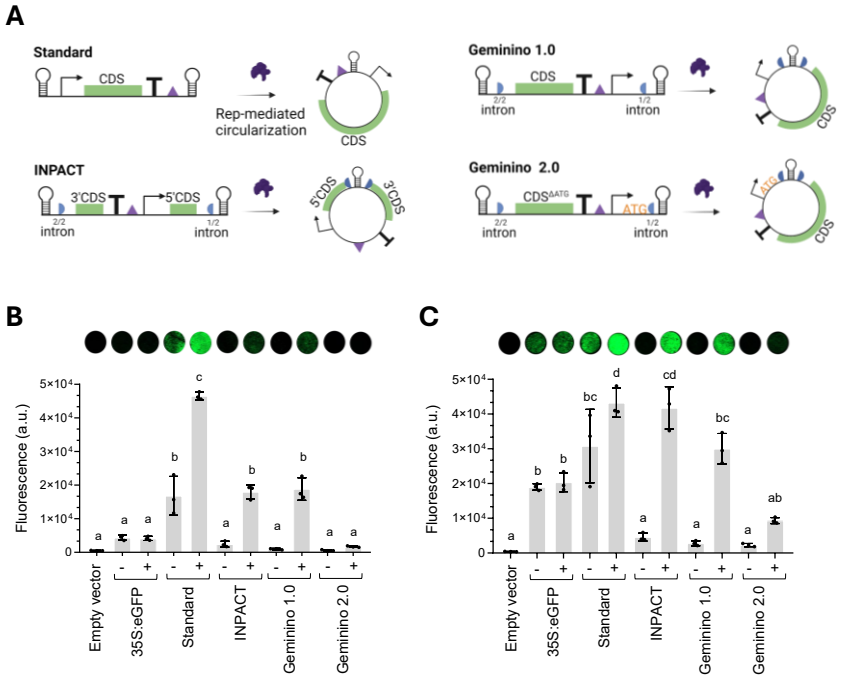



**Figure 2. Design and evaluation of the CuBe vector for regulated protein expression. A)** Schematic representation of CuBe vector configurations. The coding sequence (CDS) of

interest is rearranged with the viral elements to control pre-circularization expression levels. **B)** Transient expression of enhanced Green Fluorescent Protein (eGFP) driven by different CuBe vector configurations. All constructs were introduced into *Nicotiana benthamiana* leaves via agroinfiltration. Empty vector is the negative control, 35S:eGFP is the non-replicative control, and standard, INPACT, Geminino 1.0, and Geminino 2.0 represent configurations in B). The minus (-) and plus (+) signs indicate the absence or presence of Rep/RepA co-expressed with eGFP, respectively. **C)** Same experiment as in B) but including the silencing suppressor p19 from Tomato bushy stunt virus (TBSV). The expression of the Rep/RepA cassette was driven by a pNOS promoter and the p19 gene by a p35S promoter, with both constructs harbouring a tNOS terminator. The OD600 for each Agrobacterium strain in the final agroinfiltrated mix was $OD_{600}$ 0.05 for B) and 0.03 for C). Error bars represent standard deviation (n=3). Statistical analysis was performed using one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05). Variables with the same letters belong to the same statistical group. Figure includes images from Biorender (biorender.com).

**Optimization of a copper-regulated sensor module for the CuBe system**

Owing to its agronomically-friendly characteristics, we selected $CuSO_4$ as the chemical trigger for Rep/RepA activation and subsequent pro-replicon circularization. For this purpose, we adapted a plant copper sensor previously optimized in our laboratory (Garcia-Perez et al., 2022). The initial sensor, based on single copper regulation, comprised two transcriptional units: the first TU (pNOS:CUP2-Gal4) harboured the copper-dependent transcriptional factor CUP2-Gal4 driven by a constitutive pNOS promoter; in the second TU (CBS:minDFR:Rep/RepA), the CBS operator linked to the minDFR minimal promoter was coupled to the BeYDV Rep/RepA CDS. We first tested the ability of the copper sensor to activate a Geminino 1.0-eGFP pro-replicon in transient assays (depicted in Figure 3A). As can be observed in the red dashed line box in Figure 3B, this single-regulated circuitry resulted in copper activation of eGFP with relatively low background levels. Next, constructs containing the sensor module (comprising CBS:minDFR:Rep/RepA and pNOS:CUP2-Gal4 TUs) and the pro-replicon module (Geminino 1.0-eGFP) were stably co-transformed into *N. benthamiana* plants. However, despite our attempts, this strategy was unsuccessful. Only six regenerated shoots were obtained from 120 transformed *N. benthamiana* leaf discs, one of which lacked the CUP2-Gal4 gene, while another one was negative for Rep/RepA. From the remaining four fully equipped transgenic *N. benthamiana* plants, only one showed GFP expression but failed to produce viable seeds and proved unrecoverable by *in vitro* regeneration (Table S1). A parallel experiment with tobacco plants showed similar setbacks (data not shown).

The limitations observed in developing transgenic plants carrying copper-inducible geminiviral elements led us to speculate that basal levels of Rep/RepA may interfere with plant development. To address this concern, we introduced, as an additional layer for Rep/RepA regulation, a memory switch mediated by site-specific

recombination, as earlier proposed by Lloyd et al. (2022) (Figure 3C). An octopine synthase transcriptional terminator (OCSt) flanked by Recombination Sites (RS) was inserted in the minDFR, upstream of the Rep/RepA coding sequence. Two recombination systems were assayed: the Flp recombinase in combination with its respective FRT recognition targets, and the PhiC31 recombinase combined with its attB and attP target sites. In both systems, the expression of the recombinases was also mediated by the CBS:minDFR promoter (Figure 3C), aiming at reducing replicon activation in the absence of copper treatment.

The performance of the new double-layer regulated sensor modules (employing CBS:FRT:OCSt:FRT:minDFR:Rep/RepA or CBS:attB:OCSt:attP:minDFR:Rep/RepA switches respectively) was compared to the previous, single-layer regulation cassette (employing the CBS:minDFR:Rep/RepA construct) in the transient experiments shown in Figure 3B. Fluorescence measurements showed that eGFP expression was low in the absence of copper for all three copper-regulated strategies. Upon copper induction, fluorescence levels in all three systems raised up to levels equivalent to those obtained when employing a constitutive Rep/RepA expression construct (pNOS:Rep/RepA). Interestingly, the Flp system showed lower background levels, and it was therefore selected for a second attempt to stably integrate the CuBe system in *N. benthamiana*.

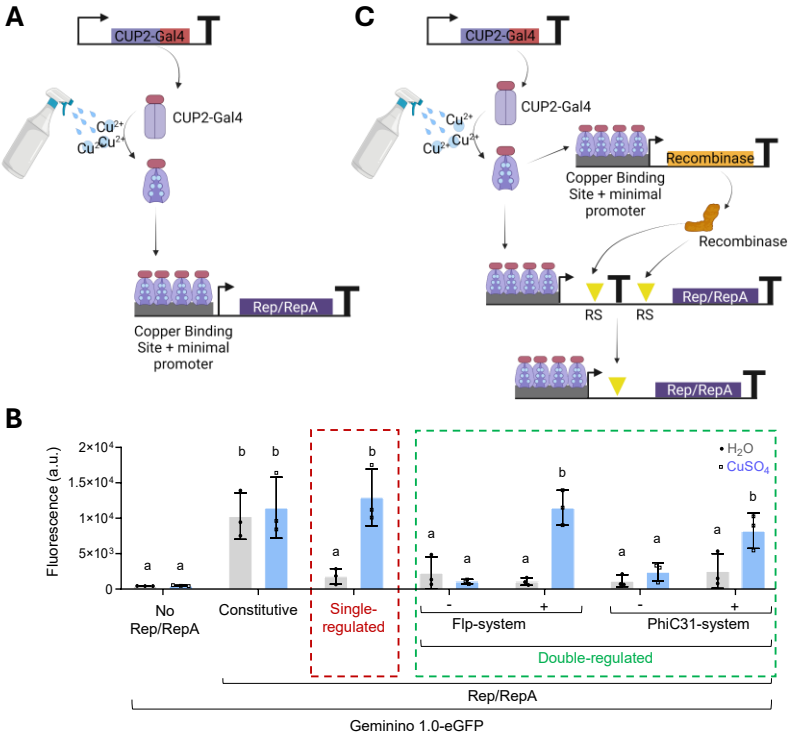


**Figure 3. Design and evaluation of copper-mediated regulation of BeYDV Rep/RepA proteins in the CuBe system**. **A)** Schematic representation of the single copper-inducible expression of BeYDV Rep/RepA proteins. The CUP2 fused to the activation domain Gal4 changes its conformation upon copper binding and binds a copper binding site (CBS) operator, allowing transcription of the downstream Rep/RepA gene. **B)** Transient expression levels of enhanced Green Fluorescent Protein (eGFP) in the Geminino 1.0 vector using different regulatory systems for replication: pNOS-driven Rep/RepA expression (Constitutive), single-layer copper regulation as depicted in A) (Single regulated); double-layer of copper regulation as is depicted in C) (Double-regulation). Both the Flippase (Flp) and the PhiC31 systems were assayed in the presence (+) or absence (-) of the respective recombinase construct. All constructs were introduced into *Nicotiana benthamiana* leaves via agroinfiltration. The $OD_{600}$ for each *Agrobacterium* strain in the final agroinfiltrated mix was $OD_{600}$ 0.03. Fluorescence was measured in leaf discs in water ($H_2O$) or 0.1 mM copper sulfate ($CuSO_4$). **C)** Schematic representation of the double-regulated system. A copper-regulated recombinase acts on the Recombination Sites (RS), flanking a transcriptional terminator (represented as a T), allowing for the de-repression of Rep/RepA transcription. Error bars represent standard deviation (n=3). Statistical analysis was performed using one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05). Variables with the same letters belong to the same statistical group. Figure includes images from Biorender (biorender.com).

### Stable *N. benthamiana* CuBe-eGFP plants

The optimized CuBe gene circuit was stably transformed in *N. benthamiana* following a co-transformation approach involving two T-DNA multigene constructs: the so-called sensor module, on the one hand, encoded the Rep/RepA, the Flp, and the CUP-Gal4; the processor module on the other hand (a.k.a. amplification or pro-replicon module) comprised the Geminino1.0-eGFP cassette (Figure 4A). The segregation of the system into two autonomous modules allows for the independent segregation of each module with the potential to generate new functional circuits through super-transformation or sexual crosses. To facilitate co-transformation, the two modules were cloned into two binary vectors with compatible replication origins. The amplification module was cloned into a mini binary vector, pLX-B3, with replication origin pBBR1 (Pasin et al., ). The sensor module was included in a pDGB3 vector (Vazquez-Vilar et al., 2017) based on pCAMBIA containing the replication origin pVS1. Both vectors were simultaneously used in a two-plasmid/one-*Agrobacterium* strain strategy for multiple delivery of T-DNAs to the plant cell. Furthermore, several additional elements were introduced in the final constructs to improve stability and facilitate the proper regulation of transgene in the genomic context. In the sensor module, each transcriptional unit was separated by intergenic regions derived from NbDFR homeologous genes. In the pro-replicon module, the Geminino 1.0-eGFP was flanked by two matrix attachment regions (MARs) from *Arabidopsis thaliana* chromosome 4 (MAR10) (Pérez-González & Caro, 2019). After being successfully tested in transient assays (Figure 4B), the transformation-ready constructs were co-transformed into *N.*

*benthamiana* plants aiming at creating first-generation CuBe plants, as graphically depicted in Figure 1B. From this co-transformation experiment, eleven T0 plants were recovered and analyzed in genetic screening, aimed at detecting the presence of the two genetic modules (Table S2), and in functional analysis, aimed at detecting eGFP in leaf discs treated with $CuSO_4$ (Figure 4C). All regenerated plants contained both the sensor and processor modules. Surprisingly, careful examination of the sensor module revealed that the OCS terminator (OCSt) originally interrupting Rep/RepA transcription was lost in all the eGFP-expressing lines except for line 9 during the transformation/regeneration steps. Plants 1 and 2 (non-functional) and plant 9 (low eGFP expression) were the only ones maintaining the OCS terminator (Figure 4C). We speculate that the presence of the OCS terminator could have served to maintain low basal expression levels of Rep/RepA during the initial stages of organ development. The basal expression of Flp during the plant regeneration may have led to the excision of the OCS terminator. Despite the absence of one of the elements of the Flp memory switch, T0 plants showed proper copper inducibility in leaf disc assays (Figure 4C). Therefore, we proceeded with the functional assessment of the next generation.

Based on optimal regulatory performance (Figure 4D), we chose four T0 plants for progression to the first transgenic generation (T1 lines 3, 4, 6, and 7). In addition, the offspring of line 9 was also analyzed, resulting in a mixed population of switchable and non-switchable plants, the latter group maintaining the OCS terminator (Figure S1). Functional analysis of 24 plants from each selected T1 line is shown in Figure 4E. Induction of eGFP was observed in all four lines after two days of copper treatment. Segregation analysis of transgene modules in the four T1 lines revealed that all lines contained a single copy of the sensor module (Tables S3 and S4). However, lines 3 and 6 contained two copies of the pro-replicon module, whereas lines 4 and 7 contained a single copy. To facilitate the interpretation of results in the subsequent analysis aimed at optimizing treatment conditions, single-copy lines 4 and 7 were selected to progress to the T2 generation.

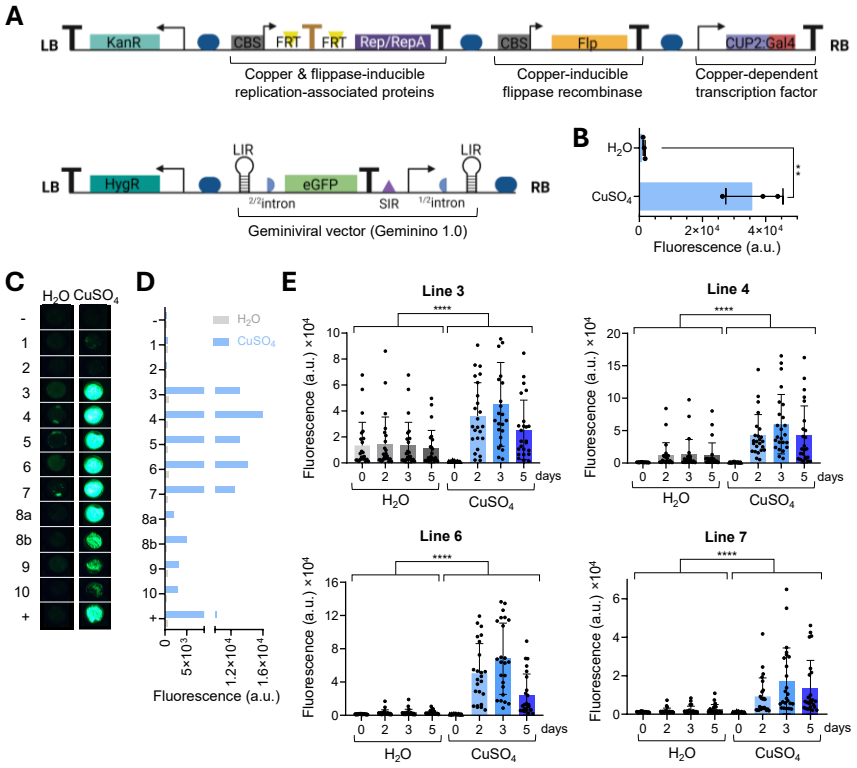



**Figure 4. Design and evaluation of the CuBe system in transgenic *Nicotiana benthamiana*. A)** Schematic representation of the two gene modules used for stable plant transformation. Flp is the flippase recombinase; CUP-Gal4 is the copper-dependent transcriptional factor; KanR is the gene nptII conferring kanamycin resistance; HygR is a hygromycin resistance gene. Arrows indicate transcriptional promoter regions; 'T' indicates transcriptional terminators; the golden 'T' is the OCS T flanked by flippase target sites (FRT); CBS is the copper binding site recognized by CUP2-Gal4; LIR and SIR are intergenic regions of BeYDV; the blue circular shapes are different intergenic regions acting as gene insulators; half circles are the initial ($^{1/2}$ intron) and the final ($^{2/2}$ intron) segments of a unique intron; LB and RB are Left and Right Borders of T-DNA, respectively. **B)** Transient evaluation of constructs in a) by agroinfiltration of *N. benthamiana* leaves and fluorescence measurement of leaf discs incubated either in water ($H_2O$) or 0.1 mM copper sulfate ($CuSO_4$) (n=3). The $OD_{600}$ for each *Agrobacterium* strain in the final agroinfiltrated mix was $OD_{600}$ 0.05. **C)** Fluorescence records at 3 days post-incubation of leaf discs from CuBe-eGFP T0 plants in $H_2O$ or 0.1 mM $CuSO_4$. **D)** Quantification of fluorescence levels shown in C). **E)** Time-course of eGFP activation in leaf discs taken from T1 CuBe plants upon incubation in $H_2O$ or 0.1 mM $CuSO_4$. Each point represents the fluorescence levels of an individual disc taken from a different T1 plant (n=24). Error bars represent standard deviation. Statistical analysis was performed using a T-student (P-value ≤ 0.05) for b) and a one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05) for f). For the ANOVA test in f), the areas under the curve of the $H_2O$ and $CuSO_4$ time-

course treatments were compared. Asterisks represent statistical significance. Figure includes images from Biorender (biorender.com).

**Copper treatment regimes**

Homozygous T2 CuBe-eGFP lines 4.11 and 7.8 (Table S5) were next assayed using different copper treatment strategies, aiming at establishing easily scalable induction regimes that could reduce operational costs in both contained and open-field conditions. We initially analyzed a spray application of plants with a $CuSO_4$ solution combined with 0.05% fluvius surfactant. Interestingly, 72h after treatment, widespread fluorescence was observable throughout the surface of the leaves, indicating appropriate $CuSO_4$ uptake. We experimentally determined optimal copper concentrations for spraying at 5 mM $CuSO_4$ (Figure 5A). Maximum eGFP expression was observed after 4 days of one-time 5 mM $CuSO_4$ application (Figure 5B), the time point at which we stopped the eGFP measurement due to the clear presence of necrosis, possibly due to copper toxicity combined with BeYDV replication (Diamos & Mason, 2019) (Figure S2).

Next, we aimed to investigate the possibility of employing systemic $CuSO_4$ treatments for the induction of the CuBe system. We first conducted a small-scale assay using seedlings grown in MS-agar plates to monitor copper movement through the plant. Seedlings were placed in MS-agar media with 0.05, 0.5, and 5 mM $CuSO_4$ (Figure S3). Remarkably, as shown in Figure 5C, a systemic copper signal was successfully established, leading to strong eGFP activation in distal leaves. Two days after copper application in agar, necrosis started in CuBe seedlings at 5 mM, and no fluorescence was reported at any recorded time point (Figure S4). Seedlings in 0.05 mM did not show necrosis nor fluorescence activation. Seedlings at 0.5 mM started to show systemic fluorescence emission on day 3 post-copper application, and they did not show necrosis symptoms except localized brown spots on day 5 (Figure S3). On day 6, after changing seedlings to new agar media without copper, the new emerging leaves did not show fluorescence, indicating that eGFP expression is totally dependent on copper presence (Figure S4).

We then attempted $CuSO_4$ systemic activation by watering plants grown in pots. However, all our efforts to reproduce activation in these conditions were unsuccessful. As an alternative strategy for large-scale systemic induction, we then assayed copper treatments in plants grown in hydroponic conditions by transferring the plants from a standard hydroponic media to an equivalent media supplemented with 1 mM $CuSO_4$. After three days in the $CuSO_4$-enriched hydroponic media, strong and widespread fluorescence levels were recorded in plant leaves, indicating successful systemic movement of the copper signal in hydroponic conditions (Figures 5D and 5E). This systemic movement of copper through the plant also produced toxicity symptoms, causing browning of roots and leaf curling. However,

no necrosis was reported in the aerial and eGFP-producing vegetative material (Figure S5). The accumulation of eGFP protein in leaves activated systemically by hydroponic induction was estimated in Coomassie-stained gels, yielding approximately 120 µg of recombinant protein per gram of fresh weight, representing a 2.6% total soluble protein (TSP) (Figure S6). Re-evaluation by Western blot of the copper-dependent induction in hydroponic conditions showed remarkably low background levels of eGFP expression in non-induced conditions (Figure 5F). Given these results, we speculate that the failure to establish systemic activation in soil-grown plants could be due to the immobilization of copper because of interactions with substrate components, although this aspect will require further investigation.

Finally, owing to the fast response of the CuBe induction system, we became interested in assaying post-harvest induction in leaves as an easily scalable option. To this end, detached leaves were incubated in water tanks containing different concentrations of $CuSO_4$ ranging from 0.01 mM to 1 mM. Using this approach, eGFP activation was low. Tissue damage was visible at 2 days post-incubation on 1 mM $CuSO_4$ and turned strong on day 5 (Figure S7). Lower copper concentrations ended up in irregular and relatively low eGFP induction at the assayed concentrations (Figure 5G). Next, we modified the post-harvest strategy, incubating the detached leaves in 1 mM $CuSO_4$ plus 0.05% fluvius for only 5 minutes and subsequently moving them to a water container for the remaining incubation time. As can be observed in Figure 5h, a much higher fluorescence signal was reported at 2 days post-copper incubation (dpci) than in the previously assessed methods and without signs of damage (Figure S8). For further optimization of the post-harvest induction, additional activation regimes were tested. Figure 5i shows that a maximum eGFP fluorescence was obtained at 4 dpci using a pulse treatment of 5 mM $CuSO_4$. However, this regime also caused initial symptoms of necrosis on day 4 (Figure S9). Therefore, we decided to set the standard treatment for post-harvest induction as a 5-min pulse incubation in 1 mM $CuSO_4$ 0.05% fluvius, followed by 3 days of incubation in water, obtaining 1.4% TSP of eGFP (Figure S10). Using these activation conditions, we also calculated the number of copies of CuBe-eGFP replicon generated per haploid genome. As can be seen in Figure 5J, under induced conditions, the estimated transgene dosage reached more than $2x10^4$ copies per haploid genome, whereas background levels in uninduced conditions reached up to 50 replicon copies per genome (Figure S11). As a final test, we used this standard regime to compare the eGFP levels produced by leaves harvested at different stages of plant development. Figure 5K shows the fluorescence emission from leaves taken from 5-week-old plants compared to 8-week-old (flowering) plants (see Figure S12 for leaf size comparison).

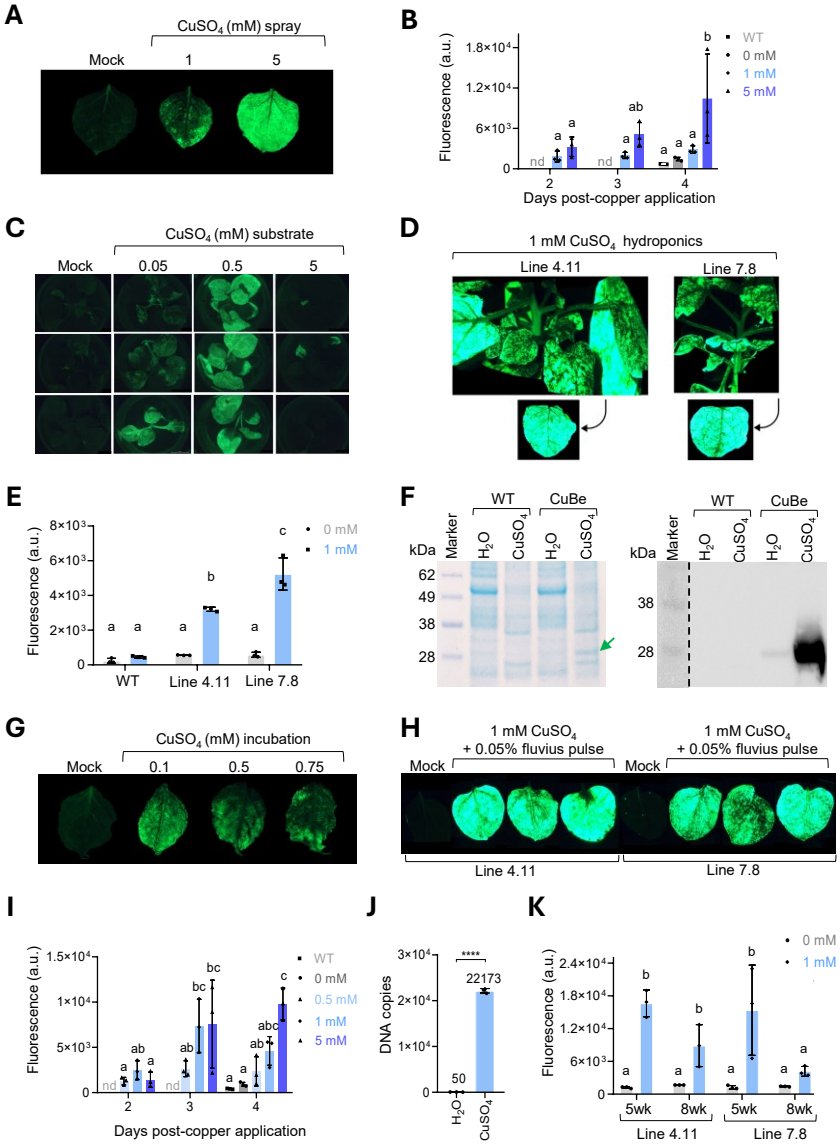



**Figure 5. Copper induction analyses of stable *Nicotiana benthamiana* CuBe-eGFP plants.** **A)** Leaves of CuBe-eGFP plants sprayed with $H_2O$ or copper sulfate ($CuSO_4$) and visualized at the fluorescent magnifier after 3 days post application (dpa). **B)** Time course of fluorescence emitted from sprayed leaves of CuBe-eGFP plants treated with different $CuSO_4$ concentrations. Each data point represents an independent leaf. **C)** Fluorescence reported at 5 dpa of CuBe-eGFP seedlings grown in agar media. **D)** Fluorescence image of plants grown in 1mM $CuSO_4$ hydroponics at 3 dpa. **E)** Quantification of fluorescence emitted in plant leaves from wildtype (WT) and transgenic lines 4.11 and 7.8 treated as in d). Each data point represents an independent plant. **F)** Coomassie-stained SDS-gel electrophoresis of total soluble proteins (left

panel) and Western blot of the same samples (right) from WT and 7.8 CuBe-eGFP plants induced hydroponically as in D). **G)** Representative fluorescence images of CuBe-eGFP leaves after 2 days of continuous post-harvest incubation in $CuSO_4$. **H)** Representative fluorescence images of CuBe-eGFP leaves pulse-treated (5 min) after harvest with a $CuSO_4$ + 0.05% fluvius solution and kept in $H_20$ for 2 days. **I)** Time-course of fluorescence levels in leaves after post-harvest pulse treatments with different $CuSO_4$ concentrations. **J)** DNA copy number estimation of Geminino 1.0 vector per plant haploid genome in CuBe-eGFP plants. Each point represents a qPCR technical replicate coming from three post-harvested leaves pulse-treated with 1 mM $CuSO_4$ + 0.05% fluvius and kept in $H_20$ for 3 days. (n = 3). **K)** Influence of plant development stage in the accumulation of eGFP. 5wk indicates 5-weeks-old plants and 8wk indicates 8-weeks-old plants. Nd stands for no data. Each data point represents a leaf from a different CuBe-eGFP plant. Error bars represent standard deviation (n=3). Statistical analysis was performed using a one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05). Variables with the same letters belong to the same statistical group. A T-student (P-value ≤ 0.05) was used for J). Figure includes images from Biorender (biorender.com).

**Production of antibody n72_hγHC by the CuBe system in transgenic *N. benthamiana* plants**

After optimizing the CuBe system using the eGFP reporter protein, we proceeded to test the versatility of the system by producing the single-chain anti-SARS-CoV-2 antibody n72_hγHC . This antibody is formed by a camelid variable heavy chain (VHH) domain that targets the receptor binding domain (RBD) of the SARS-CoV-2 spike (S) protein fused to the Fc region of a human IgG1. For developing CuBe-n72_hγHC plants, we employed the same sensor module as in CuBe-eGFP, whereas the pro-replicon module was reassembled using the n72_hγHC CDS instead of eGFP (Figure 6A). Nineteen co-transformed T0 lines were regenerated, and their functionality was screened in an ELISA test using plates coated with RBD (Figure 6B). Eleven of the 19 screened plants presented regulated production of the n72_hγHC antibody. Since a secretion signal peptide was fused to the n72_hγHC, we also analysed the anti-RBD activity of an apoplast wash obtained from T0 plants. As shown in Figure 6C, the T0 CuBe-n72_hγHC plants secrete enough n72_hγHC to detect RBD without further processing of the plant material. Six T0 plant lines were brought to T1 generation, germinated in selective media, induced by post-harvest pulse incubation of leaves in $CuSO_4$ and re-analysed in the ELISA test shown in Figure 6D. Remarkably, all T1 plants analysed were positive for copper-induced production of n72_hγHC. Mendelian segregation analysis revealed that line 8 contained two copies of the Geminino 1.0 module, which may explain its higher reactivity levels towards RBD (Tables S6 and S7). Protein extracts of the T1 8.6 plant were analysed by Western blot to confirm the tight copper-regulated expression of the n72_hγHC (Figure 6E). The Western blot analysis also indicated that leaves at different developmental stages, both fully expanded (FE) and non-fully expanded (NE), produce the antibody at similar levels. This was later confirmed by ELISA using plants treated with $CuSO_4$ following a hydroponic regime, as shown in Figure 6F.

Antibody tittering curves of crude and apoplast extracts reached end tittering points up to 3 x $10^{-4}$ (Figure 6G). In leaf post-harvest induction conditions, the calculated average number of CuBe-n72_hγHC replicons raised from 4 to almost 1x$10^4$ copies per haploid genome (Figure 6H). As shown in Figure 6I, Coomassie-stained SDS-PAGE from apoplast fluid after a post-harvest pulse induction showed a differential band attributed to n72_hγHC. Single-step protein affinity purification of the n72_hγHC from clarified protein extract produced an estimated yield of 4.5 µg/g FW purified single-chain antibody (Figure 6J and S13).

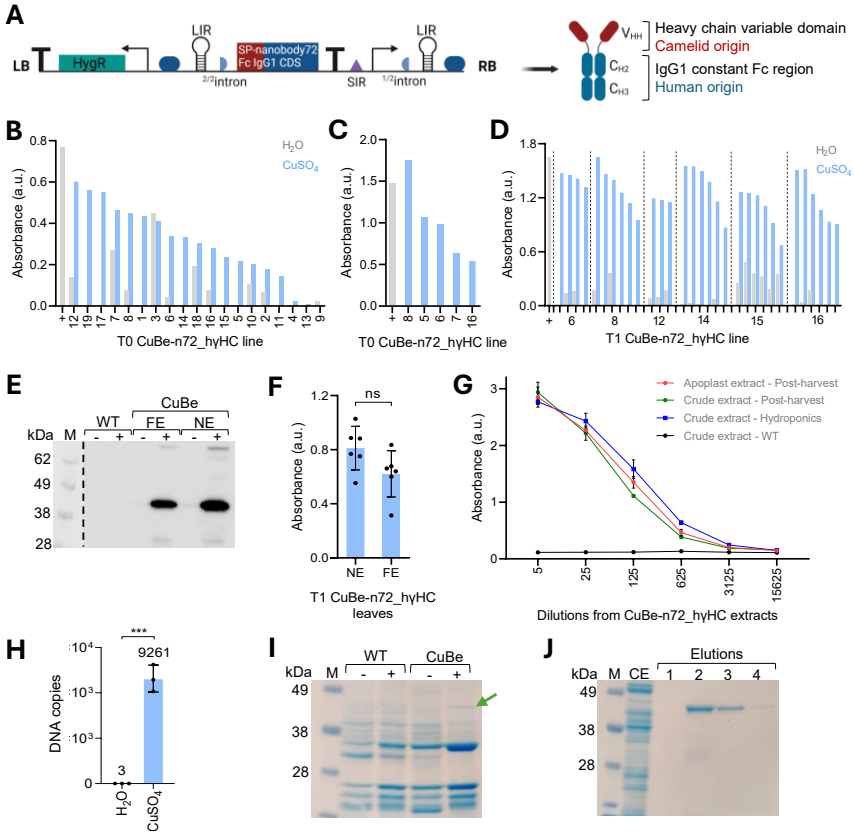



**Figure 6. Copper-induced production of n72_hγHC antibody in stable *Nicotiana benthamiana* CuBe plants. A)** Schematic representation of the pro-replicon module used for n72_hγHC expression. SP stands for signal peptide. HygR indicates the hygromycin resistance gene used as a selection marker. Arrows indicate promoter regions; LIR and SIR are intergenic regions of BeYDV; the blue circular shapes are intergenic regions acting as gene insulators; half circles are the initial (1/2 intron) and the final (2/2 intron) segments of a unique intron; LB and RB are Left and Right Borders of T-DNA, respectively. **B)** Antigen (RBD) ELISA screening of the T0 CuBe-n72_hγHC *N. benthamiana* plants. ELISA plates were coated with RBD and incubated with total soluble protein extracts from leaf discs of plants 1 to 19. Discs were previously incubated for 4 days in $H_2O$ or in 0.1 mM $CuSO_4$. **C)** Antigen (RBD) ELISA of undiluted apoplast

fluid obtained from selected post-harvested T0 leaves incubated for 2 days in 0.1 mM $CuSO_4$. **D)** Antigen (RBD) ELISA screening of different T1 CuBe-n72_hγHC plants by inducing post-harvested leaves with a 1 mM $CuSO_4$ + 0.05% fluvius pulse-treatment for 5 min and quantified 3 days post-pulse. **E)** Western Blot analysis of total soluble protein extract from fully expanded (FE) and non-fully expanded (NF) leaves of T1 8.6 CuBe-n72_hγHC plant. Plants were treated with $CuSO_4$ as indicated in d) (+) or left untreated (-). **F)** Antigen ELISA of crude extracts from different T1 plants derived from line 8, activated following a hydroponics regime with 1 mM $CuSO_4$. NE indicates non-fully expanded leaves; F, full-expanded leaves. Each data point represents a leaf from a different plant. Bars represent mean ± SD, n = 6, and asterisks result from a T-student comparison (P-value ≤ 0.05). In all the ELISA bars on b), c), d), and f), the background absorbance from wild-type crude extract samples was subtracted. As a positive control (+), extract from a wildtype plant agroinfiltrated with a constitutive n72_hγHC construct. **G)** Antigen ELISA curves of apoplast and crude extracts from T1 line 8 n72_hγHC plants to compare antibody binding activities of n72_hγHC induced and extracted by different methods. The X-axis indicates the 1:5 dilution factor of the extracts from 1:5 to 1:15625. **H)** DNA copy number estimation of Geminino 1.0 vector per plant haploid genome in CuBe-n72_hγHC plants. Each point represents a qPCR technical replicate coming from three plant leaves post-harvested treated with $CuSO_4$ as indicated in d) (+) or untreated (-) (n = 3). **I)** Coomassie-stained SDS-PAGE gel of apoplastic fluid obtained from leaves of wildtype (WT) or T1 8.6 CuBe-n72_hγHC treated with $H_2O$ (-) or with $CuSO_4$ as indicated in d). The arrow points to the n72_hγHC antibody band size. **J)** Coomassie-stained SDS-PAGE gel showing affinity purification of n72_hγHC. M stands for molecular marker (1 kDa); CE is a crude protein extract; elutions 1-4 correspond to fractions of a protein-A purification procedure. Post-harvested leaves were treated as indicated in D) for n72_hγHC expression. The figure includes images from Biorender (biorender.com).

## DISCUSSION

Plants harbouring stably integrated replicons offer a solution for the sustainable and cost-effective manufacture of recombinant products. Production yields can be optimized by segregating plant development from the activation of the pro-replicon. For this separation, conditional activation is required to control the timing, location, and amount of transgene expression. To date, two main transgenic systems, 'MagnICON' and 'INPACT,' excel at minimizing background expression in the absence of the inducer while maximizing protein production upon activation in plants. The TMV-based 'MagnICON' system (Werner et al., 2011) was implemented in *N. benthamiana* to yield up to 4.3 g GFP/kg fresh weight upon induction by a combination of root drenching and vapor of ethanol for five to seven days. The TYDV-based 'INPACT' system (Dugdale et al., 2013) reported yields of 25 mg GUS protein/g TSP (2.5% TSP) in *N. tabacum* or 0.1 g vitronectin/kg fresh weight in *N. benthamiana*. Our data show that yields obtained with the CuBe system, which reached up to 0.12 g GFP/kg fresh weight (2.6% TSP), are comparable to the INPACT system. The yield reported for the n72_hγHC antibody is more modest. However, antibody yields are highly dependent on variable chain sequence (Mason et al.,

2012), and although not suitable for yield comparison, the CuBe-n72_hγHC plants serve as an example of the consistency, versatility, and remarkable inducibility of the CuBe system. The average 50 and 4 copies of replicon per haploid genome observed under non-induced conditions in CuBe-eGFP and CuBe-n72_hγHC lines, respectively, can be probably attributed to the occasional escape of a replicon in isolated cells (Figure S14). The accumulated evidence seems to indicate that, despite the massive increase in transgene copy number associated with the replication of geminivirus-based systems, transcription and translation levels are not as high as those of, e.g., TMV-based systems. This could be due to a lack of intrinsic gene silencing suppressor activity and/or to the absence of cell-to-cell movement that, in other viral systems, such as MagnICON, expands the expression foci in case the activation signal does not reach all the cells. Silencing in CuBe might be mitigated by silencing suppressors, as reported by Zhang & Mason (2006). On the other hand, RNA virus-based systems suffer superinfection exclusion (Folimonova et al., 2010; Gleba et al., 2007; Julve et al., 2013), a property that makes it unsuitable for producing protein complexes. Therefore, the selection between geminivirus and TMV-based systems for recombinant protein production should be guided by specific application requirements.

MagnICON and INPACT, as well as other plant conditional expression systems, rely on the ethanol-responsive AlcA:AlcR gene sensor for activation (Caddick et al., 1998). However, using ethanol as an inducer entails several practical, economic, and storage drawbacks. Its high volatility makes field application practically unfeasible in open environments. Additionally, being a highly flammable compound, ethanol requires special permits for acquisition and storage tanks with high resistance to ethanol's corrosive attack (Ma et al., 2021). One of the main advantages of the CuBe system lies in its activation by $Cu^{2+}$ ions, which can be supplied, e.g., in the form of $CuSO_4$, a low-cost agrochemical suitable for indoors and open-field applications without evaporation risks and with eco-friendly validation (Lamichhane et al., 2018; Tamm et al., 2022). Although plants use $Cu^{2+}$ at physiological concentrations as a cofactor for a wide range of cell processes (Tsang et al., 2021), the CuBe system can be maintained in an OFF state when plants are grown on regular plant growth media (in our case, 2.23 mg/L $Cu_2SO_4$), becoming activated only when intracellular concentrations reach non-physiological conditions (after treatment with 1.25 g/L $CuSO_4{\cdot}5H_2O$). The origins of the $Cu^{2+}$ sensor employed here, which derives from the yeast CUP2 system that evolved to respond only to toxic $Cu^{2+}$ concentrations (Buchman et al., 1989), explain this behaviour. Furthermore, as shown in this work, under hydroponic conditions, the plant root is not able to block the uptake of the excess copper efficiently. Consequently, the high $Cu^{2+}$ signalling queue required for CuBe activation can be expanded systemically, allowing its transmission to distal tissues. Indeed, hydroponics stands out as an efficient method for copper

conditional activation in our hands, leading to one of the highest production levels among the different application methods assayed. However, a major drawback of this method is the severe necrosis observed in the roots. It is described that exposure to high copper concentrations can lead to acute toxicity, with reduced mitotic activity in the root tips, ROS production, and nutrient uptake disruption (Küpper & Andresen, 2016). Further industrial optimizations could lead to an optimal balance between toxicity and production levels using hydroponics.

Next to copper conditional expression, CuBe incorporates other specific design features that deserve consideration in the context of inducible plant biofactories. The circuity was deliberately designed to create a triple regulatory layer comprising transcriptional activation, pro-vector circularization, and site-specific recombination. Adaptation to standardized modular cloning (Vazquez-Vilar et al., 2017) facilitated parts assembly and simplified exchange of GOIs and will eventually facilitate further vector optimization by incorporation of new regulatory DNA elements, as shown by Dusek et al. (2020). The CuBe design incorporated intergenic spacers separating different TUs to enhance the isolation of the genes of interest. We also flanked the pro-replicon by MAR regions, which are described to minimize interline variation and protect transgenes from deletions (Pérez-González & Caro, 2019). Another distinctive feature is the use of the two-plasmids/one *Agrobacterium* co-transformation strategy. The use of two separated binary vector backbones with compatible origins (Pasin et al., 2017) harbouring the sensor and the amplification modules, respectively, led not only to a high transformation efficiency in our hands but, most importantly, also to independent integration of the two functional modules. In this way, once a pair of integration loci has proved optimal, offspring carrying each of the modules can be sexually segregated, and the resulting elite single-module plants can be re-used in new combinations made by super-transformation or sexual crossing.

The incorporation of a recombinase-based terminator switch (Lloyd et al., 2022), which was initially aimed at minimizing leaky pro-vector activation, turned out to facilitate the regeneration of transgenic plants harbouring the CuBe system. We obtained 11 double-transformed, functional n72_hγHC transgenic plants from 90 leaf discs. As was observed in CuBe-eGFP plants, the antibody's expression was regulated only when the terminator was at least partially eliminated (Figure S1). We were unable to regenerate transgenic plants in previous attempts with construct designs lacking this repressor switch. Furthermore, although not fully comparable, this seems to be a higher transformation rate than that reported for sequential super-transformation in similar approaches (Dugdale et al., 2013). We favour the explanation that the terminator switch worked as a robust break, repressing leaky Rep/RepA expression during early organogenesis steps, which could otherwise lead to cell death and meristem abortion. It should be noticed that, whereas

transcriptional activation follows a continuous behaviour, recombinase-based regulation follows a Boolean dynamic, with only two possible statuses (active or inactive). Consequently, although terminator removal in the absence of copper seems to have happened promiscuously during plant regeneration, some cell lineages may have kept the recombinase brake sufficiently long to survive the undifferentiated stage in which other transcriptional control mechanisms are inefficient. A similar phenomenon was reported by Guiziou et al. (2023), which led to the suggestion that integrase expression could occur during gametogenesis and then be transmitted to all cells of the next generation. The loss of the terminator at early regeneration stages did not translate into copper deregulation in differentiated tissues. Interestingly, when transgenic plants maintaining the terminator switch were agroinfiltrated with a constitutive flippase, efficient removal of the terminator was observed by PCR analysis (data not shown), indicating the proper functioning of the switch. However, the same plants were not activated under copper treatment. A possible explanation for this behaviour is the posttranslational gene silencing of the flippase transgene, as earlier reported for tyrosine integrases such as Flp operating in transgenic plants (Guiziou et al., 2023b; Liu et al., 2021).

Another remarkable feature of CuBe, as compared with other commonly used conditional systems in plant biology, is the ability to launch a sustainable activation with just a pulse (five-minute) incubation in post-harvest conditions. This effect can be related to the long stability and persistence of the Cu metal ions in the plant tissues, in contrast with other highly volatile and/or more prone to degradation (organic) signalling molecules. This method can greatly simplify and reduce the cost of industrial applications, as the same $CuSO_4$ solution can be used to induce multiple plant batches.

In summary, the CuBe system represents a successful integration of a potent inducible viral system into transgenic plants. It stands out from previous systems due to its copper-inducible activation, which allows scaling up plant-based protein production in open field conditions. We have demonstrated that the system activation is highly versatile and efficient both in vivo by spraying or hydroponic irrigation and ex vivo by pulse-incubation of leaves. This system has the potential to be integrated into several plant species since BeYDV can infect multiple dicotyledonous plants. The system's flexibility in terms of activation and the repertoire of plants it could be integrated into highlight its potential as a powerful tool in the field of Molecular Farming.

## MATERIALS AND METHODS

### Cloning and assembly of the GoldenBraid constructs

All the genetic constructs used in this work were assembled using GoldenBraid (GB) (Sarrion-Perdigones et al., 2017). All the employed expression vectors are detailed in Table S8, along with their identification labels on the GoldenBraid repository.

### Agroinfiltration experiments

Genetic constructs were transiently expressed in wild-type *Nicotiana benthamiana*. *Agrobacterium tumefaciens* C58 harbouring expression vectors were grown in LB with 50 µg/ml kanamycin/spectinomycin and 50 µg/ml rifampicin. After 16 h at 28ºC/250 rpm, bacteria cultures were resuspended in infiltration buffer (10 mM MES, 10 mM MgCl2, and 200 µM acetosyringone, pH 5.6) and infiltrated at $OD_{600}$ 0.1 into three expanded leaves of five-week-old *N. benthamiana* plants for co-expression of GB elements. The $OD_{600}$ for each *Agrobacterium* strain in the final agroinfiltrated mix was $OD_{600}$ 0.05 for Figures 2b and 4b, and $OD_{600}$ 0.03 for Figures 2c and 3b. Plants were maintained in a controlled environment (24°C light/20°C dark, 16h light/8h dark).

### Fluorescence quantification of eGFP expression

Leaf discs (5 mm diameter) were collected and placed adaxial side down in a black 96-well plate (Corning, #3792) with 250 µl distilled water per well. Fluorescence (excitation 490 nm, emission 510-570 nm) was measured using a GloMax®-Multi Detection System (Promega). For time courses, plates were incubated at 25°C, 16h light/8h dark, and 60-70% humidity. Alternatively, 50 µl of total soluble protein extract was loaded per well and measured similarly.

### Fluorescence imaging of eGFP expression

The fluorescence imaging of plant material was done by using a magnifier Leica MZ 16 F. The conditions for the imaging were 2-sec exposition, 10x gain, 0 saturation, and 1.04 gamma. Whole-leaf or plant segment images were assembled by stitching together image fragments captured with the magnifier using AutoStitch software.

### Stable transformation of *Nicotiana benthamiana*

Sensor and replicative modules were co-electroporated into *A. tumefaciens* LBA4404. Saturated cultures ($OD_{600}$ 0.2) were inoculated overnight in TY medium with antibiotics (kanamycin and spectinomycin) and 200 µM acetosyringone. Leaf discs from six-week-old surface-sterilized *N. benthamiana* leaves were placed onto co-cultivation medium (MS medium at pH 5.7, 1 mg/L 6-benzylaminopurine (BAP), 0.1 mg/L naphthalene acetic acid (NAA) and 9 g/L phytoagar) for 24 h. Explants were immersed in the *A. tumefaciens* culture for 15 minutes before being placed back onto the co-cultivation medium. After 2 days, discs were transferred to selection medium (including 100 mg/l kanamycin, 20 mg/l hygromycin, 200 mg/l carbenicillin)

for shoot regeneration. Rooted shoots were transplanted to soil and grown in a greenhouse at 18-21.5 °C and artificial light supplementation between 8:00 AM and 10:00 PM if radiation is below 150W/$m^2$.

**Copper induction of CuBe plant material**

For fluorescence screening of T0 plants, 5 mm-diameter discs were placed adaxial side down on a 96-well plate containing 250 µl 0.1 mM $CuSO_4$ per well. For spray induction, leaves were sprayed with 1 or 5 mM $CuSO_4$ + 0.05% fluvius surfactant on both sides. For agar-diffusion induction, sterilized seedlings were transferred to 12-well transparent plates (VWR, #10062-894) containing MS medium (pH 5.7) and 4.5 g/L phytoagar with 0.05, 0.5, and 5 mM $CuSO_4$. For watering-in-pot induction, the soil-perlite substrate was watered with 1 mM or 5 mM $CuSO_4$ solution (single or 9-day application). Hydroponic induction involved transferring seedlings to 3.5 cm x 3.5 cm x 4 cm rockwool cubes pre-wetted with nutrient solution. Once the seedling roots had grown through the rockwool cubes, they were transplanted to polystyrene boards placed atop trays filled with nutrient solution. An air pump was used for oxygenation. The spacing between plants was maintained at 4.5 cm. After 15 days of growth, the nutrient solution was replaced with 1 mM $CuSO_4$. Post-harvest induction used detached leaves incubated in water with 0.01-1 mM $CuSO_4$ or dipped in 1 or 5 mM $CuSO_4$ + 0.05% fluvius for 5 minutes followed by water incubation.

**Total soluble protein extraction**

Frozen leaves were ground in liquid nitrogen. 75 mg of powder was homogenized with ice-cold PBS buffer (80 mM $Na_2HPO_4{\cdot}7H_2O$, 20 mM $NaH_2PO_4{\cdot}H_2O$, 100 mM NaCl, pH 7.4) (1:3 w/v) and centrifuged (14,000 x g, 15 min, 4°C). The supernatant was used as a total soluble protein extract. For antibody purification, 8 g of ground tissue was homogenized in 20 mM phosphate buffer (7.4 mM $NaH_2PO_4$, 12.6 mM $Na_2HPO_4{\cdot}7H_2O$, pH 7) with 0.5 mM PMSF (Sigma-Aldrich, #78830) and 50 mM sodium ascorbate, centrifuged (10,000 xg, 15 min, 4°C), and filtered (0.22 µm) to remove debris.

**Antibody Purification by Affinity Chromatography**

The filtered protein extract was loaded onto a column with Protein A agarose resin (ABT Technology) for gravity-flow purification following the manufacturer's protocol. Bound antibodies were eluted with 100 mM citrate buffer (pH 3). The elution fractions (250 µl each) were immediately neutralized with 37.5 µl of 1 M Tris-HCl buffer, pH 9.

**Isolation of apoplast fluid**

Leaves expressing the apoplast-tagged protein were infiltrated with ice-cold PBS under vacuum for 1 minute, followed by slow release. After removing the excess buffer, leaves were rolled, placed into a 20 ml syringe in a 50 ml tube, and centrifuged (1000 xg, 10 min, 4°C). The apoplastic fluid was concentrated 6.5-fold

using 10 kDa Amicon Ultra-4 10K centrifugal filters (Millipore) after centrifugation (3,000 xg, 20 min, 4°C).

**SDS-PAGE and Western Blot Analysis**

Proteins were separated by SDS-PAGE using MES-SDS buffer (pH 7.3) on NuPAGE 10% Bis-Tris gels (Invitrogen) under reducing conditions, with equal protein loading confirmed by Bradford assay. Proteins were visualized by Coomassie blue staining. For Western blot, proteins were transferred by semi-wet blotting (XCell II™ Blot Module, Invitrogen, Life Technologies) to a Polyvinylidene Difluoride (PVDF) membrane (Amersham Hybond™-P, GE Healthcare). The membrane was blocked with 2% ECL Prime (GE Healthcare) in PBS-T (PBS with 0.1% (v/v) Tween-20). For eGFP detection, the membrane was probed with 1:10,000 anti-GFP (proteintech, 66002-1-Ig), and subsequently incubated with 1:10,000 HRP-labeled anti-mouse IgG secondary antibody (GE Healthcare). To detect the anti-SARS-CoV-2 antibody, the membrane was incubated with 1:20,000 HRP-conjugated anti-human IgG (Sigma-Aldrich, #A8792). ECL Prime Western blotting Detection Reagent (GE Healthcare) was used to detect the HRP-antibodies in a Fujifilm LAS-3000 imager for chemiluminescent detection of target proteins in the membrane.

**Protein quantification**

The total protein concentration in the soluble protein extraction or the apoplastic fluid was quantified by the Bio-Rad Protein following the manufacturer's guidelines. For quantifying the protein of interest, samples were separated by SDS-PAGE and stained with Coomassie Blue (as described above). A BSA standard curve was generated on the same gel for quantification. Densitometry analysis of bands using ImageJ software was used to estimate the concentration of the protein of interest. Briefly, band intensities (integrated peak areas) of BSA and the target protein were measured. The BSA standard curve then allowed the estimation of the relative abundance of the protein of interest.

**ELISA**

Costar EIA/RIA plates (Corning) were coated overnight at 4°C with 0.4 µg RBD (RayBiotech, #230-30162) in coating buffer (15 mM $Na_2CO_3$, 35 mM $NaHCO_3$, pH 9.6) per well. After washing and blocking with 2% ECL Prime blocking agent in PBS with 0.1% (v/v) Tween-20 for 2 h at RT, plates were incubated with total protein or apoplastic fluid samples for 1 h at RT. Following washes, plates were incubated with 1:2,000 HRP-conjugated anti-human IgG (Sigma-Aldrich, #A8792) for 1 h. Plates were then washed, and o-phenylenediamine dihydrochloride SIGMAFAST™ OPD tablets (Sigma-Aldrich, #P9187) were added as the substrate for the HRP colorimetric reaction. The colorimetric reaction was stopped with 3 M HCl, and 492 nm absorbance was measured. The absorbance values plotted in the graphs are the

result of subtracting the background levels of a control sample from a wildtype plant.

**Quantitative PCR analysis**

Total DNA from 100 mg leaf tissue was isolated using the DNeasy Plant Kit (Qiagen). Calibration curves for Geminino 1.0-eGFP or -n72_hγHC constructs were generated using purified plasmids mixed with 5 ng of *N. benthamiana* wildtype genomic DNA. Plasmid copy number for the curve was calculated based on the *N. benthamiana* haploid genome size (7.75 x $10^8$ bp) and DNA mass-to-length conversion (660 g/mol). DNA concentration was measured by Qubit fluorometric quantification. DNA integrity and RNA contamination were assessed by gel electrophoresis. qPCR was performed on 5 ng of total DNA (3 technical replicates) using SYBR Green dye and primers amplifying the Geminino 1.0 35S terminator-SIR region. Housekeeping gene F-BOX amplification was used for verifying equal quantities of genomic DNA in the compared samples. Ct values were interpolated against calibration curves, and the DNA copy number was normalized to the *N. benthamiana* haploid genome size.

## SUPPLEMENTARY MATERIAL

### Supplementary figures

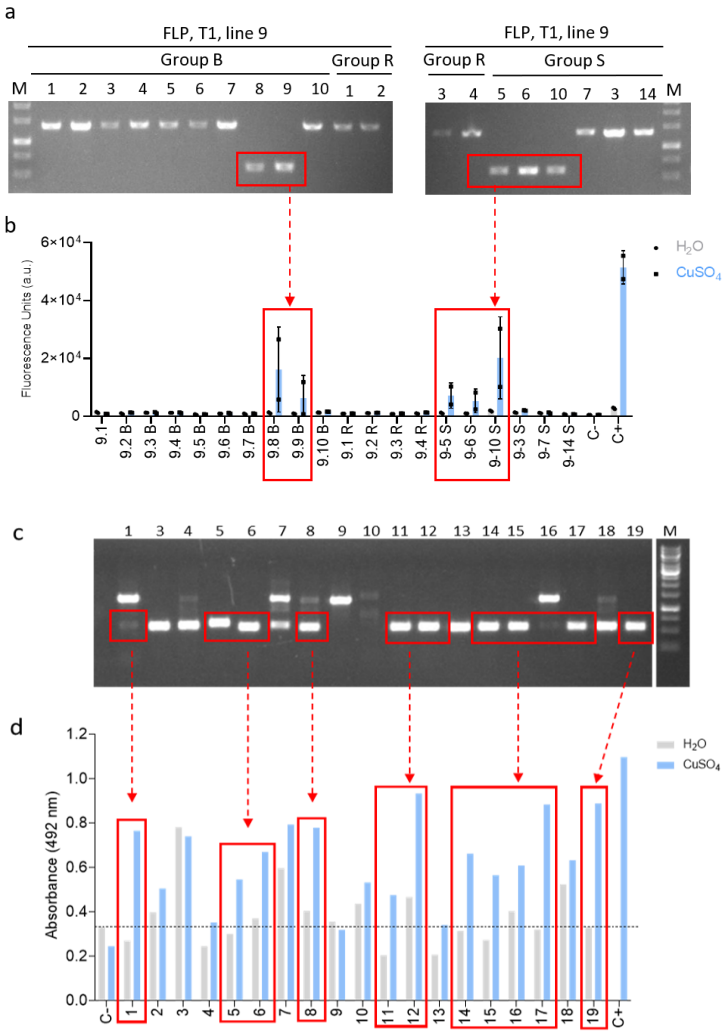



**Figure S1.** Analysis of the OCS terminator (OCSt) presence in CuBe lines. **a)** Electrophoresis gel of PCR products from amplification of the region between the start codon of Rep/RepA and the distal region before the CBS operator (see Figure 3a) in T1 line 9 CuBe-eGFP plants. Only the plants where the OCSt is gone by recombination (red square) are functional, as reported in b). **b)** Expression of eGFP in T1 line 9 plants evaluated in a) after incubation of leaf discs in water or copper sulfate. C- is negative control (empty vector infiltrated in wild-type plant), and C+ is positive control (transient expression of CuBe system in wild-type plant). **c)** Electrophoresis gel of PCR products from amplification of the region between the start codon of Rep/RepA and the distal region before the CBS operator in T0 CuBe-n72_hγHC plants. Only plants where the OCSt is partially or fully gone by recombination (red square) are functional,

as reported in d). **d)** Expression of n72_hγHC in T0 n72_hγHC plants evaluated in a) after incubation of leaf discs in water or copper sulfate. C- is negative control (empty vector infiltrated in wild-type plant), and C+ is positive control (transient expression of CuBe system in wild-type plant).

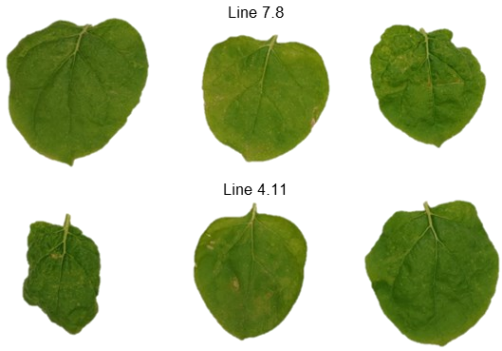



**Figure S2**. Symptoms caused by 5 mM $CuSO_4$ spraying of CuBe-eGFP leaves after four days. The first row of leaves is from transgenic line 7.8, and the second row is from line 4.11.

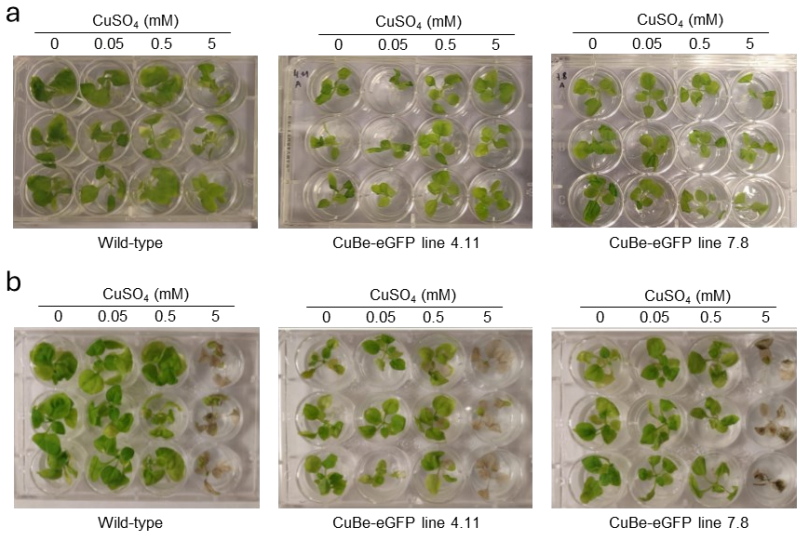
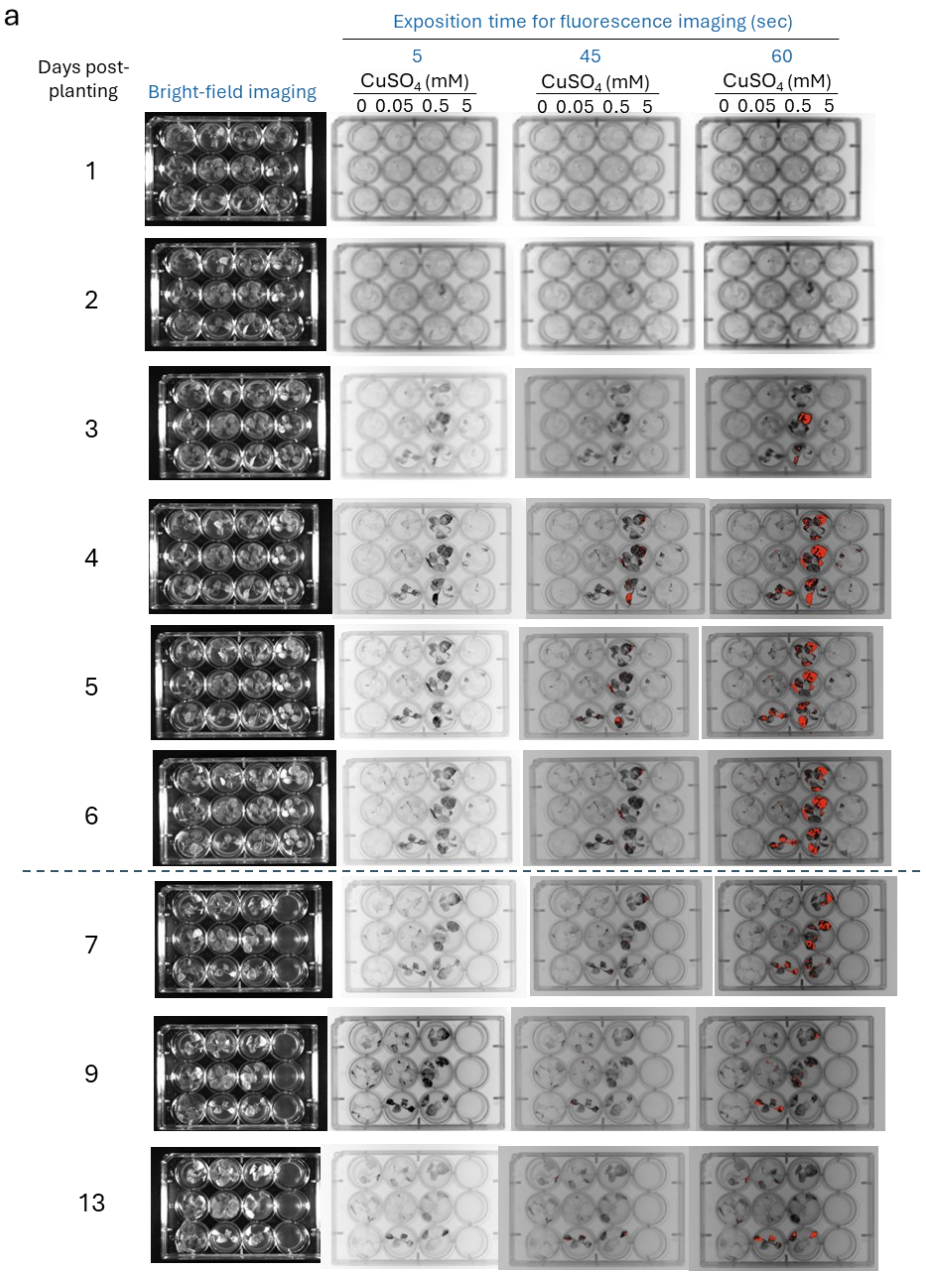



**Figure S3**. CuBe-eGFP seedlings grown in MS-agar plates to monitor copper movement through the plant at time 0 post-sowing (a) and 5 days post-sowing (b).


a
Exposition time for fluorescence imaging (sec)
5
45
60
Days post-planting
Bright-field imaging
CuSO4 (mM)
0 0.05 0.5 5
CuSO4 (mM)
0 0.05 0.5 5
CuSO4 (mM)
0 0.05 0.5 5
1
2
3
4
5
6
7
9
13

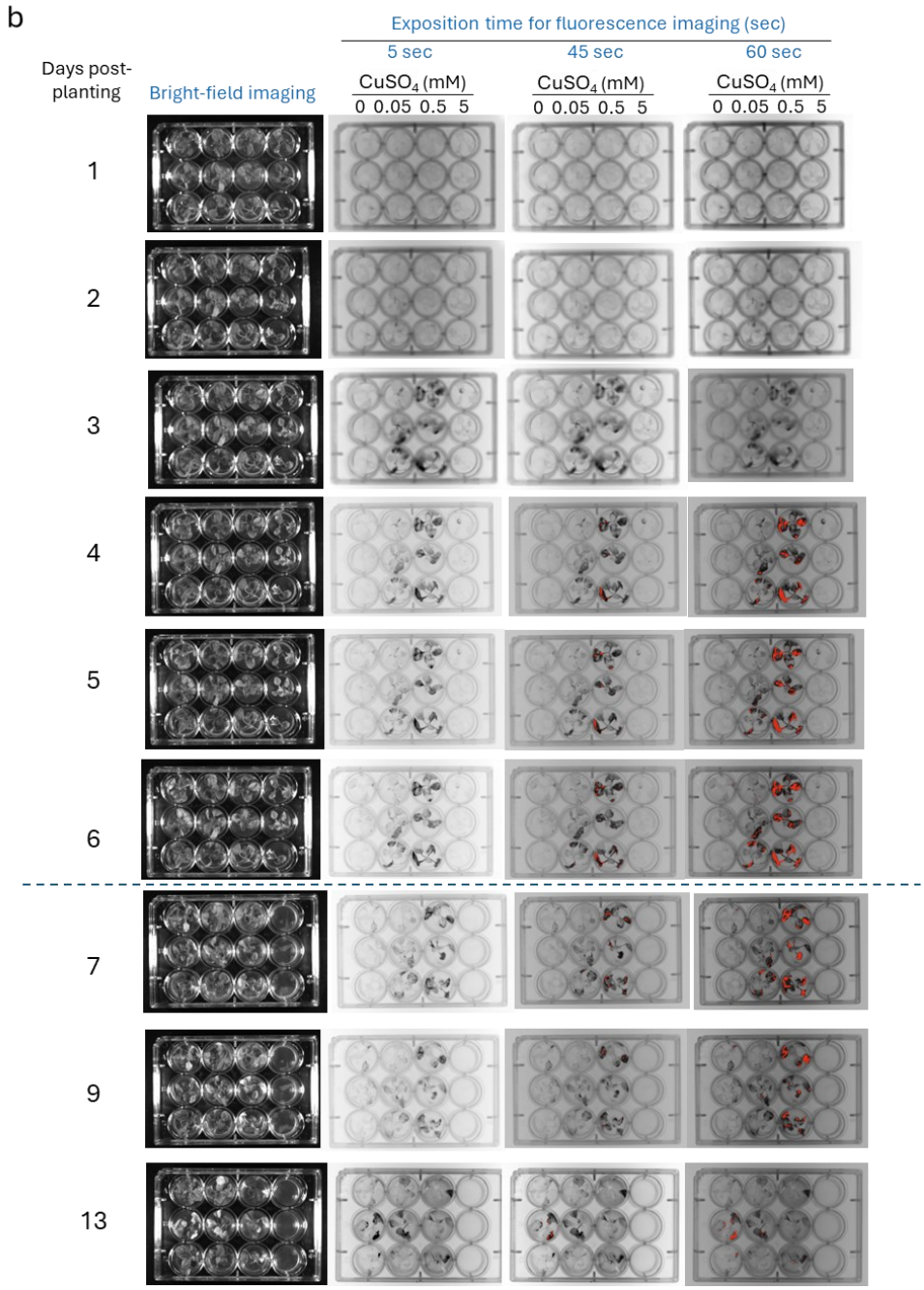



**Figure S4**. CuBe-eGFP seedlings grown in MS-agar plates to monitor copper movement through the plant. Seedlings in a) are CuBe-eGFP line 4.11, and seedlings in b) are CuBe-eGFP line 7.8. The blue line indicates a transfer of seedlings to new plates without $CuSO_4$.

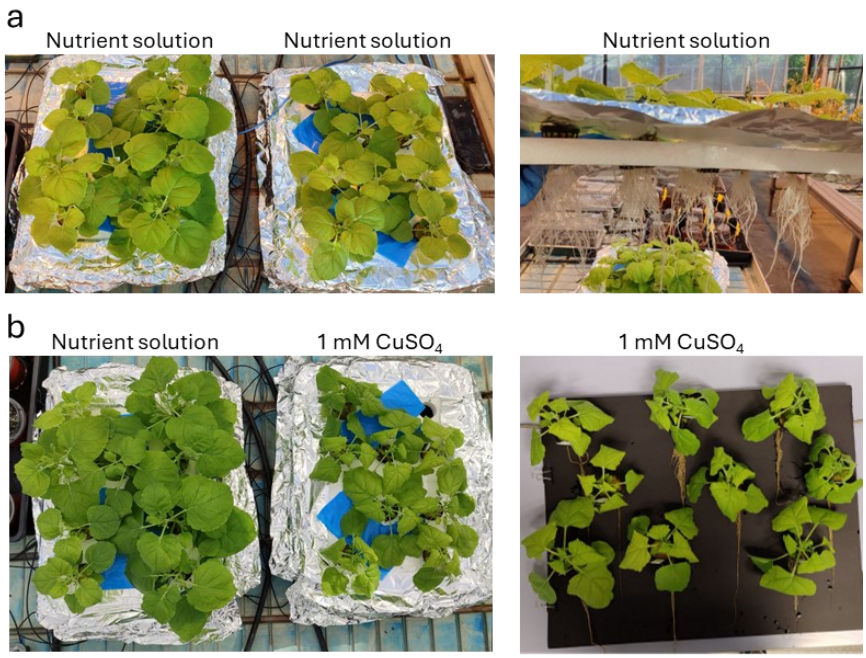
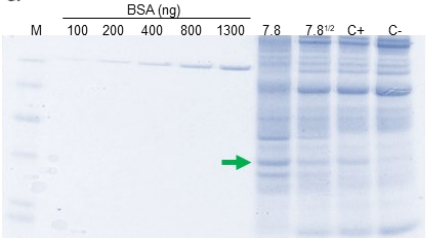
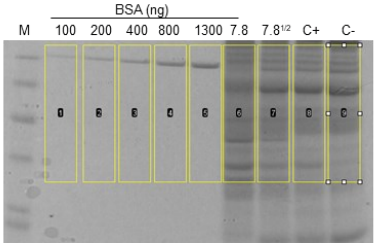



**Figure S5.** 5-week-old CuBe-eGFP plants grown in hydroponic conditions. a) Plants before exchanging the standard nutrient solution. On the right, white and healthy roots are shown. b) Plants three days after exchanging nutrient solution by 1 mM $CuSO_4$ solution in the tray on the right in the picture on the left. The picture on the right shows the browning of the roots after copper treatment.

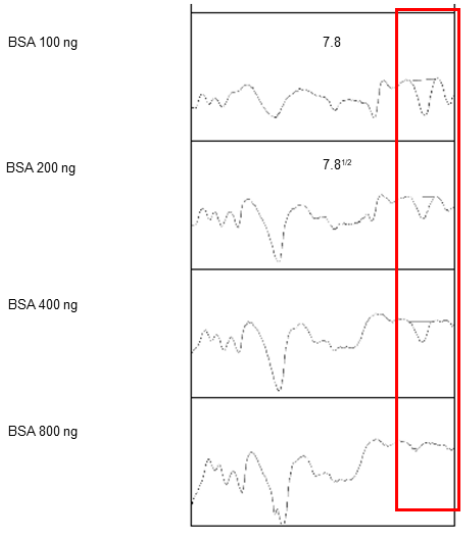
a



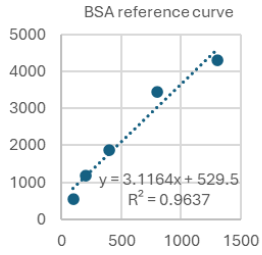

b




c




| BSA (ng) | Band intensity |
|---|---|
| 100 | 555.6 |
| 200 | 1187.7 |
| 400 | 1875.0 |
| 800 | 3452.8 |
| 1300 | 4302.1 |




d

| Sample | Band intensity | ng protein | ng eGFP/ug TSP* | % TSP |
|---|---|---|---|---|
| 7.8 | 2935.7 | 772.1 | 25.8 | 2.6 |
| 7.8 1/2 | 1671.9 | 366.6 | 12.3 | 1.2 |
| C+ | 1517.5 | 317.0 | 10.6 | 1.1 |
| C- | 39.1 | -157.4 | _ | _ |

*in 29890 ng loaded protein

e

| Sample | ng protein | µg eGFP/g FW* |
|---|---|---|
| 7.8 | 772.1 | 119 |
| 7.8 1/2 | 366.6 | 56 |

*19.5 µl of sample were loaded in the gel. As we kept a 1:3 relation FW (mg):PBS (µl), in the gel there are 6.5 mg FW

**Figure S6**. Quantification of eGFP expressed by CuBe-eGFP plants induced by hydroponic application of 1 mM $CuSO_4$. a) Coomassie-stained gel containing: BSA protein as reference for protein quantity (lanes 2-6), total soluble protein (TSP) of CuBe-eGFP line 7.8 (lane 7), same sample but diluted 1:2 with WT TSP (lane 8), positive control (C+) as TSP from transient

expression of eGFP in WT plant (lane 9), and negative control (C-) as TSP from WT plant (lane 10). The green arrow points out the band corresponding to eGFP. b) Same gel as a) imaged by Fujifilm LAS-3000 for quantification. c) Integrated peak areas (red square) corresponding to the gel bands of BSA and the target proteins. The peak area from BSA bands and their known concentrations were used to create a reference curve to estimate the relative abundance of the protein of interest. The table on the right contains the resulting correlation between protein quantity and band intensity, which is represented in the graph below. d) Percentage of total soluble protein (% TSP) corresponding to eGFP. e) Micrograms of eGFP per gram of fresh weight of plant material (µg eGFP/g FW).

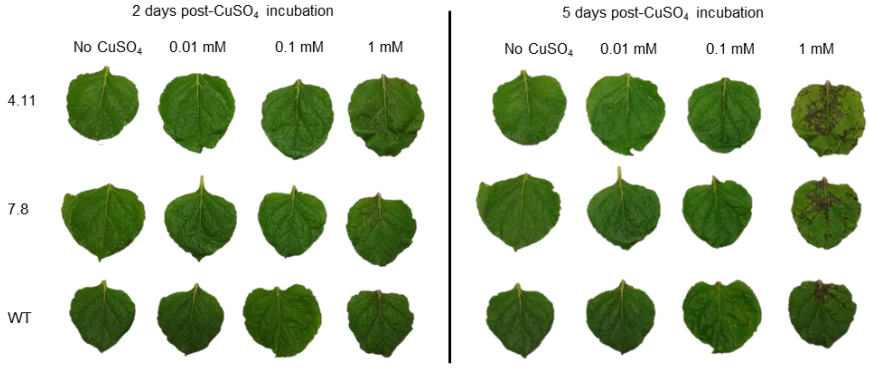
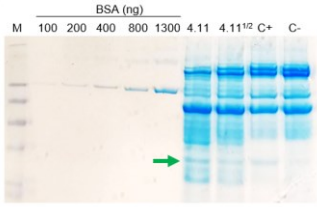



**Figure S7**. Symptoms caused by leaf incubation from CuBe-eGFP plants line 4.11 (first row) and 7.8 (second row), and WT (third row) in $CuSO_4$ solution for 2 (left) or 5 days (right).

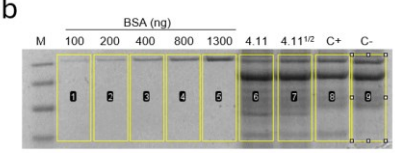

**Figure S8.** Leaves from CuBe-eGFP plants line 4.11 (first row) and 7.8 (second row), and WT (third row) after a 5-min incubation in 1 mM $CuSO_4$ + 0.05% fluvius solution and maintenance in water for 2 days.

**Figure S9**. Damage symptoms caused by incubation of leaves from CuBe-eGFP plants line 4.11 in water for 4 days after a 5-min incubation in 5 mM $CuSO_4$ + 0.05% fluvius solution.

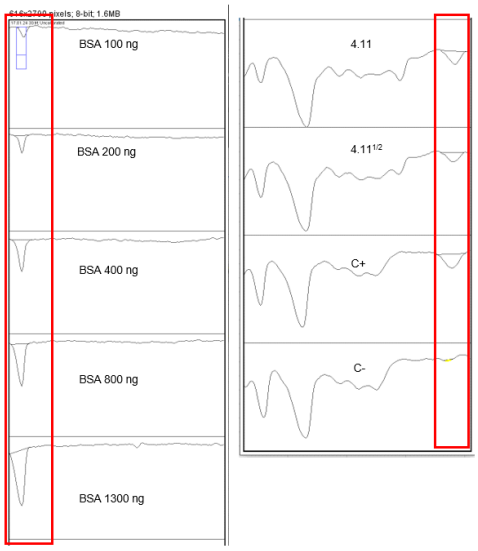
a



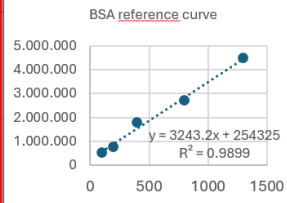

b




c




| BSA (ng) | Band intensity |
|---|---|
| 100 | 539.163 |
| 200 | 785.870 |
| 400 | 1819.355 |
| 800 | 2719.355 |
| 1300 | 4488.912 |




d

| Sample | Band intensity | ng protein | ng eGFP/ug TSP* | % TSP |
|---|---|---|---|---|
| 4.11 | 1304.6 | 323.7 | 14.0 | 1.40 |
| 4.11^1/2 | 961.2 | 217.9 | 9.4 | 0.94 |
| C+ | 1634.7 | 425.7 | 18.4 | 1.84 |
| C- | 49.4 | -63.2 | _ | _ |

*in 23105 ng loaded protein

e

| Sample | ng protein | µg eGFP/g FW* |
|---|---|---|
| 4.11 | 323.7 | 66 |
| 4.11^1/2 | 217.9 | 44 |

*14.7 µl of sample were loaded in the gel. As we kept a 1:3 relation FW (mg):PBS (µl), in the gel there are 4.9 mg FW

**Figure S10.** Quantification of eGFP expressed by CuBe-eGFP plants inducted by pulse 1 mM $CuSO_4$ copper incubation of post-harvested leaves. a) Coomassie-stained gel containing: BSA

protein as reference for protein quantity (lanes 2-6), total soluble protein (TSP) of CuBe-eGFP line 4.11 (lane 7), same sample but diluted 1:2 with WT TSP (lane 8), positive control (C+) as TSP from transient expression of eGFP in WT plant (lane 9), and negative control (C-) as TSP from WT plant (lane 10). The green arrow points out the band corresponding to eGFP. b) Same gel as a) imaged by Fujifilm LAS-3000 for quantification. c) Integrated peak areas (red square) corresponding to the gel bands of BSA and the target proteins. The peak area from BSA bands and their known concentrations were used to create a reference curve to estimate the relative abundance of the protein of interest. The table on the right contains the resulting correlation between protein quantity and band intensity, which is represented in the graph below. d) Percentage of total soluble protein (% TSP) corresponding to eGFP. e) Micrograms of eGFP per gram of fresh weight of plant material (µg eGFP/g FW).

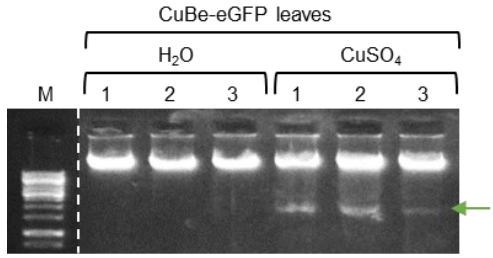



**Figure S11**. Electrophoresis gel from genomic DNA of three leaves that come from different CuBe-eGFP plants. Each number indicates one of the three assayed plants. Leaves were pulse-treated (5 min) after harvest with a $CuSO_4$ + 0.05% fluvius solution and kept in $H_20$ for 3 days. The green arrow points to the band associated with the Geminino 1.0 replicative unit whose replication is inducted by $CuSO_4$.

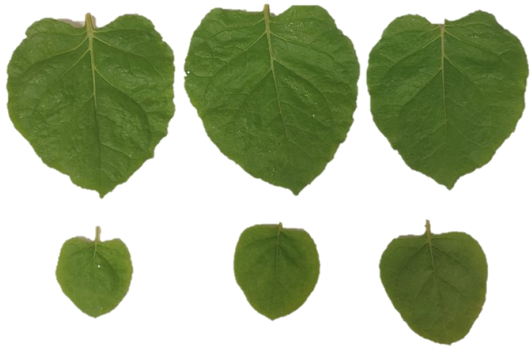
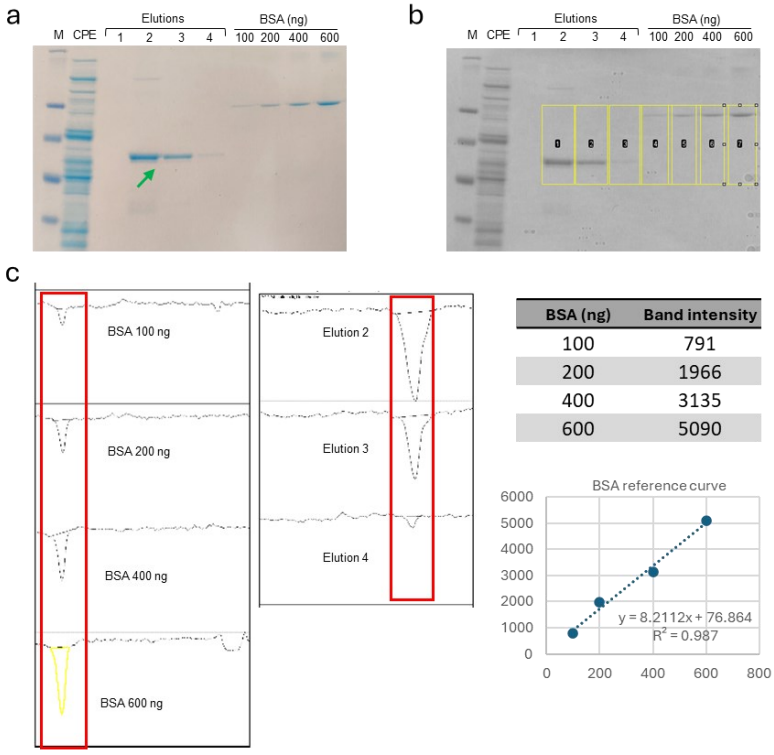


**Figure S12**. Leaves from 8-week-old (flowering) plants (first row) and 5-week-old plants (second row) of CuBe-eGFP line 4.11 after a 5 min pulse incubation in 1 mM $CuSO_4$ + 0.05% fluvius, followed by 3 days incubation in water.




| BSA (ng) | Band intensity |
|---|---|
| 100 | 791 |
| 200 | 1966 |
| 400 | 3135 |
| 600 | 5090 |

d

| Sample | Band intensity | ng protein | ng n72_hγHC /µl elution | µg in 250 µl elution | µg/g FW | Total µg/g FW |
|---|---|---|---|---|---|---|
| Elution 1 | 14537 | 1761.1 | 90.3 | 22.6 | 2.8 | |
| Elution 2 | 7835 | 944.9 | 48.5 | 12.1 | 1.5 | 4.5 |
| Elution 3 | 869 | 96.5 | 4.9 | 1.2 | 0.2 | |
| | | | *considering 19.5 µl loaded in the gel | | *considering 8 g fresh weight of plant material | |

**Figure S13**. Quantification of purified n72_hγHC expressed by CuBe-n72_hγHC plants inducted by pulse 1 mM $CuSO_4$ copper incubation of post-harvested leaves. a) Coomassie-stained gel containing BSA protein as reference for protein quantity (lanes 7-10), crude protein extract (CPE) of CuBe-n72_hγHC (lane 2), and elutions 1, 2, 3, and 4 from protein A purification (lanes 3-6). The green arrow points out the band corresponding to n72_hγHC. b) Same gel as a) imaged by Fujifilm LAS-3000 for quantification. c) Integrated peak areas (red square) corresponding to the gel bands of BSA and the target proteins. The peak area from BSA bands and their known concentrations were used to create a reference curve to estimate the relative abundance of the protein of interest. The table on the right contains the resulting correlation between protein quantity and band intensity, which is represented in the graph below. d) Micrograms of n72_hγHC per gram of fresh weight of plant material (Total µg/g FW).

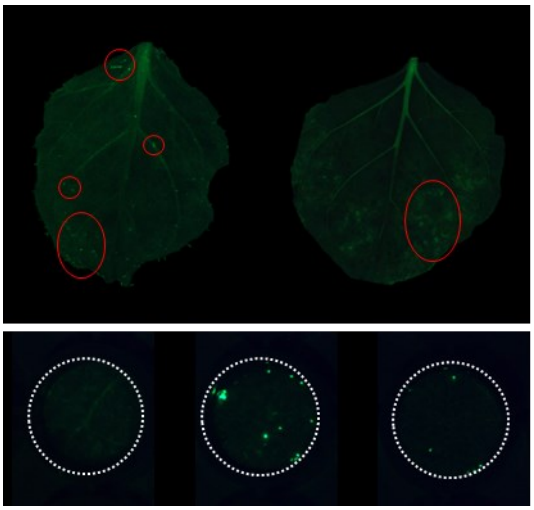


**Figure S14.** Occasional escape of replicon in isolated cells in untreated leaves from CuBe-eGFP plants. Red circles in the leaves indicate green fluorescence. The panel below shows three leaf discs where green dots might be attributed to the leaky expression of eGFP because of escaped Geminino 1.0.

**Supplementary tables**

**Table S1**. Genetic screening of transgenic plants regenerated with single-regulation for Rep/RepA and constitutive CUP-Gal4.

| Plant line | CUP PCR | Rep PCR | Functionality | Seed production |
|---|---|---|---|---|
| 1 | - | + | nd | - |
| 2 | + | + | - | - |
| 3 | + | - | nd | - |
| 4 | + | + | - | - |
| 5 | + | + | - | - |
| 6 | + | + | + | - |

For the PCR columns, '+' indicates positive PCR result for amplification of CUP-Gal4 or Rep/RepA as indicated by the presence of a specific band in electrophoresis analysis, and '-' indicates negative PCR. 'nd' indicates no data. '+' in functionality columns means the plant expressed eGFP upon copper induction. The negative symbol in the seed production column indicates that no plants produced viable seeds.

**Table S2**. Genetic screening of T0 CuBe plants.

| Plant line | CUP | eGFP | OCSt |
|---|---|---|---|
| 1 | + | + | + |
| 2 | + | + | + |
| 3 | + | + | na |
| 4 | + | + | - |
| 5 | + | + | -/+ |
| 6 | + | + | - |
| 7 | + | + | -/+ |
| 8a | + | + | -/+ |
| 8b | + | + | - |
| 9 | + | + | + |
| 10 | + | + | - |

In the CUP and GFP columns, '+' indicates positive PCR results for CUP-Gal4 and eGFP amplification, respectively, as indicated by the presence of a specific band in electrophoresis analysis, and '-' indicates negative PCR. 'nd' indicates no data. '+' in the OCSt column indicates that the PCR product matches the expected size in electrophoresis gel if OCSt is present, '-' indicates that the band matches the absence of OCSt, and '-/+' indicates that the plants contain both bands matching the presence and the absence of OCSt.

**Table S3**. Segregation analysis of Kanamycin resistance for T1 CuBe-eGFP lines.

| T1 line | Observed R | Observed S | Expected R (3:1) | Expected S (1:3) | P-value | Analysis of T-DNA insertion |
|---|---|---|---|---|---|---|
| **3** | 20 | 9 | 21.75 | 7.25 | 0.4530 | 3:1; one locus |
| **4** | 10 | 12 | 16.5 | 5.5 | 0.0014 | 1:1; one locus, insertion disrupting gene essential for gamete |
| **6** | 9 | 16 | 18.75 | 6.25 | 0.0001 | 1:2; one locus, insertion disrupting gene essential for gamete |
| **7** | 15 | 10 | 18.75 | 6.25 | 0.0833 | 3:1; one locus |

*R indicates Resistance, S indicates susceptibility

**Table S4**. Segregation analysis of Hygromycin resistance for T1 CuBe-eGFP lines.

| T1 line | Observed R | Observed S | Expected R (3:1) | Expected S (1:3) | P-value | Analysis of T-DNA insertion |
|---|---|---|---|---|---|---|
| **3** | 27 | 2 | 21.75 | 7.25 | 0.0244 | 15:1; two loci |
| **4** | 23 | 6 | 21.75 | 7.25 | 0.5919 | 3:1; one locus |
| **6** | 34 | 2 | 27 | 9 | 0.0071 | 15:1; two loci |
| **7** | 26 | 13 | 29.25 | 9.75 | 0.2294 | 3:1; one locus |

*R indicates Resistance, S indicates susceptibility

**Table S5**. Segregation analysis of Kanamycin (Kan) and Hygromycin (Hyg) resistance for T1 CuBe-eGFP lines.

| T2 line | Observed R-S Kan | Observed R-S Hyg |
|---|---|---|
| **4.11** | 38-0 | 33-0 |
| **7.8** | 25-0 | 26-0 |

*R indicates resistance, S indicates susceptibility

**Table S6.** Segregation analysis of Kanamycin resistance for T1 CuBe-n72_hγHC lines.

| T1 line | Observed R | Observed S | Expected R (3:1) | Expected S (1:3) | P-value | Analysis of T-DNA insertion |
|---|---|---|---|---|---|---|
| **6** | 24 | 10 | 25.5 | 8.5 | 0.5525 | 3:1; one locus |
| **8** | 25 | 12 | 27.75 | 9.25 | 0.2965 | 3:1; one locus |
| **14** | 38 | 6 | 33 | 11 | 0.0817 | 3:1; one locus |
| **15** | 29 | 12 | 30.75 | 10.25 | 0.5279 | 3:1; one locus |
| **16** | 33 | 7 | 30 | 10 | 0.2733 | 3:1; one locus |

*R indicates resistance, S indicates susceptibility

**Table S7**. Segregation analysis of Hygromycin resistance for T1 CuBe-n72_hγHC lines.

| T1 line | Observed R | Observed S | Expected R (3:1) | Expected S (1:3) | P-value | Analysis of T-DNA insertion |
|---|---|---|---|---|---|---|
| **6** | 36 | 4 | 30 | 10 | 0.0244 | ~ 15:1, two loci |
| **8** | 36 | 3 | 29.95 | 9.75 | 0.0285 | ~ 15:1, two loci |
| **14** | 34 | 9 | 32.25 | 10.75 | 0.5377 | 3:1; one locus |
| **15** | 32 | 14 | 34.5 | 11.5 | 0.3946 | 3:1; one locus |
| **16** | 33 | 9 | 31.5 | 10.5 | 0.5930 | 3:1; one locus |

*R indicates resistance, S indicates susceptibility

**Table S8**. GB level 1 and >1 transcriptional units and modules. Sequences can be found at https://goldenbraidpro.com/ using the GB number or GB name.

| GB number | Construct | GB name | Description |
|---|---|---|---|
| GB0107 | Empty vector | pEGB SF | Used as a stuffer construct to get the same infiltration DO for every *Agrobacterium* mix. |
| GB1541 | pNOS:CUP2-Gal4 | pEGB_3alpha2 PNos:Cup2:Gal4AD:Tnos | Transcriptional unit for constitutive expression of the transcriptional activator protein CUP2 in C-terminal fusion with GAL4 activation domain. Binds to specific DNA operator (CBS) in the presence of copper ions and promotes transcription. |
| GB4279 | 35S:eGFP | 3a1_35S:eGFP:T35S | Transcriptional unit for constitutive expression of the enhanced GFP. |
| GB3598 | pNOS:Rep/RepA | pDGB3_α1_pNOS:Rep/ RepA:tNOS | Transcriptional unit for constitutive expression of BeYDV Rep/RepA (C1:C2). |
| GB4312 | Standard replicon | 3a1_LIR1:35S:eGFP:T35S-SIR:LIR2 | BeYDV LIR1 + TU 35S:eGFP:T35S + BeYDV SIR-LIR2 in alpha1. |
| GB5177 | INPACT replicon | o1_LIR:2nd intron:3'eGFP:T35S:SIR:p35S:5'Luc:1st intron:LIR | LIR:2nd intron + 3'eGFP +T35S.SIR + p35S + 5'Luc + 1st intron:LIR for INPACT configuration. |
| GB4480 | Geminino 1.0 | a2_NEW BeYDV geminino 3.0_eGFP | BeYDV LIR1:2nd intron + enhanced GFP CDS + Ter35S + p35S + 1st intron:LIR1 in alpha2. Deconstructed replicon that needs BeYDV Rep for circularization and, thus, replication. The intronic parts have AGGT for processing. |
| GB5178 | Geminino 2.0 | a2_geminino 4.0 (no ATG)-eGFP | Geminino 2.0-no ATG on CDS. BeYDV LIR1 + 2nd half intron ICON MP + eGFP w/o ATG + t35S:SIR + P35s + 1st half of intron from ICON MP + BeYDV LIR1. |
| GB0108 | 35S:P19 | pEGB 35s:P19:tNOS | Transcriptional unit for constitutive expression of the silencing suppressor P19. |
| GB3625 | CBS:minDFR:Rep/RepA | pDGB3_α2_4xCBS:mDFR: Rep/RepA:tNOS | Transcriptional unit for copper-regulation of BeYDV Rep/RepA with 4xCBS domain+minimal promoter DFR for induction by copper. |
| GB4646 | CBS:FRT:OCSt:FRT:min DFR:Rep/RepA | 3a1_CBS-FRT-OCS-BeYDVRepRepA-tNos | Transcriptional unit for copper- and flipase-regulated BeYDV Rep/RepA. OCSt, flanked by FRT sites, is embedded in the 5'UTR of miniDFR. |

| | | | |
|---|---|---|---|
| GB4652 | CBS:minDFR:Flp | 3a2_CBS:miniDFR:FLP:tNos | Transcriptional unit for copper-regulated flipase (Flp) recombinase. |
| GB4650 | CBS:minDFR:PhiC31 | 3a2_CBS:miniDFR:PhiC31:tNos | Transcriptional unit for copper-regulated PhiC31 recombinase. |
| GB4643 | CBS:attB:OCSt:attP:minDFR:Rep/RepA | 3a1_CBS-att-OCS-BeYDVRep/RepA-tNos | Transcriptional unit for the copper and PhiC31 regulated expression of BeYDV Rep/RepA. |
| GB4684 | Processor, amplification, pro-replicon module for eGFP | pLXB3o1 HygR-MAR10-gemiGFP-MAR10 | Module for stable transformation of Geminino 1.0-eGFP flanked by two MAR10 insulators and next to a Hygromycin resistance cassette. |
| GB4697 | Sensor module | 3a1 nptII-DISTC02-CBS-FRT:Rep/RepA:tNos-DISTC19-CBS:miniDFR:FLP:tNos-PROXC02-pNos:CUP2:Gal4AD:tNos-SF | Module for copper and flippase-regulated expression of BeYDV Rep/RepA, including the TU for copper-regulated FLP, the TU for constitutive expression of CUP2-GAL4, and the nptII cassette. |
| GB4954 | Processor, amplification, pro-replicon module for n72_hγHC | pLXB3o1_hygR:MAR10:New BeYDV geminino:SP-nano72VHH-IgG1:MAR10 | Module for stable transformation of Geminino 1.0-n72_hγHC flanked by two MAR10 insulators and next to a Hygromycin resistance cassette. |

# Chapter 3

## Bioluminescence-driven optimization of geminivirus-based vectors as tools in Plant Biotechnology

---

My contribution to this chapter was essential. I contributed to the design, performing of the experiments and writing of this manuscript.

## ABSTRACT

Viral replicons are valuable tools in plant biotechnology, widely utilized to increase recombinant protein production. Their ability to amplify gene dosage in a trigger-dependent manner also opens doors to regulatory applications. This work focuses on optimizing geminivirus-based vectors for Synthetic Biology applications in plants, using bioluminescence as a sensitive, real-time reporter to characterize gene expression. Specifically, geminivirus-based synthetic replicons derived from bean yellow dwarf virus (BeYDV), tomato yellow leaf curl virus (TYLCV), and beet curly top virus (BCTV) were engineered and assessed for basal expression, inducibility, and recombinant protein co-expression potential. Our study provided insights into the strengths and limitations of each geminiviral replicon. BeYDV replicon displayed a robust activation profile suitable for complex tasks such as multi-gene expression, while TYLCV showed high expression levels despite moderate basal leakage. In contrast, BCTV demonstrated less favourable control and expression levels. Through bioluminescence-based screening, the TYLCV system was further optimized to improve regulatory precision. These findings highlight the versatility of geminivirus replicons, paving the way for future engineering of synthetic gene circuits in plants.



**Keywords**

Plant Synthetic Biology, *Nicotiana benthamiana*, Geminivirus, BeYDV, TYLCV, BCTV, bioluminescence

## INTRODUCTION

Virus-derived replicons have emerged as powerful tools in plant biotechnology due to their unique ability to hijack the cellular machinery to multiply designated pieces of genetic information (Peyret & Lomonossoff, 2015). Among the different viral replicons employed in plant biotechnology, those derived from geminiviruses have gained particular attention due to its rapid and high-yield replication and wide flexibility for cargo insertion (Lozano-Durán, 2016). The genetic amplification provided by geminiviruses replicons can serve different cellular bioengineering purposes such as improving yields of biomanufactured products and increasing template concentration for gene editing.

Geminiviruses, members of the *Geminiviridae* family, are plant-infecting viruses with small (normally ~3.0 kb) circular single-stranded DNA (ssDNA) genomes (Stanley, 1985). These viruses infect a wide range of plants and are notorious for causing significant agricultural damage worldwide (Fiallo-Olivé et al., 2021; Navas-Castillo et al., 2011). Particularly, the three most prevalent and well-characterized geminivirus genera are Mastrevirus, Curtovirus, and Begomovirus (Wu et al., 2022). The key feature of geminiviruses to multiply in the host plant is their rolling-circle replication mechanism. In this process, the viral ssDNA is first converted into a double-stranded DNA (dsDNA) intermediate by host DNA polymerase α. In mastreviruses, this synthesis begins in a region known as the Short Intergenic Region (SIR), while in Begomovirus and Curtovirus, it starts near the Intergenic Region (IR), which is referred to as the Long Intergenic Region (LIR) in mastreviruses. The dsDNA intermediate then undergoes replication through rolling-circle replication (RCR) and recombination-dependent replication (RDR) assisted by DNA polymerase δ. The RCR is initiated by the virus-encoded Replication-associated protein (Rep). Rep binds to iteron sequences, which are specific cis-acting motifs located within the IR/LIR. Additionally, the IR/LIR harbours a GC-rich repeat inverted sequence forming a stem-loop structure that contains the highly conserved nonanucleotide 5′-TAATATTAC-3′. This nonanucleotide serves as the cleavage site for Rep, whose action provides a 3'-OH group in the dsDNA genome that can be extended by polymerase δ (Wu et al., 2022). The variability of iteron sequences between different geminivirus species plays a crucial role in determining virus-specific replication, while the Rep protein contains an iteron-related domain (IRD) that also confers specificity through DNA binding (Argüello-Astorga & Ruiz-Medrano, 2001; Morilla et al., 2006).

Geminivirus-based synthetic replicons rely on two separate genetic elements, namely the Rep protein, and the replicative vector, where the gene of interest is flanked by intergenic regions (Zhang & Mason, 2006). Mechanistically, Rep protein leads to the circularization and rolling-circle replication of the vector by interacting with the flanking IR regions (Chen et al., 2011). Usually, both the Rep and the replicative vector into the plant cell are delivered transiently using Agrobacterium-mediated transient transformation (L. Ma et al., 2012). The bean yellow dwarf virus (BeYDV) (Mor et al., 2003) has served as basis for most synthetic replicons developed to date using the agroinfiltration route. This approach has successfully produced large quantities of proteins of interest like vaccine candidates, antibodies, virus-

like particles, and enzymes that mediate the heterologous enhanced expression of metabolites like triterpenoids (Regnard et al., 2010; Chen et al., 2011; Huang et al., 2010; Diamos & Mason, 2018; Romsuk et al., 2022). Moreover, the BeYDV replicative system has been also used for enhancing genome editing efficiency by the replication of the nuclease-edited DNA enzyme and the repair template in tobacco (Baltes et al., 2014), potato (Butler et al., 2016), and tomato (Dahan-Meir et al., 2018). The potential of BeYDV for enhancing protein expression is evident by the fact that very recently this system has been used in very divergent species like lettuce for the expression of virus-like proteins and antibodies (Lai et al., 2012), pepper for the expression of capsanthin (Lin et al., 2023) and microalgae to produce SARS-CoV-2 receptor antigens and basic fibroblast growth factor (Malla et al., 2021). Other geminiviruses used as basis for the development of synthetic replicons include the beet curly top virus (BCTV), a Curtovirus with exceptionally broad host range for replication (Bennett, 1971; Stenger et al., 1992). BCTV-derived replicative vectors have been used for expressing vaccine candidates in *N. benthamiana* (Chung et al., 2011), for plastid genome engineering in tobacco (Jakubiec et al., 2021) and for enhancing genome editing in *N. benthamiana* by replication of CRISPR/Cas12a components (Eini et al., 2022). Also the geminivirus wheat dwarf virus (WDV) has been used to create expression vectors that boost genome engineering in wheat (Gil-Humanes et al., 2017) and rice (Wang et al., 2017). Likewise, a sweet potato leaf curl virus (SPLCV) synthetic replicon was developed for genome editing in *N. benthamiana* (Yu et al., 2020). Very recently, Kim et al. (2024) expanded the geminiviral-based vector repertoire by generating novel vectors based on tomato yellow leaf curl virus (TYLCV), honeysuckle yellow vein virus (HYVV), and mild curly top virus (BMCTV) for the expression of turbo green fluorescence protein (tGFP) in *N. benthamiana*.

Despite recent advances, the potential of geminiviral-based systems as tools for plant biotechnology remains largely unexploited. One of the most promising future directions is the development of genome-integrated inducible replicons. The rationale behind integrated replicons is that they remain silent until a user-defined signal activates replication via Rep expression. In that way, integrated replicons would keep all the advantages of the above-described Agro-delivered synthetic replicons, while circumventing the need for cumbersome transient expression and maintaining a high biosafety profile. Ideally, the availability of a battery of integration-competent geminiviral replicons derived from different phylogenetic origins would open the way not only to extend replicon technology to more crop species, but also to combine two or more DNA amplification systems within the same plant. However, to date, maintaining an effective control on replicon activation during the silent state remains a main challenge for the development of new integrated systems. Wild type viral sequences are rich in cryptic regulatory sequences (Gong et al., 2021) that could lead to a spurious activation of any derived synthetic replicon and whose presence can only be determined experimentally.

Currently, only two inducible geminiviral-based systems have been stably integrated in the *N. benthamiana* genome, the INPACT system based on tomato yellow dwarf virus (TYDV) (Dugdale et al., 2013), and the CuBe system previously developed in our laboratory and

based on BeYDV (Garcia-Perez et al., 2024). Both the INPACT and CuBe systems enhance regulation by deconstructing the transcriptional unit of the gene of interest inserted in the geminiviral vector, preventing basal expression. In these systems, the promoter is only positioned upstream of the coding sequence once circularization occurs, a process mediated by the Rep protein. In the INPACT system, induction is triggered by ethanol via an ethanol-responsive promoter (Caddick et al., 1998) controlling Rep expression, while in the CuBe system, copper sulfate (Garcia-Perez et al., 2022) serves as the inducer for the same regulatory mechanism.

To broaden the range of geminiviral vectors for plant biotechnology uses, it is crucial to advance in viral deconstruction and to speed up the design-build-test-learn cycle (Kitano et al., 2023), focusing on reducing basal expression. Viral deconstruction involves de definition of functional modules that can be easily exchange and combined. Adaptation to modular cloning strategies, such as those developed by Dusek et al. (2020) for BeYDV and Neubauer et al. (2024) for BCTV, have provided more flexibility in assembling new geminiviral vectors. The adoption of a modular configuration facilitates not only synonymous exchange of DNA parts, but also novel rearrangement of those parts, enabling the exploration of optimized configurations that can be eventually adapted to a wider range of geminiviruses. Furthermore, achieving optimal configurations demands highly sensitive reporter systems capable of detecting minimal 'leaky' expression. Traditionally, geminiviral systems have relied on GFP or GUS reporters for characterization. Recently, a fungal bioluminescence pathway (FBP), composed of four genes, has emerged as a powerful alternative plant reporter system (Khakhar et al., 2020; Mitiouchkina et al., 2020) (Figure 1A). This pathway, which glows autonomously and therefore does not require the exogenous application of expensive luciferin substrate, is a cyclic metabolic process using caffeic acid, a common plant metabolite, as the starting substrate. The initial reaction, converting caffeic acid to hispidin, is catalysed by a large (~5.1 kilobases) polyketide synthase (HispS). HispS requires a posttranslational modification to function, typically achieved by co-expressing a phosphopantetheinyl transferase like NPGA from *Aspergillus nidulans* in plant (Khakhar et al., 2020; Mitiouchkina et al., 2020). HispS has been identified as the transcriptional limiting step of the pathway (Moreno-Giménez et al., 2022). As an alternative to HispS, smaller and non-posttranslational dependent hispidin synthases from plant origin (HmS, ASCL, PKS2) can substitute for HispS, resulting in a plant-fungal Hybrid Bioluminescence Pathway (HBP) (Palkina et al., 2024). As a practical application of the FBP in Plant Synthetic Biology, Calvache et al. (2024) developed a dual reporter system combining the luminescence of FBP with enhanced GFP (eGFP) expression to normalize for variations in transfection efficiency in transiently transformed *N. benthamiana* leaves. This system offers highly sensitive autoluminescence detection and allows normalization of expression levels against fluorescent proteins, improving data accuracy. Unlike fluorescence-based systems, bioluminescence can be tracked with consumer grade cameras, providing a faster, more straightforward method to monitor vector performance in plant tissues (Calvache et al., 2024).

In this work, we have employed a deconstruction approach, previously described for developing a modular BeYDV-based vector, to create new modular replicons based on TYLCV and BCTV. To facilitate this process, and the subsequent optimization of new replicons, we have incorporated a new reporter system based on the HBP, showing the versatility of the new reporter to speed up replicon engineering in plants. As a result of several optimization steps, we obtained an improved version of a TYLCV-derived replicon showing a low basal expression and a highly enhanced dynamic range. By leveraging these viral systems, we aim to better regulate geminiviral vector replication to engineer gene circuits based on integrated replicons, and to understand the orthogonality of different geminiviral replicons when used in combination.

## RESULTS

### Application of bioluminescence pathway as a sensitive qualitative and quantitative reporter for geminiviral systems characterization

To evaluate the potential of the autoluminescence pathway (Figure 1A) as a sensitive reporter system for characterizing geminiviral vectors, we cloned the hispidin synthase gene into the Geminino 1.0 vector based on BeYDV, as previously described in Garcia-Perez et al. (2024) (Figure 1B). The arrangement of the gene of interest in Geminino 1.0 was designed to harbour the promoter downstream of the coding sequence (CDS) in a linear configuration. The gene of interest is flanked by the Long Intergenic Region (LIR) of BeYDV. These LIRs are recognized and nicked by the Replication-associated protein (Rep), which leads to DNA circularization. In this circular configuration transcription can also proceed since the promoter relocates upstream of the CDS. We tested different hispidin synthases—HispS, HmS, PKS2, and ASCL as candidates to incorporate into our reporter system. For luminescence assays, we used three T-DNA constructs: (i) geminiviral vectors containing HispS/HmS/PKS2/ASCL, (ii) constitutive 35S CaMV-driven genes for the rest of the pathway (Luz, H3H, CPH), eGFP normalizer, and p19 silencing suppressor, and (iii) constitutive pNOS-driven bicistronic BeYDV Replication-associated proteins (Rep/RepA) gene (Figure 1B). In the HispS treatment, a 4$^{th}$ construct carrying p35S:NPGA was co-expressed. These constructs were transiently co-expressed in *N. benthamiana* leaves through agroinfiltration. Luminescence was monitored daily from one to six days post-infiltration (dpi) to analyse the effect of combining Geminino 1.0 with Rep/RepA. The normalized luminescence, expressed as luminescence/fluorescence ratios, showed that three of the tested hispidin synthases significantly increased luminescence in the presence of Rep/RepA, indicating the correlation between geminiviral replication and luminescence emission (Figure 1C). The combination of the fungal HispS and NPGA produced the highest luminescence, peaking on day 3. The plant-based synthases, HmS and ASCL, showed moderate increases, with maximum luminescence at day 2, while PKS2 had lower activity. When whole agroinfiltrated leaves were imaged at 2 days post-infiltration (dpi) (Figure 1E), HispS produced the highest luminescence, but it also exhibited strong baseline activity without Rep. Other hispidin synthases showed minimal luminescence in the absence of Rep, with HmS displaying more stable expression compared to ASCL, with less variation between leaves and greater

consistency from day 2 to day 3. Steered by simplicity of the system, we selected HmS as the preferred reporter gene for subsequent analysis of geminivirus-driven transcriptional activity: HmS is more compact than HispS, provides sufficient dynamic range, does not depend on NPGA, and probably, as a consequence, it reaches maximum expression levels at a shorter timepoint.

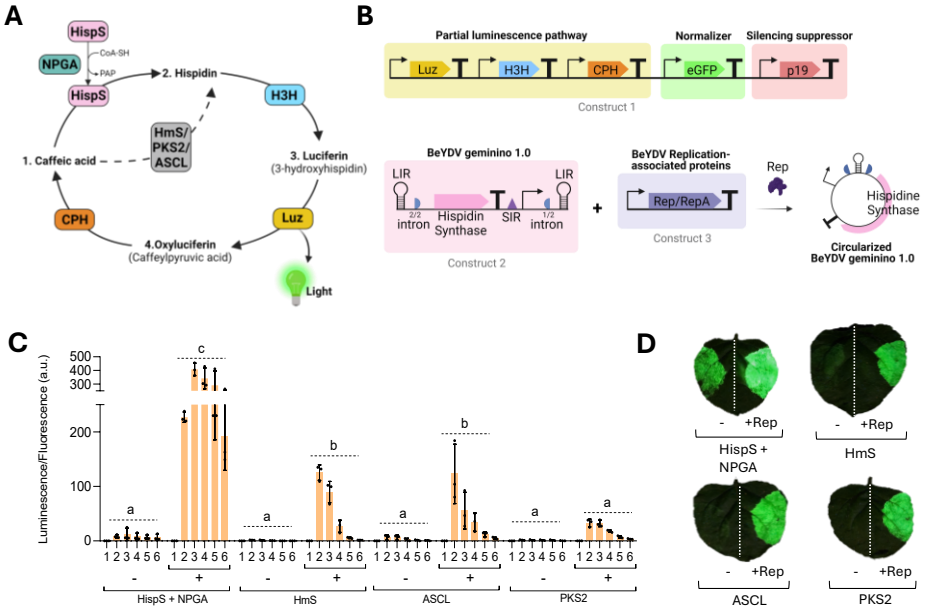



**Figure 1. Evaluation of bioluminescence in *Nicotiana benthamiana* leaves expressing different hispidin synthases through a BeYDV-based system**. **A)** Schematic representation of the bioluminescence pathway, illustrating the metabolic conversion of caffeic acid to hispidin and subsequent reactions leading to light emission. The pathway involves several enzymes, including the tested hispidin synthases (HispS, HmS, ASCL, and PKS2). **B)** Schematic representation of the genetic constructs used in the study. Construct 1 contains constitutively expressed p35S-driven genes of the luminescence pathway (Luz, H3H, CPH), along with enhanced GFP (eGFP) and the p19 silencing suppressor. Construct 2 includes the geminiviral vector carrying different hispidin synthases (HispS/HmS/ASCL/PKS2). Construct 3 is a constitutive pNOS-driven polycistronic expression cassette for BeYDV Replication-associated proteins (Rep/RepA). These constructs were co-infiltrated into *N. benthamiana* leaves to evaluate luminescence with and without Rep/RepA. In the HispS treatment, a 4$^{th}$ construct carrying p35S:NPGA was co-expressed. LIRs are the long intergenic regions (LIRs) from BeYDV, SIR is the short intergenic region, half circles are the initial (1/2 intron) and the final (2/2 intron) segments of a unique intron, T is the transcriptional terminator, and the arrow indicates the transcriptional promoter. **C)** Luminescence normalized by fluorescence (luminescence/fluorescence arbitrary units) measured over six days post-infiltration (dpi) for the different hispidin synthases—HispS + NPGA, HmS, ASCL, and PKS2— cloned in a BeYDV-based vector (Geminino 1.0) and co-expressed with or without Rep/RepA and with the rest of the genes of the bioluminescence pathway. **D)** Qualitative imaging of luminescence for the same treatments measured in C) in whole agroinfiltrated leaves at 2 dpi. Each point represents the luminescence/fluorescence values of an individual disc taken from a different leaf of the same plant (n=3). Error bars represent standard deviation. Statistical analysis was

performed using a one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05). For the ANOVA test, the areas under the curve of the time-course treatments were compared. Variables with the same letters belong to the same statistical group. The figure includes images from Biorender (biorender.com).

**Adaptation of TYLCV and BCTV replicative elements for Rep-regulated recombinant protein expression**

To expand the toolbox of geminivirus-based replicative systems, we evaluated the potential of TYLCV and BCTV-based vectors as Rep-dependent systems for recombinant protein expression. Following a modular approach parallel to that used for BeYDV (Figure 1B), we adapted the replication machinery from TYLCV and BCTV genomes (Figure 2A) by using their intergenic regions (IRs) (Figure 2B) to develop their respective Geminino 1.0 type of vectors, along with their Replication-associated proteins (Rep). To test the functionality of these new modular systems, HmS was cloned into the Geminino 1.0 vectors, with luminescence used as a reporter. Constructs were transiently expressed in *N. benthamiana* via agroinfiltration, co-infiltrated with or without the respective Rep for each geminivirus, as shown in Figure 1B. Normalized luminescence levels were measured over time to evaluate the performance of each system. As shown in Figure 2C, co-infiltration with Rep significantly increased luminescence in all three systems, with the TYLCV replicon producing the highest values. The BeYDV and BCTV systems reached their luminescence peak at 3 dpi, while TYLCV reached its maximum at 2 dpi. Notably, average basal luminescence in the absence of Rep was slightly higher for TYLCV and BCTV compared to BeYDV suggesting potential promoter activity conferred by regions within the IRs of TYLCV and BCTV (Figure 2B).

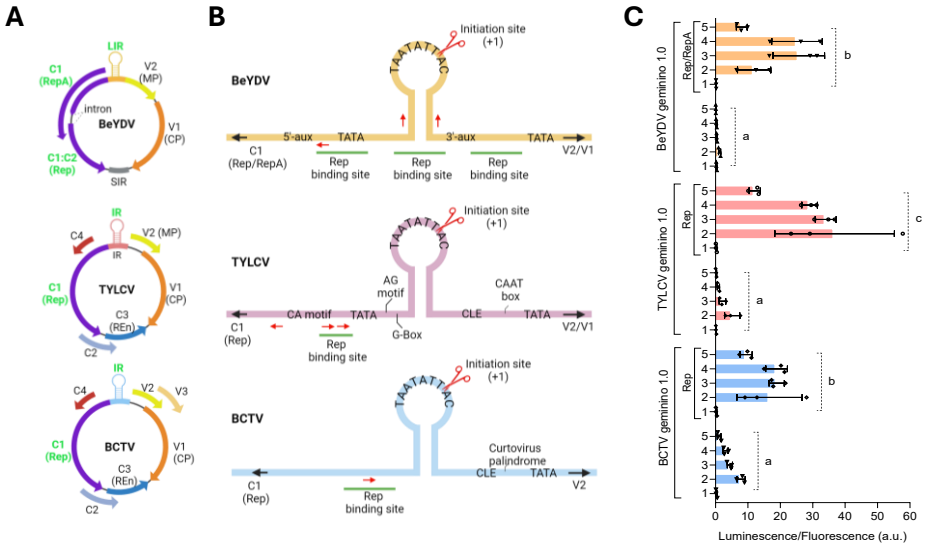



**Figure 2. Geminiviral replication elements adapted from BeYDV, TYLCV, and BCTV genomes to create geminivirus-based systems for recombinant protein expression**. **A)** Schematic representation of BeYDV, TYLCV and BCTV genomes. Names of the elements involved in replication are displayed in green: Rep/RepA (C1/C2) for BeYDV, and Rep for TYLCV and BCTV (C1) and the (long)

intergenic region (LIR/IR), which contains the origin of replication. **B)** Detailed structure of the intergenic region for each virus. The replication origin (initiation site, +1) where Rep nicks the DNA is marked by the red scissors. The Rep-binding sites are underlined, and iteron sequences are marked with red arrows. TATA box is required for transcription initiation, Auxiliary sequences (aux) and CA motifs enhance DNA replication, G-Box and AG motifs play a role in origin recognition, and conserved late element (CLE) enhance transcription. **C)** Luminescence normalized by fluorescence values (luminescence/fluorescence arbitrary units) measured over five days post-infiltration (dpi) for the different geminiviral vectors carrying HmS co-expressed with or without Rep/RepA and with the rest of the genes of the bioluminescence pathway (H3H, Luz, CPH), the normalizer eGFP, and the silencing suppressor p19 (as depicted in Figure 1B). Each point represents the luminescence/fluorescence values of an individual disc taken from a different leaf of the same plant (n=3). Error bars represent standard deviation. Statistical analysis was performed using a one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05). Variables with the same letters belong to the same statistical group. The figure includes images from Biorender (biorender.com).

**Adaptation of the new geminiviral replicative systems for Rep-regulated multiprotein expression**

One of the unique advantages of geminivirus-based systems over other RNA viral vectors for recombinant protein production is their ability to bypass replicative exclusion. This feature allows different geminivirus replicons to coexist within the same cell, making the system compatible with producing complex protein assemblies, where different subunits can be expressed from separate replicons. Taking advantage of the modularity of the system, we designed the so-called 'COMBO' configuration (Figure 3A). This configuration allows for the concatenation of multiple Geminino 1.0 replicons within a single T-DNA construct, enabling coordinated co-expression of several proteins. A newly optimized assembly syntax compatible with the GoldenBraid modular cloning system (Sarrion-Perdigones, Vazquez-Vilar, Palaci, et al., 2013) was implemented to allow a flexible assembly for the concatenation of unlimited number of replicons. We then generated new COMBO constructs for all three geminiviral systems under study (BeYDV, TYLCV and BCTV) and employed to co-express three fluorescent reporter proteins (namely GFP, BFP, and dsRed), in *N. benthamiana* by agroinfiltration (Figure 3A).

Results showed tight regulation of expression by Rep/RepA in the case of the BeYDV system, showing negligible fluorescence values in the absence of Rep/RepA for each of the three reporters, and strong fluorescence signal in the presence of Rep/RepA (Figure 3B). These results suggested a successful recircularization of the concatenated Gemininos 1.0 vectors and robust co-expression of the proteins. However, with the TYLCV and BCTV systems, basal fluorescence levels were notably higher compared to BeYDV, corroborating the previously observed basal activation of these new vectors. While the TYLCV COMBO still showed statistically significant differences with and without Rep, the BCTV COMBO results suggest that Rep might not be effective in regulating recombinant protein expression. A combination of leaky expression and suboptimal BCTV Rep processing, consistent with previous luminescence-based observations, likely prevented the efficient recircularization of the Gemininos 1.0 vectors.

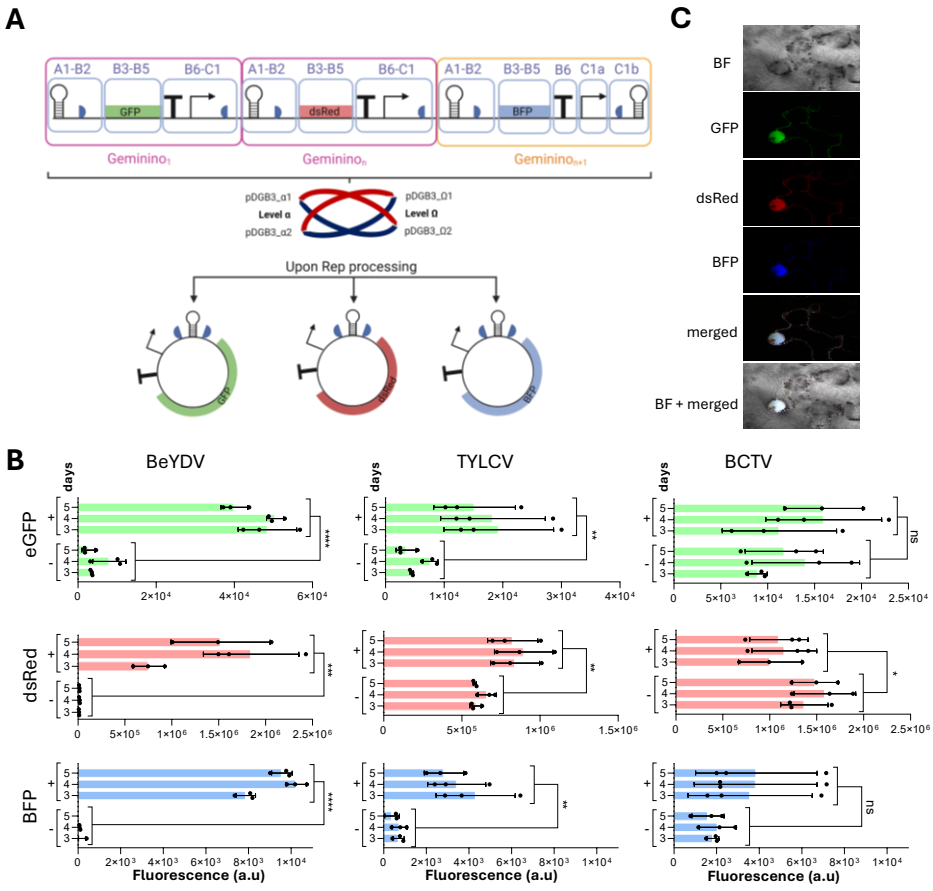



**Figure 3. Design and performance of the COMBO configuration for multi-protein expression in *N. benthamiana*. A)** Schematic representation of the COMBO configuration, enabling concatenation of multiple Geminivirus 1.0 vectors within a single T-DNA construct. The assembly process utilizes an optimized cloning grammar compatible with the GoldenBraid system (represented by the GoldenBraid icon and its cloning vectors of level alpha and omega). Each DNA part has specific overhangs that give the identity to the part (A1-B2, B3-B5, etc). *N* number of Geminivirus can be combined in this configuration. LIRs are the long intergenic regions (LIRs) from BeYDV, SIR is the short intergenic region, half circles are the initial (1/2 intron) and the final (2/2 intron) segments of a unique intron, T is the transcriptional terminator and the arrow indicates the transcriptional promoter. **B)** Fluorescence expression levels of enhanced GFP (eGFP), BFP, and dsRed in *N. benthamiana* leaves following agroinfiltration of COMBO configurations for BeYDV, TYLCV and BCTV systems. The '+' symbol indicates co-expression with the Replication-associated proteins (Rep, or Rep/RepA in the case of BeYDV) of each system, while the '-' symbol indicates the absence of Rep(/RepA). **C)** Confocal microscopy images demonstrating co-localization of eGFP, BFP, and dsRed expressed via BeYDV COMBO system in the same leaf cell. BF stands for bright field. The excitation wave lengths for each fluorescence channel were: 488 for eGFP, 560 for dsRed, and 405 for BFP. Emission wavelength fluorescence range collected for eGFP, dsRed and BFP was 488-507, 558-583, 383-448, respectively. Error bars represent standard deviation. Statistical analysis was performed using a one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05). For the ANOVA test, the areas under the curve of the time-course treatments were

compared. Variables with the same letters belong to the same statistical group. The figure includes images from Biorender (biorender.com).

For the BeYDV COMBO system, confocal microscopy analysis confirmed that eGFP, BFP, and dsRed proteins were co-expressed simultaneously in individual cells (Figure 3C). This highlights the capacity of geminivirus-based vectors to replicate and express multiple proteins in the same cell without interference. We further explored whether the position of the transcriptional units within the COMBO structure influenced protein expression levels. To do this, we tested various combinations where each fluorescent protein occupied different positions within the concatenated Gemininos 1.0. The results revealed certain positional effects: dsRed showed reduced expression when placed in the first position, BFP had lower expression when in the last position, and eGFP was less expressed when positioned in either the first or second position (Table 1). These findings suggest that the position of the transcriptional unit can influence expression levels, but the effect appears to vary depending on the specific protein. Thus, no general conclusions could be drawn from this positional analysis.

**Table 1**. Influence of position in the COMBO system for protein expression

| DsRed | | | | eGFP | | | | BFP | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| COMBO | Mean | SD | Group | COMBO | Mean | SD | Group | COMBO | Mean | SD | Group |
| G-**R**-B | 91.37 | 2.89 | b | **G**-R-B | 7.42 | 2.46 | a | G-R-**B** | 1.21 | 0.44 | a |
| **R**-G-B | 57.75 | 14.96 | a | R-**G**-B | 35.33 | 12.20 | b | R-G-**B** | 6.92 | 2.76 | b |
| **R**-B-G | 61.38 | 6.39 | a | R-B-**G** | 32.22 | 5.32 | b | R-**B**-G | 6.40 | 1.09 | b |
| B-G-**R** | 84.52 | 1.81 | b | B-**G**-R | 12.98 | 1.48 | a | **B**-G-R | 2.49 | 0.33 | a |

*COMBO column indicates the order of the transcriptional units for the fluorescent proteins inside the COMBO system.
*G=enhanced GFP, R=dsRed, B=BFP.
*Mean values are the mean fluorescence ratios of for each fluorescence protein in each configuration. Ratios were calculated as the fluorescence value of the protein of interest in the treatment divided by the total fluorescence (eGFP+dsRed+BFP) in each COMBO. SD is standard deviation, and Group indicates the statistical group. Statistical analysis was performed using a one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05) for the analysis of three replicates per sample. Variables with the same letters belong to the same statistical group.

**Exploration of orthogonality between geminiviral replicative systems**

The development of new geminiviral replicons opens the interesting possibility of combining two or more systems in the same plant. The evaluation of the potential applications of such combinations requires previous analysis of their mutual interactions. To this end, we tested the transcriptional expression levels conferred by all possible combinations of Rep proteins with the three Geminino 1.0 vectors (Figure 4A). Bioluminescence from HBP was used as a reporter, with expression levels measured 2-4 dpi. Results showed that the BeYDV system displayed strict specificity, with luminescence observed only when paired with its own Rep/RepA, and no cross-activation was observed when combined with the Reps from TYLCV or BCTV (Figure 4B). In contrast, the TYLCV and BCTV systems demonstrated some level of cross-reactivity. Although low, luminescence was detected when TYLCV and BCTV vectors were combined with Rep's from other systems, particularly in the case of BCTV. However, the baseline luminescence observed in the absence of Rep, as noted previously, likely confounds the interpretation, as leakiness allegedly due to IR promoter activity could account for the apparent lack of specificity. To

further explore this issue, luminescence images were taken from leaves infiltrated with combinations of each geminino vector and Rep, as well as controls without Rep (Figure 4C). These images confirmed the presence of significant leakiness in the TYLCV and BCTV gemininos, which appeared to be the primary cause of the observed non-specific luminescence.

To further explore the specificity of Reps and the possibility of using their combined action as a multi-input regulation of luminescence expression, we design 'hybrid' Gemininos 1.0 vectors, where two different IRs are combined in the same unit. We placed the LIR from BeYDV upstream of the geminino vector, harnessing the fact that this LIR exhibits minimal basal expression, and either the LIR of TYLCV or BCTV on the opposite side. Figure 4D illustrates this design, where the BeYDV LIR is shown in orange and the TYLCV or BCTV in red or blue, respectively. For the circularization of these hybrid vectors to happen, we were expecting that both Reps from BeYDV and TYLCV or BCTV should be present. To test these systems, we agroinfiltrated them in *N. benthamiana*, combining each hybrid Geminino 1.0 vector (BeLIR-TYIR or BeLIR-BCIR) with either one or both respective Reps. As a negative control, we used a geminino vector containing only a single BeYDV LIR at the 5' end (BeY-), and as a positive control, we used a geminino vector with two BeYDV LIRs (BeYLIR-BeYLIR). The results showed an increase in luminescence when BeLIR-TYIR or BeLIR-BCIR vectors were combined with BeYDV Rep/RepA, and TYLCV Rep or BCTV Rep, respectively. Although these increases were not statistically significant due to variability in replicate measurements (Figure 4E), the activation of luminescence was clearly visible in whole leaf imaging (Figure 4F). When only TYLCV or BCTV Rep was co-expressed with BeLIR-TYIR or BeLIR-BCIR, luminescence was lower compared to co-expression with Rep/RepA. Surprisingly, co-expression of BeYDV Rep/RepA alone with either vector produced luminescence levels as high as when both Reps were co-expressed. This suggests that BeYDV Rep might have low specificity for nicking or that Geminino 1.0 circularization can occur solely through recognition of the 5' LIR. Since the vector containing only a single LIR from BeYDV also activated luminescence, we favour the hypothesis that a single LIR is sufficient for circularization of the geminivirus vector.

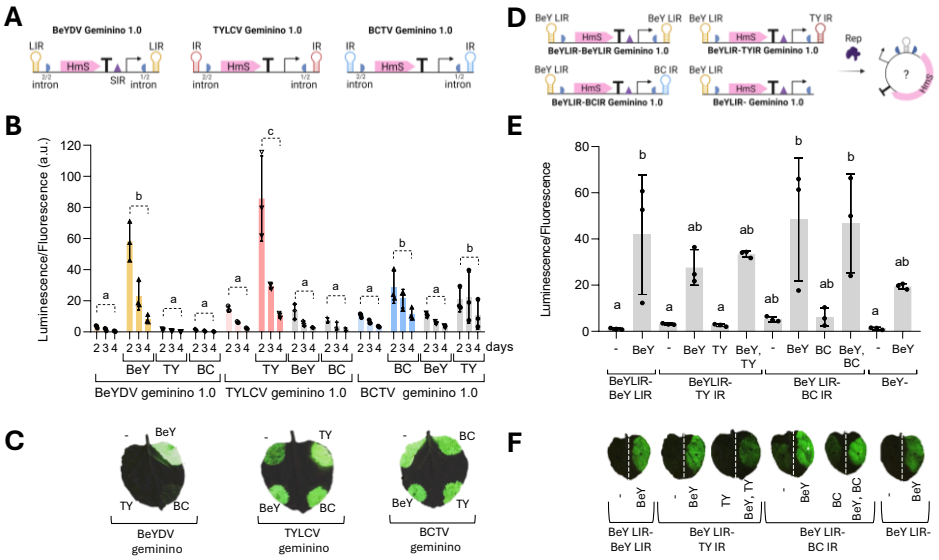



**Figure 4. Cross-compatibility of replicative systems and hybrid Geminino 1.0 vectors. A)** Schematic representation of the Geminino 1.0 vectors carrying the hispidin synthase HmS for each geminiviral system (BeYDV, TYLCV, and BCTV). LIR/IR are the (long) intergenic regions (IR/LIRs), SIR is the short intergenic region, half circles are the initial (1/2 intron) and the final (2/2 intron) segments of a unique intron, T is the transcriptional terminator, and the arrow indicates the transcriptional promoter. **B)** Luminescence normalized by fluorescence values (luminescence/fluorescence arbitrary units) resulting from the co-expression by agroinfiltration in *N. benthamiana* of BeYDV, TYLCV, or BCTV Geminino 1.0 vectors carrying the hispidin synthase HmS with the Replication-associated proteins (Rep, or Rep/RepA in the case of BeYDV) of BeYDV (BeY), TYLCV (TY), or BCTV (BC), the rest of the genes of the bioluminescence pathway (H3H, Luz, CPH), the normalizer eGFP, and the silencing suppressor p19 (as depicted in Figure 1B). Luminescence values were collected over 2-5 days post-infiltration. Each point represents the luminescence/fluorescence values of an individual disc taken from a different leaf of the same plant (n=3). **C)** Luminescence imaging of *N. benthamiana* leaves infiltrated with combinations of Geminino 1.0 vectors and the different Rep genes as described in B). **D)** Schematic representation of the hybrid Geminino 1.0 developed by combining the LIR from BeYDV with the IR of TYLCV, or BCTV. As a negative control, Geminino 1.0 vector containing only a single BeYDV LIR at the 5' end (BeY-), and as a positive control, Geminino 1.0 vector with two BeYDV LIRs (BeYLIR-BeYLIR). **E)** Luminescence normalized by fluorescence values resulting from the co-expression of hybrid vectors (BeLIR-TYIR, BeLIR-BCIR) and the controls (BeYLIR-BeYLIR, BeY-) carrying HmS by agroinfiltration in *N. benthamiana* as indicated in B). Luminescence values were collected over at day 3 post-infiltration. Each point represents the luminescence/fluorescence values of an individual disc taken from a different leaf of the same plant (n=3). **F)** Luminescence imaging of *N. benthamiana* leaves infiltrated with the treatments in E). Error bars represent standard deviation. Statistical analysis was performed using a one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05). Variables with the same letters belong to the same statistical group. The figure includes images from Biorender (biorender.com).

**Optimization of TYLCV-based vectors**

Owing to its higher activation and lower basal levels, we next decided to focus on the further optimization of TYLCV vector, taking advantage of the bioluminescence reporter and modular cloning. To address the issue of unintended basal expression in the TYLCV geminino systems, several alternative configurations were designed to reduce expression levels in the absence of Rep proteins. These configurations focused on modifying transcriptional control elements to minimize undesirable expression (Figure 5A). In the so-called Geminino 2.0, previously developed in our lab for BeYDV (Garcia-Perez et al., 2024), the ATG start codon of the HmS gene was relocated downstream of the CDS to prevent early transcription from upstream promoter elements. We also created a Geminino 3.0, where a transcriptional terminator was inserted between the first IR and the HmS CDS to block any unintended transcriptional activity from the IR. Next, a Geminino 4.0 was developed, where a predicted TATA box within the IR, which was described to promote viral gene V2 expression, was mutated to prevent transcriptional leakage from these elements. Finally, in a Geminino 5.0, HmS was oriented in the antisense strand relative to the viral genome (complementary-sense strand), aimed at preventing any influence of the 3′ IR on HmS transcription. Each of these configurations was evaluated, with luminescence serving as a proxy for HmS expression. As can be observed in Figure 5B and 5C, Geminino 2.0 produced negligible basal luminescence in the inactivated state, but also generated low luminescence levels when Rep was co-expressed, suggesting that this modification overly restricted gene expression, potentially impairing overall system functionality. Geminino 3.0 and 4.0 did not result in significant reductions in leakiness in comparison to Geminino 1.0. Both configurations continued to produce similar levels of luminescence to the original versions, indicating that the transcriptional terminator in Geminino 3.0 and the TATA box mutations in Geminino 4.0 were insufficient to fully prevent unintended transcription. By contrast, Geminino 5.0 yielded significant reduction in basal expression for the TYLCV system, while keeping luminescence levels comparable to the ones obtained with Geminino 1.0 (Figure 5B and 5C). The reverse orientation of the transcriptional unit substantially decreased luminescence, indicating that this approach effectively insulated HmS from the transcriptional influence of the TYLCV 3′ IR. The marked reduction in leakiness in the TYLCV Geminino 5.0 system highlights the potential of this approach for improving the specificity and control of geminino-based expression systems.

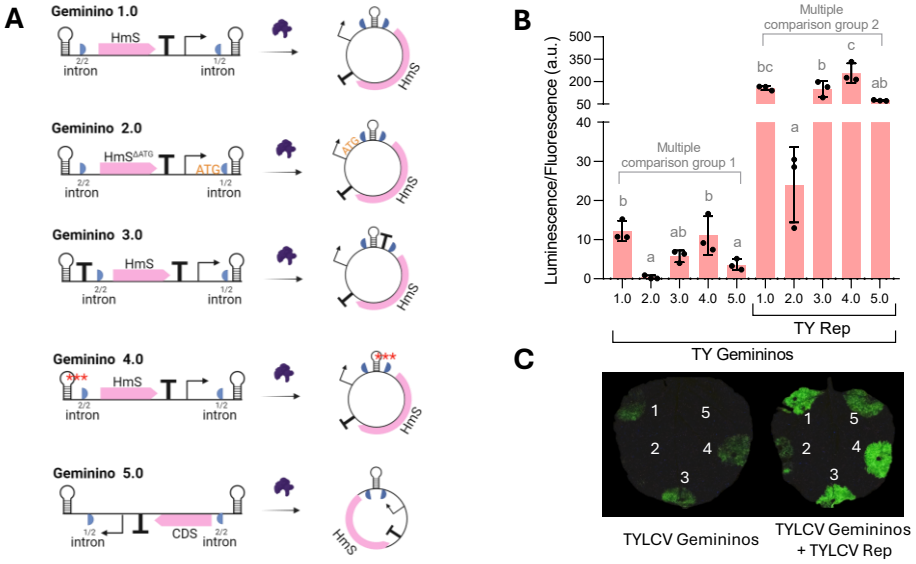



**Figure 5. Optimization of TYLCV vector for low basal expression. A)** Schematic representation of Geminino vector configurations. The coding sequence of HmS is rearranged with the viral elements (intergenic regions, IR, represented by a steam-loop) to control pre-circularization expression levels. Red asterisks in Geminino 4.0 represent 3-nucleotides point mutation in a TATA box. **B)** Luminescence normalized by fluorescence values (luminescence/fluorescence arbitrary units) resulting from the co-expression by agroinfiltration in *N. benthamiana* of TYLCV Geminino vectors in A) carrying the hispidin synthase HmS with the TYLCV Replication-associated protein (TY Rep), the rest of the genes of the bioluminescence pathway (H3H, Luz, CPH), the normalizer eGFP, and the silencing suppressor p19 (as depicted in Figure 1B). Luminescence values were collected at day 2 post-infiltration. Each point represents the luminescence/fluorescence values of an individual disc taken from a different leaf of the same plant (n=3). **C)** Luminescence imaging of *N. benthamiana* leaves infiltrated with combinations of Geminino vectors and TYLCV Rep as indicated in B). Error bars represent standard deviation. Statistical analysis was performed using a one-way ANOVA with Tukey's multiple comparisons test (P-value ≤ 0.05). Variables with the same letters belong to the same statistical group. The figure includes images from Biorender (biorender.com).

## DISCUSSION

In this study, we expanded the palette of modular geminiviral tools incorporating new elements derived from BeYDV, TYLCV, and BCTV. Our aim was to open out the utility of geminiviral vectors beyond traditional roles in Molecular Farming, moving towards more sophisticated systems capable of regulating synthetic gene circuits and facilitating multi-protein expression. Modular tools enable the exchange and optimization of individual DNA parts, allowing for the fine-tuning of vector configurations. Earlier, Dusek et al. (2020) followed this modular approach to combine in a single hyperexpression system DNA

elements coming from two previously described overexpression tools, namely the pEAQ vectors derived from Cowpea Mosaic Virus and the BeYDV replicon. Similarly, Neubauer et al. (2024) utilized the GoldenBraid modular cloning system to facilitate the optimization of a geminiviral vector. However, these studies have focused primarily on hyperexpression rather than regulation. In contrast, our study harnesses the modularity of GoldenBraid for regulatory purposes. By defining new standard overhangs for parts assembly, and by adapting DNA elements to these new overhangs, we have extended the standard "syntax" earlier proposed for modular cloning in plants (Sarrion-Perdigones et al., 2013) to new uses in geminivirus biotechnology.

One of the key findings in this study is the suitability of using the fungal-derived bioluminescent pathway as a highly sensitive reporter system for tracking geminivirus-mediated protein expression. Previous studies have often relied on traditional reporters like GFP or GUS. However, these reporters are either destructive (GUS) or offer limited dynamic range (fluorescence). While these limitations are less concerning when the goal is simply to maximize protein expression during transient expression, they become critical for more sophisticated goals as genome integration and/or synthetic gene circuit engineering, where fine-tune regulation and tight control of basal expression are required. Our results clearly demonstrate that HBP reporter is highly sensitive even when detected using consumer-grade cameras, enabling us to detect even minimal basal expression levels.

The HBP reporters allowed us to promptly assay several new geminiviral vector configurations. Specifically, in the case of TYLCV, new vector configuration involving alterations in the intergenic regions or reshuffling of transcriptional elements, were immediately reflected in the luminescence intensity in both quantitative reads and imaging acquisitions. This responsiveness underscores the utility of the bioluminescence reporter for tracking even subtle adjustments in vector optimization. Additionally, we observed that the dynamics of luminescence correlated well with the typical timeline of geminiviral Rep-mediated expression, which usually peaks around day 2 or 3 post-infiltration, followed by a decline in expression thereafter. This is consistent with previously reported results in systems like CuBe (Garcia-Perez et al., 2024), INPACT (Dugdale et al., 2013), and the BeYDV replicon studies by Zhang & Mason (2006). What makes this particularly interesting is that our bioluminescence reporter measures the output of a metabolic pathway rather than direct protein expression, reinforcing the potential of geminiviruses as vectors for modulating entire metabolic pathways, as previously demonstrated by other studies, e.g., in the production of specific metabolites (Romsuk et al., 2022).

Tight control over basal expression using BeYDV elements was already successfully demonstrated in CuBe (Garcia-Perez et al., 2024) and verified here using HBP reporter. This could potentially be attributed to our use of a mild strain of the virus (Regnard et al., 2010). By contrast, we encountered challenges in controlling basal expression when adapting TYLCV and BCTV vectors. The intergenic regions (IRs) of geminiviruses are known to contain promoter elements, including eukaryotic-like TATA boxes, which drive the expression of the

viral genome in both virion-sense and antisense directions. As shown in Figure 2B, the IRs of TYLCV and BCTV contain several annotated promoter elements, including TATA and CAAT boxes. The presence of these elements could contribute to the leaky expression of the downstream HmS gene. To investigate this, we initially targeted what appeared to be the most critical element, the TATA box, for mutation. However, we observed no significant reduction in basal expression after its mutation. Further mutation of other elements may yield more promising results, as demonstrated by Fenoll et al. (1990), who reported a 70% reduction in viral-sense promoter activity in MSV after mutating its GC-rich box. A high-throughput approach that introduces random mutations into the IR sequences may help identify the optimal combination for maintaining strong activation while minimizing basal expression. This strategy would be easy to test using the luminescence reporter system employed in our study.

For TYLCV, we eventually developed a configuration, Geminino 5.0, which achieved minimal basal expression. In this design, the CDS in reverse orientation, thus skipping any promoter activity from the IR located in the 5´end. Our results are in line with transcriptomic analysis of TYLCV showing higher transcriptional activity from the virion-sense strand (Piedra-Aguilera et al., 2019). Similarly, in BCTV, the most highly expressed genes are V2 and V3 encoded in the virion-sense strand (Standley, 2008). However, despite the success of TYLCV Geminino 5.0, this configuration was unsuccessful for BCTV (data not shown). Neubauer et al. (2024) recently reported that the inclusion of a foreign promoter upstream of a gene of interest is unnecessary in BCTV-based vectors, as the IR itself acts as a stronger promoter than the 35S promoter. BCTV is suggested to have evolved from a recombination event between mastreviruses and begomoviruses, as it contains elements related to both viral groups (Stanley, 2008; Varsani et al., 2009). This recombinant origin suggests that BCTV is highly permissive and adaptable, which may make it difficult to domesticate for Synthetic Biology applications. The presence of elements from both mastreviruses (e.g., BeYDV) and begomoviruses (e.g., TYLCV) may explain why basal expression levels are similarly high when BCTV is combined with either the Rep protein from TYLCV or BCTV itself. This observation contributes valuable insights into the field of virology, particularly regarding the cross-replication capabilities of geminiviruses. It has already been documented that when two geminiviruses co-infect the same plant, recombination and mixing of viral elements can occur (Lefeuvre & Moriones, 2015). The high infectivity and wide host range of BCTV may come with the trade-off of reduced replication efficiency, as suggested by our experiments comparing the performance of BeYDV, TYLCV, and BCTV. When it comes to achieving maximum expression levels, TYLCV appears to be the most promising geminivirus-based system for use in *N. benthamiana*, which makes sense given its adaptation to solanaceous plants and match with the results from Kim et al. (2024).

One of the most interesting and probably underutilized features of geminiviruses is their ability to support the replication of multiple viral vectors simultaneously within the same cell. We demonstrated how this capability can be harnessed for multi-protein expression by employing BeYDV-based triple replicons. Previously, Huang et al. (2010) and Dusek et al.

(2020) also demonstrated the utility of geminiviruses for co-expressing up to two proteins. Huang et al. (2010) showed robust expression of the heavy and light chains of an antibody using a single vector with tandemly arranged pro-replicons, along with the accumulation of replicating forms of separate DNA episomes in agroinfiltrated *N. benthamiana* leaves. Similarly, Dusek et al. (2020) explored the co-expression of two proteins, GFP and dsRed, but in their case, they simply positioned two transcriptional units one after the other, flanked by the LIR and SIR-LIR sequences. In this work, we went a step further by successfully co-expressing three proteins in a Rep-regulated manner. Additionally, we investigated the influence of transcriptional unit positioning within the vector configuration to determine if it impacted expression or replication. Our results showed there is not a clear order-dependent effect on expression, which contrasts with the findings from Dusek et al. (2020), where the transcriptional unit in the first position was less expressed. However, their study did not clarify whether this difference was due to transcriptional interference or replication issues. It is possible that in their system, where two standard transcriptional units were concatenated, the constitutive promoter of the first construct might have influenced the expression of the second gene, explaining the higher expression levels observed when the same protein was placed in the second position. While our COMBO triple-replicon system enables the co-expression of multiple proteins, challenges remain in developing configurations that prevent constitutive expression in the absence of the Rep protein, particularly for TYLCV and BCTV. In the case of TYLCV, it would be worth exploring the use of the COMBO configuration by concatenating Geminino 5.0 units, rather than Geminino 1.0, to improve regulation.

Another interesting finding from this work is the capability of a hybrid geminiviral vector, which combines LIRs from different viruses, to undergo circularization, as shown by HBP activation and confirmed by Sanger sequencing of the Geminino BeYLIR-BCIR (data not shown). In this hybrid vector, it appears that only the BeYDV Rep, responsible for the nicking of the 5' LIR, is necessary to mediate vector circularization. This was observed in both BeYLIR-TYIR and BeYLIR-BCIR hybrids, where BeYDV Rep alone was sufficient to initiate luminescence. Another unexpected finding in this line was that the BeYLIR- vector, initially included as a negative control, also produced luminescence when combined with Rep. This suggests that only the 5' IR may be required for circularization. This result contradicts previous findings by Neubauer et al. (2024), where a BCTV vector without the 3' IR did not show reporter expression. However, these experiments used a YFP reporter system, which may not have been as sensitive as the luminescence reporter used here. Further experiments, especially using stably integrated pro-replicons could clarify whether a single IR is sufficient for replicon activation via circularization, replication, or both.

In conclusion, this study expands the geminivirus toolbox for Plant Synthetic Biology, providing new insights into how geminiviral vectors can be optimized for controlled activation, multi-protein expression, and advanced reporter systems.

## MATERIALS & METHODS

### Cloning and assembly of GoldenBraid constructs

The DNA constructs utilized in this study were created using the GoldenBraid (GB) system, as described by Sarrion-Perdigones et al. (2013). The sequences for the constructs developed can be accessed at https://gbcloning.upv.es/search/features using the IDs listed in Table S3. Briefly, the DNA parts were cloned into the pUPD2 vector, validated by restriction analysis and sequencing, and then integrated into transcriptional units using cyclic restriction-ligation reactions. These units were further combined into multigenic constructs using BsaI (for alpha vectors) or BsmBI (for omega vectors) enzymes and T4 ligase. Constructs were transformed into *Escherichia coli* TOP 10 cells using the Zymo Research Mix&Go kit, and all assemblies were confirmed via restriction analysis.

### Agroinfiltration experiments

For transient expression, final plasmids were transformed into *Agrobacterium tumefaciens* strain EHA105. Bacterial cultures were grown for 16h at 28º and then resuspended in infiltration buffer (10 mM MES, 10 mM $MgCl_2$, 200 µM acetosyringone, pH 5.6) to a final $OD_{600}$ 0.1. After a 2-hour incubation, equal volumes of cultures were mixed when required for co-expression and they were co-infiltrated with a syringe into the abaxial side of three fully expanded leaves of five-week-old *N. benthamiana* plants. Plants were maintained in a controlled environment with 16h light/8h dark at 24°C/20°C.

### Luminescence and fluorescence measurements

For luminescence/fluorescence ratios measurements in luminescence assays, leaf discs (5 mm diameter) were collected 24 hours after infiltration and placed upside-down in a white 96-well plate (greiner bio-one, #655074) with 300 µl distilled water per well. Luminescence and fluorescence were measured in a GloMax Multi-Detection System (Promega). For time courses, plates were incubated at 25°C, under 16h light/8h dark conditions, and 60-70% humidity. Luminescence was measured with a 10-second integration time, and eGFP fluorescence using excitation 490 nm and emission 510-570 nm. Ratios of luminescence to fluorescence were calculated for each sample and timepoint.

For fluorescence measurement in the co-expression of eGFP, dsRed, and eBFP in the COMBO configuration, leaf discs (5 mm diameter) were collected at day 3 post-infiltration and placed upside-down in a black 96-well plate (Corning, #3792). Fluorescence was measured using Perkin Elmer Victor X2 Microplate Reader. Excitation wavelengths were set at 485 nm for eGFP, 530 nm for dsRed, and 405 nm for eBFP. The emission detection for fluorescence was 510 nm for eGFP, 595 nm for dsRed, and 510 nm for eBFP.

### Luminescence imaging of leaves

Leaves were harvested 24 hours post-infiltration and placed in petri dishes containing water to maintain hydration. Bioluminescence images were captured in a dark room using a SONY α7s camera set to an exposure time of 30 seconds, ISO 51200, and an aperture of f/2.8. The same ISO and aperture were used for bright-field images. To overlay dark-bioluminescence

and bright-field images, Paint.net software was used, applying a transparency level of 200 to the bioluminescence layer. Leaves remained in the dark room (at room temperature) until images were captured on days 2 or 3 post-infiltration.

**Confocal imaging to co-localize eGFP, dsRed, and BFP proteins**

Leaf discs (5 mm diameter) were harvested for imaging at 3 days post-infiltration. These sections were examined using a Zeiss LSM 510-META confocal laser scanning microscope. For fluorescence detection, an argon-ion laser emitting at 488 nm or 405 was used to capture eGFP and BFP, respectively, while a helium-neon laser at 560 nm captured dsRed fluorescence. The emission ranges were 383-448 nm for eBFP, 488-507 nm for eGFP, and 558-583 nm for dsRed. Image acquisition was managed through Zen software, and post-imaging analysis, was conducted using Fiji (ImageJ) software.

# General discussion

# GENERAL DISCUSSION

Due to their evolutionary history, plants have become highly efficient at biomass production, a trait inherited from marine ancestors adapted to low-light and low-$CO_2$ environments (Thomas & Sadras, 2001). This innate efficiency makes plants ideal for biotechnological applications, where high yields are essential for economically viable and sustainable production. Traditionally, genetic engineering has enabled plants to produce valuable heterologous compounds, such as pharmaceutical proteins and metabolites. Now, Synthetic Biology is revolutionizing this field, by turning plants from passive "chassis" into programmable "bioreactors". Using Synthetic Biology tools, we can enhance intrinsic "biocomputing" abilities of plants to sense, process, and respond to specific inputs, optimizing production with minimal interference to native regulatory networks. By assembling functionally characterized bioparts from growing repositories, self-contained synthetic gene circuits (SGCs) could be built that operate orthogonally and reliably in plant cells.

Despite recent progress, building fully operational SGCs remains challenging. This thesis presents refined sensor and processor modules and demonstrates their potential in constructing SGCs for precise gene expression control. Chapter 1 introduces a copper sensor paired with a transcriptional activator for tightly controlled endogenous gene activation. Chapter 2 presents a geminivirus-based processor coupled to a copper sensor to drive regulated recombinant protein expression in plants. Chapter 3 expands the toolkit with additional geminiviral processors, highlighting each vector's unique activation profile and potential for Plant Synthetic Biology.

Copper sulfate ($CuSO_4$) was selected as SGC-inducer due to its stability, ease of application, affordability, and scalability in laboratory and field settings. Although it is approved for organic farming under European Commission Regulation No. 2021/1165, field application must consider potential risks of soil accumulation (Tamm et al., 2022). However, the contained hydroponics and post-harvest induction methods explored in this thesis serve to reduce in-field copper usage, mitigating environmental impact.

The copper sensor optimized in this thesis, which includes a copper-responsive transcriptional factor (CUP2-Gal4) and a synthetic promoter, shows reliable and orthogonal activation. Recent improvements by other researchers demonstrate the adaptability of the sensor for fine-tuning gene expression. For example, Moreno-Giménez et al. (2022) incorporated a promising fungal-derived minimal promoter, while Chiang et al. (2024) added regulatory layers, including copper-induced splicing of a 'suicide exon' and a Cre/loxP system for controlled gene expression. In line with these refinements, this thesis introduced an additional regulatory layer using a Flp/FRT system (Chapter 2), highlighting the versatility and potential of the sensor for further fine-tuning.

Kallam et al. (2023) assessed the performance of our copper sensor with various programmable transcriptional activators (PTAs), including the transcription activator-like effector (TALE) system (Cai, Kallam, et al., 2020) and a Gal4:ΦC3 fusion protein (Bernabé-Orts et al., 2020; Cai et al., 2020). Among these, dCasEV2.1 activator (Selma et al., 2019)

demonstrated the lowest basal expression, making it ideal for achieving strong but tightly controlled expression when coupled with the copper sensor (CS/dCasEV2.1). However, Kallam et al. (2023) also observed that in transgenic plants harbouring CS/dCasEV2.1, activation levels were lower than in transient systems, likely due to limited availability of dCasEV2.1 protein components. Future improvements could include using replicative vectors, such as the geminivirus-based system developed in Chapter 2, to increase gene dosage of CS/dCasEV2.1 in stable transgenic plants.

Emerging PTAs like CRISPR-Act3.0 (Pan et al., 2021) could improve activation levels in stable systems by enhancing activator recruitment. Additionally, controlling gRNA expression through the copper sensor, while keeping dCasEV2.1 constitutively expressed, may yield stronger activation in stable plants. Although our attempt to regulate gRNA with the tRNA-gRNA approach was unsuccessful, Khan et al. (2024) demonstrated that using the Csy4 endonuclease enables precise gRNA regulation with minimal leakage. This system, designed for application in a NOR logic gate (Figure 5, Introduction) with heat and chemical inputs, illustrates the potential of regulated PTAs for precise and multi-input control of metabolic pathways. Selma et al. (2021) had already showed the utility of dCasEV2.1 in selectively activating enzymes in distinct branches of the phenylpropanoid pathway. Integrating the multiplexed dCasEV2.1 activation with multiple sensors, like the copper and temperature ones, could enable customizable modulation of plant metabolism and controlled accumulation of specific metabolites in response to external cues.

Building on the copper sensor utility demonstrated in Chapter 1, we further applied this regulation approach in Chapter 2 to enhance the production of high-value compounds. By integrating the copper sensor with a geminivirus-based vector, we developed the CuBe system, a copper-inducible platform that amplifies transgene copy number upon copper induction, increasing gene expression. Although qPCR results (Figures 5J and 6H, Chapter 2) confirmed high replicon DNA levels following copper treatment, protein yields (~120 µg/g fresh weight) were modest, similar to those of INPACT (Dugdale et al., 2013), another geminivirus-based system. These yields, however, remain below those achieved with the MagnICON system (Marillonnet et al., 2005), suggesting a saturation point in geminiviral systems where high copy numbers may overwhelm the transcriptional machinery of the plant.

To enhance expression from BeYDV vectors, incorporating alternative untranslated regions (UTRs) and matrix attachment regions (MARs), as suggested by Diamos et al. (2016), could stabilize mRNA and increase transcriptional efficiency in the CuBe system. For instance, using the UTRs employed in the pHREAC vector (Peyret et al., 2019) to outperform the pEAQ-HT vector (a widely utilized non-replicative vector Molecular Farming tool, (Sainsbury et al., 2009)), could further refine the CuBe performance. Furthermore, to enhance protein production yields, co-expression of CuBe with silencing suppressors, such as p19, can be considered (see Figure 2, Chapter 2). However, stable transformation of potent silencing suppressors can have unintended consequences, such as compromising plant defences. Therefore, using the copper sensor to also induce the expression of silencing suppressors in stable conditions might help to balance yield optimization with plant homeostasis.

Another challenge in geminivirus-based vectors is enabling cell-to-cell movement. Our attempts to confer movement by co-expressing movement proteins (MPs) from other viruses, including BeYDV and TMV, were unsuccessful, with no significant increase in eGFP-expressing cell clusters. Enhancing copper uptake and transport (this is, moving the activation signal instead of the system to be activated) within plants may provide an alternative to improve activation coverage in CuBe-based systems. Copper uptake and transport in plants are tightly regulated to avoid toxicity while ensuring proper cellular function (reviewed in Printz et al. (2016). For activation via hydroponics, strategies could involve upregulating copper uptake genes like COPT and FRO, combined with metallothioneins (MTs) or copper chaperones (e.g., ATX1) to manage root-level reactive oxygen species (ROS) while ensuring copper availability in leaves. For post-harvest induction, the physiological state of leaves could be optimized, for instance, by activating CuBe at midday when stomatal opening is at its maximum, enhancing copper uptake (Kang et al., 2007; Patel et al., 2021).

Stable transgenic plants expressing bioproducts of interest offer potential for reliable, GMP-compliant production of high-value proteins. However, the variability among regenerated lines remains a challenge (Guiziou, 2024). We addressed this by incorporating MARs in our genetic constructs to reduce positional effects (Pérez-González & Caro, 2019), yet further stability could be achieved by targeting transgene insertion into genomic "landing pads" through homologous recombination or recombinase-mediated cassette exchange. Whole-genome sequencing of high-yield CuBe lines could help identify optimal sites for predictable transgene expression. This strategy could enable the integration of other genes into prototype CuBe plants, establishing this way robust and flexible platforms for custom protein production.

Another challenge in stable expression approaches is the time-consuming transformation process. By using traditional tissue culture methods (Sparkes et al., 2006), generating stable *N. benthamiana* plants implies around six-months timeline. However, recent advances, such as using morphogenic regulator genes (Cody et al., 2023; Gordon-Kamm et al., 2019) or transforming embryonic axes and meristems (Ganguly et al., 2020; Maher et al., 2020), can significantly speed up the process. Additionally, seed generation is not a major concern in *Nicotiana* species, allowing for establishing a long-term reservoir of transformed plants once a stable line is obtained (Bally et al., 2018). Interestingly, recent studies showed the use of tools like estradiol-inducible CRISPR to promote plant regeneration (Zhang et al., 2024). Similarly, our CS/dCasEV2.1 tool could be used to further enhance transformation efficiency for the CuBe system.

Building upon our exploration of the BeYDV-based processor in Chapter 2, Chapter 3 expands on geminiviral vectors for Synthetic Biology. This chapter highlights the utility of bioluminescence as a sensitive reporter to facilitate the study of newly developed geminivirus-based vectors for precise gene expression control. Key challenges explored in this section include addressing basal expression of geminiviral vectors. While BeYDV-based vectors showed good regulation, TYLCV and BCTV vectors presented leakiness issues.

Although we were able to minimize TYLCV leakiness, BCTV still requires further refinement, potentially through targeted mutations in non-essential replication regions to conceal tight regulation with efficient gene expression. Notably in this work, we demonstrated the simultaneous co-expression of three genes by using the COMBO configuration, a promising step towards producing protein complexes. Future research might focus on optimizing protein balance between co-expressed proteins, modulating expression of each replicative unit with different strength promoters. Expanding co-expression capacity and fine-tuning expression ratios could enable advanced applications like producing complex virus-like particles decorated with immunogenic antigens for pharmaceuticals.

As an integrative vision, this thesis has significantly advanced the development of novel tools for synthetic gene circuit design in plants. The integration of the systems developed here lays the groundwork for more sophisticated gene circuits capable of executing multi-layered functions within a plant. An exciting future direction is combining CuBe with the CS/dCasEV2.1 to enable simultaneous recombinant protein expression and endogenous gene activation, potentially enhancing the production yields of plant biofactory prototypes. For example, the CS/dCasEV2.1 could upregulate chaperone proteins to enhance protein folding and stability, improving yield and quality in CuBe-driven recombinant protein production. Alternatively, activating both systems could allow plants to simultaneously produce valuable metabolites and proteins that can be sequentially extracted in a biorefinery cascade. This dual activation of independent modules with a single input demonstrates the potential of SGCs for finely tuned plant responses. Additionally, expanding this framework to include multi-input circuits responsive to diverse environmental cues would position plants as advanced bio-computing platforms. For instance, by introducing two distinct replicative systems responsive to different inducers, such as copper for a BeYDV-based Rep and temperature for a TYLCV-based Rep, plants could be engineered to produce distinct outputs based on separate inputs. This approach aligns with recent visions in plant Synthetic Biology (Guiziou, 2024) that highlight the shift toward complex, logic-based circuits capable of processing environmental cues in ways that mimic computational systems. As a final remark, while our work focused on *N. benthamiana*, extending the tools developed in this thesis to other plant species is crucial. Although *N. benthamiana* is standard for Molecular Farming, scaling to high-biomass crops like tobacco could enhance protein production for large-scale applications. Overall, this work opens pathways for next-generation applications in plant biotechnology, establishing a foundation for plant systems capable of executing self-regulating and multi-functional gene circuits to improve plant biomanufacturing.

# Conclusions

# CONCLUSIONS

1. A copper sensor resulting from the combination of the yeast CUP2 copper-responsive factor, the yeast Gal4 activation domain, and a synthetic promoter containing four tandemly repeated operators upstream of the minimal promoter of the tomato DFR gene was functionally characterized in *Nicotiana benthamiana* showing an optimized copper response.

2. The coupling of the new copper sensor to the CRISPR/Cas9-derived programmable transcriptional activator dCasEV2.1 enables fine-tuned, copper-dependent activation of endogenous *N. benthamiana* genes, as exemplified with DFR and PAL2 genes.

3. A new modular configuration of a BYDV-based geminiviral replicative vector for recombinant protein production was designed, termed as Geminino 1.0. The expression of an eGFP reporter gene using this new vector configuration resulted in a strong eGFP accumulation upon replicon activation in the presence of the Rep/RepA protein, whereas minimal reporter expression was recorded in the absence of Rep/RepA.

4. By coupling the optimized copper sensor with the newly configured Geminino 1.0 processor, a novel copper-regulated genetic circuit for the amplification of transgenes-of-interest was established. This new customizable circuit, termed CuBe, was equipped with an additional regulatory layer comprising a removable transcriptional terminator controlled by a copper-inducible Flp recombinase. When customized for transient eGFP expression in *Nicotiana benthamiana*, the CuBe circuit showed low fluorescence levels in the basal state and high eGFP expression upon copper-mediated activation.

5. Stable transgenic *N. benthamiana* lines were developed harbouring customized CuBe circuits for the overproduction of eGFP and an anti-SARS-CoV-2 antibody respectively. The new transgenic lines showed consistent conditional activation of their respective target bioproduct upon copper application using different treatments, namely foliar spraying, post-harvest leaf submergence, or the supplementation of copper in a hydroponic medium.

6. An autobiouminescence system based on a hybrid bioluminiscence pathway (HBP) was established as a sensitive and unexpensive reporter for geminiviral vector characterization. The HBP reporter facilitated the analysis of new modular viral vector configurations based on BeYDV, TYLCV, and BCTV geminivirus respectively.

# References

# REFERENCES


Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w

Ahlquist, P., Schwartz, M., Chen, J., Kushner, D., Hao, L., & Dye, B. T. (2005). Viral and host determinants of RNA virus vector replication and expression. *Vaccine*, *23*(15), 1784–1787. https://doi.org/10.1016/j.vaccine.2004.11.005

Andres, J., Blomeier, T., & Zurbriggen, M. D. (2019). Synthetic Switches and Regulatory Circuits in Plants. *Plant Physiology*, *179*(3), Article 3. https://doi.org/10.1104/pp.18.01362

Andrianantoandro, E., Basu, S., Karig, D. K., & Weiss, R. (2006). Synthetic biology: New engineering rules for an emerging discipline. *Molecular Systems Biology*, *2*(1), 2006.0028. https://doi.org/10.1038/msb4100073

Aoyama, T., & Chua, N. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. *The Plant Journal*, *11*(3), 605–612. https://doi.org/10.1046/j.1365-313X.1997.11030605.x

Argüello-Astorga, G. R., & Ruiz-Medrano, R. (2001). An iteron-related domain is associated to Motif 1 in the replication proteins of geminiviruses: Identification of potential interacting amino acid-base pairs by a comparative approach. *Archives of Virology*, *146*(8), 1465–1485. https://doi.org/10.1007/s007050170072

Baba, S., Abe, Y., Ono, C., & Hosobuchi, M. (2006). Targeted disruption of the genes, mlcR and ariB, which encode GAL4-type proteins in Penicillium citrinum. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, *1759*(8–9), 410–416. https://doi.org/10.1016/j.bbaexp.2006.08.001

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., Van Dijk, A. A., Ebrecht, A. C., … Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, *373*(6557), 871–876. https://doi.org/10.1126/science.abj8754

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

Baltes, N. J., Gil-Humanes, J., Cermak, T., Atkins, P. A., & Voytas, D. F. (2014). DNA Replicons for Plant Genome Engineering. *The Plant Cell*, *26*(1), 151–163. https://doi.org/10.1105/tpc.113.119792

Barnum, C. R., Endelman, B. J., Ornelas, I. J., Pignolet, R. M., & Shih, P. M. (2022). Optimization of Heterologous Glucoraphanin Production *In Planta*. *ACS Synthetic Biology*, *11*(5), 1865–1873. https://doi.org/10.1021/acssynbio.2c00030

Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences*, *115*(25), 6506–6511. https://doi.org/10.1073/pnas.1711842115

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, *315*(5819), 1709–1712. https://doi.org/10.1126/science.1138140

Bass, H. W., Nagar, S., Hanley-Bowdoin, L., & Robertson, D. (2000). Chromosome condensation induced by geminivirus infection of mature plant cells. *Journal of Cell Science*, *113*(7), 1149–1160. https://doi.org/10.1242/jcs.113.7.1149

## References


Belcher, M. S., Vuu, K. M., Zhou, A., Mansoori, N., Agosto Ramos, A., Thompson, M. G., Scheller, H. V., Loqué, D., & Shih, P. M. (2020). Design of orthogonal regulatory systems for modulating gene expression in plants. *Nature Chemical Biology*, *16*(8), 857–865. https://doi.org/10.1038/s41589-020-0547-4

Beltrán, J., Steiner, P. J., Bedewitz, M., Wei, S., Peterson, F. C., Li, Z., Hughes, B. E., Hartley, Z., Robertson, N. R., Medina-Cucurella, A. V., Baumer, Z. T., Leonard, A. C., Park, S.-Y., Volkman, B. F., Nusinow, D. A., Zhong, W., Wheeldon, I., Cutler, S. R., & Whitehead, T. A. (2022). Rapid biosensor development using plant hormone receptors as reprogrammable scaffolds. *Nature Biotechnology*, *40*(12), 1855–1861. https://doi.org/10.1038/s41587-022-01364-5

Benner, S. A., & Sismour, A. M. (2005). Synthetic biology. *Nature Reviews Genetics*, *6*(7), 533–543. https://doi.org/10.1038/nrg1637

Bennett, C. W. (1971). *The Curly Top Disease of Sugarbeet and Other Plants, Monograph No. 7*. The American Phytopathological Society. https://doi.org/10.1094/9780890546260

Berg, P., & Mertz, J. E. (2010). Personal Reflections on the Origins and Emergence of Recombinant DNA Technology. *Genetics*, *184*(1), 9–17. https://doi.org/10.1534/genetics.109.112144

Beritza, K., Watts, E. C., & Van Der Hoorn, R. A. L. (2024). Improving transient protein expression in agroinfiltrated *Nicotiana benthamiana*. *New Phytologist*, *243*(3), 846–850. https://doi.org/10.1111/nph.19894

Bernabé-Orts, J. M., Quijano-Rubio, A., Vazquez-Vilar, M., Mancheño-Bonillo, J., Moles-Casas, V., Selma, S., Gianoglio, S., Granell, A., & Orzaez, D. (2020a). A memory switch for plant synthetic biology based on the phage ϕC31 integration system. *Nucleic Acids Research*, *48*(6), 3379–3394. https://doi.org/10.1093/nar/gkaa104

Bernabé-Orts, J. M., Quijano-Rubio, A., Vazquez-Vilar, M., Mancheño-Bonillo, J., Moles-Casas, V., Selma, S., Gianoglio, S., Granell, A., & Orzaez, D. (2020b). A memory switch for plant synthetic biology based on the phage ϕC31 integration system. *Nucleic Acids Research*, *48*(6). https://doi.org/10.1093/nar/gkaa104

Bhattacharjee, B., & Hallan, V. (2022). Geminivirus-Derived Vectors as Tools for Functional Genomics. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.799345

Boetti, H., Chevalier, L., Denmat, L. A., Thomas, D., & Thomasset, B. (1999). Efficiency of physical (light) or chemical (ABA, tetracycline, CuSO4 or 2-CBSU)-stimulus-dependent gus gene expression in tobacco cell suspensions. *Biotechnology and Bioengineering*, *64*(1), 1–13. https://doi.org/10.1002/(SICI)1097-0290(19990705)64:1<1::AID-BIT1>3.0.CO;2-E

Bonnamy, M., Blanc, S., & Michalakis, Y. (2023). Replication mechanisms of circular ssDNA plant viruses and their potential implication in viral gene expression regulation. *mBio*, *14*(5), e01692-23. https://doi.org/10.1128/mbio.01692-23

Borah, B. K., Zarreen, F., Baruah, G., & Dasgupta, I. (2016). Insights into the control of geminiviral promoters. *Virology*, *495*, 101–111. https://doi.org/10.1016/j.virol.2016.04.033

Bortesi, L., Rademacher, T., Schiermeyer, A., Schuster, F., Pezzotti, M., & Schillberg, S. (2012). Development of an optimized tetracycline-inducible expression system to increase the accumulation of interleukin-10 in tobacco BY-2 suspension cells. *BMC Biotechnology*, *12*(1), 40. https://doi.org/10.1186/1472-6750-12-40

Boulton, M. I. (2002). Functions and interactions of mastrevirus gene products. *Physiological and Molecular Plant Pathology*, *60*(5), 243–255. https://doi.org/10.1006/pmpp.2002.0403

Brophy, J. A. N., Magallon, K. J., Duan, L., Zhong, V., Ramachandran, P., Kniazev, K., & Dinneny, J. R. (2022). Synthetic genetic circuits as a means of reprogramming plant roots. *Science*, *377*(6607), 747–751. https://doi.org/10.1126/science.abo4326

Buchman, C., Skroch, P., Welch, J., Fogel, S., & Karin, M. (1989). The CUP2 gene product, regulator of yeast metallothionein expression, is a copper-activated DNA-binding protein. *Molecular and Cellular Biology*, *9*(9), 4091–4095. https://doi.org/10.1128/mcb.9.9.4091-4095.1989

Budin, I., & Keasling, J. D. (2019). Synthetic Biology for Fundamental Biochemical Discovery. *Biochemistry*, *58*(11), 1464–1469. https://doi.org/10.1021/acs.biochem.8b00915

Burkhead, J. L., Gogolin Reynolds, K. A., Abdel-Ghany, S. E., Cohu, C. M., & Pilon, M. (2009). Copper homeostasis. *New Phytologist*, *182*(4), 799–816. https://doi.org/10.1111/j.1469-8137.2009.02846.x

Butaye, K. M. J., Goderis, I. J. W. M., Wouters, P. F. J., Pues, J. M. -T. G., Delauré, S. L., Broekaert, W. F., Depicker, A., Cammue, B. P. A., & De Bolle, M. F. C. (2004). Stable high-level transgene expression in *Arabidopsis thaliana* using gene silencing mutants and matrix attachment regions. *The Plant Journal*, *39*(3), 440–449. https://doi.org/10.1111/j.1365-313X.2004.02144.x

Butler, N. M., Baltes, N. J., Voytas, D. F., & Douches, D. S. (2016). Geminivirus-Mediated Genome Editing in Potato (Solanum tuberosum L.) Using Sequence-Specific Nucleases. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.01045

Buyel, J. F., & Fischer, R. (2012). Predictive models for transient protein expression in tobacco ( *Nicotiana tabacum* L.) can optimize process time, yield, and downstream costs. *Biotechnology and Bioengineering*, *109*(10), 2575–2588. https://doi.org/10.1002/bit.24523

Caddick, M. X., Greenland, A. J., Jepson, Ian, Krause, K.-P., Qu, N., Riddell, K. V., Salter, M. G., Schuch, W., Sonnewald, U., & Tomsett, A. B. (1998a). An ethanol inducible gene switch for plants used to manipulate carbon metabolism. *Nature Biotechnology*, *16*(2), 177–180. https://doi.org/10.1038/nbt0298-177

Caddick, M. X., Greenland, A. J., Jepson, L., Krause, K.-P., Qu, N., Riddell, K. V., Salter, M. G., Schuch, W., Sonnewald, U., & Tomsett, A. B. (1998b). An ethanol inducible gene switch for plants used to manipulate carbon metabolism. *Nature Biotechnology*, *16*(2), 177–180. https://doi.org/10.1038/nbt0298-177

Cai, Y.-M., Carrasco Lopez, J. A., & Patron, N. J. (2020). Phytobricks: Manual and Automated Assembly of Constructs for Engineering Plants. In S. Chandran & K. W. George (Eds.), *DNA Cloning and Assembly* (Vol. 2205, pp. 179–199). Springer US. https://doi.org/10.1007/978-1-0716-0908-8_11

Cai, Y.-M., Kallam, K., Tidd, H., Gendarini, G., Salzman, A., & Patron, N. J. (2020). Rational design of minimal synthetic promoters for plants. *Nucleic Acids Research*, *48*(21), 11845–11856. https://doi.org/10.1093/nar/gkaa682

Calvache, C., Vazquez-Vilar, M., Moreno-Giménez, E., & Orzaez, D. (2024). A quantitative autonomous bioluminescence reporter system with a wide dynamic range for Plant Synthetic Biology. *Plant Biotechnology Journal*, *22*(1), 37–47. https://doi.org/10.1111/pbi.14146

Calvache, C., Vazquez-Vilar, M., Selma, S., Uranga, M., Fernández-del-Carmen, A., Daròs, J., & Orzáez, D. (2022). Strong and tunable anti-CRISPR/Cas activities in plants. *Plant Biotechnology Journal*, *20*(2), 399–408. https://doi.org/10.1111/pbi.13723

Cameron, D. E., Bashor, C. J., & Collins, J. J. (2014). A brief history of synthetic biology. *Nature Reviews Microbiology*, *12*(5), 381–390. https://doi.org/10.1038/nrmicro3239

Cantú-Iris, M., Pastor-Palacios, G., Mauricio-Castillo, J. A., Bañuelos-Hernández, B., Avalos-Calleros, J. A., Juárez-Reyes, A., Rivera-Bustamante, R., & Argüello-Astorga, G. R. (2019). Analysis of a new begomovirus unveils a composite element conserved in the CP gene promoters of several Geminiviridae genera: Clues to comprehend the complex regulation of late genes. *PLOS ONE*, *14*(1), e0210485. https://doi.org/10.1371/journal.pone.0210485

Carbonell, P., Jervis, A. J., Robinson, C. J., Yan, C., Dunstan, M., Swainston, N., Vinaixa, M., Hollywood, K. A., Currin, A., Rattray, N. J. W., Taylor, S., Spiess, R., Sung, R., Williams, A. R., Fellows, D., Stanford,

N. J., Mulherin, P., Le Feuvre, R., Barran, P., … Scrutton, N. S. (2018). An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Communications Biology*, *1*(1), 66. https://doi.org/10.1038/s42003-018-0076-9

Carey, M., Leatherwood, J., & Ptashne, M. (1990). A Potent GAL4 Derivative Activates Transcription at a Distance in Vitro. *Science*, *247*(4943), 710–712. https://doi.org/10.1126/science.2405489

Casas-Mollano, J. A., Zinselmeier, M. H., Erickson, S. E., & Smanski, M. J. (2020). CRISPR-Cas Activators for Engineering Gene Expression in Higher Eukaryotes. *The CRISPR Journal*, *3*(5), 350–364. https://doi.org/10.1089/crispr.2020.0064

Castillo, A. G., Collinet, D., Deret, S., Kashoggi, A., & Bejarano, E. R. (2003). Dual interaction of plant PCNA with geminivirus replication accessory protein (Ren) and viral replication protein (Rep). *Virology*, *312*(2), 381–394. https://doi.org/10.1016/S0042-6822(03)00234-4

Castro, C., Arnold, J. J., & Cameron, C. E. (2005). Incorporation fidelity of the viral RNA-dependent RNA polymerase: A kinetic, thermodynamic and structural perspective. *Virus Research*, *107*(2), 141–149. https://doi.org/10.1016/j.virusres.2004.11.004

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., & Prasher, D. C. (1994). Green Fluorescent Protein as a Marker for Gene Expression. *Science*, *263*(5148), 802–805. https://doi.org/10.1126/science.8303295

Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C. D., Wiegand, D. J., Ter-Ovanesyan, D., Braff, J. L., Davidsohn, N., Housden, B. E., Perrimon, N., Weiss, R., Aach, J., Collins, J. J., & Church, G. M. (2015a). Highly efficient Cas9-mediated transcriptional programming. *Nature Methods*, *12*(4), 326–328. https://doi.org/10.1038/nmeth.3312

Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C. D., Wiegand, D. J., Ter-Ovanesyan, D., Braff, J. L., Davidsohn, N., Housden, B. E., Perrimon, N., Weiss, R., Aach, J., Collins, J. J., & Church, G. M. (2015b). Highly efficient Cas9-mediated transcriptional programming. *Nature Methods*, *12*(4), Article 4. https://doi.org/10.1038/nmeth.3312

Chen, B.-S., Chang, C.-H., & Lee, H.-C. (2009). Robust synthetic biology design: Stochastic game theory approach. *Bioinformatics*, *25*(14), 1822–1830. https://doi.org/10.1093/bioinformatics/btp310

Chen, M., & Qi, L. S. (2017). Repurposing CRISPR System for Transcriptional Activation. In L.-C. Li (Ed.), *RNA Activation* (Vol. 983, pp. 147–157). Springer Singapore. https://doi.org/10.1007/978-981-10-4310-9_10

Chen, Q., He, J., Phoolcharoen, W., & Mason, H. S. (2011). Geminiviral vectors based on bean yellow dwarf virus for production of vaccine antigens and monoclonal antibodies in plants. *Human Vaccines*, *7*(3), 331–338. https://doi.org/10.4161/hv.7.3.14262

Chiang, B., Lin, K., Chen, Y., Huang, C., Goh, F., Huang, L., Chen, L., & Wu, C. (2024). Development of a tightly regulated copper-inducible transient gene expression system in *Nicotiana benthamiana* incorporating a suicide exon and Cre recombinase. *New Phytologist*, *244*(1), 318–331. https://doi.org/10.1111/nph.20021

Chi-Ham, C. L., Boettiger, S., Figueroa-Balderas, R., Bird, S., Geoola, J. N., Zamora, P., Alandete-Saez, M., & Bennett, A. B. (2012). An intellectual property sharing initiative in agricultural biotechnology: Development of broadly accessible technologies for plant transformation. *Plant Biotechnology Journal*, *10*(5), 501–510. https://doi.org/10.1111/j.1467-7652.2011.00674.x

Chowdhury, R., Bouatta, N., Biswas, S., Floristean, C., Kharkar, A., Roy, K., Rochereau, C., Ahdritz, G., Zhang, J., Church, G. M., Sorger, P. K., & AlQuraishi, M. (2022). Single-sequence protein structure prediction using a language model and deep learning. *Nature Biotechnology*, *40*(11), 1617–1623. https://doi.org/10.1038/s41587-022-01432-w

Chung, H. Y., Lee, H. H., Kim, K. I., Chung, H. Y., Hwang-Bo, J., Park, J. H., Sunter, G., Kim, J. B., Shon, D. H., Kim, W., & Chung, I. S. (2011). Expression of a recombinant chimeric protein of hepatitis A virus

VP1-Fc using a replicating vector based on Beet curly top virus in tobacco leaves and its immunogenicity in mice. *Plant Cell Reports*, *30*(8), 1513–1521. https://doi.org/10.1007/s00299-011-1062-6

Chung, Y. H., Church, D., Koellhoffer, E. C., Osota, E., Shukla, S., Rybicki, E. P., Pokorski, J. K., & Steinmetz, N. F. (2022). Integrating plant molecular farming and materials research for next-generation vaccines. *Nature Reviews Materials*, *7*(5), 372–388. https://doi.org/10.1038/s41578-021-00399-5

Chylinski, K., Makarova, K. S., Charpentier, E., & Koonin, E. V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Research*, *42*(10), 6091–6105. https://doi.org/10.1093/nar/gku241

Cody, J. P., Maher, M. F., Nasti, R. A., Starker, C. G., Chamness, J. C., & Voytas, D. F. (2023). Direct delivery and fast-treated Agrobacterium co-culture (Fast-TrACC) plant transformation methods for Nicotiana benthamiana. *Nature Protocols*, *18*(1), 81–107. https://doi.org/10.1038/s41596-022-00749-9

Cohen, S. N., Chang, A. C. Y., & Hsu, L. (1972). Nonchromosomal Antibiotic Resistance in Bacteria: Genetic Transformation of *Escherichia coli* by R-Factor DNA. *Proceedings of the National Academy of Sciences*, *69*(8), 2110–2114. https://doi.org/10.1073/pnas.69.8.2110

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823. https://doi.org/10.1126/science.1231143

Contag, C. H., & Bachmann, M. H. (2002). Advances in In Vivo Bioluminescence Imaging of Gene Expression. *Annual Review of Biomedical Engineering*, *4*(1), 235–260. https://doi.org/10.1146/annurev.bioeng.4.111901.093336

Corrado, & Karali. (2009). Inducible gene expression systems and plant biotechnology. *Biotechnology Advances*, *27*(6), 733–743. https://doi.org/10.1016/j.biotechadv.2009.05.006

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

DalCorso, G., Manara, A., Piasentin, S., & Furini, A. (2014). Nutrient metal elements in plants. *Metallomics*, *6*(10), 1770–1788. https://doi.org/10.1039/C4MT00173G

Davies, J. A. (2018). *Synthetic Biology: A Very Short Introduction*. Oxford Academic.

De Lorenzo, V., & Danchin, A. (2008). Synthetic biology: Discovering new worlds and new words: The new and not so new aspects of this emerging research field. *EMBO Reports*, *9*(9), 822–827. https://doi.org/10.1038/embor.2008.159

Del Valle, I., Fulk, E. M., Kalvapalle, P., Silberg, J. J., Masiello, C. A., & Stadler, L. B. (2021). Translating New Synthetic Biology Advances for Biosensing Into the Earth and Environmental Sciences. *Frontiers in Microbiology*, *11*, 618373. https://doi.org/10.3389/fmicb.2020.618373

Diamos, A. G., & Mason, H. S. (2018a). Chimeric 3′ flanking regions strongly enhance gene expression in plants. *Plant Biotechnology Journal*, *16*(12), 1971–1982. https://doi.org/10.1111/pbi.12931

Diamos, A. G., & Mason, H. S. (2018b). High-level expression and enrichment of norovirus virus-like particles in plants using modified geminiviral vectors. *Protein Expression and Purification*, *151*, 86–92. https://doi.org/10.1016/j.pep.2018.06.011

Diamos, A. G., & Mason, H. S. (2019). Modifying the Replication of Geminiviral Vectors Reduces Cell Death and Enhances Expression of Biopharmaceutical Proteins in Nicotiana benthamiana Leaves. *Frontiers in Plant Science*, *9*, 1974. https://doi.org/10.3389/fpls.2018.01974

Diamos, A. G., Rosenthal, S. H., & Mason, H. S. (2016). 5′ and 3′ Untranslated Regions Strongly Enhance Performance of Geminiviral Replicons in Nicotiana benthamiana Leaves. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.00200

## References


Diego-Martin, B., González, B., Vazquez-Vilar, M., Selma, S., Mateos-Fernández, R., Gianoglio, S., Fernández-del-Carmen, A., & Orzáez, D. (2020). Pilot Production of SARS-CoV-2 Related Proteins in Plants: A Proof of Concept for Rapid Repurposing of Indoor Farms Into Biomanufacturing Facilities. *Frontiers in Plant Science*, *11*, 612781. https://doi.org/10.3389/fpls.2020.612781

Diego-Martin, B., González, B., Vazquez-Vilar, M., Selma, S., Mateos-Fernández, R., Gianoglio, S., Fernández-Del-Carmen, A., & Orzáez, D. (2020). Pilot Production of SARS-CoV-2 Related Proteins in Plants: A Proof of Concept for Rapid Repurposing of Indoor Farms Into Biomanufacturing Facilities. *Frontiers in Plant Science*, *11*, 612781. https://doi.org/10.3389/fpls.2020.612781

Dudley, Q. M., Jo, S., Guerrero, D. A. S., Chhetry, M., Smedley, M. A., Harwood, W. A., Sherden, N. H., O'Connor, S. E., Caputi, L., & Patron, N. J. (2022). Reconstitution of monoterpene indole alkaloid biosynthesis in genome engineered Nicotiana benthamiana. *Communications Biology*, *5*(1), 949. https://doi.org/10.1038/s42003-022-03904-w

Dugdale, B., Mortimer, C. L., Kato, M., James, T. A., Harding, R. M., & Dale, J. L. (2013). In Plant Activation: An Inducible, Hyperexpression Platform for Recombinant Protein Production in Plants. *The Plant Cell*, *25*(7), 2429–2443. https://doi.org/10.1105/tpc.113.113944

Dusek, J., Plchova, H., Cerovska, N., Poborilova, Z., Navratil, O., Kratochvilova, K., Gunter, C., Jacobs, R., Hitzeroth, I. I., Rybicki, E. P., & Moravec, T. (2020). Extended Set of GoldenBraid Compatible Vectors for Fast Assembly of Multigenic Constructs and Their Use to Create Geminiviral Expression Vectors. *Frontiers in Plant Science*, *11*, 522059. https://doi.org/10.3389/fpls.2020.522059

Eagle, P. A., Orozco, B. M., & Hanley-Bowdoin, L. (1994). A DNA sequence required for geminivirus replication also mediates transcriptional regulation. *The Plant Cell*, *6*(8), 1157–1170. https://doi.org/10.1105/tpc.6.8.1157

Eidenberger, L., Kogelmann, B., & Steinkellner, H. (2023). Plant-based biopharmaceutical engineering. *Nature Reviews Bioengineering*, *1*(6), 426–439. https://doi.org/10.1038/s44222-023-00044-6

Eini, O., Schumann, N., Niessen, M., & Varrelmann, M. (2022). Targeted mutagenesis in plants using Beet curly top virus for efficient delivery of CRISPR/Cas12a components. *New Biotechnology*, *67*, 1–11. https://doi.org/10.1016/j.nbt.2021.12.002

Elliott, D. A., & Brand, A. H. (2008). The GAL4 System: A Versatile System for the Expression of Genes. In C. Dahmann (Ed.), *Drosophila* (Vol. 420, pp. 79–95). Humana Press. https://doi.org/10.1007/978-1-59745-583-1_5

Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, *403*(6767), 335–338. https://doi.org/10.1038/35002125

Engler, C., Gruetzner, R., Kandzia, R., & Marillonnet, S. (2009). Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS ONE*, *4*(5), Article 5. https://doi.org/10.1371/journal.pone.0005553

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J. D. G., Patron, N. J., & Marillonnet, S. (2014). A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synthetic Biology*, *3*(11), Article 11. https://doi.org/10.1021/sb4001504

Eslami, M., Adler, A., Caceres, R. S., Dunn, J. G., Kelley-Loughnane, N., Varaljay, V. A., & Martin, H. G. (2022). Artificial intelligence for synthetic biology. *Communications of the ACM*, *65*(5), 88–97. https://doi.org/10.1145/3500922

Evdokimov, A. G., Pokross, M. E., Egorov, N. S., Zaraisky, A. G., Yampolsky, I. V., Merzlyak, E. M., Shkoporov, A. N., Sander, I., Lukyanov, K. A., & Chudakov, D. M. (2006). Structural basis for the fast maturation of Arthropoda green fluorescent protein. *EMBO Reports*, *7*(10), 1006–1012. https://doi.org/10.1038/sj.embor.7400787

Fenoll, C., Schwarz, J. J., Black, D. M., Schneider, M., & Howell, S. H. (1990). The intergenic region of maize streak virus contains a GC-rich element that activates rightward transcription and binds maize nuclear factors. *Plant Molecular Biology*, *15*(6), 865–877. https://doi.org/10.1007/BF00039426

Ferreira, S. S., Anderson, C. E., & Antunes, M. S. (2023). A logical way to reprogram plants. *Biochemical and Biophysical Research Communications*, *654*, 80–86. https://doi.org/10.1016/j.bbrc.2023.02.080

Fesenko, E., & Edwards, R. (2014). Plant synthetic biology: A new platform for industrial biotechnology. *Journal of Experimental Botany*, *65*(8), 1927–1937. https://doi.org/10.1093/jxb/eru070

Fiallo-Olivé, E., Lett, J.-M., Martin, D. P., Roumagnac, P., Varsani, A., Zerbini, F. M., & Navas-Castillo, J. (2021a). ICTV Virus Taxonomy Profile: Geminiviridae 2021: This article is part of the ICTV Virus Taxonomy Profiles collection. *Journal of General Virology*, *102*(12). https://doi.org/10.1099/jgv.0.001696

Fiallo-Olivé, E., Lett, J.-M., Martin, D. P., Roumagnac, P., Varsani, A., Zerbini, F. M., & Navas-Castillo, J. (2021b). ICTV Virus Taxonomy Profile: Geminiviridae 2021: This article is part of the ICTV Virus Taxonomy Profiles collection. *Journal of General Virology*, *102*(12). https://doi.org/10.1099/jgv.0.001696

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

Fischer, R., Schillberg, S., Hellwig, S., Twyman, R. M., & Drossard, J. (2012). GMP issues for recombinant plant-derived pharmaceutical proteins. *Biotechnology Advances*, *30*(2), 434–439. https://doi.org/10.1016/j.biotechadv.2011.08.007

Folimonova, S. Y., Robertson, C. J., Shilts, T., Folimonov, A. S., Hilf, M. E., Garnsey, S. M., & Dawson, W. O. (2010). Infection with Strains of *Citrus Tristeza Virus* Does Not Exclude Superinfection by Other Strains of the Virus. *Journal of Virology*, *84*(3), 1314–1325. https://doi.org/10.1128/JVI.02075-09

Fondong, V. N. (2013). Geminivirus protein structure and function. *Molecular Plant Pathology*, *14*(6), 635–649. https://doi.org/10.1111/mpp.12032

Forestier, E. C. F., Cording, A. C., Loake, G. J., & Graham, I. A. (2023). An Engineered Heat-Inducible Expression System for the Production of Casbene in Nicotiana benthamiana. *International Journal of Molecular Sciences*, *24*(14), 11425. https://doi.org/10.3390/ijms241411425

Ganguly, S., Ghosh, G., Ghosh, S., Purohit, A., Chaudhuri, R. K., Das, S., & Chakraborti, D. (2020). Plumular meristem transformation system for chickpea: An efficient method to overcome recalcitrant tissue culture responses. *Plant Cell, Tissue and Organ Culture (PCTOC)*, *142*(3), 493–504. https://doi.org/10.1007/s11240-020-01873-8

Gao, X. J., Chong, L. S., Kim, M. S., & Elowitz, M. B. (2018). Programmable protein circuits in living cells. *Science*, *361*(6408), 1252–1258. https://doi.org/10.1126/science.aat5062

Garcia-Perez, E., Diego-Martin, B., Quijano-Rubio, A., Moreno-Giménez, E., Selma, S., Orzaez, D., & Vazquez-Vilar, M. (2022). A copper switch for inducing CRISPR/Cas9-based transcriptional activation tightly regulates gene expression in Nicotiana benthamiana. *BMC Biotechnology*, *22*(1), 12. https://doi.org/10.1186/s12896-022-00741-x

Garcia-Perez, E., Vazquez-Vilar, M., Lozano-Duran, R., & Orzaez, D. (2024). *CuBe: A geminivirus-based copper-regulated expression system suitable for post-harvest activation*. https://doi.org/10.1101/2024.04.30.591823

Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature*, *403*(6767), 339–342. https://doi.org/10.1038/35002131

Garner, K. L. (2021). Principles of synthetic biology. *Essays in Biochemistry*, *65*(5), 791–811. https://doi.org/10.1042/EBC20200059

## References


Gatz, C., & Quail, P. H. (1988). Tn10-encoded tet repressor can regulate an operator-containing plant promoter. *Proceedings of the National Academy of Sciences*, *85*(5), 1394–1397. https://doi.org/10.1073/pnas.85.5.1394

Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C. V., Sánchez-León, S., Baltes, N. J., Starker, C., Barro, F., Gao, C., & Voytas, D. F. (2017). High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR /Cas9. *The Plant Journal*, *89*(6), 1251–1262. https://doi.org/10.1111/tpj.13446

Giraud, G., Stadhouders, R., Conidi, A., Dekkers, D. H. W., Huylebroeck, D., Demmers, J. A. A., Soler, E., & Grosveld, F. G. (2014). NLS-tagging: An alternative strategy to tag nuclear proteins. *Nucleic Acids Research*, *42*(21), e163–e163. https://doi.org/10.1093/nar/gku869

Gleba, Y., Klimyuk, V., & Marillonnet, S. (2007). Viral vectors for the expression of proteins in plants. *Current Opinion in Biotechnology*, *18*(2), 134–141. https://doi.org/10.1016/j.copbio.2007.03.002

Gleba, Y. Y., Tusé, D., & Giritch, A. (2013). Plant Viral Vectors for Delivery by Agrobacterium. In K. Palmer & Y. Gleba (Eds.), *Plant Viral Vectors* (Vol. 375, pp. 155–192). Springer Berlin Heidelberg. https://doi.org/10.1007/82_2013_352

Golubova, D., Tansley, C., Su, H., & Patron, N. J. (2024). Engineering Nicotiana benthamiana as a platform for natural product biosynthesis. *Current Opinion in Plant Biology*, *81*, 102611. https://doi.org/10.1016/j.pbi.2024.102611

Gong, P., Tan, H., Zhao, S., Li, H., Liu, H., Ma, Y., Zhang, X., Rong, J., Fu, X., Lozano-Durán, R., Li, F., & Zhou, X. (2021). Geminiviruses encode additional small proteins with specific subcellular localizations and virulence function. *Nature Communications*, *12*(1), 4278. https://doi.org/10.1038/s41467-021-24617-4

Gordon-Kamm, B., Sardesai, N., Arling, M., Lowe, K., Hoerster, G., Betts, S., & Jones, T. (2019). Using Morphogenic Genes to Improve Recovery and Regeneration of Transgenic Plants. *Plants*, *8*(2), 38. https://doi.org/10.3390/plants8020038

Goulet, C., Khalf, M., Sainsbury, F., D'Aoust, M., & Michaud, D. (2012). A protease activity–depleted environment for heterologous proteins migrating towards the leaf cell apoplast. *Plant Biotechnology Journal*, *10*(1), 83–94. https://doi.org/10.1111/j.1467-7652.2011.00643.x

Granger, C. L., & Cyr, R. J. (2000). Microtubule reorganization in tobacco BY-2 cells stably expressing GFP-MBD. *Planta*, *210*(3), 502–509. https://doi.org/10.1007/s004250050037

Granger, C. L., & Cyr, R. J. (2001). Characterization of the yeast copper-inducible promoter system in Arabidopsis thaliana. *Plant Cell Reports*, *20*(3), 227–234. https://doi.org/10.1007/s002990000308

Guiziou, S. (2024). Biocomputing in plants, from proof of concept to application. *Current Opinion in Biotechnology*, *87*, 103146. https://doi.org/10.1016/j.copbio.2024.103146

Guiziou, S., Maranas, C. J., Chu, J. C., & Nemhauser, J. L. (2023a). An integrase toolbox to record gene-expression during plant development. *Nature Communications*, *14*(1), 1844. https://doi.org/10.1038/s41467-023-37607-5

Guiziou, S., Maranas, C. J., Chu, J. C., & Nemhauser, J. L. (2023b). An integrase toolbox to record gene-expression during plant development. *Nature Communications*, *14*(1), 1844. https://doi.org/10.1038/s41467-023-37607-5

Gurdo, N., Volke, D. C., McCloskey, D., & Nikel, P. I. (2023). Automating the design-build-test-learn cycle towards next-generation bacterial cell factories. *New Biotechnology*, *74*, 1–15. https://doi.org/10.1016/j.nbt.2023.01.002

Gutierrez, C. (1999). Geminivirus DNA replication. *Cellular and Molecular Life Sciences (CMLS)*, *56*(3–4), 313–329. https://doi.org/10.1007/s000180050433

HamediRad, M., Chao, R., Weisberg, S., Lian, J., Sinha, S., & Zhao, H. (2019). Towards a fully automated algorithm driven platform for biosystems design. *Nature Communications*, *10*(1), 5150. https://doi.org/10.1038/s41467-019-13189-z

Hanley-Bowdoin, L., Bejarano, E. R., Robertson, D., & Mansoor, S. (2013). Geminiviruses: Masters at redirecting and reprogramming plant processes. *Nature Reviews Microbiology*, *11*(11), 777–788. https://doi.org/10.1038/nrmicro3117

Hanley-Bowdoin, L., Settlage, S. B., & Robertson, D. (2004). Reprogramming plant gene expression: A prerequisite to geminivirus DNA replication. *Molecular Plant Pathology*, *5*(2), 149–156. https://doi.org/10.1111/j.1364-3703.2004.00214.x

He, Y., Zhang, T., Sun, H., Zhan, H., & Zhao, Y. (2020). A reporter for noninvasively monitoring gene expression and plant transformation. *Horticulture Research*, *7*(1), 152. https://doi.org/10.1038/s41438-020-00390-1

Hefferon, K. (2017). Plant Virus Expression Vectors: A Powerhouse for Global Health. *Biomedicines*, *5*(4), 44. https://doi.org/10.3390/biomedicines5030044

Heim, R., Cubitt, A. B., & Tsien, R. Y. (1995). Improved green fluorescence. *Nature*, *373*(6516), 663–664. https://doi.org/10.1038/373663b0

Hiratsu, K., Matsui, K., Koyama, T., & Ohme-Takagi, M. (2003). Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. *The Plant Journal*, *34*(5), 733–739. https://doi.org/10.1046/j.1365-313X.2003.01759.x

Hirschi, S., Ward, T. R., Meier, W. P., Müller, D. J., & Fotiadis, D. (2022). Synthetic Biology: Bottom-Up Assembly of Molecular Systems. *Chemical Reviews*, *122*(21), 16294–16328. https://doi.org/10.1021/acs.chemrev.2c00339

Huang, Z., Phoolcharoen, W., Lai, H., Piensook, K., Cardineau, G., Zeitlin, L., Whaley, K. J., Arntzen, C. J., Mason, H. S., & Chen, Q. (2010). High-level rapid production of full-size monoclonal antibodies in plants by a single-vector DNA replicon system. *Biotechnology and Bioengineering*, *106*(1), 9–17. https://doi.org/10.1002/bit.22652

Hummel, N. F. C., Zhou, A., Li, B., Markel, K., Ornelas, I. J., & Shih, P. M. (2023). The trans-regulatory landscape of gene networks in plants. *Cell Systems*, *14*(6), 501-511.e4. https://doi.org/10.1016/j.cels.2023.05.002

Hur, J., Buckley, K. J., Lee, S., & Davis, K. R. (2007). Transcriptional Activator Elements for Curtovirus C1 Expression Reside in the 3′ Coding Region of ORF C1. *Molecules and Cells*, *23*(1), 80–87. https://doi.org/10.1016/S1016-8478(23)07392-2

Ikeda, M., & Ohme-Takagi, M. (2009). A Novel Group of Transcriptional Repressors in Arabidopsis. *Plant and Cell Physiology*, *50*(5), 970–975. https://doi.org/10.1093/pcp/pcp048

Iram, A., Dong, Y., & Ignea, C. (2024). Synthetic biology advances towards a bio-based society in the era of artificial intelligence. *Current Opinion in Biotechnology*, *87*, 103143. https://doi.org/10.1016/j.copbio.2024.103143

Jakubiec, A., Sarokina, A., Choinard, S., Vlad, F., Malcuit, I., & Sorokin, A. P. (2021). Replicating minichromosomes as a new tool for plastid genome engineering. *Nature Plants*, *7*(7), 932–941. https://doi.org/10.1038/s41477-021-00940-y

Jedličková, V., Ebrahimi Naghani, S., & Robert, H. S. (2022). On the trail of auxin: Reporters and sensors. *The Plant Cell*, *34*(9), 3200–3213. https://doi.org/10.1093/plcell/koac179

Jefferson, R. A., Kavanagh, T. A., & Bevan, M. W. (1987). GUS fusions: Beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal*, *6*(13), 3901–3907. https://doi.org/10.1002/j.1460-2075.1987.tb02730.x

Jeske, H. (2001). DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. *The EMBO Journal*, *20*(21), 6158–6167. https://doi.org/10.1093/emboj/20.21.6158

## References


Jores, T., Tonnies, J., Wrightsman, T., Buckler, E. S., Cuperus, J. T., Fields, S., & Queitsch, C. (2021). Synthetic promoter designs enabled by a comprehensive analysis of plant core promoters. *Nature Plants*, *7*(6), 842–855. https://doi.org/10.1038/s41477-021-00932-y

Julve, J. M., Gandía, A., Fernández-del-Carmen, A., Sarrion-Perdigones, A., Castelijns, B., Granell, A., & Orzaez, D. (2013). A coat-independent superinfection exclusion rapidly imposed in Nicotiana benthamiana cells by tobacco mosaic virus is not prevented by depletion of the movement protein. *Plant Molecular Biology*, *81*(6), 553–564. https://doi.org/10.1007/s11103-013-0028-1

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kallam, K., Moreno-Giménez, E., Mateos-Fernández, R., Tansley, C., Gianoglio, S., Orzaez, D., & Patron, N. (2023). Tunable control of insect pheromone biosynthesis in *Nicotiana benthamiana*. *Plant Biotechnology Journal*, *21*(7), 1440–1453. https://doi.org/10.1111/pbi.14048

Kang, Y., Outlaw, W. H., Andersen, P. C., & Fiore, G. B. (2007). Guard-cell apoplastic sucrose concentration – a link between leaf photosynthesis and stomatal aperture size in the apoplastic phloem loader *Vicia faba* L. *Plant, Cell & Environment*, *30*(5), 551–558. https://doi.org/10.1111/j.1365-3040.2007.01635.x

Kenyon, L., Tsai, W.-S., Shih, S.-L., & Lee, L.-M. (2014). Emergence and diversity of begomoviruses infecting solanaceous crops in East and Southeast Asia. *Virus Research*, *186*, 104–113. https://doi.org/10.1016/j.virusres.2013.12.026

Khakhar, A., Starker, C. G., Chamness, J. C., Lee, N., Stokke, S., Wang, C., Swanson, R., Rizvi, F., Imaizumi, T., & Voytas, D. F. (2020). Building customizable auto-luminescent luciferase-based reporters in plants. *eLife*, *9*, e52786. https://doi.org/10.7554/eLife.52786

Khalil, A. S., & Collins, J. J. (2010). Synthetic biology: Applications come of age. *Nature Reviews Genetics*, *11*(5), 367–379. https://doi.org/10.1038/nrg2775

Khan, M. A., Herring, G., Zhu, J. Y., Oliva, M., Fourie, E., Johnston, B., Zhang, Z., Potter, J., Pineda, L., Pflueger, J., Swain, T., Pflueger, C., Lloyd, J. P. B., Secco, D., Small, I., Kidd, B. N., & Lister, R. (2024). CRISPRi-based circuits to control gene expression in plants. *Nature Biotechnology*. https://doi.org/10.1038/s41587-024-02236-w

Kim, N.-S., Lee, K.-R., Lee, J., Kil, E.-J., Lee, J., & Lee, S.-K. (2024). High production of recombinant protein using geminivirus-based deconstructed vectors in Nicotiana benthamiana. *Frontiers in Plant Science*, *15*, 1407240. https://doi.org/10.3389/fpls.2024.1407240

Kitano, S., Lin, C., Foo, J. L., & Chang, M. W. (2023). Synthetic biology: Learning the way toward high-precision biological design. *PLOS Biology*, *21*(4), e3002116. https://doi.org/10.1371/journal.pbio.3002116

Knapp, D. J. H. F., Michaels, Y. S., Jamilly, M., Ferry, Q. R. V., Barbosa, H., Milne, T. A., & Fulga, T. A. (2019). Decoupling tRNA promoter and processing activities enables specific Pol-II Cas9 guide RNA expression. *Nature Communications*, *10*(1), 1490. https://doi.org/10.1038/s41467-019-09148-3

Knapp, S. (2019). People and plants: The unbreakable bond. *PLANTS, PEOPLE, PLANET*, *1*(1), 20–26. https://doi.org/10.1002/ppp3.4

Kocaoglan, E. G., Andreou, A., Nirkko, J., Ochoa-Villarreal, M., Loake, G., & Nakayama, N. (2024). *Mobius Assembly for Plant Systems highlights promoter-coding sequences-terminator interaction in gene regulation*. https://doi.org/10.1101/2024.07.10.602858

Koide, T., Lee Pang, W., & Baliga, N. S. (2009). The role of predictive modelling in rationally re-engineering biological systems. *Nature Reviews Microbiology*, *7*(4), 297–305. https://doi.org/10.1038/nrmicro2107

Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., Hsu, P. D., Habib, N., Gootenberg, J. S., Nishimasu, H., Nureki, O., & Zhang, F. (2015a). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, *517*(7536), 583–588. https://doi.org/10.1038/nature14136

Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., Hsu, P. D., Habib, N., Gootenberg, J. S., Nishimasu, H., Nureki, O., & Zhang, F. (2015b). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. https://doi.org/10.1038/nature14136

Koo, J. C., Asurmendi, S., Bick, J., Woodford-Thomas, T., & Beachy, R. N. (2004). Ecdysone agonist-inducible expression of a coat protein gene from tobacco mosaic virus confers viral resistance in transgenic *Arabidopsis*. *The Plant Journal*, *37*(3), 439–448. https://doi.org/10.1046/j.1365-313X.2003.01869.x

Kortemme, T. (2024). De novo protein design—From new structures to programmable functions. *Cell*, *187*(3), 526–544. https://doi.org/10.1016/j.cell.2023.12.028

Koukara, J., & Papadopoulou, K. K. (2023). Advances in plant synthetic biology approaches to control expression of gene circuits. *Biochemical and Biophysical Research Communications*, *654*, 55–61. https://doi.org/10.1016/j.bbrc.2023.02.061

Kovalskaya, N., Foster-Frey, J., Donovan, D. M., Bauchan, G., & Hammond, R. W. (2016). Antimicrobial Activity of Bacteriophage Endolysin Produced in Nicotiana benthamiana Plants. *Journal of Microbiology and Biotechnology*, *26*(1), 160–170. https://doi.org/10.4014/jmb.1505.05060

Ku, H.-M., Tan, C.-W., Su, Y.-S., Chiu, C.-Y., Chen, C.-T., & Jan, F.-J. (2012). The effect of water deficit and excess copper on proline metabolism in Nicotiana benthamiana. *Biologia Plantarum*, *56*(2), 337–343. https://doi.org/10.1007/s10535-012-0095-1

Kumar, Pandita, & Sharma. (2021). Copper bioavailability, uptake, toxicity and tolerance in plants: A comprehensive review. *Chemosphere*, *262*, 127810. https://doi.org/10.1016/j.chemosphere.2020.127810

Kunii, A., Hara, Y., Takenaga, M., Hattori, N., Fukazawa, T., Ushijima, T., Yamamoto, T., & Sakuma, T. (2018). Three-Component Repurposed Technology for Enhanced Expression: Highly Accumulable Transcriptional Activators via Branched Tag Arrays. *The CRISPR Journal*, *1*(5), 337–347. https://doi.org/10.1089/crispr.2018.0009

Küpper, H., & Andresen, E. (2016). Mechanisms of metal toxicity in plants. *Metallomics*, *8*(3), 269–285. https://doi.org/10.1039/C5MT00244C

Kusunoki, K., & Yamamoto, Y. Y. (2017). Plant Promoter Database (PPDB). In A. D. J. Van Dijk (Ed.), *Plant Genomics Databases* (Vol. 1533, pp. 299–314). Springer New York. https://doi.org/10.1007/978-1-4939-6658-5_18

Lai, H., He, J., Engle, M., Diamond, M. S., & Chen, Q. (2012). Robust production of virus-like particles and monoclonal antibodies with geminiviral replicon vectors in lettuce. *Plant Biotechnology Journal*, *10*(1), 95–104. https://doi.org/10.1111/j.1467-7652.2011.00649.x

Lamichhane, J. R., Osdaghi, E., Behlau, F., Köhl, J., Jones, J. B., & Aubertot, J.-N. (2018). Thirteen decades of antimicrobial copper compounds applied in agriculture. A review. *Agronomy for Sustainable Development*, *38*(3), 28. https://doi.org/10.1007/s13593-018-0503-9

Lazarowitz, S. G., & Shepherd, R. J. (1992). Geminiviruses: Genome structure and gene function. *Critical Reviews in Plant Sciences*, *11*(4), 327–349. https://doi.org/10.1080/07352689209382350

Leduc, S. (1912). *La biologie synthétique*. Poinat.

Lee, L.-Y., & Gelvin, S. B. (2008). T-DNA Binary Vectors and Systems. *Plant Physiology*, *146*(2), 325–332. https://doi.org/10.1104/pp.107.113001

Lefeuvre, P., & Moriones, E. (2015). Recombination as a motor of host switches and virus emergence: Geminiviruses as case studies. *Current Opinion in Virology*, *10*, 14–19. https://doi.org/10.1016/j.coviro.2014.12.005

Lescot, M. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, *30*(1), 325–327. https://doi.org/10.1093/nar/30.1.325

Li, J., Blue, R., Zeitler, B., Strange, T. L., Pearl, J. R., Huizinga, D. H., Evans, S., Gregory, P. D., Urnov, F. D., & Petolino, J. F. (2013). Activation domains for controlling plant gene expression using designed transcription factors. *Plant Biotechnology Journal*, *11*(6), 671–680. https://doi.org/10.1111/pbi.12057

Li, R., Jia, X., & Mao, X. (2005). Ethanol-inducible gene expression system and its applications in plant functional genomics. *Plant Science*, *169*(3), 463–469. https://doi.org/10.1016/j.plantsci.2005.04.006

Li, Y. (2015). Split-inteins and their bioapplications. *Biotechnology Letters*, *37*(11), 2121–2137. https://doi.org/10.1007/s10529-015-1905-2

Li, Z., Wang, Y., Liu, J., Rawding, P., Bu, J., Hong, S., & Hu, Q. (2021). Chemically and Biologically Engineered Bacteria-Based Delivery Systems for Emerging Diagnosis and Advanced Therapy. *Advanced Materials*, *33*(38), 2102580. https://doi.org/10.1002/adma.202102580

Li, Z., Zhang, D., Xiong, X., Yan, B., Xie, W., Sheen, J., & Li, J.-F. (2017). A potent Cas9-derived gene activator for plant and mammalian cells. *Nature Plants*, *3*(12), 930–936. https://doi.org/10.1038/s41477-017-0046-0

Liao, C.-Y., Smet, W., Brunoud, G., Yoshida, S., Vernoux, T., & Weijers, D. (2015). Reporters for sensitive and quantitative measurement of auxin response. *Nature Methods*, *12*(3), 207–210. https://doi.org/10.1038/nmeth.3279

Lin, Z., Ali, M. M., Yi, X., Zhang, L., & Wang, S. (2023). Fast and High-Efficiency Synthesis of Capsanthin in Pepper by Transient Expression of Geminivirus. *International Journal of Molecular Sciences*, *24*(19), 15008. https://doi.org/10.3390/ijms241915008

Lindbo, J. A. (2007). TRBO: A High-Efficiency Tobacco Mosaic Virus RNA-Based Overexpression Vector. *Plant Physiology*, *145*(4), 1232–1240. https://doi.org/10.1104/pp.107.106377

Liu, C., Yu, H., Zhang, B., Liu, S., Liu, C., Li, F., & Song, H. (2022). Engineering whole-cell microbial biosensors: Design principles and applications in monitoring and treatment of heavy metals and organic pollutants. *Biotechnology Advances*, *60*, 108019. https://doi.org/10.1016/j.biotechadv.2022.108019

Liu, J., Li, H., Hong, C., Lu, W., Zhang, W., & Gao, H. (2024). Quantitative *RUBY* reporter assay for gene regulation analysis. *Plant, Cell & Environment*, pce.14947. https://doi.org/10.1111/pce.14947

Liu, L., Davies, J. W., & Stanley, J. (1998). Mutational analysis of bean yellow dwarf virus, a geminivirus of the genus Mastrevirus that is adapted to dicotyledonous plants. *Journal of General Virology*, *79*(9), 2265–2274. https://doi.org/10.1099/0022-1317-79-9-2265

Liu, R., Long, Q., Zou, X., Wang, Y., & Pei, Y. (2021). DNA methylation occurring in Cre-expressing cells inhibits loxP recombination and silences loxP-sandwiched genes. *New Phytologist*, *231*(1), 210–224. https://doi.org/10.1111/nph.17353

Liu, W., & Stewart, C. N. (2015a). Plant synthetic biology. *Trends in Plant Science*, *20*(5), 309–317. https://doi.org/10.1016/j.tplants.2015.02.004

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

Liu, W., & Stewart, N. (2015b). Plant synthetic biology. *Trends in Plant Science*, *20*(5), 309–317. https://doi.org/10.1016/j.tplants.2015.02.004

Liu, Y., Jin, W., Wang, L., & Wang, X. (2014). Replication-associated proteins encoded by Wheat dwarf virus act as RNA silencing suppressors. *Virus Research*, *190*, 34–39. https://doi.org/10.1016/j.virusres.2014.06.014

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif.)*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262

Lloyd, J. P. B., Ly, F., Gong, P., Pflueger, J., Swain, T., Pflueger, C., Fourie, E., Khan, M. A., Kidd, B. N., & Lister, R. (2022a). Synthetic memory circuits for stable cell reprogramming in plants. *Nature Biotechnology*, *40*(12), 1862–1872. https://doi.org/10.1038/s41587-022-01383-2

Lloyd, J. P. B., Ly, F., Gong, P., Pflueger, J., Swain, T., Pflueger, C., Fourie, E., Khan, M. A., Kidd, B. N., & Lister, R. (2022b). Synthetic memory circuits for stable cell reprogramming in plants. *Nature Biotechnology*, *40*(12), 1862–1872. https://doi.org/10.1038/s41587-022-01383-2

Lozano-Durán, R. (2016). Geminiviruses for biotechnology: The art of parasite taming. *New Phytologist*, *210*(1), 58–64. https://doi.org/10.1111/nph.13564

Lozano-Juste, J., Infantes, L., Garcia-Maquilon, I., Ruiz-Partida, R., Merilo, E., Benavente, J. L., Velazquez-Campoy, A., Coego, A., Bono, M., Forment, J., Pampín, B., Destito, P., Monteiro, A., Rodríguez, R., Cruces, J., Rodriguez, P. L., & Albert, A. (2023). Structure-guided engineering of a receptor-agonist pair for inducible activation of the ABA adaptive response to drought. *Science Advances*, *9*(10), eade9948. https://doi.org/10.1126/sciadv.ade9948

Luna, A. P., & Lozano-Durán, R. (2020). Geminivirus-Encoded Proteins: Not All Positional Homologs Are Made Equal. *Frontiers in Microbiology*, *11*, 878. https://doi.org/10.3389/fmicb.2020.00878

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

Ma, L., Lukasik, E., Gawehns, F., & Takken, F. L. W. (2012). The Use of Agroinfiltration for Transient Expression of Plant Resistance and Fungal Effector Proteins in Nicotiana benthamiana Leaves. In M. D. Bolton & B. P. H. J. Thomma (Eds.), *Plant Fungal Pathogens* (Vol. 835, pp. 61–74). Humana Press. https://doi.org/10.1007/978-1-61779-501-5_4

Ma, Y., Wang, X. R., Li, T., Zhang, J., Gao, J., & Sun, Z. Y. (2021). Hydrogen and ethanol: Production, storage, and transportation. *International Journal of Hydrogen Energy*, *46*(54), 27330–27348. https://doi.org/10.1016/j.ijhydene.2021.06.027

Maher, M. F., Nasti, R. A., Vollbrecht, M., Starker, C. G., Clark, M. D., & Voytas, D. F. (2020). Plant gene editing through de novo induction of meristems. *Nature Biotechnology*, *38*(1), 84–89. https://doi.org/10.1038/s41587-019-0337-2

Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., Charpentier, E., Cheng, D., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Scott, D., Shah, S. A., Siksnys, V., Terns, M. P., Venclovas, Č., White, M. F., Yakunin, A. F., … Koonin, E. V. (2020). Evolutionary classification of CRISPR–Cas systems: A burst of class 2 and derived variants. *Nature Reviews Microbiology*, *18*(2), 67–83. https://doi.org/10.1038/s41579-019-0299-x

Mali, P., Esvelt, K. M., & Church, G. M. (2013). Cas9 as a versatile tool for engineering biology. *Nature Methods*, *10*(10), 957–963. https://doi.org/10.1038/nmeth.2649

Malla, A., Rosales-Mendoza, S., Phoolcharoen, W., & Vimolmangkang, S. (2021). Efficient Transient Expression of Recombinant Proteins Using DNA Viral Vectors in Freshwater Microalgal Species. *Frontiers in Plant Science*, *12*, 650820. https://doi.org/10.3389/fpls.2021.650820

Mandegar, Huebsch, & Frolov. (2016). CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell Stem Cell*, *18*(4), 541–553. https://doi.org/10.1016/j.stem.2016.01.022

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

Marillonnet, S., & Grützner, R. (2020). Synthetic DNA Assembly Using Golden Gate Cloning and the Hierarchical Modular Cloning Pipeline. *Current Protocols in Molecular Biology*, *130*(1), e115. https://doi.org/10.1002/cpmb.115

Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V., & Gleba, Y. (2005). Systemic Agrobacterium tumefaciens–mediated transfection of viral replicons for efficient transient expression in plants. *Nature Biotechnology*, *23*(6), 718–723. https://doi.org/10.1038/nbt1094

Markets & Markets. (2023). *Synthetic Biology Market by Tools (Oligonucleotides, Enzymes, Synthetic Cells), Technology (Genome Engineering, Bioinformatics), Applications (Tissue Regeneration, Biofuel, Food, Agriculture, Consumer Care, Environmental) & Region—Global Forecast to 2027* (Report Code: BT 2910). https://www.marketsandmarkets.com/Market-Reports/synthetic-biology-market-889.html

Martin, L. B. B., Kikuchi, S., Rejzek, M., Owen, C., Reed, J., Orme, A., Misra, R. C., El-Demerdash, A., Hill, L., Hodgson, H., Liu, Y., Keasling, J. D., Field, R. A., Truman, A. W., & Osbourn, A. (2024). Complete biosynthesis of the potent vaccine adjuvant QS-21. *Nature Chemical Biology*, *20*(4), 493–502. https://doi.org/10.1038/s41589-023-01538-5

Matzat, L. H., & Lei, E. P. (2014). Surviving an identity crisis: A revised view of chromatin insulators in the genomics era. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1839*(3), 203–214. https://doi.org/10.1016/j.bbagrm.2013.10.007

Medford, J. I., & Prasad, A. (2016). Towards programmable plant genetic circuits. *The Plant Journal*, *87*(1), 139–148. https://doi.org/10.1111/tpj.13235

Mett, V. L., Lochhead, L. P., & Reynolds, P. H. (1993). Copper-controllable gene expression system for whole plants. *Proceedings of the National Academy of Sciences*, *90*(10), 4567–4571.

Mitiouchkina, T., Mishin, A. S., Somermeyer, L. G., Markina, N. M., Chepurnyh, T. V., Guglya, E. B., Karataeva, T. A., Palkina, K. A., Shakhova, E. S., Fakhranurova, L. I., Chekova, S. V., Tsarkova, A. S., Golubev, Y. V., Negrebetsky, V. V., Dolgushin, S. A., Shalaev, P. V., Shlykov, D., Melnik, O. A., Shipunova, V. O., … Sarkisyan, K. S. (2020). Plants with genetically encoded autoluminescence. *Nature Biotechnology*, *38*(8), Article 8. https://doi.org/10.1038/s41587-020-0500-9

Mohamed, R., Meilan, R., & Strauss, S. H. (2001). Complex behavior of a copper-inducible gene expression system in transgenic poplar. *International Journal of Forest Genetics*. https://agris.fao.org/agris-search/search.do?recordID=SK2001000875

Mojica, F. J. M., Diez-Villaseñor, C., Garcia-Martinez, J., & Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*, *60*(2), 174–182. https://doi.org/10.1007/s00239-004-0046-3

Mojica, F. J. M., & Montoliu, L. (2016). On the Origin of CRISPR-Cas Technology: From Prokaryotes to Mammals. *Trends in Microbiology*, *24*(10), 811–820. https://doi.org/10.1016/j.tim.2016.06.005

Moon, T. S. (2023). Seven governing principles in biology. *Frontiers in Synthetic Biology*, *1*, 1296513. https://doi.org/10.3389/fsybi.2023.1296513

Mor, T. S., Moon, Y., Palmer, K. E., & Mason, H. S. (2003). Geminivirus vectors for high-level expression of foreign proteins in plant cells. *Biotechnology and Bioengineering*, *81*(4), 430–437. https://doi.org/10.1002/bit.10483

Moradpour, M., & Abdulah, S. N. A. (2020). CRISPR/dCas9 platforms in plants: Strategies and applications beyond genome editing. *Plant Biotechnology Journal*, *18*(1), 32–44. https://doi.org/10.1111/pbi.13232

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

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

Morilla, G., Castillo, A. G., Preiss, W., Jeske, H., & Bejarano, E. R. (2006). A Versatile Transreplication-Based System To Identify Cellular Proteins Involved in Geminivirus Replication. *Journal of Virology*, *80*(7), 3624–3633. https://doi.org/10.1128/JVI.80.7.3624-3633.2006

Mortimer, C. L., Dugdale, B., & Dale, J. L. (2015). Updates in inducible transgene expression using viral vectors: From transient to stable expression. *Current Opinion in Biotechnology*, *32*, 85–92. https://doi.org/10.1016/j.copbio.2014.11.009

Munnelly, K. (2013). Engineering for the 21st Century: Synthetic Biology. *ACS Synthetic Biology*, *2*(5), 213–215. https://doi.org/10.1021/sb400039g

Nagar, S., Hanley-Bowdoin, L., & Robertson, D. (2002). Host DNA Replication Is Induced by Geminivirus Infection of Differentiated Plant Cells. *The Plant Cell*, *14*(12), 2995–3007. https://doi.org/10.1105/tpc.005777

Nandagopal, N., & Elowitz, M. B. (2011). Synthetic Biology: Integrated Gene Circuits. *Science*, *333*(6047), Article 6047. https://doi.org/10.1126/science.1207084

Navas-Castillo, J., Fiallo-Olivé, E., & Sánchez-Campos, S. (2011). Emerging Virus Diseases Transmitted by Whiteflies. *Annual Review of Phytopathology*, *49*(1), 219–248. https://doi.org/10.1146/annurev-phyto-072910-095235

Neubauer, M., Vollen, K., Ascencio-Ibanez, J. T., Hanley-Bowdoin, L., Stepanova, A. N., & Alonso, J. M. (2024). *Development of modular geminivirus-based vectors for high cargo expression and gene targeting in plants*. https://doi.org/10.1101/2024.06.29.601216

Nnachi, R. C., Sui, N., Ke, B., Luo, Z., Bhalla, N., He, D., & Yang, Z. (2022). Biosensors for rapid detection of bacterial pathogens in water, food and environment. *Environment International*, *166*, 107357. https://doi.org/10.1016/j.envint.2022.107357

Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H., & Ohme-Takagi, M. (2001). Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *The Plant Cell*, *13*(8), 1959–1968. PubMed. https://doi.org/10.1105/tpc.010127

Olsen, K. M., Lea, U. S., Slimestad, R., Verheul, M., & Lillo, C. (2008). Differential expression of four Arabidopsis PAL genes; PAL1 and PAL2 have functional specialization in abiotic environmental-triggered flavonoid synthesis. *Journal of Plant Physiology*, *165*(14), 1491–1499. https://doi.org/10.1016/j.jplph.2007.11.005

Omelina, E. S., Yushkova, A. A., Motorina, D. M., Volegov, G. A., Kozhevnikova, E. N., & Pindyurin, A. V. (2022). Optogenetic and Chemical Induction Systems for Regulation of Transgene Expression in Plants: Use in Basic and Applied Research. *International Journal of Molecular Sciences*, *23*(3), 1737. https://doi.org/10.3390/ijms23031737

Opgenorth, P., Costello, Z., Okada, T., Goyal, G., Chen, Y., Gin, J., Benites, V., De Raad, M., Northen, T. R., Deng, K., Deutsch, S., Baidoo, E. E. K., Petzold, C. J., Hillson, N. J., Garcia Martin, H., & Beller, H. R. (2019). Lessons from Two Design–Build–Test–Learn Cycles of Dodecanol Production in *Escherichia coli* Aided by Machine Learning. *ACS Synthetic Biology*, *8*(6), 1337–1351. https://doi.org/10.1021/acssynbio.9b00020

Ordiz, M. I., Barbas, C. F., & Beachy, R. N. (2002). Regulation of transgene expression in plants with polydactyl zinc finger transcription factors. *Proceedings of the National Academy of Sciences*, *99*(20), 13290–13295. https://doi.org/10.1073/pnas.202471899

Paine, J. A., Shipton, C. A., Chaggar, S., Howells, R. M., Kennedy, M. J., Vernon, G., Wright, S. Y., Hinchliffe, E., Adams, J. L., Silverstone, A. L., & Drake, R. (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nature Biotechnology*, *23*(4), 482–487. https://doi.org/10.1038/nbt1082

Palkina, K. A., Karataeva, T. A., Perfilov, M. M., Fakhranurova, L. I., Markina, N. M., Somermeyer, L. G., Garcia-Perez, E., Vazquez-Vilar, M., Rodriguez-Rodriguez, M., Vazquez-Vilriales, V., Shakhova, E. S., Mitiouchkina, T., Belozerova, O. A., Kovalchuk, S. I., Alekberova, A., Malyshevskaia, A. K., Bugaeva, E. N., Guglya, E. B., Balakireva, A., … Sarkisyan, K. S. (2024). A hybrid pathway for self-sustained luminescence. *Science Advances*, *10*(10), eadk1992. https://doi.org/10.1126/sciadv.adk1992

Pan, C., Wu, X., Markel, K., Malzahn, A. A., Kundagrami, N., Sretenovic, S., Zhang, Y., Cheng, Y., Shih, P. M., & Qi, Y. (2021). CRISPR–Act3.0 for highly efficient multiplexed gene activation in plants. *Nature Plants*, *7*(7), 942–953. https://doi.org/10.1038/s41477-021-00953-7

Park, J.-J., Dempewolf, E., Zhang, W., & Wang, Z.-Y. (2017). RNA-guided transcriptional activation via CRISPR/dCas9 mimics overexpression phenotypes in Arabidopsis. *PLOS ONE*, *12*(6), e0179410. https://doi.org/10.1371/journal.pone.0179410

Pasin, F., Bedoya, L. C., Bernabé-Orts, J. M., Gallo, A., Simón-Mateo, C., Orzaez, D., & García, J. A. (2017a). Multiple T-DNA Delivery to Plants Using Novel Mini Binary Vectors with Compatible Replication Origins. *ACS Synthetic Biology*, *6*(10), 1962–1968. https://doi.org/10.1021/acssynbio.6b00354

Pasin, F., Bedoya, L. C., Bernabé-Orts, J. M., Gallo, A., Simón-Mateo, C., Orzaez, D., & García, J. A. (2017b). Multiple T-DNA Delivery to Plants Using Novel Mini Binary Vectors with Compatible Replication Origins. *ACS Synthetic Biology*, *6*(10), 1962–1968. https://doi.org/10.1021/acssynbio.6b00354

Patel, P., Kumar Jatav, P., Jain, R., Gupta, S., Kothari, S., & Kachhwaha, S. (2021). Humidity induced opening of stomata leads to enhanced uptake of copper nanoparticles in Triticum aestivum L. *Materials Today: Proceedings*, *43*, 3191–3196. https://doi.org/10.1016/j.matpr.2021.01.712

Patil, M., & Dhar, P. K. (2015). A Brief Introduction to Synthetic Biology. In V. Singh & P. K. Dhar (Eds.), *Systems and Synthetic Biology* (pp. 229–240). Springer Netherlands. https://doi.org/10.1007/978-94-017-9514-2_12

Patil, P. L., Gharat, S. K., Jadhav, K. R., & Kadam, V. J. (2023). Engineered Bacteria: General Overview as Therapeutic Agent and a NovelDrug Delivery System. *Current Pharmaceutical Biotechnology*, *24*(11), 1351–1364. https://doi.org/10.2174/1389201024666221220113517

Patron, N. J., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., Raitskin, O., Leveau, A., Farré, G., Rogers, C., Smith, A., Hibberd, J., Webb, A. A. R., Locke, J., Schornack, S., Ajioka, J., Baulcombe, D. C., Zipfel, C., Kamoun, S., … Haseloff, J. (2015). Standards for plant synthetic biology: A common syntax for exchange of DNA parts. *New Phytologist*, *208*(1), 13–19. https://doi.org/10.1111/nph.13532

Paul, M., Reljic, R., Klein, K., Drake, P. M., Van Dolleweerd, C., Pabst, M., Windwarder, M., Arcalis, E., Stoger, E., Altmann, F., Cosgrove, C., Bartolf, A., Baden, S., & Ma, J. K.-C. (2014). Characterization of a plant-produced recombinant human secretory IgA with broad neutralizing activity against HIV. *mAbs*, *6*(6), 1585–1597. https://doi.org/10.4161/mabs.36336

Peretó, J. (2016). Erasing Borders: A Brief Chronicle of Early Synthetic Biology. *Journal of Molecular Evolution*, *83*(5–6), 176–183. https://doi.org/10.1007/s00239-016-9774-4

Pérez-González, A., & Caro, E. (2016). Hindrances to the Efficient and Stable Expression of Transgenes in Plant Synthetic Biology Approaches. In S. Singh (Ed.), *Systems Biology Application in Synthetic Biology* (pp. 79–89). Springer India. https://doi.org/10.1007/978-81-322-2809-7_7

Pérez-González, A., & Caro, E. (2019a). Benefits of using genomic insulators flanking transgenes to increase expression and avoid positional effects. *Scientific Reports*, *9*(1), 8474. https://doi.org/10.1038/s41598-019-44836-6

Pérez-González, A., & Caro, E. (2019b). Benefits of using genomic insulators flanking transgenes to increase expression and avoid positional effects. *Scientific Reports*, *9*(1), 8474. https://doi.org/10.1038/s41598-019-44836-6

Persad, R., Reuter, D. N., Dice, L. T., Nguyen, M.-A., Rigoulot, S. B., Layton, J. S., Schmid, M. J., Poindexter, M. R., Occhialini, A., Stewart, C. N., & Lenaghan, S. C. (2020). The Q-System as a Synthetic Transcriptional Regulator in Plants. *Frontiers in Plant Science*, *11*, 245. https://doi.org/10.3389/fpls.2020.00245

Peyret, H., & Lomonossoff, G. P. (2013). The pEAQ vector series: The easy and quick way to produce recombinant proteins in plants. *Plant Molecular Biology*, *83*(1–2), 51–58. https://doi.org/10.1007/s11103-013-0036-1

Peyret, H., & Lomonossoff, G. P. (2015). When plant virology met *Agrobacterium*: The rise of the deconstructed clones. *Plant Biotechnology Journal*, *13*(8), 1121–1135. https://doi.org/10.1111/pbi.12412

Peyret, H., Brown, J. K., & Lomonossoff, G. P. (2019). Improving plant transient expression through the rational design of synthetic 5′ and 3′ untranslated regions. *Plant methods*, *15*, 1-13. https://doi.org/10.1186/s13007-019-0494-9

Peyret, H., Ponndorf, D., Meshcheriakova, Y., Richardson, J., & Lomonossoff, G. P. (2020). Covalent protein display on Hepatitis B core-like particles in plants through the in vivo use of the SpyTag/SpyCatcher system. *Scientific Reports*, *10*(1), 17095. https://doi.org/10.1038/s41598-020-74105-w

Pfotenhauer, A. C., Occhialini, A., Nguyen, M.-A., Scott, H., Dice, L. T., Harbison, S. A., Li, L., Reuter, D. N., Schimel, T. M., Stewart, C. N., Beal, J., & Lenaghan, S. C. (2022). Building the Plant SynBio Toolbox through Combinatorial Analysis of DNA Regulatory Elements. *ACS Synthetic Biology*, *11*(8), 2741–2755. https://doi.org/10.1021/acssynbio.2c00147

Pfotenhauer, A. C., Reuter, D. N., Clark, M., Harbison, S. A., Schimel, T. M., Stewart, C. N., & Lenaghan, S. C. (2024). Development of new binary expression systems for plant synthetic biology. *Plant Cell Reports*, *43*(1), 22. https://doi.org/10.1007/s00299-023-03100-y

Pickar-Oliver, A., & Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. *Nature Reviews Molecular Cell Biology*, *20*(8), 490–507. https://doi.org/10.1038/s41580-019-0131-5

Piedra-Aguilera, Á., Jiao, C., Luna, A. P., Villanueva, F., Dabad, M., Esteve-Codina, A., Díaz-Pendón, J. A., Fei, Z., Bejarano, E. R., & Castillo, A. G. (2019). Integrated single-base resolution maps of transcriptome, sRNAome and methylome of Tomato yellow leaf curl virus (TYLCV) in tomato. *Scientific Reports*, *9*(1), 2863. https://doi.org/10.1038/s41598-019-39239-6

Pina, A. S., Batalha, Í. L., Dias, A. M. G. C., & Roque, A. C. A. (2021). Affinity Tags in Protein Purification and Peptide Enrichment: An Overview. In N. E. Labrou (Ed.), *Protein Downstream Processing* (Vol. 2178, pp. 107–132). Springer US. https://doi.org/10.1007/978-1-0716-0775-6_10

Piskacek, M., Havelka, M., Rezacova, M., & Knight, A. (2016). The 9aaTAD Transactivation Domains: From Gal4 to p53. *PLOS ONE*, *11*(9), e0162842. https://doi.org/10.1371/journal.pone.0162842

Plahar, H. A., Rich, T. N., Lane, S. D., Morrell, W. C., Springthorpe, L., Nnadi, O., Aravina, E., Dai, T., Fero, M. J., Hillson, N. J., & Petzold, C. J. (2021). BioParts—A Biological Parts Search Portal and Updates to the ICE Parts Registry Software Platform. *ACS Synthetic Biology*, *10*(10), 2649–2660. https://doi.org/10.1021/acssynbio.1c00263

Pouvreau, B., Blundell, C., Vohra, H., Zwart, A. B., Arndell, T., Singh, S., & Vanhercke, T. (2020). A Versatile High Throughput Screening Platform for Plant Metabolic Engineering Highlights the Major Role of ABI3 in Lipid Metabolism Regulation. *Frontiers in Plant Science*, *11*, 288. https://doi.org/10.3389/fpls.2020.00288

Printz, B., Lutts, S., Hausman, J.-F., & Sergeant, K. (2016). Copper Trafficking in Plants and Its Implication on Cell Wall Dynamics. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.00601

Qi, H., Blanchard, A., & Lu, T. (2013). Engineered genetic information processing circuits. *WIREs Systems Biology and Medicine*, *5*(3), 273–287. https://doi.org/10.1002/wsbm.1216

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, *152*(5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

Qiu, Y., Hu, J., & Wei, G.-W. (2021). Cluster learning-assisted directed evolution. *Nature Computational Science*, *1*(12), 809–818. https://doi.org/10.1038/s43588-021-00168-y

Radivojević, T., Costello, Z., Workman, K., & Garcia Martin, H. (2020). A machine learning Automated Recommendation Tool for synthetic biology. *Nature Communications*, *11*(1), 4879. https://doi.org/10.1038/s41467-020-18008-4

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Ranawaka, B., An, J., Lorenc, M. T., Jung, H., Sulli, M., Aprea, G., Roden, S., Llaca, V., Hayashi, S., Asadyar, L., LeBlanc, Z., Ahmed, Z., Naim, F., De Campos, S. B., Cooper, T., De Felippes, F. F., Dong, P., Zhong, S., Garcia-Carpintero, V., … Waterhouse, P. M. (2023). A multi-omic Nicotiana benthamiana resource for fundamental research and biotechnology. *Nature Plants*, *9*(9), 1558–1571. https://doi.org/10.1038/s41477-023-01489-8

Rattanapisit, K., Bulaon, C. J. I., Strasser, R., Sun, H., & Phoolcharoen, W. (2023). In vitro and in vivo studies of plant-produced Atezolizumab as a potential immunotherapeutic antibody. *Scientific Reports*, *13*(1), 14146. https://doi.org/10.1038/s41598-023-41510-w

Regnard, G. L., Halley-Stott, R. P., Tanzer, F. L., Hitzeroth, I. I., & Rybicki, E. P. (2010). High level protein expression in plants through the use of a novel autonomously replicating geminivirus shuttle vector. *Plant Biotechnology Journal*, *8*(1), 38–46. https://doi.org/10.1111/j.1467-7652.2009.00462.x

Rodríguez-Negrete, E., Lozano-Durán, R., Piedra-Aguilera, A., Cruzado, L., Bejarano, E. R., & Castillo, A. G. (2013). Geminivirus R ep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. *New Phytologist*, *199*(2), 464–475. https://doi.org/10.1111/nph.12286

Romsuk, J., Yasumoto, S., Fukushima, E. O., Miura, K., Muranaka, T., & Seki, H. (2022). High-yield bioactive triterpenoid production by heterologous expression in Nicotiana benthamiana using the Tsukuba system. *Frontiers in Plant Science*, *13*. https://www.frontiersin.org/articles/10.3389/fpls.2022.991909

Roslan, H. A., Salter, M. G., Wood, C. D., White, M. R. H., Croft, K. P., Robson, F., Coupland, G., Doonan, J., Laufs, P., Tomsett, A. B., & Caddick, M. X. (2001). Characterization of the ethanol-inducible alc gene-expression system in Arabidopsis thaliana. *The Plant Journal*, *28*(2), 225–235. https://doi.org/10.1046/j.1365-313X.2001.01146.x

Ruhel, R., & Chakraborty, S. (2019). Multifunctional roles of geminivirus encoded replication initiator protein. *VirusDisease*, *30*(1), 66–73. https://doi.org/10.1007/s13337-018-0458-0

Saijo, T., & Nagasawa, A. (2014a). Development of a tightly regulated and highly responsive copper-inducible gene expression system and its application to control of flowering time. *Plant Cell Reports*, *33*(1), 47–59. https://doi.org/10.1007/s00299-013-1511-5

Saijo, T., & Nagasawa, A. (2014b). Development of a tightly regulated and highly responsive copper-inducible gene expression system and its application to control of flowering time. *Plant Cell Reports*, *33*(1), 47–59. https://doi.org/10.1007/s00299-013-1511-5

Sainsbury, F., & Lomonossoff, G. P. (2008). Extremely High-Level and Rapid Transient Protein Production in Plants without the Use of Viral Replication. *Plant Physiology*, *148*(3), 1212–1218. https://doi.org/10.1104/pp.108.126284

Sainsbury, F., Thuenemann, E. C., & Lomonossoff, G. P. (2009). pEAQ: Versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnology Journal*, *7*(7), 682–693. https://doi.org/10.1111/j.1467-7652.2009.00434.x

Salazar-González, J. A., Bañuelos-Hernández, B., & Rosales-Mendoza, S. (2015). Current status of viral expression systems in plants and perspectives for oral vaccines development. *Plant Molecular Biology*, *87*(3), 203–217. https://doi.org/10.1007/s11103-014-0279-5

Sarrion-Perdigones, A., Falconi, E. E., Zandalinas, S. I., Juárez, P., Fernández-del-Carmen, A., Granell, A., & Orzaez, D. (2011). GoldenBraid: An iterative cloning system for standardized assembly of reusable genetic modules. *PloS One*, *6*(7), Article 7. https://doi.org/10.1371/journal.pone.0021622

Sarrion-Perdigones, A., Vazquez-Vilar, M., Palaci, J., Castelijns, B., Forment, J., Ziarsolo, P., Blanca, J., Granell, A., & Orzaez, D. (2013). GoldenBraid 2.0: A Comprehensive DNA Assembly Framework for Plant Synthetic Biology. *PLANT PHYSIOLOGY*, *162*(3), 1618–1631. https://doi.org/10.1104/pp.113.217661

Sarrion-Perdigones, A., Vazquez-Vilar, M., Palací, J., Castelijns, B., Forment, J., Ziarsolo, P., Blanca, J., Granell, A., & Orzaez, D. (2013a). GoldenBraid 2.0: A Comprehensive DNA Assembly Framework for Plant Synthetic Biology. *Plant Physiology*, *162*(3), 1618–1631. https://doi.org/10.1104/pp.113.217661

Sarrion-Perdigones, A., Vazquez-Vilar, M., Palací, J., Castelijns, B., Forment, J., Ziarsolo, P., Blanca, J., Granell, A., & Orzaez, D. (2013b). GoldenBraid 2.0: A Comprehensive DNA Assembly Framework for Plant Synthetic Biology1[C][W][OA]. *Plant Physiology*, *162*(3), 1618–1631. https://doi.org/10.1104/pp.113.217661

Schillberg, S., Raven, N., Spiegel, H., Rasche, S., & Buntru, M. (2019). Critical Analysis of the Commercial Potential of Plants for the Production of Recombinant Proteins. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.00720

Selma, S., Bernabé-Orts, J. M., Vazquez-Vilar, M., Diego-Martin, B., Ajenjo, M., Garcia-Carpintero, V., Granell, A., & Orzaez, D. (2019). Strong gene activation in plants with genome-wide specificity using a new orthogonal CRISPR/Cas9-based programmable transcriptional activator. *Plant Biotechnology Journal*, *17*(9), Article 9. https://doi.org/10.1111/pbi.13138

Selma, S., Sanmartín, N., Espinosa-Ruiz, A., Gianoglio, S., Lopez-Gresa, M., Vázquez-Vilar, M., Flors, V., Granell, A., & Orzaez, D. (2021). *Custom-made design of metabolite composition in* N. benthamiana *leaves using CRISPR activators* [Preprint]. Synthetic Biology. https://doi.org/10.1101/2021.07.12.452005

Serrano, L. (2007). Synthetic biology: Promises and challenges. *Molecular Systems Biology*, *3*(1), 158. https://doi.org/10.1038/msb4100202

Shafiq, M., Qurashi, F., Mushtaq, S., Hussain, M., Hameed, A., & Saleem Haider, M. (2020). DNA plant viruses: Biochemistry, replication, and molecular genetics. In *Applied Plant Virology* (pp. 169–182). Elsevier. https://doi.org/10.1016/B978-0-12-818654-1.00013-X

Shahmuradov, I. A. (2003). PlantProm: A database of plant promoter sequences. *Nucleic Acids Research*, *31*(1), 114–117. https://doi.org/10.1093/nar/gkg041

Shakir, S., Mubin, M., Nahid, N., Serfraz, S., Qureshi, M. A., Lee, T.-K., Liaqat, I., Lee, S., & Nawaz-ul-Rehman, M. S. (2023). REPercussions: How geminiviruses recruit host factors for replication. *Frontiers in Microbiology*, *14*, 1224221. https://doi.org/10.3389/fmicb.2023.1224221

Shanmugaraj, B., I. Bulaon, C. J., & Phoolcharoen, W. (2020). Plant Molecular Farming: A Viable Platform for Recombinant Biopharmaceutical Production. *Plants*, *9*(7), 842. https://doi.org/10.3390/plants9070842

Skerker, J. M., Lucks, J. B., & Arkin, A. P. (2009). Evolution, ecology and the engineered organism: Lessons for synthetic biology. *Genome Biology*, *10*(11), 114. https://doi.org/10.1186/gb-2009-10-11-114

Smith, H. O., Hutchison, C. A., Pfannkoch, C., & Venter, J. C. (2003). Generating a synthetic genome by whole genome assembly: φX174 bacteriophage from synthetic oligonucleotides. *Proceedings of the National Academy of Sciences*, *100*(26), 15440–15445. https://doi.org/10.1073/pnas.2237126100

Sparkes, I. A., Runions, J., Kearns, A., & Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols*, *1*(4), 2019–2025. https://doi.org/10.1038/nprot.2006.286

Stanley, J. (1985). The Molecular Biology of Geminiviruses. In *Advances in Virus Research* (Vol. 30, pp. 139–177). Elsevier. https://doi.org/10.1016/S0065-3527(08)60450-9

Stanley, J. (2008). Beet Curly Top Virus. In *Encyclopedia of Virology* (pp. 301–307). Elsevier. https://doi.org/10.1016/B978-012374410-4.00696-8

Stenger, D. C., Stevenson, M. C., Hormuzdi, S. G., & Bisaro, D. M. (1992). A number of subgenomic DNAs are produced following agroinoculation of plants with beet curly top virus. *Journal of General Virology*, *73*(2), 237–242. https://doi.org/10.1099/0022-1317-73-2-237

Stepanova, A. N., Yun, J., Likhacheva, A. V., & Alonso, J. M. (2007). Multilevel Interactions between Ethylene and Auxin in *Arabidopsis* Roots. *The Plant Cell*, *19*(7), 2169–2185. https://doi.org/10.1105/tpc.107.052068

Sun, L., & Chen, Z. J. (2004). The novel functions of ubiquitination in signaling. *Current Opinion in Cell Biology*, *16*(2), 119–126. https://doi.org/10.1016/j.ceb.2004.02.005

Szybalski, W. (1974). In Vivo and in Vitro Initiation of Transcription. In A. Kohn & A. Shatkay (Eds.), *Control of Gene Expression* (Vol. 44, pp. 23–24). Springer US. https://doi.org/10.1007/978-1-4684-3246-6_3

Takasugi, P. R., Wang, S., Truong, K. T., Drage, E. P., Kanishka, S. N., Higbee, M. A., Bamidele, N., Ojelabi, O., Sontheimer, E. J., & Gagnon, J. A. (2022). Orthogonal CRISPR-Cas tools for genome editing, inhibition, and CRISPR recording in zebrafish embryos. *Genetics*, *220*(1), iyab196. https://doi.org/10.1093/genetics/iyab196

Tamm, L., Thuerig, B., Apostolov, S., Blogg, H., Borgo, E., Corneo, P. E., Fittje, S., de Palma, M., Donko, A., Experton, C., Alcázar Marín, É., Morell Pérez, Á., Pertot, I., Rasmussen, A., Steinshamn, H., Vetemaa, A., Willer, H., & Herforth-Rahmé, J. (2022). Use of Copper-Based Fungicides in Organic Agriculture in Twelve European Countries. *Agronomy*, *12*(3), 673. https://doi.org/10.3390/agronomy12030673

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., & Vale, R. D. (2014a). A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell*, *159*(3), 635–646. https://doi.org/10.1016/j.cell.2014.09.039

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., & Vale, R. D. (2014b). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. https://doi.org/10.1016/j.cell.2014.09.039

Tang, W., Luo, X., & Samuels, V. (2004). Regulated gene expression with promoters responding to inducers. *Plant Science*, *166*(4), 827–834. https://doi.org/10.1016/j.plantsci.2003.12.003

Thiruppathi, D. (2020). Mix, Match, and Maize: A Synthetic System for Maize Nuclear Auxin Response Circuits. *Plant Physiology*, *183*(2), 416–417. https://doi.org/10.1104/pp.20.00489

Thomas, H., & Sadras, V. O. (2001). The capture and gratuitous disposal of resources by plants. *Functional Ecology*, *15*(1), 3–12. https://doi.org/10.1046/j.1365-2435.2001.00488.x

Thompson, M. G., Moore, W. M., Hummel, N. F. C., Pearson, A. N., Barnum, C. R., Scheller, H. V., & Shih, P. M. (2020). *Agrobacterium tumefaciens*: A Bacterium Primed for Synthetic Biology. *BioDesign Research*, *2020*, 8189219. https://doi.org/10.34133/2020/8189219

Tian, C., Li, J., Wu, Y., Wang, G., Zhang, Y., Zhang, X., Sun, Y., & Wang, Y. (2024). An integrative database and its application for plant synthetic biology research. *Plant Communications*, *5*(5), 100827. https://doi.org/10.1016/j.xplc.2024.100827

Tian, C., Zhang, Y., Li, J., & Wang, Y. (2022). Benchmarking Intrinsic Promoters and Terminators for Plant Synthetic Biology Research. *BioDesign Research*, *2022*, 9834989. https://doi.org/10.34133/2022/9834989

Tian, J., Chen, M., Zhang, J., Li, K., Song, T., Zhang, X., & Yao, Y. (2017). Characteristics of dihydroflavonol 4-reductase gene promoters from different leaf colored Malus crabapple cultivars. *Horticulture Research*, *4*(1), 17070. https://doi.org/10.1038/hortres.2017.70

Tiwari, S. B., Belachew, A., Ma, S. F., Young, M., Ade, J., Shen, Y., Marion, C. M., Holtan, H. E., Bailey, A., Stone, J. K., Edwards, L., Wallace, A. D., Canales, R. D., Adam, L., Ratcliffe, O. J., & Repetti, P. P. (2012a). The EDLL motif: A potent plant transcriptional activation domain from AP2/ERF transcription factors. *The Plant Journal*, *70*(5), 855–865. https://doi.org/10.1111/j.1365-313X.2012.04935.x

Tiwari, S. B., Belachew, A., Ma, S. F., Young, M., Ade, J., Shen, Y., Marion, C. M., Holtan, H. E., Bailey, A., Stone, J. K., Edwards, L., Wallace, A. D., Canales, R. D., Adam, L., Ratcliffe, O. J., & Repetti, P. P. (2012b). The EDLL motif: A potent plant transcriptional activation domain from AP2/ERF transcription factors. *The Plant Journal*, *70*(5), 855–865. https://doi.org/10.1111/j.1365-313X.2012.04935.x

Tsang, T., Davis, C. I., & Brady, D. C. (2021). Copper biology. *Current Biology*, *31*(9), R421–R427. https://doi.org/10.1016/j.cub.2021.03.054

Turnbull, C., Lillemo, M., & Hvoslef-Eide, T. A. K. (2021). Global Regulation of Genetically Modified Crops Amid the Gene Edited Crop Boom – A Review. *Frontiers in Plant Science*, *0*. https://doi.org/10.3389/fpls.2021.630396

Urbino, C., Jammes, M., Belabess, Z., Troadec, E., Autechaud, A., & Peterschmitt, M. (2022). Invasive tomato yellow leaf curl virus recombinants challenge virus diagnosis and disease management. In *Geminivirus: Detection, Diagnosis and Management* (pp. 497–511). Elsevier. https://doi.org/10.1016/B978-0-323-90587-9.00004-3

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

Varma, A., & Malathi, V. G. (2003). Emerging geminivirus problems: A serious threat to crop production. *Annals of Applied Biology*, *142*(2), 145–164. https://doi.org/10.1111/j.1744-7348.2003.tb00240.x

Varsani, A., Shepherd, D. N., Dent, K., Monjane, A. L., Rybicki, E. P., & Martin, D. P. (2009). A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. *Virology Journal*, *6*(1), 36. https://doi.org/10.1186/1743-422X-6-36

Vazquez-Vilar, M., Bernabé-Orts, J. M., Fernandez-del-Carmen, A., Ziarsolo, P., Blanca, J., Granell, A., & Orzaez, D. (2016). A modular toolbox for gRNA–Cas9 genome engineering in plants based on the GoldenBraid standard. *Plant Methods*, *12*(1), 10. https://doi.org/10.1186/s13007-016-0101-2

Vazquez-Vilar, M., Quijano-Rubio, A., Fernandez-Del-Carmen, A., Sarrion-Perdigones, A., Ochoa-Fernandez, R., Ziarsolo, P., Blanca, J., Granell, A., & Orzaez, D. (2017). GB3.0: A platform for plant bio-design that connects functional DNA elements with associated biological data. *Nucleic Acids Research*, *45*(4), Article 4. https://doi.org/10.1093/nar/gkw1326

Vazquez-Vilar, M., Quijano-Rubio, A., Fernandez-del-Carmen, A., Sarrion-Perdigones, A., Ochoa-Fernandez, R., Ziarsolo, P., Blanca, J., Granell, A., & Orzaez, D. (2017). GB3.0: A platform for plant bio-design that connects functional DNA elements with associated biological data. *Nucleic Acids Research*, gkw1326. https://doi.org/10.1093/nar/gkw1326

Vazquez-Vilar, M., Selma, S., & Orzaez, D. (2023). The design of synthetic gene circuits in plants: New components, old challenges. *Journal of Experimental Botany*, *74*(13), 3791–3805. https://doi.org/10.1093/jxb/erad167

Vilanova, C., & Porcar, M. (2014). iGEM 2.0—Refoundations for engineering biology. *Nature Biotechnology*, *32*(5), 420–424. https://doi.org/10.1038/nbt.2899

Vinoth Kumar, R., & Shivaprasad, P. V. (2020). Plant-virus-insect tritrophic interactions: Insights into the functions of geminivirus virion-sense strand genes. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1936), 20201846. https://doi.org/10.1098/rspb.2020.1846

Vollen, K., Zhao, C., Alonso, J. M., & Stepanova, A. N. (2024). Sourcing DNA parts for synthetic biology applications in plants. *Current Opinion in Biotechnology*, *87*, 103140. https://doi.org/10.1016/j.copbio.2024.103140

Wang, C., Fan, S., Xu, N., Li, Z., Zhang, S., & Zhu, S. (2022). Structural basis of DNA recognition of tomato yellow leaf curl virus replication-associated protein. *International Journal of Biological Macromolecules*, *205*, 316–328. https://doi.org/10.1016/j.ijbiomac.2022.02.106

Wang, M., Lu, Y., Botella, J. R., Mao, Y., Hua, K., & Zhu, J. (2017). Gene Targeting by Homology-Directed Repair in Rice Using a Geminivirus-Based CRISPR/Cas9 System. *Molecular Plant*, *10*(7), 1007–1010. https://doi.org/10.1016/j.molp.2017.03.002

Wang, W., Zheng, G., & Lu, Y. (2021). Recent Advances in Strategies for the Cloning of Natural Product Biosynthetic Gene Clusters. *Frontiers in Bioengineering and Biotechnology*, *9*, 692797. https://doi.org/10.3389/fbioe.2021.692797

Wang, Y., & Demirer, G. S. (2023). Synthetic biology for plant genetic engineering and molecular farming. *Trends in Biotechnology*, *41*(9), 1182–1198. https://doi.org/10.1016/j.tibtech.2023.03.007

Wang, Y.-H., Wei, K. Y., & Smolke, C. D. (2013). Synthetic Biology: Advancing the Design of Diverse Genetic Systems. *Annual Review of Chemical and Biomolecular Engineering*, *4*(1), 69–102. https://doi.org/10.1146/annurev-chembioeng-061312-103351

Wani, K. I., & Aftab, T. (2022). Limitations, Biosafety, Ethics, Regulatory Issues in Molecular Farming in Plants. In K. I. Wani & T. Aftab (Eds.), *Plant Molecular Farming: Applications and New Directions* (pp. 61–74). Springer International Publishing. https://doi.org/10.1007/978-3-031-12794-6_5

Waters, V. L., Hirata, K. H., Pansegrau, W., Lanka, E., & Guiney, D. G. (1991). Sequence identity in the nick regions of IncP plasmid transfer origins and T-DNA borders of Agrobacterium Ti plasmids. *Proceedings of the National Academy of Sciences*, *88*(4), 1456–1460. https://doi.org/10.1073/pnas.88.4.1456

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., Wang, J., Venkatesh, P., Sappington, I., Torres, S. V., … Baker, D. (2023). De novo design of protein structure and function with RFdiffusion. *Nature*, *620*(7976), 1089–1100. https://doi.org/10.1038/s41586-023-06415-8

Wawrzyniak, P., Płucienniczak, G., & Bartosik, D. (2017). The Different Faces of Rolling-Circle Replication and Its Multifunctional Initiator Proteins. *Frontiers in Microbiology*, *8*, 2353. https://doi.org/10.3389/fmicb.2017.02353

Weber, E., Engler, C., Gruetzner, R., Werner, S., & Marillonnet, S. (2011). A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS ONE*, *6*(2), e16765. https://doi.org/10.1371/journal.pone.0016765

Weinmann, P., Gossen, M., Hillen, W., Bujard, H., & Gatz, C. (1994). A chimeric transactivator allows tetracycline-responsive gene expression in whole plants. *The Plant Journal*, *5*(4), 559–569. https://doi.org/10.1046/j.1365-313X.1994.5040559.x

Werner, S., Breus, O., Symonenko, Y., Marillonnet, S., & Gleba, Y. (2011). High-level recombinant protein expression in transgenic plants by using a double-inducible viral vector. *Proceedings of the National Academy of Sciences*, *108*(34), 14061–14066. https://doi.org/10.1073/pnas.1102928108

Wilbers, R. H. P., Westerhof, L. B., Van Noort, K., Obieglo, K., Driessen, N. N., Everts, B., Gringhuis, S. I., Schramm, G., Goverse, A., Smant, G., Bakker, J., Smits, H. H., Yazdanbakhsh, M., Schots, A., & Hokke, C. H. (2017). Production and glyco-engineering of immunomodulatory helminth glycoproteins in plants. *Scientific Reports*, *7*(1), 45910. https://doi.org/10.1038/srep45910

Williams, J., Regedanz, E., Lucinda, N., Nava Fereira, A. R., Lacatus, G., Berger, M., O'Connell, N., Coursey, T., Ruan, J., Bisaro, D. M., & Sunter, G. (2024). Mutation of the conserved late element in geminivirus CP promoters abolishes Arabidopsis TCP24 transcription factor binding and decreases H3K27me3 levels on viral chromatin. *PLOS Pathogens*, *20*(7), e1012399. https://doi.org/10.1371/journal.ppat.1012399

Wittmann, B. J., Yue, Y., & Arnold, F. H. (2021). Informed training set design enables efficient machine learning-assisted directed protein evolution. *Cell Systems*, *12*(11), 1026-1045.e7. https://doi.org/10.1016/j.cels.2021.07.008

Wu, M., Bejarano, E. R., Castillo, A. G., & Lozano-Durán, R. (2022a). Geminivirus DNA replication in plants. In *Geminivirus: Detection, Diagnosis and Management* (pp. 323–346). Elsevier. https://doi.org/10.1016/B978-0-323-90587-9.00038-9

Wu, M., Bejarano, E. R., Castillo, A. G., & Lozano-Durán, R. (2022b). Geminivirus DNA replication in plants. In *Geminivirus: Detection, Diagnosis and Management* (pp. 323–346). Elsevier. https://doi.org/10.1016/B978-0-323-90587-9.00038-9

Wu, M., Wei, H., Tan, H., Pan, S., Liu, Q., Bejarano, E. R., & Lozano-Durán, R. (2021). Plant DNA polymerases α and δ mediate replication of geminiviruses. *Nature Communications*, *12*(1), 2780. https://doi.org/10.1038/s41467-021-23013-2

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

Yang, Y., Chaffin, T. A., Ahkami, A. H., Blumwald, E., & Stewart, C. N. (2022). Plant synthetic biology innovations for biofuels and bioproducts. *Trends in Biotechnology*, *40*(12), 1454–1468. https://doi.org/10.1016/j.tibtech.2022.09.007

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., Beyer, P., & Potrykus, I. (2000). Engineering the Provitamin A (β-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science*, *287*(5451), 303–305. https://doi.org/10.1126/science.287.5451.303

## References


Yu, Y., Wang, X., Sun, H., Liang, Q., Wang, W., Zhang, C., Bian, X., Cao, Q., Li, Q., Xie, Y., Ma, D., Li, Z., & Sun, J. (2020). Improving CRISPR-Cas-mediated RNA targeting and gene editing using SPLCV replicon-based expression vectors in *Nicotiana benthamiana*. *Plant Biotechnology Journal*, *18*(10), 1993–1995. https://doi.org/10.1111/pbi.13384

Zalatan, J. G., Lee, M. E., Almeida, R., Gilbert, L. A., Whitehead, E. H., La Russa, M., Tsai, J. C., Weissman, J. S., Dueber, J. E., Qi, L. S., & Lim, W. A. (2015). Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell*, *160*(1–2), 339–350. https://doi.org/10.1016/j.cell.2014.11.052

Zerbini, F. M., Briddon, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., Rivera-Bustamante, R., Roumagnac, P., Varsani, A., & ICTV Report Consortium. (2017). ICTV Virus Taxonomy Profile: Geminiviridae. *Journal of General Virology*, *98*(2), 131–133. https://doi.org/10.1099/jgv.0.000738

Zhang, C., Tang, Y., Tang, S., Chen, L., Li, T., Yuan, H., Xu, Y., Zhou, Y., Zhang, S., Wang, J., Wen, H., Jiang, W., Pang, Y., Deng, X., Cao, X., Zhou, J., Song, X., & Liu, Q. (2024). An inducible CRISPR activation tool for accelerating plant regeneration. *Plant Communications*, *5*(5), 100823. https://doi.org/10.1016/j.xplc.2024.100823

Zhang, F. (2019). Development of CRISPR-Cas systems for genome editing and beyond. *Quarterly Reviews of Biophysics*, *52*. https://doi.org/10.1017/S0033583519000052

Zhang, X., & Mason, H. (2006). Bean Yellow Dwarf Virus replicons for high-level transgene expression in transgenic plants and cell cultures. *Biotechnology and Bioengineering*, *93*(2), 271–279. https://doi.org/10.1002/bit.20695

Zhang, Y., Butelli, E., Alseekh, S., Tohge, T., Rallapalli, G., Luo, J., Kawar, P. G., Hill, L., Santino, A., Fernie, A. R., & Martin, C. (2015). Multi-level engineering facilitates the production of phenylpropanoid compounds in tomato. *Nature Communications*, *6*, 1–11. https://doi.org/10.1038/ncomms9635

Zhang, Y., Yin, C., Zhang, T., Li, F., Yang, W., Kaminski, R., Fagan, P. R., Putatunda, R., Young, W.-B., Khalili, K., & Hu, W. (2015). CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs. *Scientific Reports*, *5*(1), 16277. https://doi.org/10.1038/srep16277

Zhao, C., Yaschenko, A., Alonso, J. M., & Stepanova, A. N. (2021). Leveraging synthetic biology approaches in plant hormone research. *Current Opinion in Plant Biology*, *60*, 101998. https://doi.org/10.1016/j.pbi.2020.101998

Zhou, Y., Li, J., Pan, Y., Zheng, Y., Pan, Y., Ding, Y., Su, C., & Zhang, X. (2017). Establishment of a tetracycline-off and heat shock-on gene expression system in tobacco. *Journal of Integrative Agriculture*, *16*(5), 1112–1119. https://doi.org/10.1016/S2095-3119(16)61514-9

Cover illustrated by Xenja Ploeger.