Article Type: Methods Paper

Loop assembly: a simple and open system for recursive fabrication of DNA circuits

Bernardo Pollak[1], Ariel Cerda[2], Mihails Delmans[1], Simón Álamos[3], Tomás Moyano[2], Anthony West[4], Rodrigo A. Gutiérrez[2], Nicola Patron[4], Fernán Federici[1,2,*] & Jim Haseloff[1,*]

[1] Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK.

[2] Departamento de Genética Molecular y Microbiología, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile 8331150, Santiago, Chile, Fondo de Desarrollo de Áreas Prioritarias, Center for Genome Regulation, Millennium Institute for Integrative Biology (iBio) 8331150, Santiago, Chile.

[3] Department of Plant and Microbial Biology, University of California, Berkeley, California CA94720, USA.

[4] Earlham Institute, Norwich Research Park, Norwich NR4 7UZ, UK.

[*] Corresponding authors.

Jim Haseloff jh295@cam.ac.uk.

Telephone: +44-1223-766546

Fernán Federici

Telephone: +44-1223-766546

ffederici@bio.puc.cl

Bernardo Pollak (0000-0003-2329-7401), Nicola Patron (0000-0002-8389-1851) and Jim Haseloff (0000-0003-4793-8058).

Received: 3 June 2018

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/nph.15625

Accepted: 20 September 2018

Keywords: Loop assembly, standardised DNA assembly, Type IIS, unique nucleotide sequences (UNS)**, common syntax, recursive assembly, *LoopDesigner*, OpenMTA.**

Summary

- High efficiency methods for DNA assembly have enabled routine assembly of synthetic DNAs of increased size and complexity. However, these techniques require customisation, elaborate vector sets or serial manipulations for the different stages of assembly.
- We have developed Loop assembly based on a recursive approach to DNA fabrication. The system makes use of two Type IIS restriction endonucleases and corresponding vector sets for efficient and parallel assembly of large DNA circuits. Standardised level 0 parts can be assembled into circuits containing 1, 4, 16 or more genes by looping between the two vector sets. The vectors also contain modular sites for hybrid assembly using sequence overlap methods.
- Loop assembly enables efficient and versatile DNA fabrication for plant transformation. We show construction of plasmids up to 16 genes and 38 Kb with high efficiency (>80%). We have characterized Loop assembly on over 200 different DNA constructs and validated the fidelity of the method by high-throughput Illumina plasmid sequencing.
- Our method provides a simple generalised solution for DNA construction with standardised parts. The cloning system is provided under an OpenMTA license for unrestricted sharing and open access.

Introduction

Standardised approaches to the assembly of large DNAs have played an important role in the development of systematic strategies for reprogramming biological systems. This began with the implementation of idempotent assembly methods based on DNA digestion/ligation using standardised nested restriction endonuclease (RE) sites, such as the BioBrick assembly method (Knight, 2003; Shetty *et al.,* 2008). More recently, assembly techniques that enabled the parallel assembly of multiple components in a single reaction have been established.

These include methods that utilise long-sequence overlaps (Bitinaite *et al.,* 2007; Li and Elledge, 2007; Zhu *et al.,* 2007; Gibson *et al.,* 2009; Bryksin and Matsumura, 2010; Zhang *et al.,* 2012; Beyer *et al.,* 2015; Jin *et al.,* 2016), systems reliant on *in vivo* recombination (Ma *et al.,* 1987; Gibson *et al.,* 2008b; Joska *et al.,* 2014), and Golden Gate (Engler *et al.,* 2008) based methods that rely on selective digestion and re-ligation of plasmid DNAs with Type IIS RE (Sarrion-Perdigones *et al.,* 2011; Weber *et al.,* 2011; Sarrion-Perdigones *et al.,* 2013; Engler *et al.,* 2014; Storch *et al.,* 2015; Iverson *et al.,* 2016; Moore *et al.,* 2016). Type IIS and long-overlap based methods have allowed increased scale and efficiency of DNA circuit assembly, while *in vivo* recombination remains the method of choice for genome-scale manipulations (Gibson *et al.,* 2008a,b; Benders *et al.,* 2010; Gibson *et al.,* 2010a; Karas *et al.,* 2012, 2013).

Gibson assembly, a sequence overlap-based method, was developed for the synthesis and assembly of *Mycoplasma* genomes (Gibson *et al.,* 2008a, 2010a) and enabled assembly of DNAs up to several hundred kb in one-pot isothermal reactions (Gibson *et al.,* 2009). This method has been widely adopted by the synthetic biology community, being scar-free, versatile and relatively efficient. However, Gibson assembly generally relies on the use of oligonucleotides to perform *in vitro* amplification of DNA fragments, which can be error-prone (Keohavong & Thilly, 1989; Gibson *et al.*, 2010b; Potapov & Ong, 2017). The method is also sensitive to sequence composition and repeats, hence efforts have been made to standardise and streamline Gibson assembly by including flanking unique nucleotide sequences (UNS) that can be used as long overlaps for cloning of transcription units (TUs) into larger constructs (Torella *et al.,* 2013). Perhaps due to the flexible nature of Gibson assembly, a standard for composing elemental parts into TUs has not been proposed yet. Laboratories that employ Gibson assembly rely on their own set of rules and templates for DNA parts, and there has been no community-wide effort to develop a common standard.

In contrast, Type IIS assembly systems are virtually free of *ad hoc* design, and are highly efficient for both assembly of TUs and assembly of elementary parts into TUs (Patron, 2016). These methods do not require PCR amplification or fragment isolation, and allow parallel assembly of a large number of DNA parts (Potapov *et al.*, 2018). Instead of PCR, these methods exploit Type IIS RE to generate fragments with short complimentary overhangs that can be ligated in a one-pot reaction. While this approach can be scarless, the application of standard overhangs (fusion sites) for DNA parts with a defined function (e.g. promoter, CDS, terminators) allows the same DNA parts to be re-assembled into multiple constructs without

redesign or modification (Engler *et al.,* 2008; Sarrion-Perdigones *et al*., 2011; Weber *et al.* 2011; Lampropoulos *et al.*, 2013; Sarrion-Perdigones *et al.* 2013; Binder *et al.,* 2014). Recently, a common syntax has been proposed by developers and adopters of Type IIS cloning methods. This standard defines an unambiguous arrangement of 12 Type IIS overhangs that form boundaries between functional domains found within a generalised eukaryote gene (Patron *et al.,* 2015). The common syntax is based on the widely used MoClo and GoldenBraid standards, and has found acceptance in the plant field (Patron *et al.,* 2015), and iGEM in the form of PhytoBricks standard parts. The common syntax ensures that these Type IIS assembly systems can share a common stock of standardised DNA parts to be shared and used in an off-the-shelf manner. The establishment of a common standard for stock DNA parts also provides a prevailing syntax that enhances transferability and reproducibility for compiling genetic instructions in different labs. Assembly of an exact copy of a genetic construct is possible simply by knowing its composition, eliminating unnecessary *ad hoc* design and enabling simple abstract descriptions that contain a precise implied sequence. However, Type IIS assembly systems require the refactoring or 'domestication' of DNA parts, generally performed through PCR or DNA synthesis. Domestication refers to the elimination of RE sites present in the DNA sequence prior to its use in the assembly system. To date, the most commonly used REs have been BsaI, BsmBI and BpiI, which have 6 bp recognition sites that, while not frequent on average, are regularly encountered in DNA sequences (Lin & O'Callaghan, 2018). Type IIS REs such as SapI and AarI with 7 bp recognition sites can be used to lower the probability of finding sites requiring domestication, and are used in the Electra™ (ATUM) and GeneArt™ (ThermoFisher) kits, respectively. Type IIS based systems have found rapid acceptance in the synthetic biology field due to the need for robustness, scalability and compatibility with automated assembly methods. Since synthetic biology is already at the point where constructs can consist of multiple logic gates (Nielsen *et al.,* 2016), entire biosynthetic pathways (Temme *et al.,* 2012) or engineered genomic DNA (Richardson *et al.,* 2017), robust assembly methods such as Type IIS assembly are essential to enable fabrication of higher-order genetic constructs.

Despite much progress in the technical aspects of DNA construction and part reusability, restrictive intellectual property (IP) practices and material transfer agreements (MTA) can hinder the sharing of DNA components in both the public and private sectors, delaying experimental work through paperwork and legal consultation. For this purpose, an international effort is underway to establish the OpenMTA (http://www.openmta.org) as a way of expediting the sharing of biological materials. The OpenMTA provides a legal template for free and unrestricted distribution of materials, providing a formal mechanism for effectively placing materials in the public domain, in a way that extends existing practices. Open sharing of DNA assembly systems and parts through the OpenMTA will facilitate the

engineering of new solutions for problems in human health, agriculture and the environment, such as those identified as Sustainable Development Goals by the United Nations (http://www.un.org/sustainabledevelopment) and Global Grand Challenges by the Gates Foundation (https://gcgh.grandchallenges.org).

Here we present Loop assembly, a versatile, simple and efficient DNA fabrication system based on recursive DNA assembly. It combines all the benefits of Type IIS assembly, but requires only a set of eight plasmids to build constructs with theoretically unlimited length. As well as Type IIS assembly, the system integrates long-overlap assembly methods. In this way, four TUs can be assembled into multiple transcription units by using alternative methods such as Gibson Assembly via flanking UNS (Torella *et al.,* 2013). In our method, Type IIS assemblies are performed through iterated 'loops'. Two sets of four plasmid vectors are provided, which allow alternating assembly cycles. First, Level 0 parts, defined by the PhytoBrick common syntax, are assembled into Level 1 transcription units in each of four odd-numbered vectors using BsaI. Second, four Level 1 modules can then be assembled into a Level 2 construct in each of the four even-numbered vectors using SapI. Following this, Level 2 constructs can be combined by cloning back into odd-numbered vectors, using BsaI, to create Level 3 assemblies containing up to 16 transcription units each. The iterative process of combining genetic modules, four at a time, can be continued without theoretical limit, alternating assembly steps between odd and even Loop vectors. Since levels are used recursively, it is possible to create hybrid levels that can contain a mixture of parts from different levels of the same parity (i.e. Level 2 vectors combined with elements from Level 0 vectors). In addition, we have developed *LoopDesigner*, a software framework for *in silico* sequence handling and assembly design. The software tools are open source and available through *Github*, and Loop assembly vectors are provided through the OpenMTA for unrestricted use. We have developed and tested the Loop assembly system in different laboratories and provide data to support the efficiency and robustness of the method. We have assembled over 200 constructs with up to 16 TUs and over 38 kb in size. We have tested Loop constructs *in planta* and validated their function in transgenic *Marchantia polymorpha,* and through transient expression in *Arabidopsis thaliana* protoplasts.

Materials and Methods

Construction of Loop assembly backbones. Loop assembly vectors were constructed using Gibson assembly (Gibson *et al.,* 2009). Several changes were made to a pGreenII vector (Hellens, et al., 2000) to obtain a basic plasmid backbone for the Loop assembly vectors: BsaI and SapI sites were removed from the plasmid using silent mutations when possible. In order to reduce issues with stability of large constructs in bacteria (Moore *et al.* 2016, Watson *et al.*, 2016), two nucleotides of the pGreenII ColEI-derived origin of replication were mutated, reversing it into the medium-low copy number pBR322 origin of replication. A region extending from the T-DNA left border to the hygromycin resistance gene cassette was replaced with the sequence of the pET15 vector (Haseloff, 1999) from the nptII nosT terminator to the UAS_{GAL4} promoter (bases 2851-3527). A spectinomycin resistance was cloned to replace the nptI cassette to provide a microbial selection marker for the pEven plasmids. UNSes were cloned into the kanamycin and spectinomycin version of vector backbones after the 3' end of the pET15 vector sequence and the right border (RB). Finally, the Loop restriction enzyme sites (BsaI and SapI), overhangs and the lacZ cassette were cloned in between the UNS, yielding the pOdd and pEven vectors. L0 plasmids used for Loop Type IIS assembly were assembled using Gibson assembly into a modified pUDP2 (BBa_P10500) plasmid, which contained a 20 bp random sequence (5'-TAGCCGGTCGAGTGATACACTGAAGTCTC-3') downstream of the 3' convergent BsaI site and upstream of the BioBrick suffix, to provide non-homologous flanking regions for correct orientation during overlap assembly.

DNA Spacers. Random DNA sequences were retrieved from Random DNA Sequence Generator (http://www.faculty.ucr.edu/~mmaduro/random.htm), ordered as dsDNA fragments from IDT and assembled using Gibson assembly.

Plasmids and construct design. L0 parts used for DNA construction are described in **Table S1,** their sequences included in **Supplementary Files** and are available through Addgene. Sequences for Loop plasmids and resulting multigene assemblies are included in **Supplementary Files.**

The design of the constructs was performed using *LoopDesigner* software, installed on a local machine. The software was configured to use Loop assembly backbones together with BsaI and SapI RE, as well as A-B and - overhangs. In addition, definition of 12 L0 part

types were added to the software, based on the overhangs specified by the common syntax. The sequences of the L0 parts were added to the *LoopDesigner* database, assigning one of the defined part types, and assembled consequently into Level 1 and Level 2 constructs *in silico.* The concentration of L0 parts and Level 1 constructs were adjusted to those suggested by the LoopDesigner for 10 µL reactions.

Loop Type IIS assembly protocol. The Loop Type IIS assembly protocol was adapted from Patron, 2016, and can be found in https://www.protocols.io/view/loop-assembly-pyqdpvw. 15 fmol of each part to be assembled were mixed with 7.5 fmol of the receiver plasmid in a final volume of 5 µL with dH_20 (**Table S2**). The reaction mix containing 3 µL of dH_20, 1 µL of T4 DNA ligase buffer 10x (NEB cat. B0202), 0.5 uL of 1 mg/mL purified bovine serum albumin (1:20 dilution in dH_20 of BSA, Molecular Biology Grade 20 mg/mL, NEB cat. B9000), 0.25 µL of T4 DNA ligase at 400 U/µL (NEB cat. M0202) and 0.25 µL of corresponding restriction enzyme at 10 U/µL (BsaI NEB cat. R0535 or SapI NEB cat. R0569) was prepared on ice. Then, 5 µL of the reaction mix was combined with the 5 µL of DNA mix for a reaction volume of 10 µL (**Table S3)** by pipetting and incubated in a thermocycler using the program described in **Table S4.** For SapI reactions, T4 DNA ligase buffer was replaced by CutSmart buffer (NEB cat. B7204S) supplemented with 1 mM ATP. 1 µL of the reaction mix was added to 50 µL of chemically competent TOP10 cells (ThermoFisher cat. C4040100) and following incubation at 42°C for 30 seconds, samples were left on ice for 5 minutes, 250 µL of SOC media was added and cells incubated at 37 ºC for 1 hour. Finally, 5 µL of 25 mg/mL of X-Gal (Sigma-Aldrich cat. B4252) dissolved in DMSO, was added and the cells were plated onto selective LB-agar plates supplemented with 1mM IPTG (Sigma cat. I6758). Assembly reactions were also automated. The assembly reactions were identical except scaled down to a total volume of 1 µL. Reactions were set up on a Labcyte Echo in 384 well plates and incubated on a thermal cycling machine using the same conditions as described above. Reactions were transformed into 4 µL competent XL10-Gold® Ultracompetent Cells (Agilent Technologies, Santa Clara, CA, USA) and plated onto eight-well selective LB-agar plates. Colonies were picked for growth in 1 mL of media in 96-well plates on a Hamilton STARplus® platform.

Standardised PCR of transcriptional units. PCR using UNS oligonucleotides was performed with an annealing temperature of 60 ºC, 35 cycles using Phusion High-Fidelity DNA polymerase (ThermoFisher cat. F-530) in 50 µL reactions, according to manufacturer's instructions. Template was added to a final concentration of 20 pg/µL. DNA fragments were visualised using SYBR Safe DNA Gel Stain (ThermoFisher cat. S33102) on a blue LED transilluminator (IORodeo). DNA purification was performed using NucleoSpin Gel and PCR Clean-up purification kit (Macherey-Nagel, cat. 740609.250). UNS primers used in TUs amplification are listed in **Table S5**.

Validation by sequencing. The sequences of assembled plasmids were verified by complete sequencing using 150 base pair paired-end reads on an Illumina MiSeq platform, and can be found in the EMBL-ENA database grouped under study PRJEB29863. Libraries were prepared using the Nextera XT DNA Library Prep Kit (Illumina cat. FC-131-1096), using the manufacturer's protocol modified to a one in four dilution. Reads were filtered and trimmed for low-quality bases and mapped to plasmids using the 'map to reference tool' from the Geneious 8.1.8 software (http://www.geneious.com, Kearse *et al.,* 2012), with standard parameters. Sequence fidelity was determined manually.

***Agrobacterium*-mediated *Marchantia* transformation.** *Agrobacterium*-mediated transformation was carried out as described previously (Ishizaki *et al.,* 2008), with the following exceptions: half of an archegonia-bearing sporangium (spore-head) was used for each transformation. Dried spore-heads were crushed in a 50 mL Falcon tube with a 15 mL Falcon tube and resuspended in 1 mL of water per spore-head. Resuspended spores were filtered through a 40 µm mesh (Corning cat. 352340) and 1 mL of suspension was aliquoted into a 1.5 mL Eppendorf tube and centrifuged at 13,000 xg for 1 min at room temperature. Supernatant was discarded and spores were resuspended in 1 mL of sterilisation solution, and incubated at room temperature for 20 min at 150 RPM on an orbital shaker. The sterilisation solution was prepared by dissolving one Milton mini-sterilising tablet (Milton Pharmaceutical UK Company, active ingredient: Sodium dichloroisocyanurate CAS: 2893-78-9: 19.5% w/w) in 25 mL of sterile water. Samples were centrifuged at 13,000 xg for 1 min, washed once with sterile water and resuspended in 100 µL of sterile water per each

spore-head used. One hundred µL of sterilised spores were inoculated onto half strength Gamborg's B5 1 % (w/v) agar plates and grown under constant fluorescent lighting (50-60 mol photons/m^2s) upside down for 5 days until co-cultivation. Sporelings were co-cultivated with previously transformed and induced *Agrobacterium* GV2260 transformed with the pSoup plasmid (Hellens *et al*. 2000) in 250 mL flasks containing 25 mL of half strength Gamborg's B5 media supplemented with 5 % (w/v) sucrose, 0.1 % (w/v) N-Z Amine A (Sigma cat. C7290), 0.03 % (w/v) L-glutamine (Sigma cat. G8540) and 100 µM acetosyringone (Sigma-Aldrich cat. D134406) for 36 h, until washed and plated onto selective media.

Laser-scanning confocal microscopy. A microscope slide was fitted with a 65 µL Gene Frame (ThermoFisher cat. AB0577) and 65 µL of dH_20 was placed in the center. *Marchantia gemmae* were carefully deposited on the drop of dH_20 using a small inoculation loop and a #0 coverslip was attached to the Gene Frame. Slides were examined on a Leica TCS SP8 confocal microscope platform equipped with a white-light laser (WLL) device. Imaging was conducted using a Leica HC PL APO 20x CS2 air objective with a sequential scanning mode with laser wavelengths of 405 nm, 488 nm and 515 nm, capturing emitted fluorescence at 450-482 nm, 492-512 nm and 520-550 nm windows in each sequential scan, respectively. Z-stacks were collected every 5 µm for the complete volume range and maximum intensity projections were processed using ImageJ software. Fluorescence bleedthrough from the blue pseudocoloured channel (membrane localized eGFP) into the green pseudocoloured channel (nuclear localized Venus) was eliminated using custom *Python* scripts which subtracted a 20% of the value of pixels present in the blue channel to the green channel. Images were edited to scale the pixel intensity to the full 8-bit range and a merged image was processed.

Transient expression in Arabidopsis mesophyll protoplasts. Well-expanded leaves from 3-4 weeks old Arabidopsis plants (Columbia-0) were used for protoplast transfection. Plants were grown at 22ºC, low light (75 µmol/m^2s) and short photoperiod (12 h light/12 h dark) conditions. Protoplasts were isolated and PEG- transfected according to Yoo *et al.* 2007. For transfection, 6 µL of Loop L2 plasmids (2 µg/µL) isolated by NucleoBond Xtra Midi/Maxi

purification kit (Macherey-Nagel cat. 740410.50), were used. Transfected protoplasts were incubated for 12 h in light and then visualized by epifluorescent microscopy in a Neubauer chamber (Hirschmann).

Epifluorescence microscopy. Transfected protoplasts were visualized using a Nikon Ni microscope equipped with a 49021 ET - EBFP2/Coumarin/Attenuated DAPI filter cube (ex: 405/20 nm, dichroic: 425 nm, em: 460/50 nm), 96227 AT-EYFP filter cube (ex: 495/20 nm, dichroic: 515 nm, em: 540/30 nm), 96223 AT-ECFP/C filter cube (ex: 495/20 nm, dichroic: 515 nm, em: 540/30 nm) and a 96312 G-2E/C filter cube (ex: 540/20 nm, dichroic: 565 nm, em: 620/60 nm).

***LoopDesigner*.** In order to implement an object oriented model for Loop assembly, we built a PartsDB library (https://github.com/HaseloffLab/PartsDB) to define several interlinked classes, each of which is associated with a table in a relational SQL database. The structure of *LoopDesigner* is built around a *Part* class, which either represents an ordered collection of children parts it is assembled from, or a DNA sequence in case of L0 parts. In this way we ensured that the actual DNA sequence is only stored once, while the sequences of L1 and higher parts are constructed on demand from the relational links. In addition, each *Part* is associated with one of the *Backbone* instances, which together with a *Part* sequence represents a complete Loop assembly plasmid. Every instance of a *Backbone* class is a combination of a *Base Sequence* and a donor *Restriction Enzyme Site*, e.g. pOdd 1-4 and pEven 1-4 are *Backbone* instances in the schema described in this paper. *Base Sequence* represents a type of a receiver plasmid, e.g. pOdd and pEven, and is composed of a DNA sequence of the plasmid and an instance of a receiver *Restriction Enzyme Site*. Finally, *Restriction Enzyme Site* class is composed of a *Restriction Enzyme* instance, which stores restriction enzyme recognition sequence, and a pair of overhang sequences, which can be either receiver or donor overhangs.

Results

Loop assembly. Loop assembly kit consists of two sets of plasmids that participate in a cyclic assembly process. Type IIS restriction endonucleases, BsaI and SapI, are used alternately for recursive assembly of genetic modules into a quartet of either odd (L1, L3, ...) or even (L2, L4, ...) receiver plasmids. At each step in the assembly 'loop', four genetic modules are combined into a receiver plasmid (**Fig. 1a**). Odd and even-level plasmids use alternating types of antibiotic selection, kanamycin resistance for odd-levels (pOdd plasmids) and spectinomycin resistance for even-levels (pEven plasmids), to enable the use of one-pot digestion-ligation assembly reaction (Engler *et al.,* 2008). At each level (except for TU assembly from L0 parts), four parental plasmids are required, leading to an exponential increase in the number of TUs by a factor of 4 per level (**Fig. 1b)**.

Plasmids in Loop assembly act both as donors and receivers due to the special arrangement of the RE sites. The odd receiver plasmids contain a pair of divergent BsaI sites that are removed in the cloning reaction, while a pair of convergent SapI sites, flanking the BsaI sites, allows the odd plasmids to act as donors for assembly into the following level. Similarly, the even plasmids contain a pair of divergent SapI sites flanked by convergent BsaI sites (**Fig. 2a**). Upon digestion, donor plasmids release DNA fragments (between the convergent RE sites) with specific overhangs that define the direction and position in the assembly, while the receiver plasmids release the divergent RE sites allowing for assembly of the donor fragments.

The overhangs created by the BsaI digestion of the odd receivers allow the construction of transcription units from any parts that are compatible with the PhytoBrick standard (Patron *et al.,* 2015), such as L0 parts derived from MoClo and GoldenBraid plasmid libraries (if free of SapI sites). BsaI overhang sequences are termed A, B, C, E and F, with A and F designated as flanking terminal-overhangs, and SapI overhang sequences are termed □, □, □, □ and □, with □ and □ designated as flanking terminal-overhangs. Examples of odd and even-level assemblies are shown in (**Fig. 2b,c**).

Each reaction requires four donor plasmids (or DNA spacers) for successful assembly into a receiver of the next level. In order to provide a replacement for assemblies with less than four fragments,, we designed 200 bp long 'universal spacer' parts comprised of random DNA sequence free of BsaI and SapI sites. Plasmids containing spacers with flanking terminal-

overhangs are provided for odd (pOdd-spacer) and even levels (pEven-spacer). They can be used for direct assembly into any of the four receiver plasmids of their corresponding level (**Fig. S1**).

Assembly of synthetic promoters. The recursive nature of Loop assembly allows one to mix parts from different levels but with the same parity. For example, a multimeric promoter might be constructed from elemental parts through recursive assembly. Figure 3 shows the generation of synthetic promoters by cloning L0 functional domains (e.g. TF recognition sites and minimum promoter sequences) with flanking terminal-overhangs into specific L1 plasmid positions, which determine the order of motif arrangement in the following L2 assembly. Different TF recognition sites can be used in positions 1 (and overhangs), 2 (and overhangs) and 3 (and overhangs), while a minimal promoter sequence is placed in position 4 (and overhangs) of L1 receiver plasmids. These elements can then be composed in specific order. In this example, different combinations of TF binding sites and minimal promoter were cloned into positions 1 (A and B overhangs) and 2 (B and C overhangs) of L2 receiver plasmids. The resulting composite promoter elements could be mixed with standard L0 gene parts, to create a customised hybrid gene assembly in an odd-level plasmid (**Fig. 3a**).

Using this approach, we assembled three fluorescent reporters with synthetic promoters comprised of multimeric binding sites. The promoters included binding domains for the transcription factors GAL4 (Guarente *et al.,* 1982; West *et al.,* 1984; Giniger *et al.,* 1985), HAP1 (Zhang and Guarente, 1994), a cytokinin operator (Müller and Sheen, 2008) and a minimal CaMV 35S promoter (Benfey and Chua, 1990) derivative (F. Federici and J. Haseloff, unpublished results) driving a Venus fluorescent protein (Nagai *et al*., 2002). Resulting reporters were composed of the same elements but with differing motif arrangements, containing 13 nucleotide scars between the motifs. Each reporter contained three dimeric binding domains for GAL4, three dimeric binding domains for HAP1, one dimeric CK operator binding domain and the minimal CaMV 35S promoter. (see **Notes S1**). The composite synthetic promoters, which were the result of 20 different assembly reactions, were verified through sequencing and showed no sequence errors. The sequences of the final constructs (pL3-1_PC1, pL3-1_PC2 and pL3-1_PC3) can be found in **SI**.

The recursive nature of Loop assembly also enables hybrid assemblies of multiple TUs derived from donor plasmids from different levels (i.e. three Level 1 and one Level 3 plasmids). These can be assembled into a hybrid even receiver plasmid, providing further flexibility in the fabrication of genetic constructs (**Fig. 3b**).

UNS for standardised overlap assembly. Apart from their capacity for Type IIS assembly, Loop vectors were designed for long overlap assembly techniques. Loop plasmids contain UNS that allow the use of standard primers for the amplification of TUs derived from Type IIS DNA parts (PhytoBricks, MoClo and GoldenBraid), since these can be assembled into UNS-flanked TUs by BsaI-mediated Type IIS assembly. Alternatively, TUs can be assembled from PCR-fragments or DNA synthesis into Loop plasmids by overlap assembly methods such as Gibson assembly (**Fig. 4a**). Each Loop plasmid contains two flanking UNS and a terminal UNS_X. TUs can be assembled into a multi-TU destination plasmid (pUNSDest) by using overlap assembly methods (**Fig. 4b**). UNS have been designed following a number of guidelines to provide enhanced performance in PCR reactions and overlap assembly. Design rules are listed in **Methods S1** and sequences provided in **Table S6**. Forward and reverse standard primers correspond to the first 20 bp of each UNS in both forward and reverse complement orientations, respectively, and are provided in **Table S5**. UNS have the advantage that they are designed for highly efficient PCR with standard conditions (60 ºC, 35 cycles), resulting in single amplicons with high yields (**Fig. S2**). This eliminates the need for gel purification during the workflow of Gibson assembly, if appropriate on-column purification is performed.

Reliability of Loop assembly. To evaluate the reliability of the technique, we tested L1 Type IIS Loop assemblies in different laboratories, and obtained consistent results (**Table 1**). We assembled over 200 plasmids using the Type IIS pathway for Levels 1-3 and obtained average assembly efficiency between 83 to 97 % depending on the level of assembly and complexity of constructs (**Table 1, Notes S2**). This was evaluated though DNA profiling by means of restriction endonuclease digestion (**Fig. S3**). Further, we performed Illumina sequencing of 92 Level 2 and Level 3 assembled constructs to validate Loop assembly fidelity at the sequence level, to determine if the reaction had produced correct assemblies and if mutations had been introduced by our method. We found that 95.4 % of constructs

assembled correctly with 98.8 % of overhang scars present at expected junctions. Overall, 99.8 % of nucleotides were correctly assembled, and the few incorrect constructs showed missing regions due to misassembly, rather than sequence errors *per se* (**Table S7).**

***In planta* activity of Loop plasmids.** Loop vectors were derived from the pGreenII (Hellens *et al.,* 2000) plant binary transformation vector, but decoupled from plant selection markers (see Materials and Methods) to enable their introduction during assembly. As in pGreenII, Loop plasmids contain elements for propagation in *Agrobacterium tumefaciens* and are capable of *Agrobacterium*-mediated plant transformation. We have tested the application of Loop constructs in plant developmental biology by assembling TU composed of fluorescent proteins, localisation tags and endogenous promoters. This allowed us to highlight cellular features and track patterns of gene expression *in planta*. A Level 2 construct (pL2-1_TPL) containing four TU composed of a HygR selectable marker, a mTurquoise2-N7 nuclear-localised reporter driven by a Mp*EF1*α constitutive promoter (Nagaya *et al.,* 2011; Althoff *et al.,* 2014), a Venus-N7 nuclear-localised reporter driven by a Mp*TPL* tissue specific promoter (Flores-Sandoval *et al.,* 2015) and a eGFP-Lti6b membrane-localised marker driven by a Mp*EF1*α promoter was assembled from L0 parts (**Table S1**) using Loop assembly and transformed into *Marchantia polymorpha* (Marchantia). Regenerated transformants were obtained and clonal propagules called *gemmae* were examined using confocal microscopy. All three fluorescent protein reporter genes were expressed and allowed visualisation of distinct cellular and subcellular features across the tissue (**Fig. 5**).

In addition, four L1 TUs that had been constructed by Type IIS Loop assembly were assembled into a multi-TU destination plasmid using Gibson assembly. Transfected protoplasts showed expression of the engineered fluorescent reporters in their expected localizations (**Fig. S4**), providing a fast and efficient system to evaluate functionality of Loop constructs. Plasmid maps for constructs are provided in **Fig. S5.**

Loop assembly design automation. We have developed software tools to aid Loop assembly experiments. We developed *LoopDesigner*, a web application that facilitates (i) the sequence design and domestication of Level 0 DNA parts, (ii) generation of a Loop assembly parts database, and (iii) simulation of Loop assembly reactions and the resulting plasmid maps and sequences (**Fig. 6**). An input L0 sequences are domesticated by identifying unwanted RE sites and removing them by the introduction of synonymous mutations. Appropriate BsaI overhangs are added according to the rules of the common syntax for DNA parts and stored

in the part database. See methods for detailed description of the *LoopDesigner* implementation. We invite readers to visit the *LoopDesigner* web tool available at loopdesigner.herokuapp.com (supported in Google Chrome) for exploring Loop assembly techniques. The source code of *LoopDesigner* is available at *GitHub* (https://github.com/HaseloffLab/LoopDB *LoopDesigner* branch), and provided under an MIT license.

Discussion

The design of Loop assembly was inspired by existing assembly methods such as GoldenBraid, MoClo, and standardised Gibson assembly. We attempted to integrate. these techniques in a general-purpose DNA assembly system. Loop assembly combines recursive use of two restriction enzymes and plasmid sets, which together create a simple and versatile Type IIS assembly platform. Type IIS RE sites are employed in head-to-head configurations, eliminating the requirement for end-linkers used in MoClo systems. Instead, restriction sites for successive levels are integrated in receiver plasmids, as in GoldenBraid, but using quaternary assembly parity instead of binary. This enables faster assembly of large constructs, but demands all four positions to be filled by either TUs or by spacers. Fixing the number of donor parts allows systematisation without increasing the complexity of assembly, with standardised reactions containing determined number of DNA parts and overhangs. Further, the recursive nature of Loop assembly enables the usage of a compact plasmid library whilst providing versatile construction strategies. We show the use of recursive assembly for the fabrication of complex DNA such as synthetic promoters composed of repetitive sequences and hybrid levels. Type IIS restriction sites in the Loop vectors are flanked by standardised UNS, enabling the use of Loop vectors with overlap assembly methods and the reuse of oligonucleotides for assembly. This provides users of Gibson and overlap assembly methods with the capacity to tap into libraries of domesticated DNA parts already available. We have demonstrated the high efficiency of Loop assembly by generating a variety of constructs with different number of TUs, achieving reliable assembly of constructs up to 16 TUs composed of 56 individual parts. Additionally, we have used Loop assembly for the generation of multispectral reporter constructs and show their activity *in planta.*

Use and characterisation of the products of Loop assembly demonstrated that it is a robust and reliable DNA assembly system regardless of levels and types of parts. Loop assemblies varying in size, total number of fragments and DNA composition were performed in order to provide an accurate estimate of the method performance in routine use. The high rate of successful assemblies, even in the absence of cPCR pre-selection, considerably decreases the effort and time required for DNA construction. Further, the system takes advantage of (i) a common syntax for DNA parts, (ii) a simple, recursive assembly scheme, (iii) a small set of plasmid vectors and (iv) streamlined protocols, to provide a streamlined and logical framework for assembly that will enable rapid adoption by students and non-specialists. As Loop assembly integrates Type IIS and overlap assembly, it encourages the development of a community around a DNA construction system, yielding a growing collection of DNA parts and composites. The wide compatibility of Loop assembly facilitates proper curation and improvement of DNA part collections through collaboration, easier exchange and transfer of genetic modules between labs, and cross-validation. The ability to use either overlap or Type IIS assembly provides further flexibility in making DNA constructions where sequence alterations introduced by removal of illegal RE sites are not desirable (such as for experiments involving native genetic sequences), or when the assembly fails by one of the pathways.

Although the falling costs of DNA synthesis suggest that DNA synthesis of transcriptional units or even chromosomes might eventually be time and cost-effective, synthetic biology requires the capacity for rapid, high-throughput and combinatorial assemblies. This is necessary for characterisation and troubleshooting of smaller DNA parts and circuits before compiling high-level devices and systems. Assembly systems that are tailored to exploit the opportunities provided by automation technologies will undoubtedly benefit from robotics platforms. Automated design and liquid handling platforms for fabrication DNA constructs have already been adopted by some and the technologies are rapidly expanding: at the high end of the market platforms such as the Echo (Labcyte) are enabling miniaturization and increasing throughput, yielding a substantial reduction in reaction costs (Kanigowska *et al.,* 2016), while low-cost platforms such as the OT-One S (OpenTrons) are aiming to make automated pipetting affordable in every laboratory.

To enable facile design of constructs we developed *LoopDesigner*, a software framework that provides an interface between digital design and experimentation. We have demonstrated the usefulness of the *LoopDesigner* by implementing a simple web tool where users can design assembly strategy and run virtual reactions before stepping into the lab. The *LoopDesigner* framework allows definition of Loop assembly schemas of arbitrary complexity, with any number of levels and plasmids per level, as well as with any possible restriction enzymes and overhangs. In this sense, *LoopDesigner* generalises the concept of the assembly, so that the assembly schema presented in this paper become a single instance of many possible implementations of the Loop assembly, allowing for the exploration of novel ways of assembling DNA parts through Type IIS strategies.

DNA construction has been traditionally coupled to the concurrent use of plasmids in model organisms. Loop assembly provides additional throughput and versatility for working with general-purpose backbones, to which users can add specific new functions e.g. parts for transfection. Vectors could be decoupled from specific uses by modularising replication origins and selection markers as basic DNA parts and introducing host-specific elements during the assembly process. This would provide higher flexibility during design, and allow switching selection markers when super-transformation is required, for instance. Such approaches would make the DNA fabrication process host-agnostic, promoting the development of universal DNA assembly systems using standards such as the common syntax, which would provide unprecedented exchange of DNA components within the biological sciences.

Until recently, the majority of materials for research were exchanged under a Uniform Biological Material Transfer Agreement (UB-MTA). This is a bilateral legal agreement that, in its standard form, does not allow redistribution, exchange or use with those outside of educational and research organisations. At the same time, in scientific publishing and in software, there is a trend toward openness to facilitate collaboration and translation of basic research. An excellent example of how the open source philosophy has powered and enabled innovation is exemplified by community-based coding projects such as the ones hosted by *Github* (https://github.com). *Git* was originally developed for the purpose of distributed software development, and nowadays most collaborative projects, both in the public and the private sector, use *Git* as an underlying framework. It is unlikely that we will see similar success in DNA engineering and synthetic biology, unless new forms of unrestricted DNA

sharing and assembly are established under more open frameworks such as the OpenMTA. We support the adoption of an open-source inspired L0 elemental part exchange by providing Loop assembly for the higher-level construction of these L0 components under an OpenMTA framework. Work to establish the OpenMTA will ensure access to the Loop assembly system for work in both the public and the private sector.

Acknowledgements We would like to acknowledge Diego Orzaez from Instituto de Biología Molecular y Celular de Plantas for helpful suggestions and advice regarding SapI and Type IIS restriction enzymes. Oleg Raitskin for advice on Type IIS assembly protocols, H. Ghareeb from Göttingen University, T. W. J. Gadella, and J. Goedhart from University of Amsterdam for providing mTurquoise2 DNA, Jen Sheen from Harvard University for providing TCS promoter DNA, Teva Vernaux from of Ecole Normale Supérieure de Lyon for Venus N7 DNA, and Jeff Lichtman from Harvard University for mTagBFP2 DNA. Susana Sauret-Gueto, Eftychis Frangedakis and Marta Tomaselli from the University of Cambridge for troubleshooting SapI reactions. We thank Linda Kahl and Drew Endy, BioBricks Foundation for discussions about implementation of the OpenMTA. Support for the authors was provided by Becas Chile and the Cambridge Trust (to B.P.), University of Cambridge BBSRC DTP programme (to M.D.), and the Biotechnology and Biological Sciences Research Council and Engineering and Physical Sciences Research Council [OpenPlant Grant No. BB/L014130/1] (to N.P., F.F. and J.H.). Laboratory automation, next-generation sequencing and library construction was delivered via the BBSRC National Capability in Genomics (BB/CCG1720/1) at Earlham Institute. F.F. acknowledges funding from CONICYT Fondecyt Iniciación 11140776. F.F. and R.A.G. acknowledge funding from Fondo de Desarrollo de Areas Prioritarias (FONDAP) Center for Genome Regulation (15090007) and Millennium Nucleus Center for Plant Systems and Synthetic Biology (NC130030). The authors declare that they have no competing financial interests.

Author contributions BP, FF & JH conceived the study. BP & FF developed the Loop assembly logic. BP designed and implemented the Loop assembly schema, built the Loop assembly library, optimised reaction conditions, validated recursive assembly and performed characterisation of Loop assembly efficiency. BP & MD built the Loop assembly constructs. TM, SA and FF designed the UNS for plants. SA did preliminary experiments to validate UNS-mediated assembly. AC assembled Loop constructs through UNS-overlap assembly and collected expression data on *Arabidopsis thaliana* protoplasts. MD & BP

designed LoopDesigner and MD developed PartsDB & LoopDesigner. AW performed HT sequencing on Loop assembly constructs. NP advised regarding Type IIS restriction endonucleases. BP performed the *Marchantia* work and collected the microscopy data. FF and JH supervised the project. FF, RAG & JH secured funding for the development of the project. BP, MD, RAG, NP, FF and JH wrote the manuscript. All authors revised the manuscript.

References

Althoff F, Kopischke S, Zobell O, Ide K, Ishizaki K, Kohchi T, Zachgo S. 2014. Comparison of the MpEF1 and CaMV35 promoters for application in *Marchantia polymorpha* overexpression studies. *Transgenic Research* **23**:235–244.

Benders GA, Noskov VN, Denisova EA, Lartigue C, Gibson DG, Assad-Garcia N, Chuang RY, Carrera W, Moodie M, Algire MA *et al.* 2010. Cloning whole bacterial genomes in yeast. *Nucleic Acids Research* **38**:2558–2569.

Benfey PN, Chua NH. 1990. The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. *Science* **250**:959–966.

Beyer HM, Gonschorek P, Samodelov SL, Meier M, Weber W, Zurbriggen MD. 2015. AQUA cloning: a versatile and simple enzyme-free cloning approach. *PLoS ONE* **10**:e0137652.

Binder A, Lambert J, Morbitzer R, Popp C, Ott T, Lahaye T, Parniske M. 2014. A modular plasmid assembly kit for multigene expression, gene silencing and silencing rescue in plants. *PLoS ONE* **9,** e88218.

Bitinaite J, Rubino M, Varma KH, Schildkraut I, Vaisvila R, Vaiskunaite R. 2007. USER™ friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Research* **35**:1992–2002.

Bryksin AV, Matsumura I. 2010. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *BioTechniques* **48**:463–465.

Engler C, Kandzia R, Marillonnet S. 2008. A one pot one step precision cloning method with high throughput capability. *PLoS ONE* **3**:e3647.

Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JDG, Patron NJ, Marillonnet S. 2014. A Golden Gate modular cloning toolbox for plants. *ACS Synthetic Biology* **3**:839–843.

Flores-Sandoval E, Eklund DM, Bowman JL. 2015. A simple auxin transcriptional response system reg- ulates multiple morphogenetic processes in the liverwort *Marchantia polymorpha*. *PLoS Genetics* **11**:e1005207.

Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA *et al.* 2008a. Complete chemical synthesis assembly and cloning of a *Mycoplasma genitalium* genome. *Science* **319**:1215–1220.

Gibson DG, Benders GA, Axelrod KC, Zaveri J, Algire MA, Moodie M, Montague MG, Venter JC, Smith HO, Hutchison CA. 2008b. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proceedings of the National Academy of Sciences USA* **105**:20404–20409.

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**:343–345.

Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM *et al.* 2010a. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**:52–56.

Gibson DG, Smith HO, Hutchison CA, Venter JC, Merryman C. 2010b. Chemical synthesis of the mouse mitochondrial genome. *Nat Meth* **7,** 901–903.

Giniger E, Varnum SM, Ptashne M. 1985. Specific DNA binding of GAL4 a positive regulatory protein of yeast. *Cell* **40**:767–774.

Guarente L, Yocum RR, Gifford P. 1982. A GAL10-CYC1 hybrid yeast promoter identifies the GAL4 regulatory region as an upstream site. *Proceedings of the National Academy of Sciences USA* **79**:7410–7414.

Haseloff J. 1999. GFP variants for multispectral imaging of living cells. *Methods Cell Biology* **58**:139–151.

Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM. 2000. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Molecular Biology* **42**:819–832.

Ishizaki K, Chiyoda S, Yamato KT, Kohchi T. 2008. *Agrobacterium*-mediated transformation of the haploid liverwort *Marchantia polymorpha* L. an emerging model for

plant biology. *Plant Cell Physiology* **49**:1084–1091.

Iverson SV, Haddock TL, Beal J, Densmore DM. 2016. CIDAR MoClo: improved MoClo assembly standard and new *E. coli* part library enable rapid combinatorial design for synthetic and traditional biology. *ACS Synthetic Biology* **5**:99–103.

Jin P, Ding W, Du G, Chen J, Kang Z. 2016. DATEL: a scarless and sequence-independent DNA assembly method using thermostable exonucleases and ligase. *ACS Synthetic Biology* **5**:1028–1032.

Joska TM, Mashruwala A, Boyd JM, Belden WJ. 2014. A universal cloning method based on yeast homologous recombination that is simple efficient and versatile. *Journal of Microbiological Methods* **100**:46–51.

Kanigowska P, Shen Y, Zheng Y, Rosser S, Cai Y. 2016. Smart DNA fabrication using sound waves: applying acoustic dispensing technologies to synthetic biology. *Journal of Laboratory Automation* **21**:49–56.

Karas BJ, Molparia B, Jablanovic J, Hermann WJ, Lin YC, Dupont CL, Tagwerker C, Yonemoto IT, Noskov VN, Chuang RY *et al.* 2013. Assembly of eukaryotic algal chromosomes in yeast. *Journal of Biological Engineering* **7**:30.

Karas BJ, Tagwerker C, Yonemoto IT, Hutchison CAIII, Smith HO. 2012. Cloning the *Acholeplasma laidlawii* PG-8A genome in *Saccharomyces cerevisiae* as a yeast centromeric plasmid. *ACS Synthetic Biology* **1**:22–28.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C *et al.* 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**:1647–1649.

Keohavong P, Thilly WG. 1989. Fidelity of DNA polymerases in DNA amplification. *Proceedings of the National Academy of Sciences* **86,** 9253–9257.

Knight T. 2003. *Idempotent vector design for standard assembly of Biobricks*. Technical report MIT Synthetic Biology Working Group MIT Artificial Intelligence Laboratory Boston MA USA.

Lampropoulos A, Sutikovic Z, Wenzl C, Maegele I, Lohmann JU, Forner J. 2013. GreenGate---a novel, versatile, and efficient cloning system for plant transgenesis. *PLoS ONE* **8,** e83043.

Li MZ, Elledge SJ. 2007. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nature Methods* **4**:251–256.

Lin D. O'Callaghan CA. 2018. MetClo: methylase-assisted hierarchical DNA assembly using a single type IIS restriction enzyme. *Nucleic Acids Research*. doi:10.1093/nar/gky596

Ma H, Kunes S, Schatz PJ, Botstein D. 1987. Plasmid construction by homologous recombination in yeast. *Gene* **58**:201–216.

Moore SJ, Lai HE, Kelwick RJR, Chee SM, Bell DJ, Polizzi KM, Freemont PS. 2016. EcoFlex: a multifunctional MoClo kit for *E. coli* synthetic biology. *ACS Synthetic Biology* **5**:1059–1069.

Müller B, Sheen J. 2008. Cytokinin and auxin interaction in root stem-cell specification during early embryogenesis. *Nature* **453**:1094–1097.

Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. 2002. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature Biotechnology* **20,** 87–90.

Nagaya S, Takemura M, Ohyama K. 2011. Endogenous promoter 5 ′-UTR and transcriptional terminator enhance transient gene expression in a liverwort *Marchantia polymorpha* L. *Plant Biotechnology Journal* **28**:493– 496.

Nielsen AAK, Der BS, Shin J, Vaidyanathan P, Paralanov V, Strychalski EA, Ross D, Densmore D, Voigt CA. 2016. Genetic circuit design automation. *Science* **352**:aac7341–aac7341.

Patron NJ. 2016. *DNA Assembly for Plant Biology* volume 16 of *DNA Assembly for Plant Biology*. John Wiley & Sons Inc. Hoboken NJ USA.

Patron NJ, Orzaez D, Marillonnet S, Warzecha H, Matthewman C, Youles M, Raitskin O, Leveau A, Farré G, Rogers C *et al.* 2015. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytologist* **208**:13–19.

Potapov V, Ong JL. 2017. Examining Sources of Error in PCR by Single-Molecule Sequencing. *PLoS ONE* **12,** e0169774.

Potapov V, Ong JL, Kucera RB, Langhorst BW, Bilotti K, Pryor JM, Cantor EJ, Canton B, Knight TF, Evans TC *et al.* 2018. Optimization of Golden Gate assembly through application of ligation sequence-dependent fidelity and bias profiling. *bioRxiv* 322297. DOI:10.1101/322297.

Richardson SM, Mitchell LA, Stracquadanio G, Yang K, Dymond JS, DiCarlo JE, Lee D, Huang CLV, Chandrasegaran S, Cai Y, *et al.* 2017. Design of a synthetic yeast genome. *Science* **355**:1040–1044.

Sarrion-Perdigones A, Falconi EE, Zandalinas SI, 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**:e21622.

Sarrion-Perdigones A, Vazquez-Vilar M, Palací 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**:1618–1631.

Shetty RP, Endy D, Knight TF. 2008. Engineering BioBrick vectors from BioBrick parts. *Journal of Biological Engineering* **2**:5.

Storch M, Casini A, Mackrow B, Fleming T, Trewhitt H, Ellis T, Baldwin GS. 2015. BASIC: a new biopart assembly standard for idempotent cloning provides accurate single-tier DNA assembly for synthetic biology. *ACS Synthetic Biology* **4**:781–787.

Temme K, Zhao D, and Voigt CA. 2012. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proceedings of the National Academy of Sciences USA* **109**:7085–7090.

Torella JP, Boehm CR, Lienert F, Chen JH, Way JC, Silver PA. 2013. Rapid construction of insulated genetic circuits via synthetic sequence-guided isothermal assembly. *Nucleic Acids Research* **42**:gkt860– 689.

Watson MR, LinYF, Hollwey E, Dodds RE, Meyer P, McDowall KJ. 2016. An Improved Binary Vector and Escherichia coli Strain for Agrobacterium tumefaciens-Mediated Plant

Transformation. *G3* **6,** 2195–2201.

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S. 2011. A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* **6**:e16765–11.

West RW Jr, Yocum RR, Ptashne M. 1984. *Saccharomyces cerevisiae* GAL1-GAL10 divergent promoter region: location and function of the upstream activating sequence UASG. *Molecular and Cellular Biology* **4**:2467– 2478.

Yoo SD, Cho YH, Sheen J. 2007. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols* **2**:1565–1572.

Zhang L, Guarente L. 1994. The yeast activator HAP1–a GAL4 family member–binds DNA in a directly repeated orientation. *Genes & Development.* **8**:2110–2119.

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

Zhu B, Cai G, Hall E, Freeman G. 2007. In-Fusion™ assembly: seamless engineering of multidomain fusion proteins modular vectors and mutations. *BioTechniques* **43**:354–359.

Fig. 1 Overview of Loop assembly. (a) Loop assembly workflow, L0 parts are assembled to L1 transcription units (TUs) into one pOdd receiver by BsaI-mediated Type IIS assembly. L1 TUs are assembled to L2 multi-TUs into one pEven receiver by SapI-mediated Type IIS assembly. This workflow is then repeated for higher-level assemblies. Only four odd level and four even level receiver plasmids are required for Loop assembly. (b) Combinatorial and exponential assembly. L0 parts can be assembled to L1 TUs into any of the four positions of odd receivers. Genetic modules can be easily be swapped in each TU arrangement and receiver position. L1 TUs can then be assembled into L2 multi-TUs with variable combinations of the L1 TUs, also into any of the four positions of the even receivers. Each round of assembly generates four assembled plasmids and consequent rounds of assembly increase the number of TUs by a factor of four, leading to an exponential increase in TU number.

Fig. 2 Loop assembly schema. (a) Loop receiver plasmids. Each of the four pOdd and pEven receivers plasmids has a specific set of SapI (3 bp) and BsaI (4 bp) convergent overhangs respectively, required for higher level assembly. Odd receivers contain diverging BsaI restriction sites and terminal-overhangs according to the common syntax, making them compatible for cloning L0 parts into pOdd plasmids. They contain SapI converging sites with donor-overhangs for directing SapI-mediated Type IIS assembly into even-level receivers. pEven plasmids have SapI diverging restriction sites and terminal-overhangs to receive parts from pOdd plasmids. For higher-level assemblies, pEven plasmids contain converging BsaI sites with donor-overhangs for BsaI-mediated Type IIS assembly into pOdd plasmids. (b) Loop odd-level assembly. L0 DNA parts containing overhangs defined in the common syntax are assembled into a Loop odd-level receiver. BsaI digestion releases the DNA modules, which are assembled into an even-level receiver by directional assembly defined by 4 bp overhangs. pOdd plasmids contain A and F overhangs as terminal-overhangs for receiving parts, which are flanked by convergent SapI restriction sites with 3 bp donor-overhangs for further assembly. (c) Loop even-level assembly. Four previously assembled pL1 transcription units (TUs) are assembled into a pEven plasmid. SapI digestion releases TUs from pL1 plasmids, which are assembled into an even-level receiver by directional assembly defined by 3 bp overhangs. pEven plasmids contain and overhangs as terminal-overhangs, which are

flanked by convergent BsaI restriction sites with donor-overhangs defined in the common syntax required for further assembly.

Fig. 3 Hybrid assembly. (a) Synthetic promoter assembly. L0 functional domains flanked by terminal overhangs are assembled into odd level receivers in any given position. L1 motifs are then assembled into L2 composites with differing arrangements into positions 1 and 2. L2 composites in positions 1 and 2 are used in a hybrid assembly with L0 parts to generate a hybrid odd level transcription unit (TU) with a synthetic promoter composed of the L0 functional domains in the defined arrangement. (b) Mixed level assembly. L3 and L1 parts are assembled into an even level receiver generating a hybrid even level multi-TU plasmid.

Fig. 4 Loop overlap assembly. (a) Transcription unit (TU) assembly for overlap assembly. Unique nucleotide sequences (UNS) flanked transcription units (TUs) can be generated either by standard L0 BsaI-mediated Type IIS assembly or by overlap assembly methods using PCR-fragments or DNA synthesis. TUs produced by overlap assembly are only compatible with the overlap assembly pathway but do not require domestication. (b) Standardised overlap assembly. Linear UNS flanked TUs are amplified by PCR or excised from plasmids by digestion by uncommon restriction enzymes. Linear UNS flanked TUs are then assembled to the destination plasmid pUNSDest by overlap assembly methods.

Fig. 5 *In planta* activity of a Loop assembly construct. *Marchantia gemmae* transformed with a L2 construct were imaged through a Leica SP8 laser-scanning confocal microscope to assess expression of fluorescent markers. mTurquoise2-N7, Venus-N7 and eGFP-Lti6b were excited with appropriate wavelengths and fluorescence was captured in their respective emission windows in sequential scanning mode. Images shown are Z-stack maximum intensity projections.

Fig. 6 Design automation. A DNA sequence is submitted to *LoopDesigner*, which screens for BsaI and SapI sites and domesticates them to silent mutations where possible. A *part* type is specified for the assembly schema to save the *part* to the database library. To perform an *in silico* assembly, a receiver plasmid is selected which displays the compatible *parts* that can be placed in the current position of the assembly schema. As *parts* are included, the next compatible *parts* are displayed. When the assembly schema finds that the all *parts* required to complete the assembly are selected, the assembly simulation is performed. Then, *LoopDesigner* outputs the resulting plasmid map with its concurrent highlighted sequence and a protocol for Loop Type IIS reaction setup or export of GenBank sequence. Instructions to robots can be outputted if an API is provided with the required information (plasmid positions, ID mappings, robot functions) to produce the concurrent instruction file using *Python* scripting. The assembled *part* is then saved into the *part* library database for further assembly.

Table 1 Loop assembly efficiency.

Level	Constructs (Nº)	TU (Nº)	Avr. length (bp)	O. efficiency* (%)	Avr. efficiency† (%)
Lab 1					
L1	104	1	6243	96	97
L2	79	4	13519	82	88
L3	23	16	26731	81	83
Hybrid	3	Var.	5473	100	100
Lab 2					
L1	14	1	5570	91	91
UNS overlap	5	4	12548	71	71

*Overall efficiency calculated as total number of samples with correct RD patterns over total samples tested.

†Average efficiency calculated as the mean of correct RD patterns over samples tested per construct.

UNS, unique nucleotide sequences; TU, transcription unit.

Supporting Information