A framework for the modular and combinatorial assembly of synthetic gene circuits

Javier Santos-Moreno* and Yolanda Schaerli*

Department of Fundamental Microbiology, University of Lausanne, Biophore Building, 1015 Lausanne, Switzerland

ABSTRACT: Synthetic gene circuits emerge from iterative design-build-test cycles. Most commonly, the time-limiting step is the circuit construction process. Here, we present a hierarchical cloning scheme based on the widespread Gibson assembly method and make the set of constructed plasmids freely available. Our two-step modular cloning scheme allows for simple, fast, efficient and accurate assembly of gene circuits and combinatorial circuit libraries in *Escherichia coli*. The first step involves Gibson assembly of transcriptional units from constituent parts into individual intermediate plasmids. In the second step, these plasmids are digested with specific sets of restriction enzymes. The resulting flanking regions have overlaps that drive a second Gibson assembly into a single plasmid to yield the final circuit. This approach substantially reduces time and sequencing costs associated with gene circuit construction and allows for modular and combinatorial assembly of circuits. We demonstrate the usefulness of our framework by assembling a CRISPR-based double-inverter circuit and a combinatorial library of 3-node networks.

Synthetic biology relies on a design-build-test process. Decreasing costs of DNA synthesis have favored the construction of increasingly complex synthetic gene circuits[1, 2]. However, this is often accompanied by a rise in time spent for molecular cloning, which can actually account for a substantial fraction of researchers' activity. Due to technical limitations in the number of fragments that are efficiently assembled in a single reaction, complex gene networks usually cannot be constructed in a single step. Sequential cloning of circuit sub-components becomes tedious and time-consuming with increasing circuit complexity. Therefore, fast and efficient DNA assembly schemes are needed more than ever.

For most of current cloning methods a trade-off exists between modularity and seamless assembly. Golden Gate cloning[3] and Gibson assembly[4] are two popular cloning schemes that illustrate this trade-off. Golden Gate cloning and its derivatives, as well as other restriction-based methods such as the BioBrick standard[5], offer modular assembly of parts and the possibility to easily build combinatorial libraries. However, the assembly process often leaves behind "scar" sequences between adjacent parts, which can affect the performance of the final construct, especially if present in coding sequences (CDSs) or untranslated regions (UTRs). Besides, these methods require a substantial *a priori* effort to obtain the required set of modular parts in the appropriate format. On the other hand, Gibson assembly and its derivatives, as well as other overlap-directed DNA assembly methods (e.g. SLIC[6], SLiCE[7], CPEC[8]), allow for seamless cloning. Nevertheless, this comes at the cost of modularity, since the interface between two assembled parts is usually unique and needs to be re-designed for each new combination. The need to design and synthesize unique overlapping sequences increases the assembly cost and delays the design-

build-test cycle, and it makes these approaches less suitable for combinatorial assembly.

Recently, efforts have been made to address this trade-off, for example in the form of a modular variant of Gibson assembly known as MODAL[9] or the Start-Stop assembly[10] method that aims at modifying Golden Gate for seamless cloning. However, these new approaches are not exempt from the modularity-seamless trade-off, and thus increased modularity intrinsically requires linker sequences between parts[9], and similarly a lack of between-part linkers comes at the cost of modularity[10]. These constraints might limit the widespread adoption of these assembly methods as compared to the parental Gibson and Golden Gate approaches.

For the assembly of complex gene circuits, a hierarchical cloning scheme is probably the best choice in terms of efficiency and speed[11-13]. Cloning all parts in parallel within a single reaction soon becomes unrealistic as the number of parts increases. A sequential scheme divides the process in consecutive steps, which becomes lengthy as the number of steps increases. The strength of hierarchical cloning schemes relies on the division of the assembly process into discrete steps or levels, in which parallel assembly reactions are performed within a given level before their products are used as building blocks for the next level. For instance, MoClo[13] is a hierarchical extension of Golden Gate in which parts such as promoters, UTRs, CDSs and terminators are first assembled into transcriptional units, which are subsequently combined into multigene constructs. 3G assembly[14] combines Golden Gate and Gibson assembly in a hierarchical scheme that allows for single-day construction of complex circuits.
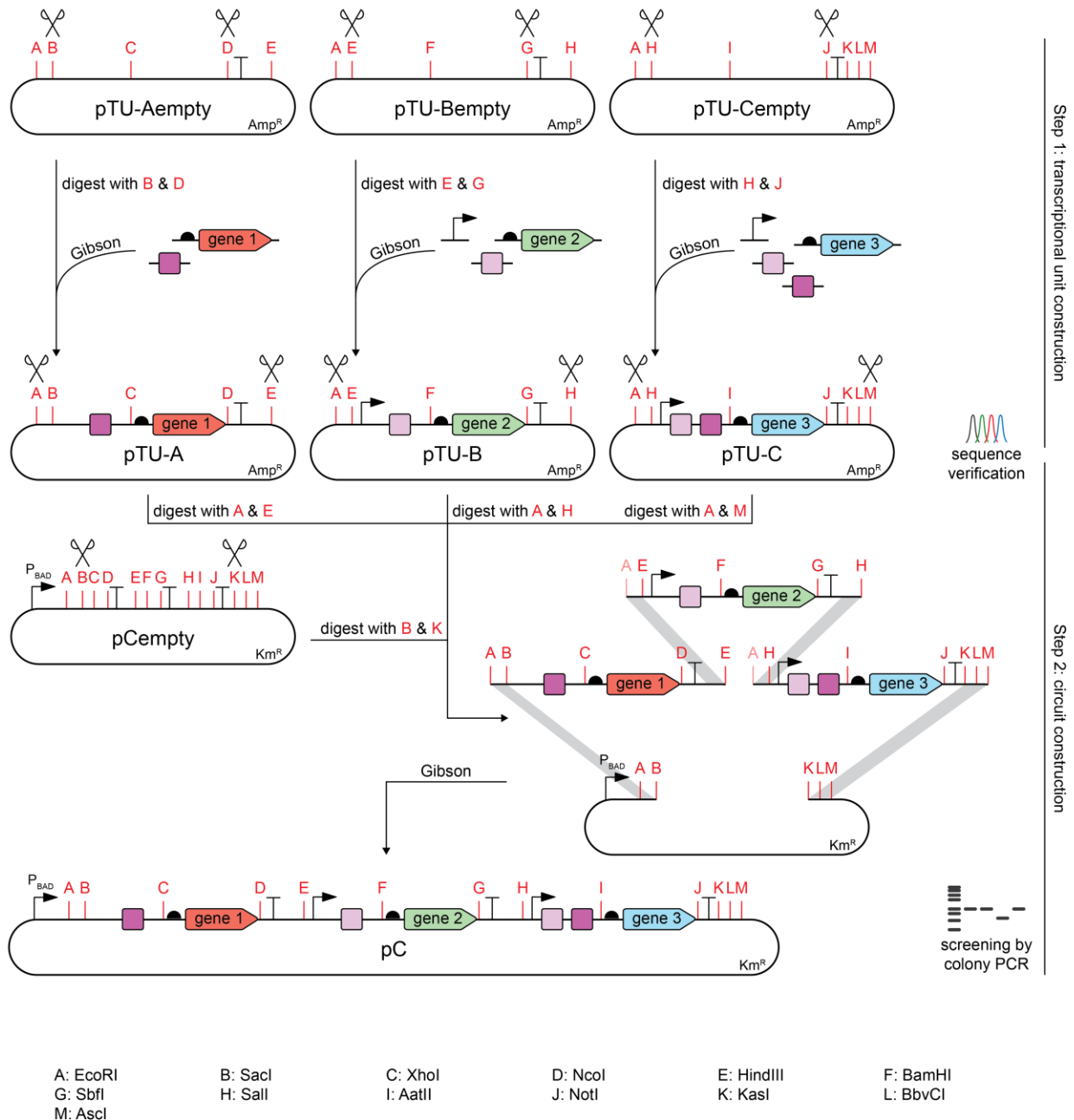
Despite modularity and hierarchical assembly of Golden Gate variants, Gibson assembly remains the most popular cloning method among synthetic biologists[15]. The reasons

behind this may include a need or preference for seamless assembly, or the complex design and implementation of Golden Gate schemes, which can be daunting especially during the initial setup.

Here, we present a hierarchical cloning scheme based on the widespread Gibson assembly method. Starting from basic parts such as promoters, regulators, regulator binding sites or coding sequences, we first assemble transcriptional units (TUs) into intermediate plasmids. Next, we combine transcriptional units to clone gene circuits into the final recipient vector. Our hierarchical assembly scheme allows for modular and combinatorial assembly of gene circuits in only two steps in *Escherichia coli (E. coli)*. We demonstrate its usefulness by assembling a double-inverter circuit with almost 100% accuracy and a combinatorial library of 3-node networks.

RESULTS AND DISCUSSION

First, we designed a multiple cloning site (MCS) for the construction of 3-node gene regulatory networks, and cloned it into a plasmid containing AraC, P_{BAD}, a ColA replicon and kanamycin resistance[16] to yield pCempty (Figure 1). The MCS contains three "slots" flanked by unique restriction enzyme (RE) sites for the insertion and removal of transcriptional units. Additionally, each slot contains an extra internal RE site, resulting in a total of three unique RE sites per slot. This allows us to keep a part permanently in the plasmid (e.g. a fluorescent reporter) and easily add extra parts up- or downstream. Slots are followed by strong Rho-independent transcriptional terminators

A: EcoRI   B: SacI   C: XhoI   D: NcoI   E: HindIII   F: BamHI
G: SbfI   H: SalI   I: AatII   J: NotI   K: KasI   L: BbvCI
M: AscI

Figure 1. Schematic representation of a 3-node circuit constructing framework. The assembly occurs in two steps: during step 1 the individual transcriptional units (TUs) are assembled from their constituent parts into the intermediate plasmids (pTU-A, pTU-B and pTU-C), and in step 2 all the transcriptional units are combined into the receiver vector pCempty to yield the plasmid with the final circuit (pC). Both steps rely on RE digestion for vector linearization and on Gibson assembly for the directional, seamless cloning, and both allow for modular and combinatorial assembly. Note that during step 2 the EcoRI sites in TU-B and TU-C are not reconstituted in the final construct pC. Sanger sequencing verification is required at the end of step 1, but not after step 2. Bent arrows: promoters; squares: regulator binding sites; pointed rectangles: genes; semicircles: ribosome binding sites (RBS); and "T"-s: transcriptional terminators.
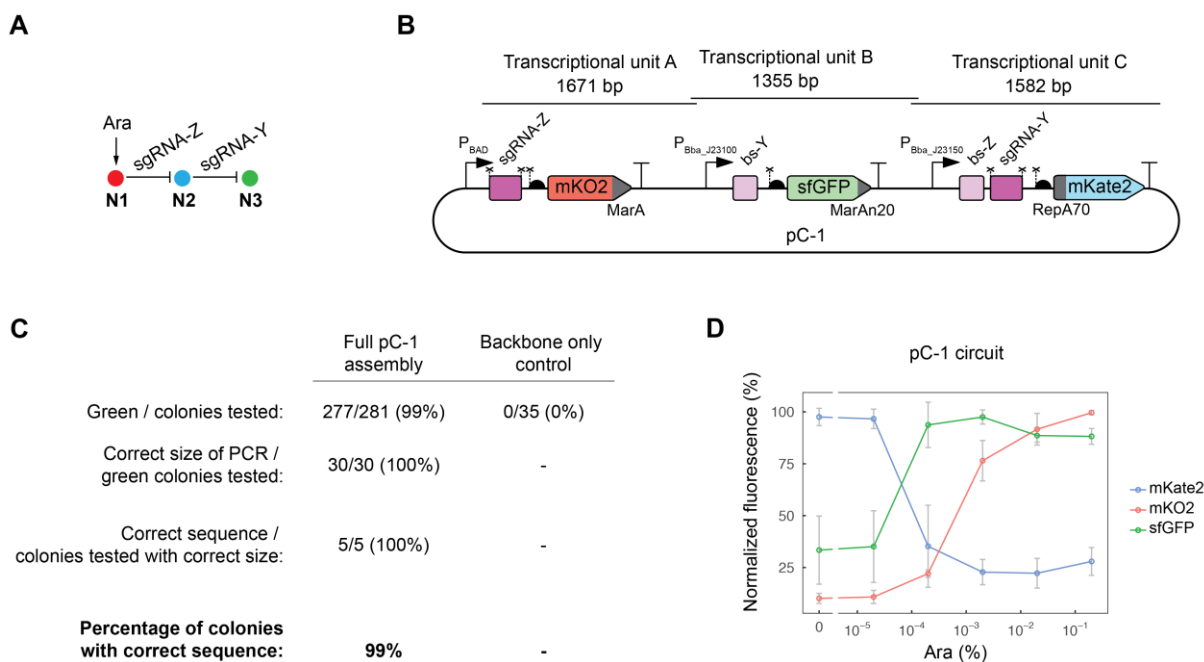
Figure 2. Implementation of the 3-node assembly strategy. A. The double-repression logic displayed by the 3-node circuit is depicted. B. Main relevant features of its implementation in pC-1, with the length of transcriptional units A, B and C shown. Bent arrows: promoters; squares: sgRNA binding sites; rectangles: sgRNAs; crosses: csy4 recognition sites; semicircles: RBSs; pointed rectangles: reporter genes with degradation tags; and "T"-s: transcriptional terminators. C. Accuracy of pC-1 assembly from its constituent single-node plasmids, as well as a control for backbone re-circularization. Three readouts with different throughput were used for determining accuracy: the fraction of green colonies (indicative of sfGFP presence), the fraction of green colonies carrying an insert of the correct length, and the fraction of colonies carrying a good-sized insert that showed no change at the nucleotide sequence level. Combining these data, we calculated the percentage of transformants that carried the correct sequence. Data based on two independent experiments. D. Assessment of the behavior of six non-sequenced transformants carrying the pC-1 circuit. Fluorescence levels in response to different arabinose concentrations indicate that all six transformants bear the pC-1 circuit that displays the expected double-repression logic. Mean and s.d. from six biological replicates.

to avoid read-through, and isolated from each other by biologically neutral 200bp spacers (designed with R2oDNA[17]) to minimize compositional context. The RE sites are separated by unique short spacers (12 or 30bp) that are lost during standard column purification of DNA, thus avoiding re-ligation. To allow for the hierarchical combinatorial assembly of networks with three transcriptional units, we split the MCS in three units, each harboring a slot, and each unit was cloned into plasmids with ampicillin resistance, giving the starting plasmids pTU-Aempty, pTU-Bempty and pTU-Cempty (Figure 1).

To construct a 3-node circuit, we carry out two steps (Figure 1). In step 1, the starting plasmids pTU-Aempty, pTU-Bempty and pTU-Cempty are opened with the appropriate REs, and the parts that will compose the transcriptional units are inserted using Gibson assembly. As a result, we obtain intermediate plasmids pTU-A, pTU-B and pTU-C, each one carrying

one of the transcriptional units of the final circuit. Step 1 is modular in the sense that different promoters can be combined with different operators and coding sequences. This step also allows for the combinatorial cloning of a library of parts. Importantly, it is also possible to keep specific parts, such as fluorescent reporters, permanently in the starting plasmids, so that only variable parts (e.g. promoters and operators) need to be cloned between the remaining restriction sites.

In step 2, all three intermediate plasmids are digested with EcoRI plus an additional RE (HindIII for pTU-A, SalI for pTU-B and AscI for pTU-C), while the receiver vector pCempty is digested with SacI and KasI. As a result, transcriptional units A, B and C are flanked by overlapping regions that drive their directional A-B-C assembly into the linearized pCempty receiver plasmid. Digestions are column-purified and assembled in a Gibson reaction, resulting in the pC plasmid

harboring the final user-designed circuit. Importantly, intermediate pTU-A, pTU-B and pTU-C vectors carry a different selection marker (Amp^R) than the receiver pCempty plasmid (Km^R) to avoid the recovery of non-digested/re-ligated pTU vectors at the end of step 2. Like step 1, step 2 is also modular (any pTU-A can be combined with any pTU-B and any pTU-C) and allows for the combinatorial cloning of transcriptional unit libraries. Each of the steps is performed in a single day; considering the time needed for assembly, bacterial growth and construct screening, a new synthetic gene circuit can be obtained from its basic constituent parts within 5 days. Once a library of TUs is available, composed circuits are assembled and screened within 24 hours.

To demonstrate the usefulness of our assembly method, we decided to assemble a 3-node gene regulatory network. The circuit is activated by arabinose (Ara) and deploys a double-inverter logic in which the first node (N1) represses the second node (N2), which subsequently represses the third node (N3) (Figure 2A and 2B). Repression is implemented through CRISPR interference (CRISPRi)[18], for which a second plasmid carrying dCas9 (pJ1996_v2) is needed. This second plasmid also carries Csy4, a CRISPR endonuclease that recognizes and cleaves short RNA sequences[19]. We use csy4 recognition sites to "isolate" functional parts transcribed together within the same TU, so that once transcribed and physically separated by Csy4 they can act
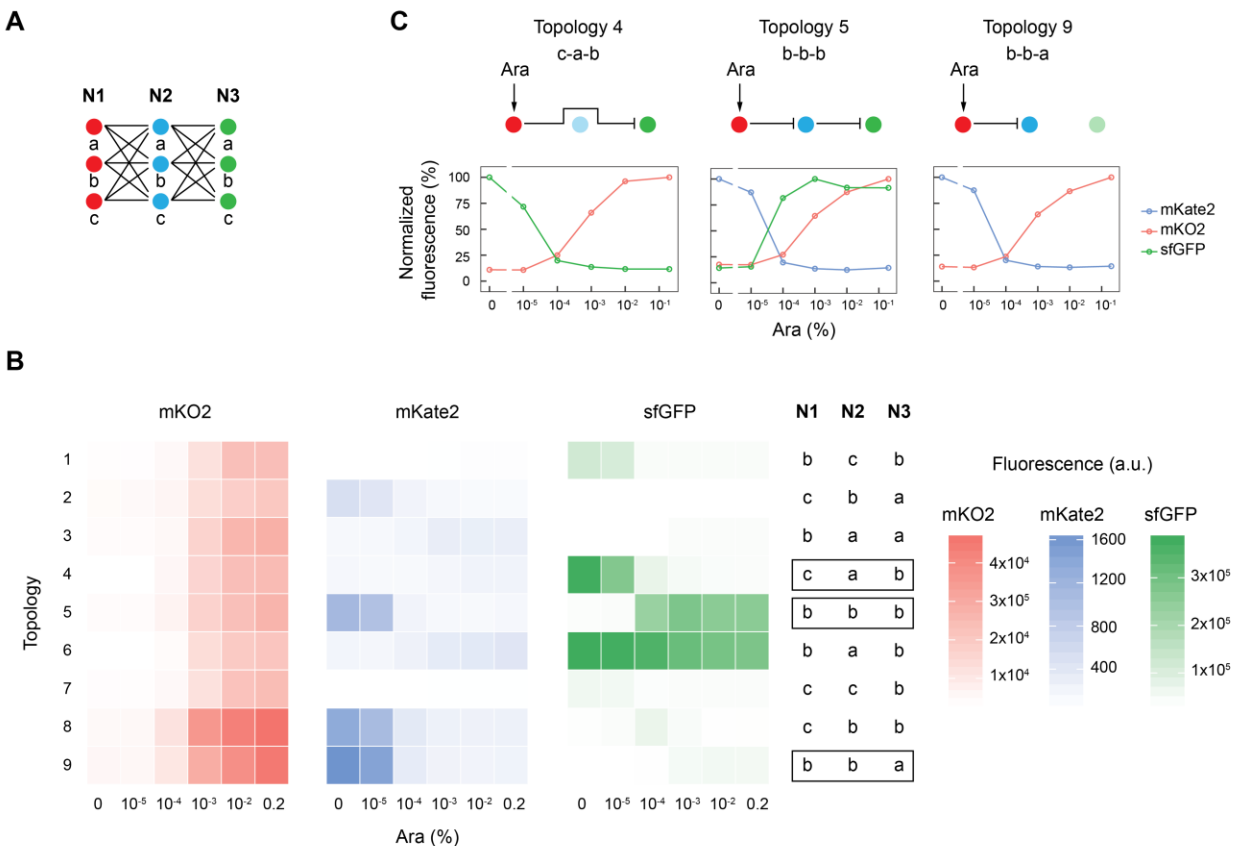
Figure 3. Combinatorial assembly of 3-node networks. A. Schema depicting all 27 different constructs that could result from the combinatorial assembly of three different N1 nodes, together with three N2 and three N3. B. Fluorescence levels of the 9 different topologies characterized in response to six arabinose concentrations. The identity of each construct is indicated on the right. C. Three topologies are shown in detail. The construct identity and the network connectivity are indicated, and normalized fluorescence levels are plotted below. N1 is always under the control of P$_{BAD}$ promoter. N2 and N3 are either under the control of constitutive promoters (strong shading) or lack any promoter (weak shading).

independently of each other. All three nodes carry a fluorescent reporter (mKO2, sfGFP and mKate2)[20-22] tagged with a degradation tag (MarA, MarAn20 and RepA70[23], respectively). N1 is under the control of a P$_{BAD}$ promoter and produces a single-guide RNA (sgRNA-Z) that represses N2 by binding downstream of the constitutive promoter (BBa_J23150). N2 in turn produces sgRNA-Y that represses expression of N3 controlled by the BBa_J23100 constitutive promoter. Thus, when subjected to increasing concentrations of Ara, N1 levels should rise while N2 levels decrease and N3 levels, in turn, increase.

In step 1, we used pTU-Aempty, pTU-Bempty and pTU-Cempty variants with fluorescent reporter genes (mKO2,

sfGFP, mKate2 in pTU-A-005, pTU-B-005 and pTU-C-005, respectively) already present within the slots. We digested the three plasmids and inserted promoters, binding sites and sgRNAs upstream of the reporters using Gibson assembly. The sgRNAs were amplified from a storage vector and the promoters and binding sites were amplified from oligonucleotides. We used the MODAL strategy[9] to facilitate the modular insertion of parts in step 1, but this is not a prerequisite. In step 2, the resulting intermediate plasmids containing the transcriptional units were digested with appropriate REs (Figure 1) and assembled into a RE-linearized receiver vector (the pCempty variant pC-0) to give pC-1 (Figure 2B).

Since only colonies incorporating TU-B should show green fluorescence, we used a blue light transilluminator for a pre-screen for green colonies. 99% of the colonies were green (Figure 2C), indicating that they contained TU-B. We performed colony PCR on green colonies using primers that anneal to the backbone and amplify all three nodes and observed that all (100%) green colonies carried the expected A-B-C insert (Figure 2C). Among those with the expected insert, five were randomly chosen for Sanger sequencing and all contained the correct sequence (100%). In summary, starting from a plate containing transformant colonies and in the absence of any pre-screening, the chances of picking a colony with the intended sequence were close to 100% (Figure 2C).

When building circuits of long sequence length, their verification through Sanger sequencing can become expensive and time-consuming. The cloning scheme presented here reduces sequencing-related time- and money-costs, since sequence verification is only needed after step 1 – at the intermediate plasmid level (Figure 1). The large majority of the mutations detected originated from synthesized oligonucleotides used as ssDNA templates to generate short parts (e.g. promoters) or as primers. Once an intermediate plasmid has been sequence-verified, it can be used repeatedly to build multiple circuits. Gibson assembly uses a high-fidelity DNA polymerase to fill the short 3' overhangs, and thus the probability of mutations arising during Gibson is low[4]. Since step 2 does not involve the use of synthetic oligonucleotides, the accuracy of the sequence of the final constructs is virtually 100% (no mutations found in >41 kb sequenced). Sequencing of the final construct is thus not necessary if the presence of all the TUs can be confirmed by other means (e.g. by colony PCR) (Figure 2C).

To confirm that this assembly framework can be used to efficiently build complex circuits that behave as intended with only one sequencing round (at the end of step 1), we assessed the behavior of colonies carrying the double-inverter pC-1 circuit. Briefly, plasmid DNA was extracted from six random pC-1 transformants verified by colony PCR, and transformed into E. coli MK01[24] containing the plasmid carrying dCas9 and Csy4 (pJ1996_v2). Cells were grown in the presence of different Ara concentrations, and fluorescence was recorded. All six clones exhibited the expected circuit behavior consisting of a rise in mKO2 levels with increasing Ara concentration, concomitant with an mKate2 decrease and a sfGFP increase (Figure 2D).

We next sought to investigate whether our method could be used to assemble combinatorial libraries of 3-node networks. The potential for combinatorial assembly is a desirable feature of DNA assembly methods, since it enables the building of large numbers of constructs without the need to design and build them individually. These libraries can then be rapidly tested provided that an appropriate high-throughput screening method is put in place[25, 26].

We set up a one-pot step 2 reaction in which we added 3 different pTU-A variants, together with 3 pTU-B and 3 pTU-C, which could result in a maximum of 27 different constructs (Figure 3A and Supplementary Figure S1). Among these, some constructs may display varying repressive interactions between the three nodes, including the double-inverter (Figure 2), while others may lack any promoter and connectivity for some of the nodes. We determined the identity of 24 randomly picked transformants through Sanger sequencing, which revealed the presence of 9 different topologies, indicating the intended

exploration of design space (Figure 3B). We measured fluorescence levels of the 9 different topologies and verified that they behave as expected according to their topology (Figure 3B and C). Therefore, our method provides a powerful platform for combinatorial assembly of gene circuits.

The choice of the assembly strategy is a multifactorial decision that includes, among others, the complexity of the intended construct, time and money constraints, the potential need for seamless constructs and the lab expertise. The number of available cloning schemes is large and steadily increasing, and there is no approach clearly superior to all the others: it is the specificity of the research project that determines the final choice. Here, we extend the palette of available cloning schemes by presenting a straightforward framework that allows for modular and combinatorial assembly of gene circuits.

Our method is based on the popular Gibson assembly, and provides time- and cost-effective, efficient and accurate construction of 3-node circuits using simple and robust molecular biology methods. The assembly reactions presented here involved inserts ranging from ~120 bp to ~1700 bp and up to 4 parts (1 vector + 3 inserts). However, Gibson assembly is also suited for larger inserts and a higher number of parts per reaction[4], and we have indeed successfully cloned 5- and 6-part constructs.

Compared to type IIS restriction enzyme-based multistep methods such as MoClo[13], 3G[14] or Loop assembly[27], our approach allows for a higher flexibility in the order of intra-TU parts, which is normally fixed in the former. Although we employed the MODAL strategy[9] to increase speed and modularity of TU assembly, this is not required, and a standard Gibson approach effectively yields seamless TUs, as opposed to Golden Gate-derived approaches[13, 14, 27]. New users might also find our approach more straightforward than other available methods.

We provide a set of 18 E. coli starting and receiver plasmids freely via Addgene (# 124409-124426; Table S1). In addition to the plasmids used for the examples described here, it also contains vectors with alternative resistance markers, origins of replication and without AraC and P_BAD promoter. While the system is optimized for 3-node circuits, an extra slot between RE sites KasI and BbvCI allows for the addition of a fourth TU to a 3-node construct. Alternatively, the system may be extended for simultaneous e.g. 4-node assembly by building a new starting plasmid (pTU-Dempty) and adapting pCempty so that it overlaps with the pTU-D-derived fragment. However, the use of large plasmids often poses technical challenges, and therefore users may choose to split complex circuits into two compatible vectors. In addition to the receiver vector pC-0 (kanamycin[R]), we developed and provide compatible variants pC-0_v3 and pC-0_v5 (ampicillin[R]). This allows for the easy construction of 6-node circuits (or even 8-node ones, if the 4th slots are used), without the need of further modifying the system. Moreover, the framework is not restricted to gene circuits, but could also be used to assemble other synthetic systems, such as metabolic pathways.

In summary, we present here a time- and cost-efficient cloning scheme based on the popular Gibson assembly method that allows for modular and combinatorial assembly of synthetic circuits with an almost 100% accuracy. We believe that it is thus of interest for other synthetic biologists.

5

METHODS

DNA fragment preparation. sfGFP was amplified from pET-23100-lacO(SymR+1)-GFP-LVA plasmid[16]. mKO2 was purchased as a gBlock from IDT. mKate2 was amplified from plasmid pLPT107[28], which was a gift from Johan Paulsson (Addgene plasmid # 85525). Csy4 was amplified from plasmid pCsy4[29], which was a gift from Adam Arkin & Stanley Qi (Addgene plasmid # 44252). dCas9 was amplified from pJWV102-dCas9sp plasmid, which was a gift from Jan-Willem Veening[30]. Protein coding genes and sgRNAs[31] were stored in a pJET1.2/blunt vector (Thermo Scientific) and amplified using KOD high-fidelity polymerase (Merck). For promoters and sgRNA binding sites, oligonucleotides were used as templates for KOD PCRs. All parts carried the same Prefix (CAGCCTGCGGTCCGG) and Suffix (TCGCTGGGACGCCCG) sequences for modular Gibson assembly using MODAL [9]. Forward and reverse primers annealed to Prefix and Suffix sequences, respectively, and added unique linkers to the parts. Primers (desalted) were purchased from Microsynth or Sigma-Aldrich. All annealing steps were performed at 60°C. PCR amplifications were column-purified using the Monarch PCR & DNA Cleanup Kit (NEB) following manufacturer's instructions. A list of primers can be found in the Supporting Information (Table S2).

DNA assembly and transformation. All digestions were performed for 1-2h at 37° with REs purchased from NEB, and purified using the Monarch PCR & DNA Cleanup Kit (NEB). Note that a complete digestion of receiver vectors in both step 1 and 2 is key to avoid the recovery of empty backbones. For step 1, plasmids were linearized using two REs (Figure 1) and PCR-amplified parts were assembled into the linear vectors using the NEBuilder HiFi DNA Assembly Master Mix (NEB) for 1h at 50°C, using a 2:1 insert:vector ratio or a 5:1 ratio when the insert was smaller than 250bp. For step 2, plasmids were digested as depicted in Figure 1, and purification was performed as mentioned above. Gibson assembly was performed with the NEBuilder HiFi DNA Assembly Master Mix (NEB) for 3h at 50°C using 50ng of receiver vector and a 2:1 insert:vector ratio. For the combinatorial assembly, variants of the same transcriptional unit were pooled for insert:vector ratio calculation. Of note, incubation of the Gibson reaction for shorter periods (1h) also results in very high assembly efficiency. 1µl of non-purifed Gibson reaction was transformed into 50µl of electrocompetent NEB5α cells, and 2/5 of them were plated onto selective agar plates.

Colony PCR. Colony PCRs were run using Taq polymerase (NEB) with Thermopol buffer. All annealing steps were performed at 54°C. PCR products with the expected size were column-purified with the Monarch PCR & DNA Cleanup Kit (NEB) and Sanger sequenced by GATC-Eurofins or Microsynth. Transformants with the correct sequences were grown in 5ml selective lysogeny broth (LB), and plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (QIAGEN).

Fluorescence measurements. MK01[24] electrocompetent cells were transformed with pJ1996_v2 plasmid (bearing dCas9 and Csy4) as well as with pC-1 or a pC plasmid derived from the combinatorial assembly of 3-node networks (Figure 3). Single colonies were used to inoculate 2ml of selective LB, and after ~6h cells were centrifuged at 4000rcf and resuspended in selective EZ medium (Teknova) with 0.4% glycerol as carbon source. 120µl of 0.05 OD600 bacterial suspensions were added per well on a 96-well CytoOne plate (Starlab), and 2.4µl of the indicated concentrations of L-arabinose (Sigma) were added. Plates were incubated at 37°C with double-orbital shaking in a Synergy H1 microplate reader (Biotek). Fluorescence was determined after 6h (for double-inverter) or 16h (for combinatorial library) with the following settings: mKO2: Ex. 542nm, Em. 571nm; sfGPF: Ex. 479nm, Em. 520nm; mKate2: Ex. 588nm, Em. 633nm. Fluorescence levels were corrected for fluorescence signal in a blank sample and for bacterial autofluorescence using a strain with no reporter genes and divided by OD600 to correct for differences in bacterial concentration.

ASSOCIATED CONTENT

Supporting Information

The following Supporting Information is available free of charge on the publisher website:

List of plasmids available through Addgene (Table S1).

List of primers for assembly, colony PCR and sequencing (Table S2) .

Supplementary Figure S1.

Annotated sequence of pC-1 circuit.

AUTHOR INFORMATION

Corresponding Author

* E-mail: javier.santosmoreno@unil.ch
* E-mail: yolanda.schaerli@unil.ch

Author Contributions

JSM and YS designed research. JSM performed experiments, analyzed data and prepared the figures. JSM and YS wrote the manuscript. All authors have given approval to the final version of the manuscript.

ABBREVIATIONS

Ara, Arabinose. CDS, Coding Sequence. MCS, Multiple Cloning Site. MODAL, Modular Overlap-Directed Assembly with Linkers. RBS, Ribosome Binding Site. RE, Restriction Enzyme. sgRNA, single guide RNA. TU, Transcriptional Unit. UTR, Untranslated Region.

REFERENCES

[1] Hughes, R. A., and Ellington, A. D. (2017) Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology, *Cold Spring Harb Perspect Biol 9*, a023812.

[2] Santos-Moreno, J., and Schaerli, Y. (2019) Using Synthetic Biology to Engineer Spatial Patterns, *Advanced Biosystems 3*, 1800280.

[3] Engler, C., Kandzia, R., and Marillonnet, S. (2008) A One Pot, One Step, Precision Cloning Method with High Throughput Capability, *Plos One 3*, e3647.

[4] Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nat Methods 6*, 343-345.

[5] Shetty, R. P., Endy, D., and Knight, T. F., Jr. (2008) Engineering BioBrick vectors from BioBrick parts, *J Biol Eng 2*, 5.

[6] Li, M. Z., and Elledge, S. J. (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC, *Nat Methods 4*, 251-256.

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

[8] Quan, J. Y., and Tian, J. D. (2009) Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways, *Plos One 4*, e6441.

[9] Casini, A., MacDonald, J. T., De Jonghe, J., Christodoulou, G., Freemont, P. S., Baldwin, G. S., and Ellis, T. (2014) One-pot DNA construction for synthetic biology: the Modular Overlap-Directed Assembly with Linkers (MODAL) strategy, *Nucleic Acids Res. 42*, e7.

[10] Taylor, G. M., Mordaka, P. M., and Heap, J. T. (2018) Start-Stop Assembly: a functionally scarless DNA assembly system optimized for metabolic engineering, *Nucleic Acids Res. 47*, e17.

[11] Guye, P., Li, Y., Wroblewska, L., Duportet, X., and Weiss, R. (2013) Rapid, modular and reliable construction of complex mammalian gene circuits, *Nucleic Acids Res. 41*, e156-e156.

[12] Torella, J. P., Boehm, C. R., Lienert, F., Chen, J.-H., Way, J. C., and Silver, P. A. (2013) Rapid construction of insulated genetic circuits via synthetic sequence-guided isothermal assembly, *Nucleic Acids Res. 42*, 681-689.

[13] Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011) A Modular Cloning System for Standardized Assembly of Multigene Constructs, *Plos One 6*, e16765.

[14] Halleran, A. D., Swaminathan, A., and Murray, R. M. (2018) Single Day Construction of Multigene Circuits with 3G Assembly, *ACS Synth Biol 7*, 1477-1480.

[15] Kahl, L. J., and Endy, D. (2013) A survey of enabling technologies in synthetic biology, *J Biol Eng 7*, 13.

[16] Schaerli, Y., Munteanu, A., Gili, M., Cotterell, J., Sharpe, J., and Isalan, M. (2014) A unified design space of synthetic stripe-forming networks, *Nat Commun 5*, 4905.

[17] Casini, A., Christodoulou, G., Freemont, P. S., Baldwin, G. S., Ellis, T., and MacDonald, J. T. (2014) R2oDNA designer: computational design of biologically neutral synthetic DNA sequences, *ACS Synth Biol 3*, 525-528.

[18] Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression, *Cell 152*, 1173-1183.

[19] Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J. A. (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease, *Science 329*, 1355-1358.

[20] Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., Imamura, T., Ogawa, M., Masai, H., and Miyawaki, A. (2008) Visualizing spatiotemporal dynamics of multicellular cell-cycle progression, *Cell 132*, 487-498.

[21] Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006) Engineering and characterization of a superfolder green fluorescent protein, *Nat. Biotechnol. 24*, 79-88.

[22] Shcherbo, D., Murphy, C. S., Ermakova, G. V., Solovieva, E. A., Chepurnykh, T. V., Shcheglov, A. S., Verkhusha, V. V., Pletnev, V. Z., Hazelwood, K. L., Roche, P. M., Lukyanov, S., Zaraisky, A. G., Davidson, M. W., and Chudakov, D. M. (2009) Far-red fluorescent tags for protein imaging in living tissues, *Biochem. J 418*, 567-574.

[23] Butzin, N. C., and Mather, W. H. (2018) Crosstalk between Diverse Synthetic Protein Degradation Tags in Escherichia coli, *ACS Synth Biol 7*, 54-62.

[24] Kogenaru, M., and Tans, S. J. (2014) An improved Escherichia coli strain to host gene regulatory networks involving both the AraC and Lacl inducible transcription factors, *J Biol Eng 8*, 2.

[25] Schaerli, Y., and Isalan, M. (2013) Building synthetic gene circuits from combinatorial libraries: screening and selection strategies, *Mol Biosyst 9*, 1559-1567.

[26] Duarte, J. M., Barbier, I., and Schaerli, Y. (2017) Bacterial Microcolonies in Gel Beads for High-Throughput Screening of Libraries in Synthetic Biology, *ACS Synth Biol. 6*, 1988-1995.

[27] Pollak, B., Cerda, A., Delmans, M., Alamos, S., Moyano, T., West, A., Gutierrez, R. A., Patron, N. J., Federici, F., and Haseloff, J. (2019) Loop assembly: a simple and open system for recursive fabrication of DNA circuits, *New Phytol 222*, 628-640.

[28] Potvin-Trottier, L., Lord, N. D., Vinnicombe, G., and Paulsson, J. (2016) Synchronous long-term oscillations in a synthetic gene circuit, *Nature 538*, 514-517.

[29] Qi, L., Haurwitz, R. E., Shao, W., Doudna, J. A., and Arkin, A. P. (2012) RNA processing enables predictable programming of gene expression, *Nat. Biotechnol. 30*, 1002-1006.

[30] Liu, X., Gallay, C., Kjos, M., Domenech, A., Slager, J., van Kessel, S. P., Knoops, K., Sorg, R. A., Zhang, J. R., and Veening, J. W. (2017) High-throughput CRISPRi phenotyping identifies new essential genes in Streptococcus pneumoniae, *Mol Syst Biol 13*, 931.

[31] Didovyk, A., Borek, B., Hasty, J., and Tsimring, L. (2016) Orthogonal Modular Gene Repression in Escherichia coli Using Engineered CRISPR/Cas9, *ACS Synth Biol 5*, 81-88.

For Table of Contents Use Only

A framework for the modular and combinatorial assembly of synthetic gene circuits

Javier Santos-Moreno* and Yolanda Schaerli*

Department of Fundamental Microbiology, University of Lausanne, Biophore Building, 1015 Lausanne, Switzerland

javier.santosmoreno@unil.ch

yolanda.schaerli@unil.ch