Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system

Stefan Werner, Carola Engler, Ernst Weber, Ramona Gruetzner and Sylvestre Marillonnet*
Icon Genetics GmbH; Halle, Germany

Keywords: cloning, DNA assembly, genetic engineering, synthetic biology, metabolic engineering, directed evolution
Submitted: 06/30/11
Revised: 09/23/11
Accepted: 09/26/11
http://dx.doi.org/10.4161/bbug.3.1.18223
*Correspondence to: Sylvestre Marillonnet; Email: marillonnet@icongenetics.de

Recent progress in the field of synthetic biology has led to the creation of cells containing synthetic genomes. Although these first synthetic organisms contained copies of natural genomes, future work will be directed toward engineering of organisms with modified genomes and novel phenotypes. Much work, however, remains to be done to be able to routinely engineer novel biological functions. As a tool that will be useful for such purpose, we have recently developed a modular cloning system (MoClo) that allows high throughput assembly of multiple genetic elements. We present here new features of this cloning system that allow to increase the speed of assembly of multigene constructs. As an example, 68 DNA fragments encoding basic genetic elements were assembled using three one-pot cloning steps, resulting in a 50 kb construct containing 17 eukaryotic transcription units. This cloning system should be useful for generating the multiple construct variants that will be required for developing gene networks encoding novel functions, and fine-tuning the expression levels of the various genes involved.

The field of synthetic biology is progressing at a rapid pace, with the recent generation of organisms containing synthetic genomes. While assembling complete genomes and rebooting cells with these is now possible, designing gene networks that encode new biological functions is still very challenging, even at a small scale. In fact, most genetic engineering work that has been performed to date has dealt with fewer than ten genes. There are several reasons why this is the case. One reason is that until recently, the tools available for cloning and assembly of DNA fragments were still very limited (DNA fragments were, and are still, subcloned one or two fragments at a time), and required the design of construct-specific cloning strategies for each project, limiting the number of constructs that could effectively be made. Once cloning strategies were designed, multiple cloning steps were often needed to obtain the desired construct. Cloning steps were then performed manually in a slow and inefficient process that could not be easily automated. One additional factor limiting engineering work requiring multiple genes is that the design of constructs that would provide a desired phenotype cannot be determined entirely from sequence alone. Therefore, multiple construct variants have to be made, with the number of variants needed growing exponentially with the number of genetic elements involved.

In the past few years, several advances have nevertheless facilitated cloning. Recombinase-based systems allow DNA fragments of interest to be subcloned easily from one vector to the next using a process that can be automatized. Recombinase-based cloning is however limited by the presence of recombination sites left in the constructs, and by the limited number of fragments that can be subcloned in one step. Use of rare-cutting endonucleases facilitates the design of cloning strategies for constructs containing multiple genes, but still relies on standard cloning methods. Development of the BioBrick system simplified the design of cloning...
strategies as the same strategy could be repeated from one step to the next to get larger and larger constructs; BioBrick assembly is however limited by the number of fragments that can be assembled in one step. Finally, sequence and ligation-independent cloning has allowed up to ten fragments to be assembled in one step, but requires PCR, which is error-prone. Ligation-independent cloning, but also methods based on homologous recombination, requires homology between the fragments to assemble, limiting the ability to freely combine sets of basic genetic elements in combinatorial libraries.

Recently, we developed a system to assemble multiple DNA fragments in a one-pot reaction. This cloning method, called Golden Gate cloning, is based on the use of type IIS restriction enzymes combined with restriction-ligation. Extremely high cloning efficiency is obtained using a simple one-pot incubation of multiple undigested entry constructs and a destination vector in the presence of restriction enzyme and ligase. In order to eliminate the need for construct-specific cloning strategies, we also developed a modular cloning system (MoClo) that allows any multigene construct of choice to be made by using a defined set of pre-made vectors and a defined assembly strategy. Basic genetic elements such as promoters, coding sequences or terminators (commonly called basic parts in the synthetic biology community) are cloned as level 0 modules. Each module is flanked by two BsaI restriction sites specific for each module type, with cleavage sites designed to allow sets of compatible modules to be assembled in one step in a level 1 destination vector (using the enzyme BpiI). The resulting level 1 constructs contain transcriptional units, which are then assembled six at a time in level 2 constructs using a second type IIS enzyme (BpiI). The cloning process can then be repeated to add more transcription units to the resulting construct by alternating the type IIS enzyme(s) used for cloning. We have previously shown that 44 basic genetic elements could be assembled in a defined order with three one-pot cloning steps resulting in a construct containing 11 transcription units. The first step consisted of the parallel assembly of the 11 transcription units from basic genetic parts. The second step consisted of assembly of five of these transcription units, and the third step added the remaining six transcription units.

We have now performed a further round of cloning to add another six transcription units. Since the final construct was expected to be 50 kb in size and the construct that we made previously had a high copy origin of replication (a colE1 ori derived from pUC19), we thought that it would be preferable to use a plasmid backbone with a lower copy number. We therefore made a plasmid vector backbone containing the low copy origin of replication p15a (derived from pACYC184, from New England Biolabs). Since a new vector backbone was used, the two cloning steps needed to assemble the first 11 transcription units were repeated as described previously. Two thousand colonies of the expected color were obtained for the first cloning step, and 446 for the second step (Fig. 1A). For both cloning steps, three out of four colonies contained a correct construct. Six new transcription units (cL1-12 to cL1-17) were also assembled from basic parts, including for each transcription unit a promoter, a 5' UTR, a protein coding sequence and a terminator (the structure of the six transcription units and their constituent parts are shown as part of the final construct cL2-21 in Fig. 1B). To test the efficiency of cloning, six reactions were performed to add from one to six transcription units to the construct already containing 11 transcription units. Cloning was performed by incubation of the destination vector (cL2-15), one to six transcription unit constructs (cL1-12 to cL1-17), and a compatible end-linker (pELB-5 to 3, containing a LacZa fragment) in the presence of BpiI, Esp3I and T4 DNA ligase. Colonies of the expected color (blue) were obtained for all reactions, and four out of four colonies screened for each cloning reaction were found to contain correct constructs. Therefore, 17 transcription units can be easily cloned using the MoClo system. Additional transcription units could then be added with further rounds of level 2 cloning.

One goal for the MoClo system would be to allow assembly of constructs in the range of 100 kb or more. With an approximate size of 2.7 kb per transcription unit (as was the case for the first 17 transcription units) and a vector backbone of 4.3 kb (size of the particular vector that we used), a 100 kb construct would contain approximately 36 transcription units and would require assembly of 144 to 180 basic parts, depending on whether four or five basic parts are used for each transcription unit. Using the MoClo system as presented earlier, this would take seven cloning steps (a few more if six transcription units cannot be cloned in one step for larger constructs). A first step would assemble in parallel the 36 transcription units, and six following steps would be necessary to assemble the 36 transcription units, six at a time. Although this approach is feasible, a faster cloning strategy can also be used. For example, a new cloning level, that we named M (for multiplication), was designed as an intermediate step between level 1 and level 2 (Fig. 2). Level M is used for preassembly of several transcription units, which are then subcloned as preassembled groups in level 2 constructs. Level M requires a set of seven destination vectors and seven end-linkers (Level M, Fig. 3). Level M vectors and end-linkers are designed in a way that preassembled transcription units become flanked by two BsaI sites after cloning (cLM-1 to 3, Fig. 4). This is in contrast with level 2 constructs in which two type IIS restriction sites are made available for further cloning (BsaI or Esp3I), not flanking the assembled transcription units, but in the end-linker that follows them (cL2-22, Fig. 4). Using level M, 36 transcription units could theoretically be cloned in three steps (more steps if cloning becomes not efficient enough for assembly of six DNA fragments at a time for larger constructs), with 6 level M constructs made in parallel in one step (each containing six transcription units), and assembly of these to a final construct in a following cloning step. As an example for cloning using level M, a construct containing the same 17 transcription units as in cL2-21 was cloned. Three level M constructs containing six, six and five transcription units each (cLM1 to 3, Figs. 4 and 5) were first assembled using three parallel one-pot cloning reactions. All four colonies screened for each cloning
Figure 2. Relationship between cloning levels 0, 1, 2, M and P. Level 1 is used for assembly of transcription units from basic level 0 modules. Several level 1 transcription units can then be assembled in level 2 multigene constructs. Level 2-x constructs (boxed in gray) do not contain any type IIS restriction site and cannot be used for further constructs. Level 2i-x (i for intermediate) contain two type IIS restriction sites after the last transcription unit for addition of further transcription units. Several transcription units can also be preassembled in a level M construct. Assembled transcription units in level M constructs are flanked by two type IIS restriction sites, allowing them to be further subcloned either in level 2 or level P constructs. Level P constructs can be further assembled in level M constructs (also see Figure 4).
reaction contained correct constructs. The 17 transcription units were then assembled using a second cloning step to make a level 2i (cL2-22; i for intermediate). Therefore, using level M, 17 transcription units can be assembled from basic modules in three cloning steps.

Level 2i constructs contain two type IIS restriction sites located after the last transcription unit cloned. This provides the possibility for adding additional transcription units into this construct, but does not allow for subcloning the assembled transcription units in a new vector backbone. Instead of a level 2i, a new cloning level, level P, could be made such that assembled transcription units are again flanked by two type IIS restriction sites (Level P, Fig. 4). This structure would allow several level P constructs to be further assembled in level M cloning vectors (or in any other plasmid with compatible restriction sites), and the process repeated between level M and P until constructs become too large for successful cloning or transformation in recipient cells (Fig. 2). As such, level M and P are functionally similar to level alpha and omega described for the GoldenBraid system except that level M and P allow subcloning of more than two DNA fragments at a time. Cloning of level P constructs requires an additional set of destination vectors and end-linkers (Fig. 3). To test level P cloning, the 3 level M constructs described above were cloned in a level P destination vector (cLP-1, Figs. 4 and 5).

The examples given above illustrate the versatility of the MoClo system. In cases where a moderate number of genes need to be cloned and the composition of a final construct is not necessarily established at the beginning of a project, the system of vectors designed for repeated use of level 2 cloning steps provide flexibility for adding transcription units to an ongoing assembly as needed. In cases where many genes or DNA fragments need to be rapidly assembled in a hierarchical way, level M and/or level P or a combination with level 2 can provide additional speed. A common feature of all these vector sets is the continual re-use of the same seven cleavage sites (fusion sites) that are present in both level 2, level M and level P destination vectors and end-linkers (TGCC, GCAA, ACTA, TTAC, CAGA, TGTG and GAGC). This feature allows a finite number of cloning vectors to add an unlimited (except ultimately by construct size) number of transcription units (or any other genetic elements) in a final construct, but also provides the flexibility to choose the optimal path for DNA construction for a particular project.

Golden Gate cloning and the MoClo system are not designed for genome scale assembly, since large pieces of DNA ultimately become difficult to handle in vitro. For such purpose, and if needed, other technologies will be required such as those described by Gibson et al., or Itaya et al. What Golden Gate cloning and the MoClo system are designed to achieve is to allow combinatorial assembly of multiple genetic elements to make libraries of variant multigene constructs from which to select appropriate phenotypes. Other technologies such as Multiplex...
Automated Genome Engineering (MAGE) can also be very useful to introduce large amounts of genetic variation at multiple selected positions in a genome. In fact, MoClo and MAGE are complementary technologies (for use in prokaryotes), since MoClo can be used to combine multiple genes from one or several source organisms into synthetic operons/multigene constructs, while MAGE could be used on organisms containing these to optimize the level of expression of endogenous genes.

**Figure 4.** Structure of level 1, M, 2i and P constructs. Level 1 constructs contain a transcription unit flanked by two BpiI sites in opposite orientations. Cloning of the first six transcription units together with end-linker pELM-6 in cloning vector pLM-1 leads to construct cLM-1. Construction of cLM-2 and cLM-3 requires different and appropriate destination vectors and end-linkers (not shown). cLM-1, 2 and 3 can be subcloned in either level 2 or level P vectors. The resulting constructs cL2-22 or cLP-1 differ principally by the location of type II restriction sites required for the next step of cloning: both sites are located after the assembled transcription units in cL2-22, but flank the assembled transcription units in cLP-1.

**Figure 5.** Cloning of 17 transcription units using level M and P. Five or six transcription units were first assembled in level M constructs (cLM-1 to 3). The 3 level M constructs were then assembled in level 2 or P cloning vectors resulting in constructs cL2-22 and cLP-1, respectively.
genes that may affect or regulate the level of expression of the introduced genes or pathways.

Thanks to new DNA assembly/modification technologies such as MoClo, MAGE or any other DNA assembly/shuffling technologies, generation of combinatorial libraries with the desired genetic variability should not be a limiting factor anymore. Engineering of new biological functions and phenotypes will also require increasingly predictive computer programs to guide genetic engineers and to narrow the amount of genetic variability that has to be constructed so that it can actually be screened in order to obtain useful phenotypes.

Acknowledgments
This work was funded by Icon Genetics GmbH.

References
1. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. Science 2010; 329:52-6; PMID:20488990; http://dx.doi.org/10.1126/science.1190719
2. Baker M. Synthetic genomes: The next step for the synthetic genome. Nature 2011; 473:403-8; PMID:21593873; http://dx.doi.org/10.1038/473403a
3. Marsischky G, LaBaer J. Many paths to many clones: a comparative look at high-throughput cloning methods. Genome Res 2004; 14:2020-8; PMID:15489321; http://dx.doi.org/10.1101/gr.2528804
4. Dafny-Yelin M, Trifita T. Delivery of multiple transgenes to plant cells. Plant Physiol 2007; 145:1118-28; PMID:18056862; http://dx.doi.org/10.1104/pp.107.106104
5. Knight TF. Idempotent Vector Design for Standard Assembly of BioBricks. MIT Synthetic Biology Working Group Technical Report 2003.
6. Li MZ, Elledge SJ. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods 2007; 4:251-6; PMID:17293868; http://dx.doi.org/10.1038/nmeth1010
7. Gibson DG, Benders GA, Axelrod KC, Zaveri J, Algire MA, Moodie M, et al. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome. Proc Natl Acad Sci USA 2008; 105:20404-9; PMID:19073939; http://dx.doi.org/10.1073/pnas. 0811011010
8. Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. PLoS ONE 2008; 3:e3647; PMID:18985154; http://dx.doi.org/10.1371/journal.pone.0003647
9. Engler C, Gruetzmacher R, Kandzia R, Marillonnet S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS ONE 2009; 4:e5553; PMID:19436741; http://dx.doi.org/10.1371/journal.pone.0005553
10. Weber E, Engler C, Gruetzmacher R, Werner S, Marillonnet S. A modular cloning system for standardized assembly of multigene constructs. PLoS ONE 2011; 6:e16765; PMID:21356738; http://dx.doi.org/10.1371/journal.pone.0016765
11. Sarrion-Perdigones A, Falkoni EE, Zandalinas SI, Juarez P, Fernandez-del-Carmen A, Granell A, et al. GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. PLoS ONE 2011; 6:e21622; PMID:21750718; http://dx.doi.org/10.1371/journal.pone.0021622
12. Itaya M, Fujita K, Kuroki A, Tsuge K. Bottom-up genome assembly using the Bacillus subtilis genome vector. Nat Methods 2007; 4:51-3; PMID:17293868; http://dx.doi.org/10.1038/nmeth1143
13. Itaya M, Fujita K, Kuroki A, Tsuge K. Bottom-up genome assembly using the Bacillus subtilis genome vector. Nat Methods 2007; 4:51-3; PMID:18066072; http://dx.doi.org/10.1038/nmeth1143
14. Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, et al. Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. Science 2008; 319:1215-20; PMID:18218864; http://dx.doi.org/10.1126/science.1151721
15. Benders GA, Noskov VN, Denisova EA, Lartigue C, Gibson DG, Assad-Garcia N, et al. Cloning whole bacterial genomes in yeast. Nucleic Acids Res 2010; 38:2558-69; PMID:20211840; http://dx.doi.org/10.1093/nar/gkq119
16. Ellis T, Adie T, Baldwin GS. DNA assembly for synthetic biology: from parts to pathways and beyond. Integr Biol (Camb) 2011; 3:109-18; PMID:21246151; http://dx.doi.org/10.1039/c0ib00070a
17. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, et al. Programming cells by multiplex genome engineering and accelerated evolution. Nature 2009; 460:894-8; PMID:19633652; http://dx.doi.org/10.1038/nature08187
18. MacDonald JT, Barnes C, Kinney RJ, Freemont PS, Stan GB. Computational design approaches and tools for synthetic biology. Integr Biol (Camb) 2011; 3:97-108; PMID:21258712; http://dx.doi.org/10.1039/c0ib00077a