Golden Gate Assembly system dedicated to complex pathway manipulation in *Yarrowia lipolytica*

Ewelina Celińska, Rodrigo Ledesma-Amaro, Macarena Larroude, Tristan Rossignol, Cyrille Pauthenier and Jean-Marc Nicaud

1 Department of Biotechnology and Food Microbiology, Poznan University of Life Sciences, ul. Wojska Polskiego 48, 60-627 Poznan, Poland.
2 Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France.
3 Institute of System and Synthetic Biology, Université d’Evry vald’Essonne, Bt. Geneavenir 6 Genopole Campus 1, 5 rue Henry Desbrueres 91000, Evry, France.

Summary

In this study, we have adopted Golden Gate modular cloning strategy to develop a robust and versatile yeast *Yarrowia lipolytica*. To this end, a broad set of destination vectors and interchangeable building blocks have been constructed. The DNA modules were assembled on a scaffold of predesigned 4 nt overhangs covering three transcription units (each bearing promoter, gene and terminator), selection marker gene and genomic integration targeting sequences, constituting altogether thirteen elements. Previously validated DNA modules (regulatory elements and selection markers) were adopted as the Golden Gate bricks. The system’s operability was demonstrated based on synthetic pathway of carotenoid production. This technology greatly enriches a molecular biology toolbox dedicated to this industrially relevant microorganism enabling fast combinatorial cloning of complex synthetic pathways.

Introduction

Synthetic biology, and in particular combinatorial cloning, derives from engineering concepts of standardization, modularity of building blocks and simplification of assembly lines. Traditional protocols employing restriction digestion and one-by-one element cloning are both time- and cost-inefficient (Celińska and Grajek, 2013; Matthaus et al., 2014). Due to the recent development of DNA assembly techniques for metabolic pathway engineering, a great worldwide effort is now being pursued towards establishing such cloning platforms for an individual organism of interest. In the last years, several DNA modular assembly platforms have been developed and new standards have been defined (Sands and Brent, 2001). Golden Gate (GG) modular cloning system, relying on type II restriction enzymes, appears as one of the most robust techniques within this field (Engler et al., 2008; Gao et al., 2013). The core of GG strategy lies in establishing a library of standardized and interchangeable DNA parts, which can be subsequently assembled in a single-step, one-pot reaction. Examples of such GG platforms have been recently reported for *Escherichia coli*, yeast or plant species (Engler et al., 2014; Terfrüchte et al., 2014; Agmon et al., 2015; Kakui et al., 2015; Lee et al., 2015; Mitchell et al., 2015; Iversen et al., 2016; Moore et al., 2016). Nevertheless, many of valuable biotech workhorses are still queuing the line for a customized GG platform, amongst these the non-conventional yeast *Y. lipolytica*, which is a well-established biotechnological chassis for the production of numerous valuable bioproducts (Nicaud, 2012; Ledesma-Amaro et al., 2015; Madzak, 2015; Ledesma-Amaro and Nicaud, 2016). Establishing a GG combinatorial cloning platform for *Y. lipolytica* appears as an urgent need, enabling more time- and cost-efficient assembly of complex DNA constructions and standardization of DNA modules, which can be easily exchanged between different laboratories.
Here, we present a robust modular GG system for expression of one, two or three customizable transcription units (TU) in a versatile cassette for *Y. lipolytica*.

**Results and discussion**

**Defining standards and modular parts of the GG system for *Y. lipolytica***

The ultimate goal of this study was the development of extensive and versatile system of building blocks (Golden Gate Fragments; GGFs) easily fitted into the general scaffold of GGA (Golden Gate Assembly), enabling fast construction of complex expression vectors dedicated to metabolic engineering of *Y. lipolytica*. To complete this task, a set of thirteen 4 nt overhangs was assigned a letter, to enable unambiguous referring to the respective position in the GGA. Moreover, the overhangs flanking individual genes (in position G1, G2 or G3; see Fig. 1) were designed to

---

**Fig. 1. Golden Gate Assembly platform dedicated to *Yarrowia lipolytica* in a one-, two- and three-transcription unit (TU)-bearing format.**

A. Schematic representation of complete GG destination vectors in a one-, two- and three-TU-bearing format. Each of GGA covers at least one TU (composed of promoter – P; gene as ORF – G; and terminator – T) and selection marker (M), which are flanked with integration targeting sequences (InsUP and InsDOWN), constructed on a destination vector backbone, either pSB1A/C/K3-RFP/amiliCP/amiliGFP for ampicillin (A), chloramphenicol (C) or kanamycin (K) resistance marker respectively; pSB vector series offers exploitation of one of the following reporter genes in *E. coli*: RFP – red chromophore; amiliCP – blue chromophore; amiliGFP – green fluorophore. Building blocks are flanked with pre-designed 4 nt overhang, each assigned a letter from A–M. Sequences of the overhangs are provided next to the respective position. Abbreviations: InsUP/DOWN – sequence targeting insertion in the 5’/3’ region; M – selection marker; P1/2/3 – promoter governing expression of the gene of interest in transcription unit 1, 2 and 3 respectively; G1/2/3 – gene of interest in transcription unit 1, 2 and 3 respectively; T1/2/3 – terminator located in the transcription unit 1, 2 and 3 respectively. NotI/SfiI: NotI/SfiI recognition sites for release of the expression cassettes prior to *Y. lipolytica* transformation. B. Donor vectors constructed on a backbone of pCRT II-Blunt TOPO vector (Invitrogen/Life Technologies) bearing kanamycin resistance gene. Each building block is flanked with respective BsaI recognition sites and predefined 4 nt overhangs. Exemplary building blocks (Golden Gate Fragments; GGFs) constructed in this study are provided as a list underneath the donor vectors’ schemes. For reference to nucleotide sequences, see Supporting Information.

© 2017 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology, *Microbial Biotechnology*, 10, 450–455
minimize the scar between a coding sequence and the regulatory elements. Thus, the overhang prior to the coding sequence was designed to maximally fit into the consensus sequence preceding significant fraction of ORFs in *Y. lipolytica* genome – CACA – or include the start codon (Gasmii et al., 2011).

The GG modular system was designed to include two flanking insertion sites, a selection marker and one, two or three TUs in each assembly. The number of TUs can be easily manipulated by amplification of ‘hybrid’ terminators, provided that the same sequence was used in the merged positions (T1, T2, T3; compare Fig. 1). These ‘hybrid’ terminators should be amplified with forward primer with E’ overhang and reverse primer with ‘L’ overhang to render T1-3 ‘hybrid’ terminator, skipping the second and third transcription units, or with ‘H’ overhang bearing forward primer and ‘L’ overhang bearing reverse primer, to gain T2-3 ‘hybrid’ terminator, skipping the third transcription unit.

The destination vector backbones were equipped with a chromophore (RFP or amilCP) or fluorophore (amiGFP) gene flanked with BsaI sites with pre-designed overhangs to define a colour-based negative cloning selection marker. The BsaI recognition sites were positioned outwardly with respect to the backbone of the destination vector, so that after digestion with BsaI, both the reporter gene and the BsaI recognition sites were released, leaving the vector’s backbone flanked with protruding 4 nt overhangs. A set of different antibiotic resistance genes was incorporated into the system of the destination vectors, to enable greater flexibility of experimental set-up. In the proof-of-concept experiments (with three TUs bearing cassette; described below), we have used a pSB1K3-RFP variant from the iGEM collection (http://parts.igem.org/ Collections).

It is commonly accepted that only integrative expression cassettes constitute truly potent tools in *Y. lipolytica* genetic engineering, as the so far developed episomal plasmids tend to be unstable and low-copy (Liu et al., 2014). This fact ruled the construction of both the destination plasmids as well as the scaffold of GGA, flanked with GGFs targeting integration of the expression cassette into the host genome (InsUP and InsDOWN). Thus, the destination vector’s backbone containing solely bacterial elements can be discarded prior to transformation of *Y. lipolytica* cells. To enable the release of the expression cassette composed of transcription units, selection marker and targeting sequences, from the bacterial backbone, rare-cutting restriction endonuclease recognition sites NotI or SfiI, depending on the targeted GGA structure, were added at the borders of the expression cassette in the InsUP and InsDOWN fragments (Fig. 1).

**Generation of GG library dedicated to *Y. lipolytica***

All the sequences to be used as building blocks of the envisioned GGA were extracted from *Y. lipolytica* W29 genome sequence or one of the previously constructed vectors from our own collection. All the sequences were analysed in silico to find internal BsaI site, to be eventually eliminated using assembly PCR technique (Sambrook and Russell, 2001). Subsequently, an extensive set of primers (referred to as GGP, for Golden Gate Primers) equipped with the predesigned system of the 4 nt overhangs and externally located BsaI recognition sites was designed (Supporting Information; Pauthenier and Faulon, 2014). The set of GGF presented in this work, the interchangeable parts for promoter, terminator, selection marker and integration site (Fig. 1B), constitutes the response to our current needs, but can be easily expanded on the other building blocks of interest by simply adding the required prefixes and suffixes within the primers at the stage of the element amplification. As presented in Fig. 1, the major focus of this study was to develop a system of easily tractable phenotype of positive transformants. Therefore, for the genes of interest we have chosen the complete pathway for carotenoid production, for convenient traceability of positive clones.

With the same concept in mind, LIP2 and URA3 loci were selected as the targeting platforms, and thus, correct integration could be verified on tributyrin- or sucrose-containing plate respectively. A LIP2 deletion results in a reduced halo of triglyceride hydrolysis (Pignède et al., 2000), while replacement of ura3-320 allele bearing *S. cerevisiae* SUC2 gene gives transformants unable to grow on sucrose (Nicaud et al., 1989; Barth and Gaillardin, 1996). Zeta sequences were chosen to reach random integrations in Po1d strain (JMY195) or at a zeta docking platform in Po1d derivative JMY1212, bearing a zeta sequence at the *leu2-270* locus (Bordes et al., 2007).

All the amplicons bearing respective GGFs were cloned in the donor vectors (pCR Blunt II TOPO vectors; Thermo Fisher Scientific, Villebon sur Yvette, France), verified by BsaI digestions and sequencing and deposited in our collection. The structure of the targeted GGA cassette ruled the composition of the GG reaction mixture. The Golden Gate reaction conditions were designed based on the previously published protocols (Engler et al., 2008; Pautenier et al., 2012; Agmon et al., 2015). The reaction mixture contained precalculated equimolar amount of each GGF and the destination vector (50 pmoles of ends), 2 μl of T4 DNA ligase buffer (NEB), 5 U of BsaI, 200 U of T4 and ddH2O up to 20 μl. The following thermal profile was applied: [37°C for 5 min, 16°C for 2 min] × 60, 55°C for 5 min, 80°C for 5 min, 15°C ∞. Subsequently, the reaction mixture was...
used for *E. coli* DH5α transformation (Sambrook and Russell, 2001). White colonies were screened for identification of complete GGA through plasmid isolation, restriction digestion and multiplex PCR. Complete GGA was subsequently linearized and used for transformation of either *Y. lipolytica* Po1d *ura-* *leu-* (JMY195 *MatA* *leu2-270* *ura3-302* *xpr2-322*) or Po1d *ura-* *leu+* bearing a zeta docking platform (JMY1212; *MatA* *ura3-302* *xpr2-322*, *LEU2*, *zeta*; Bordes et al., 2007) via the lithium acetate transformation protocol (Chen et al., 1997). Derivative strain Po1d *leu+* *ura+* (JMY2900; *MatA* *ura3-302* *xpr2-322*) was used as prototroph control. Clones bearing complete GGA integrated with the host genome exhibited red–orange phenotype due to carotenoid production. All the GGF building blocks included in the GGA have been proven functional in previous experiments (see Table S1).

**Proof of concept based on carotenoid-pathway-containing Golden Gate Assembly**

To validate robustness and efficiency of the here developed GG system, we have constructed a complex assembly structure comprising twelve GGFs and the destination vector, covering three TUs, accompanied by selection marker and integration targeting sequences (Fig. 2A).

The three genes within this assembly were native GGS1 (geranylgeranyl diphosphate synthase; YALI0D17050g) in G1 position and heterologous carB (phytoene dehydrogenase; AJ238028.1) and carRP (phytoene synthase; AJ250827.1) coding sequences from *Mucor circinelloides* in G2 and G3 positions respectively (sequences were kindly provided by DSM Company). Such a set of genes was previously proved to be efficiently expressed in *Y. lipolytica* cells, constructed using DNA assembler strategy (Gao et al., 2014) or classical cloning techniques (Matthaus et al., 2014). All the genes were assembled with corresponding regulatory elements, zeta insertion sites and *URA3* selection marker. The correct complete GGA was first preselected in *E. coli* clones through multiplex PCR and restriction digestion and then linearized and transformed into competent *Y. lipolytica* JMY195 cells. On average, more than $4 \times 10^2$ colonies could be obtained from a single transformation reaction. About 10% of the transformants were white (exemplary results for one transformation run: 50/517, 91/1052, 41/396). The remaining clones exhibited different intensity of the orange colour development (Fig. 2B). The observed variability in the carotenoid-producer phenotype could result from varying number of either complete or partial GGA copies integrated within the host genome as observed by Gao et al. (2014), although this would be unlikely with the amount of DNA we used (Bordes et al., 2007). More likely, the random distribution of the cassette within the genomic DNA due to the zeta targeting sequences, as investigated in Pignede et al. (2000), may be responsible for the colour variability. This hypothesis is supported by the results obtained when the same GGA was used to transform strain JMY1212, bearing a zeta docking platform at the *leu2-270* locus enabling homologous recombination. Although the number of white colonies after transformation of this strain was higher (33%; 33/98), the orange colonies did not present the colour variability that was found amongst JMY195-based transformants. In a previous study with DNA assembler strategy, when the same carotenoid-pathway genes were used, the efficiency of

![Fig. 2. Golden Gate Assembly bearing ‘carotenoid synthesis’ pathway dedicated to integration with *Yarrowia lipolytica* genome.](image-url)
reaching desired phenotype was 20% (Gao et al., 2014). In the current study, the efficiency was largely improved, as 67% to 90% of the obtained Y. lipolytica clones exhibited desired red-orange phenotype, for JMY1212 and JMY195 strains respectively. Improved efficiency was accompanied by time- and workload savings, as well as versatility and standardization, being indigenous for modular DNA cloning techniques, altogether making this approach a method of choice for synthetic biology endeavours.

Here, we present a Golden Gate standard specifically designed to modify Y. lipolytica in a versatile manner. Robustness and reliability of the system were validated for thirteen-element-bearing GGA, with high efficiency of the desired phenotype recovery in the recombinants. This technology greatly enriches the molecular biology toolbox dedicated to this industrially relevant microorganism, permitting faster and accurate multiple target engineering, a limitless expansion of the building blocks library from endogenous or heterologous origins, or the use of interchangeable modules in combinatorial approaches followed by screening desired phenotypes.

Acknowledgements

This study was financially supported by a programme with the reference ANR-11-BTBR-0003, Agence Nationale de la Recherche, under the ‘Investissements d’avenir’. RL would like to thank AgreenSkills Fellowship (FP7-267196), funded by European Union Marie-Curie FP7 COFUND People Programme and the Synthetic Genome Summer Course 2016. EC (EMBO_ASTF_117-2016) would like to thank European Molecular Biology Organization for supporting her fellowship at the BIMLip team.

Authors’ contribution

EC and RLA contributed equally to this work. EC, RLA, ML and TR contributed experimental data. CP participated in the selection of the 4 nt overhangs for the GGA scaffold and provided destination vectors. EC, RLA and JMN wrote the manuscript. The work was performed in the laboratory of TR and JMN.

Conflict of interest

The authors do not have any conflict of interest.

References

Agmon, N., Mitchell, L.A., Cai, Y., Ikushima, S., Chuang, J., Zheng, A., et al. (2015) Yeast Golden Gate (yGG) for the efficient assembly of S. cerevisiae transcription units. ACS Synth Biol 4: 853–859.

Barth, G., and Gaillardin, C. (1996) Yarrowia lipolytica. In Nonconventional Yeasts in Biotechnology A Handbook. Wolf, K. (ed). Berlin Heidelberg: Springer, pp. 313–388.

Bordes, F., Fudalej, F., Dossat, V., Nicaud, J.-M., and Marty, A. (2007) A new recombinant protein expression system for high-throughput screening in the yeast Yarrowia lipolytica. J Microbiol Methods 70: 493–502.

Celinski, E., and Grajek, W. (2013) A novel multigene expression construct for modification of glycerol metabolism in Yarrowia lipolytica. Microb Cell Fact 12: 102.

Chen, D.C., Beckerich, J.M., and Gaillardin, C. (1997) One-step transformation of the dimorphic yeast Yarrowia lipolytica. Appl Microbiol Biotechnol 48: 232–235.

Engler, C., Kandzia, R., and 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, T.-M., Werner, S., Jones, J.D.G., et al. (2014) A Golden Gate modular cloning toolbox for plants. ACS Synth Biol 3: 839–843.

Gao, X., Yan, P., Shen, W., Li, X., Zhou, P., and Li, Y. (2013) Modular construction of plasmids by parallel assembly of linear vector components. Anal Biochem 437: 172–177.

Gao, S., Han, L., Zhu, L., Ge, M., Yang, S., Jiang, Y., and Chen, D. (2014) One-step integration of multiple genes into the oleaginous yeast Yarrowia lipolytica. Biotechnol Lett 36: 2523–2528.

Gasmi, N., Fudalej, F., Kallel, H., and Nicaud, J.-M. (2011) A molecular approach to optimize hIFN-2b expression and secretion in Yarrowia lipolytica. Appl Microbiol Biotechnol 89: 109–119.

Iverson, S.V., Haddock, T.L., Beal, J., and Densmore, D.M. (2016) CIDAR MoClo: improved MoClo assembly standard and new E. coli part library enable rapid combinatorial design for synthetic and traditional biology. ACS Synth Biol 5: 99–103.

Kakul, Y., Sunaga, T., Ara, K., Dodgson, J., Ji, L., Csinkas-Nagy, A., et al. (2015) Module-based construction of plasmids for chromosomal integration of the fission yeast Schizosaccharomyces pombe. Open Biol 5: 150054.

Ledesma-Amaro, R., and Nicaud, J.M. (2016) Yarrowia lipolytica as a biotechnological chassis to produce usual and unusual fatty acids. Prog Lipid Res 61: 40–50.

Ledesma-Amaro, R., Dulermo, T., and Nicaud, J.M. (2015) Engineering Yarrowia lipolytica to produce biodiesel from raw starch. Biotechnol Biofuels 8: 148.

Lee, M.E., DeLoache, W.C., Cervantes, B., and Dueber, J.E. (2015) A highly characterized yeast toolkit for modular, multipart assembly. ACS Synth Biol 4: 975–986.

Liu, L., Otoupal, P., Pan, A., and Alper, H.S. (2014) Increasing expression level and copy number of a Yarrowia lipolytica plasmid through regulated centromere function. FEMS Yeast Res 14: 1124–1127.

Madzak, C. (2015) Yarrowia lipolytica: recent achievements in heterologous protein expression and pathway engineering. Appl Microbiol Biotechnol 99: 4559–4577.

Matthaus, F., Ketelhut, M., Gatter, M., and Barth, G. (2014) Production of lycopene in the non-carotenoid-producing

© 2017 The Authors. Microbial Biotechnology published by John Wiley & Sons Ltd and Society for Applied Microbiology, Microbial Biotechnology, 10, 450–455
york Yarrowia lipolytica. Appl Environ Microbiol 80: 1660–1669.
Mitchell, L.A., Chuang, J., Agmon, N., Khunsriraksakul, C., Phillips, N.A., Cai, Y., et al. (2015) Versatile genetic assembly system (VEGAS) to assemble pathways for expression in S. cerevisiae. Nucleic Acids Res 43: 6620–6630.
Moore, S.J., Lai, H., Kelwick, R., Chee, S.M., Bell, D., Polizzi, K.M., and Freemont, P.S. (2016) EcoFlex – a multifunctional MoClo kit for E. coli synthetic biology. ACS Synth Biol 5: 1059–1069.
Nicaud, J.M. (2012) Yarrowia lipolytica. Yeast 29: 409–418. Nicaud, J., Fabre, E., and Gaillardin, C. (1989) Expression of invertase activity in Yarrowia lipolytica and its use as a selective marker. Curr Genet 16: 253–260.
Pauthenier, C., and Faulon, J.L. (2014) PrecisePrimer: an easy-to-use web server for designing PCR primers for DNA library cloning and DNA shuffling. Nucleic Acids Res 42: 205–209.
Pauthenier, C., El-Sayyed, H., Rostain, W., Cerisy, T., Gallo, C., Ferreira, R., et al. (2012) BBF RFC 92: The GoldenBricks assembly: A standardized one-shot cloning technique for complete cassette assembly.
Pignède, G., Wang, H.J., Fudalej, F., Seman, M., Gaillardin, C., and Nicaud, J.M. (2000) Autocloning and amplification of LIP2 in Yarrowia lipolytica. Appl Environ Microbiol 66: 3283–3289.
Pignède, G., Wang, H., Fudalej, F., Gaillardin, C., Seman, M., and Nicaud, J.M. (2000) Characterization of all extracellular lipase encoded by LIP2 in Yarrowia lipolytica. J Bacteriol 182: 2802–2810.
Sambrook, J., and Russell, D. (2001) Molecular Cloning: A Laboratory Manual, 3rd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
Sands, B., and Brent, R. (2016) Overview of post Cohen-Boyer methods for single segment cloning and for multisegment DNA assembly. Current Protocols in Molecular Biology 113: 3.26.1–3.26.20.
Terfrüchte, M., Joehnk, B., Fajardo-Somera, R., Braus, G.H., Riquelme, M., Schipper, K., and Feldbrügge, M. (2014) Establishing a versatile Golden Gate cloning system for genetic engineering in fungi. Fungal Genet Biol 62: 1–10.

Supporting information
Additional Supporting Information may be found online in the supporting information tab for this article:
Table S1. A list of plasmids and oligonucleotides used in this study.