Standards for plant synthetic biology: a common syntax for exchange of DNA parts

Summary

Inventors in the field of mechanical and electronic engineering can access multitudes of components and, thanks to standardization, parts from different manufacturers can be used in combination with each other. The introduction of BioBrick standards for the assembly of characterized DNA sequences was a landmark in microbial engineering, shaping the field of synthetic biology. Here, we describe a standard for Type IIS restriction endonuclease-mediated assembly, defining a common syntax of 12 fusion sites to enable the facile assembly of eukaryotic transcriptional units. This standard has been developed and agreed by representatives and leaders of the international plant science and synthetic biology communities, including inventors, developers and adopters of Type IIS cloning methods. Our vision is of an extensive catalogue of standardized, characterized DNA parts that will accelerate plant bioengineering.

Introduction

The World Bank estimates that almost 40% of land mass is used for cultivation of crop, pasture or forage plants (World Development Indicators, The World Bank 1960–2014). Plants also underpin production of building and packing materials, medicines, paper and decorations, as well as food and fuel. Plant synthetic biology offers the means and opportunity to engineer plants and algae for new roles in our environment, to produce therapeutic compounds and to address global problems such as food insecurity and the contamination of ecosystems with agrochemicals and macronutrients. The adoption of assembly standards will greatly accelerate the pathway from product design to market, enabling the full potential of plant synthetic biology to be realized.

The standardization of components, from screw threads to printed circuit boards, drives both the speed of innovation and the economy of production in mechanical and electronic engineering. Products as diverse as ink-jet printers and airplanes are designed and constructed from component parts and devices. Many of these components can be selected from libraries and catalogues of standard parts in which specifications and performance characteristics are described. The agreement and implementation of assembly standards that allow parts, even those from multiple manufacturers, to be assembled together has underpinned invention in these fields.

This conceptual model is the basis of synthetic biology, with the same ideal being applied to biological parts (DNA fragments) for the engineering of biological systems. The first widely-adopted biological standard was the BioBrick, for which sequences and performance data are stored in the Registry of Standard Biological Parts (Knight, 2003). BioBrick assembly standard 10 (BBF RFC 10) was the first biological assembly standard to be introduced. Its key feature is that the assembly reactions are idempotent: each reaction retains the key structural elements of the constituent parts so that resulting assemblies can be used as input in identical assembly processes (Knight, 2003; Shetty et al., 2008). Over the years, several other BioBrick assembly standards have been developed that diminish some of the limitations of standard 10 (Phillips & Silver, 2006; Anderson et al., 2010). Additionally, several alternative technologies have been developed that confer the ability to assemble multiple parts in a single reaction (Engler et al., 2008; Gibson et al., 2009; Quan & Tian, 2009; Li & Elledge, 2012; Kok et al., 2014).

While overlap-dependent methods are powerful and generally result in ‘scarless’ assemblies, their lack of idempotency and the requirement for custom oligonucleotides and amplification of even well characterized standard parts for each new assembly are considerable drawbacks (Ellis et al., 2011; Liu et al., 2013; Patron, 2014). Assembly methods based on Type IIS restriction enzymes, known widely as Golden Gate cloning, are founded on standard parts that can be characterized, exchanged and assembled cheaply, easily and in an automatable way without proprietary tools and reagents (Engler et al., 2009, 2014; Sarrion-Perdigones et al., 2011; Werner et al., 2012).

Type IIS assembly methods have been widely adopted in plant research laboratories with many commonly used sequences being adapted for Type IIS assembly and subsequently published and shared through public plasmid repositories such as AddGene (Sarrion-Perdigones et al., 2011; Weber et al., 2011; Emami et al., 2013; Lampropoulos et al., 2013; Bindet et al., 2014; Engler et al., 2014; Vafaee et al., 2014). Type IIS assembly systems have also been adopted for the engineering of fungi (Terfrüchte et al., 2014) and ‘IP-Free’ host expression systems have been developed for bacteria, mammals and yeast (Whitman et al., 2013).

To reap the benefits of the exponential increase in genomic information and DNA assembly technologies, bioengineers require assembly standards to be agreed for multicellular eukaryotes. A standard for plants must be applicable to the diverse taxa that comprise Archaeplastida and also be capable of retaining the features that minimize the need to reinvent common steps such as transferring genetic material into plant genomes. In this Viewpoint article, the authors of which include inventors, developers and
adopter of Golden Gate cloning methods from multiple international institutions, we define a Type IIS genetic grammar for plants, extendible to all eukaryotes. This sets a consensus for establishing a common language across the plant field, putting in place the framework for a sequence and data repository for plant parts.

**Golden Gate cloning**

Golden Gate cloning is based on Type IIS restriction enzymes and enables parallel assembly of multiple DNA parts in a one-pot, one-step reaction. Contrary to Type II restriction enzymes, Type IIS restriction enzymes recognize nonpalindromic sequence motifs and cleave outside of their recognition site (Fig. 1a). These features enable the production of user-defined overhangs on either strand, which in turn allow multiple parts to be assembled in a predetermined order and orientation using only one restriction enzyme. Parts are released from their original plasmids and assembled into a new plasmid backbone in the same reaction, bypassing time-consuming steps such as custom primer design, PCR amplification and gel purification (Fig. 1b).

The one-step digestion–ligation reaction can be performed with any collection of plasmid vectors and parts providing that:

1. Parts are housed in plasmids flanked by a convergent pair of Type IIS recognition sequences;
2. The accepting plasmid has a divergent pair of recognition sequences for the same enzyme, between which the part or parts will be assembled;
3. The parts themselves, and all plasmid backbones, are otherwise free of recognition sites for this enzyme;
4. None of the parts are housed in a plasmid backbone with the same antibiotic resistance as the accepting plasmid into which parts will be assembled;
5. The overhangs created by digestion with the Type IIS restriction enzymes are unique and nonpalindromic.

To date, several laboratories have converted ‘in-house’ and previously published plasmids for use with Golden Gate cloning and have assigned compatible overhangs to standard elements such as promoters, coding sequences and terminators found in eukaryotic genes (Sarrion-Perdigones et al., 2011; Weber et al., 2011; Emami et al., 2013; Lampropoulos et al., 2013; Binder et al., 2014; Engler et al., 2014). The GoldenBraid2.0 (GB2.0) and Golden Gate Modular Cloning (MoClo) assembly standards, the main features of which are described later, are both widely used having been adopted by large communities of plant research laboratories such as the European Cooperation in Science and Technology (COST) network for plant metabolic engineering, the Engineering Nitrogen Symbiosis for Africa (ENSA) project, the C₄ Rice project and the Realizing Increased Photosynthetic Activity (RIPE) project. MoClo and GB2.0 are largely, though not entirely, interoperable.

![Fig. 1](image-url)

Fig. 1 (a) Type IIS restriction enzymes such as *Bsa*I are directional, cleaving outside of their nonpalindromic recognition sequences. (b) Providing compatible overhangs are produced on digestion, standard parts cloned in plasmid backbones flanked by a pair of convergent Type IIS restriction enzyme recognition sites can be assembled in a single digestion–ligation reaction into an acceptor plasmid with divergent Type IIS restriction enzyme recognition sites and a unique bacterial selection cassette.
compatible. Other standards have been developed independently resulting in parts that are noninterchangeable with laboratories using MoClo or GB2.0. Even small variations prevent the exchange of parts and hinder the creation of a registry of standard, characterized, exchangeable parts for plants. The standard syntax defined later addresses these points, establishing a common grammar to enable the sharing of parts throughout the plant science community, whilst maintaining substantial compatibility with the most widely adopted Type IIS-based standards.

A standard Type IIS syntax for plants

Plasmid backbones of standard parts

For sequences to be assembled reliably in a desired order and in a single step, all internal instances of the Type IIS restriction enzyme recognition sequence must be removed. The removal of such sites and the cloning into a compatible backbone, flanked by a convergent pair of Type IIS restriction enzyme recognition sequences, is described as ‘domestication’. Assembly of standard parts into a complete transcriptional unit uses the enzyme BsdI. Standard parts for plants must minimally, therefore, be domesticated for BsdI (Fig. 2). Parts must also be housed in plasmid backbones that, apart from the convergent pair of BsdI recognition sites flanking the part, are otherwise free from this motif. The plasmid backbone should also not contain bacterial resistance to ampicillin/carbenicillin or kanamycin as these are commonly utilized in the plasmids in which standard parts will be assembled into complete transcriptional units (Sarrion-Perdigones et al., 2013; Engler et al., 2014) (Fig. 2). When released from its plasmid backbone by BsdI, each part will contain specific, four-base-pair, 5’ overhangs, known as fusion sites (Fig. 2).

For assembly of transcriptional units into multigene constructs MoClo and GB2.0 require that parts are free of at least one other enzyme. In both systems transcriptional units can be used directly or may be assembled with other transcriptional units to make multigene assemblies. MoClo uses BpiI to assemble multiple transcriptional units in a single step. These can be reassembled into larger constructs using either BsdI and BsmBI (Weber et al., 2011) or by an iterative, fast-track method that alternates between BsdI and BpiI (Werner et al., 2012). GB2.0 uses BsdI and BsmBI for iterative assembly of transcriptional units into multigene constructs (Sarrion-Perdigones et al., 2013). All three enzymes recognize six base-pair sequences and produce four-base-pair 5’ overhangs. Compatibility with MoClo and GB2.0 multigene assemble plasmid systems can therefore be obtained by domesticating BpiI and BsmBI as well as BsdI recognition sequences (Fig. 2).

Standard parts

A standard syntax for eukaryotic genes has been defined and 12 fusion points assigned (Fig. 3). Such complexity allows for the complex and precise engineering of genes that is becoming increasingly important for plant synthetic biology. Standard parts are sequences that have been cloned into a compatible backbone (described earlier) and are flanked by a convergent pair of BsdI recognition sequences and two of the defined fusion sites. The sequence can comprise just one of the 10 defined parts of genetic

---

**Fig. 2** (a) Standard parts for plants are free from BsdI recognition sequences. To be compatible with Golden Gate Modular Cloning (MoClo) and GoldenBraid2.0 (GB2.0) they must also be free from BpiI and BsmBI recognition sequences. (b) Standard parts are housed in plasmid backbones flanked by convergent BsdI recognition sequences. The plasmid backbones are otherwise free from BsdI recognition sites. The plasmid backbone should not confer bacterial resistance to ampicillin, carbenicillin or kanamycin. When released from their backbone by BsdI, parts are flanked by four-base-pair 5’ overhangs, known as fusion sites.
syntax bounded by an adjacent pair of adjacent fusion sites. However, when the full level of complexity is unnecessary, or if particular functional elements such as amino (N)- or carboxyl (C)-terminal tags are not required, standard parts can comprise sequences that span multiple fusion sites (Fig. 3).

The sequences that comprise the fusion sites have been selected both for maximum compatibility in the one-step digestion–ligation reaction and to maximize biological functionality. The 5' nontranscribed region is separated into core, proximal and distal promoter sequences, with the core region containing the transcriptional start site (TSS). The transcribed region is separated into coding parts and 5' and 3' untranslated parts. For maximum flexibility, an ATG codon for methionine is wholly or partially encoded into two fusion sites. The translated region, therefore, may be divided into three or four parts. The 3' nontranslated region is followed by the 3' nontranscribed region, which contains the polyadenylation sequence (PAS). Amino acids coded by fusion sites within the coding region have been rationally selected: neutral, nonpolar amino acids, methionine and alanine, are encoded in the 3' overhangs of parts that may be used to house signal and transit peptides in order to prevent interference with recognition and cleavage. An alternative overhang, encoding a glycine, is also included to give greater flexibility for the fusion of noncleaved coding parts. Serine, a small amino acid commonly used to link peptide and reporter tags, is encoded in the overhang that will fuse C-terminal tag parts to coding sequences.

Universal acceptor plasmids (UAPs)

Universal acceptor plasmids (UAPs) allow the conversion of any sequence to a standard part in a single step (Fig. 4). This is achieved by PCR amplification of desired sequences as a single fragment or, if restriction sites need to be domesticated, as multiple fragments (Fig. 4). The oligonucleotide primers used for amplification add 5’ sequences to allow cloning into the UAP, add the standard fusion sites that the sequence will be flanked with when released from the UAP as a standard part with BsaI and can also introduce mutations (Fig. 4). Two UAPs, pUPD2 (https://gbcloning.org/feature/GB0307) and pUAP1 (AddGene No. 63674) can be used to create new standard parts in the chloramphenicol resistant pSB1C3 backbone, in which the majority of BioBricks housed at the Registry of Standard Parts are cloned. A spectinomycin resistant UAP, pAGM9121 has been published previously (AddGene No. 51833; Engler et al., 2014).

Compatibility with multigene assembly systems

Standard parts are assembled into transcriptional units in plasmid vectors that contain the features and sequences required for delivery to the cell, for example Left border (LB) and Right border (RB) sequences and an origin of replication for Agrobacterium-mediated delivery. Subsequently, transcriptional units can be assembled into multigene constructs in plasmid acceptors that also contain these features. It is important that a standard Type IIS syntax be
compatible with the plasmid vector systems that are in common use such as GB2.0 and MoClo while also allowing space for further innovation in Type IIS-mediated multigene assembly methodologies and the development of plasmid vectors with features required for delivery to other species and by other delivery methods. The definition of a standard Type IIS syntax for plants is therefore timely and will allow the growing plant synthetic biology community access to an already large library of standard parts.

Conclusions

Synthetic biology aims to simplify the process of designing, constructing and modifying complex biological systems. Plants provide an ideal chassis for synthetic biology, are amenable to genetic engineering and have relatively simple requirements for growth (Cook et al., 2014; Fesenko & Edwards, 2014). However, their eukaryotic gene structure and the methods commonly used for transferring DNA to their genomes demand specific plasmid vectors and a tailored assembly standard. Here, we have defined a Type IIS genetic syntax that employs the principles of part reusability and standardization. The standard has also been submitted as a Request for Comments (BBF RFC 106) (Rutten et al., 2015) at the BioBrick Foundation to facilitate iGEM teams working on plant chassis. Using the standards described here, new standard parts for plants can be produced and exchanged between laboratories enabling the facile construction of transcriptional units. We invite the plant science and synthetic biology communities to build on this work by adopting this standard to create a large repository of characterized standard parts for plants.

Fig. 4 (a) Universal acceptor plasmids (UAPs) comprise a small plasmid backbone conferring resistance to spectinomycin or chloramphenicol in bacteria. They contain a cloning site consisting of a pair of divergent Type IIS recognition sequences (e.g. BpiI, as depicted, or BsmBI) flanked by overlapping convergent Bsal recognition sequences. (b) A sequence containing an illegal BsaI recognition sequence can be amplified in two fragments using oligonucleotide primers with 5' overhangs (red dashed lines) that (i) introduce a mutation to destroy the illegal site (reversed type), (ii) add Type IIS recognition sequences (e.g. BpiI, as depicted, or BsmBI) and fusion sites to allow one step digestion–ligation into the universal acceptor, and (iii) add the desired fusion sites (green numbers) that will define the type of standard part and that will flank the part when released from the backbone with BsaI. (c) When the resulting amplicons are cloned into a UAP, the new standard part will be flanked by a pair of convergent BsaI recognition sequences capable of releasing the part with the desired fusion sites (green numbers).

Acknowledgements

This work was supported by the UK Biotechnological and Biological Sciences Research Council (BBSRC) Synthetic Biology Research Centre ‘OpenPlant’ award (BB/L014130/1), BBSRC grant no. BB/K005952/1 (A.O. and A.L.), BBSRC grant no. BB/L02182X/1 (A.A.R.W.), the Spanish MINECO grant no. BIO2013-42193-R (D.O.), the BBSRC Institute Strategic

© 2015 The Authors
New Phytologist © 2015 New Phytologist Trust
Programme Grants ‘Understanding and exploiting plant and microbial metabolism’ and ‘Biotic interactions for crop productivity’, the John Innes Foundation and the Gatsby Foundation. Supported by the Engineering and Physical Sciences Research Council (EPSRC) project, through a grant to the John Innes Centre from The Bill & Melinda Gates Foundation, the DOE Early Career Award and the DOE Joint BioEnergy Institute supported by the U.S. Department of Energy, Office of Biological and Environmental Research through contract DE-AC02-05CH1123. The authors also acknowledge the support of COST Action FA1006, PlantEngine.

Nicola J. Patron1,2,4, Diego Orzaez3, Sylvester Marillonnet4, Heribert Warzecha5, Colette Matthewman2,6, Mark Youles1, Oleg Raitskin1,2, Aymeric Leveau6, Gemma Farre6, Christian Rogers6, Alison Smith7, Julian Hibberd1,2, Alex A. R. Webb2,7, James Locke2,8, Sebastian Schornack2,8, Jim Ajoka2,5,6, David C. Baulcombe2,7, Cyril Zipfel1, Sophie Kamoun1, Jonathan D. G. Jones1, Hannah Kuhn1, Silke Robatzek1, H. Peter Van Esse1, Dale Sanders2,6, Giles Oldroyd2,6, Cathie Martin2,6, Rob Field2,6, Sarah O’Connor2,6, Samantha Fox2,6, Ben Miller6, Andy Breakspear6, Guru Radhakrishnan6, Pierre-Marc Delaux5, Dominique Loque10,11, Antonio Granell2,10,11, Alain Tissier1, Patrick Shih12, Thomas P. Brutnell12, W. Paul Quick13, Heiko Rischer14, Paul D. Fraser15, Asaph Aharoni16, Christine Raines17, Paul F. South18, Jean-Michel Ane19, Björn R. Hamberger20, Jane Langdale21, Jens Stougaard22, Harro Bouwmeester23, Michael Udvardi24, James A. H. Murray25, Vardis Ntoukakis26, Patrick Schüler26, Katherine Denby26, Keith J. Edwards27, Anne Osbourn2,6 and Jim Haseloff2,7

1The Sainsbury Laboratory, Norwich Research Park, Norfolk, NR4 7RG, UK;
2OpenPlant Consortium: The University of Cambridge, The John Innes Centre and The Sainsbury Laboratory, Norwich, NR4 7UH, UK;
3Instituto de Biología Molecular y Celular de Plantas (IBMCP), Consejo Superior de Investigaciones Científicas, Universidad Politécnica de Valencia, Avda Tarongers SN, Valencia, Spain;
4Leibniz-Institut für Pflanzenbiochemie, Weinberg 3, 06120, Halle (Saale), Germany;
5Plant Biotechnology and Metabolic Engineering, Technische Universität Darmstadt, Schnittspahnstrasse 4, Darmstadt 64287, Germany;
6The John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK;
7Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EA, UK;
8The Sainsbury Laboratory, Cambridge University, Bateman Street, Cambridge, CB2 1LR, UK;
9Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QF, UK;
10Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA;
11Joint BioEnergy Institute, EmeryStation East, 5885 Hollis St, 4th Floor, Emeryville, CA 94608, USA;
12The Donald Danforth Plant Science Center, St Louis, MO 63132, USA;
13Department of Animal and Plant Sciences, University of Sheffield, Sheffield, S10 2TN, UK;
14VTT Technical Research Centre of Finland, Espoo 02044, Finland;
15School of Biological Sciences, Royal Holloway, University of London, Egham Hill, Egham, TW20 0EX, UK;
16Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel;
17School of Biological Sciences, University of Essex, Colchester, CO4 3SQ, UK;
18United States Department of Agriculture, Global Change and Photosynthesis Research Unit, ARS 1206 West Gregory Drive, Urbana, IL 61801, USA;
19Departments of Bacteriology and Agronomy, University of Wisconsin, 1575 Linden Drive, Madison, WI, USA;
20Biochemistry Laboratory, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, Frederiksberg C, Denmark;
21Department of Plant Sciences, University of Oxford, Oxford, OX1 3RB, UK;
22Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10C, Aarhus, Denmark;
23Wageningen UR, Wageningen University, Wageningen 6700 AA, the Netherlands;
24Plant Biology Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA;
25School of Biosciences, Sir Martin Evans Building, Cardiff University, Museum Avenue, Cardiff, CF10 3AX, UK;
26Warwick Integrative Synthetic Biology Centre and School of Life Sciences, University of Warwick, Coventry, CV4 7AL, UK;
27BrisSynBio, Life Sciences Building, University of Bristol, Tyndall Avenue, Bristol, BS8 1TQ, UK (*)Author for correspondence: tel: +44 1603 450527; email nicola.patron@tsl.ac.uk

References


Key words: cloning, DNA assembly, genetic syntax, GoldenGate, synthetic biology, Type IIS restriction endonucleases.