View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Apollo

1 **Standards for Plant Synthetic Biology: A Common Syntax for Exchange of DNA**
2 **Parts**
3

4 Nicola J Patron*[1,2], Diego Orzaez[3], Sylvestre Marillonnet[4], Heribert Warzecha[5],
5 Colette Matthewman[2,6], Mark Youles[1], Oleg Raitskin[1,2], Aymeric Leveau[6], Gemma
6 Farré[6], Christian Rogers[6], Alison Smith[2,7], Julian Hibberd[2,7], Alex AR Webb[2,7], James
7 Locke[2,8], Sebastian Schornack[2,8], Jim Ajioka[2,7], David C Baulcombe[2,7], Cyril Zipfel[1],
8 Sophien Kamoun[1], Jonathan DG Jones[1],Hannah Kuhn[1], Silke Robatzek[1], H Peter Van
9 Esse[1], Giles Oldroyd[2,6], Cathie Martin[2,6], Rob Field [2,6], Sarah O'Connor[2,6], Samantha
10 Fox[2], Brande Wulff [2], Ben Miller[2], Andy Breakspear[2], Guru Radhakrishnan[2], Pierre-
11 Marc Delaux[2], Dominique Loque[9], Antonio Granell[3], Alain Tissier[4], Patrick Shih[9],
12 Thomas P Brutnell[10], Paul Quick W[11], Heiko Rischer[12], Paul D Fraser[13], Asaph
13 Aharoni[14], Christine Raines[15], Paul South[16], Jean-Michel Ané[17], Björn R Hamberger[18],
14 Jane Langdale[19], Jens Stougaard[20], Harro Bouwmeester[21], Michael Udvardi[22], James
15 AH Murray[23], Anne Osbourn[2,6], Jim Haseloff[2,7]
16
17 *nicola.patron@tsl.ac.uk +44 1603 450527
18
19 **Affiliations**
20 [1] The Sainsbury Laboratory, Norwich Research Park, Norfolk, UK
21 [2] OpenPlant Consortium: The University of Cambridge, The John Innes Centre and The
22 Sainsbury Laboratory.
23 [3] Instituto de Biología Molecular y Celular de Plantas (IBMCP), Consejo Superior de
24 Investigaciones Científicas, Universidad Politécnica de Valencia. Avda Tarongers SN,
25 Valencia, Spain
26 [4] Leibniz-Institut für Pflanzenbiochemie, Weinberg 3, 06120 Halle (Saale), Germany
27 [5] Technische Universität Darmstadt, Plant Biotechnology and Metabolic Engineering,
28 Schnittspahnstrasse 4, Darmstadt, Germany
29 [6] The John Innes Centre, Norwich Research Park, Norfolk, UK
30 [7] Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, UK
31 [8] The Sainsbury Laboratory, Cambridge University, Bateman Street, Cambridge, UK
32 [9] Joint BioEnergy Institute, Emeryville, CA, USA
33 [10] The Donald Danforth Plant Science Center, St. Louis, Missouri, USA
34 [11] Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK
35 [12] VTT Technical Research Centre of Finland, Finland

36 [13] School of Biological Sciences, Royal Holloway, University of London, Egham Hill, Egham,
37 Surrey, UK
38 [14] Department of Plant Sciences, Weizmann Institute of Science, Rehovot, Israel
39 [15] School of Biological Sciences, University of Essex, Colchester, UK
40 [16] United States Department Of Agriculture, ARS 1206 West Gregory Drive, Urbana, IL,
41 USA
42 [17] Departments of Bacteriology and Agronomy, University of Wisconsin Madison, 1575
43 Linden Drive, Madison, WI ,USA
44 [18] Biochemistry Laboratory, Department of Plant and Environmental Sciences University of
45 Copenhagen, Thorvaldsensvej 40, Frederiksberg C, Denmark
46 [19] Department of Plant Sciences, University of Oxford, Oxford, United Kingdom.
47 [20] Centre for Carbohydrate Recognition and Signalling, Department of Molecular Biology and
48 Genetics, Aarhus University, Gustav Wieds Vej 10C, Aarhus, Denmark.
49 [21] Wageningen UR, Wageningen University, Wageningen, the Netherlands.
50 [22] Plant Biology Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway,
51 Ardmore, OK, USA
52 [23] School of Biosciences, Sir Martin Evans Building, Cardiff University, Museum Avenue,
53 Cardiff, Wales, UK
54
55 ***Total word count main body of the text: 2337***
56 *Abstract: 129*
57 *Introduction: 689*
58 *Golden Gate Cloning: 438*
59 *A Type IIS Syntax for Plants: 912*
60 *Summary: 169*
61 ***Number of figures:***
62 *Four color figures*
63

Summary

Inventors in the field of mechanical and electronic engineering can access multitudes of components and, thanks to standardisation, parts from different manufacturers can be used in combination with each other. The introduction of BioBrick standards for the assembly of characterised 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 twelve 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 standardised, characterised 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. 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 realised.

The standardisation 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.

98 This conceptual model is the basis of synthetic biology, with the same ideal being
99 applied to biological parts (DNA fragments) for the engineering of biological
100 systems. The first widely-adopted biological standard was the BioBrick, for which
101 sequences and performance data are stored in the Registry of Standard Biological
102 Parts (Knight, 2003). BioBrick assembly standard 10 (BBF RFC 10) was the first
103 biological assembly standard to be introduced. Its key feature is that the assembly
104 reactions are idempotent: each reaction retains the key structural elements of
105 the constituent parts so that resulting assemblies can be used as input in identical
106 assembly processes (Knight, 2003; Shetty *et al.*, 2008). Over the years, several other
107 BioBrick assembly standards have been developed that diminish some of the
108 limitations of standard 10 (Phillips & Silver, 2006; Anderson *et al.*, 2010).
109 Additionally, several alternative technologies have been developed that confer the
110 ability to assemble multiple parts in a single reaction (Engler *et al.*, 2008; Gibson *et*
111 *al.*, 2009; Quan & Tian, 2009; Li & Elledge, 2012; Kok *et al.*, 2014).
112
113 While overlap-dependent methods are powerful and generally result in 'scarless'
114 assemblies, their lack of idempotency and the requirement for custom
115 oligonucleotides and amplification of even well characterised standard parts for each
116 new assembly are considerable drawbacks (Ellis *et al.*, 2011; Liu *et al.*, 2013; Patron,
117 2014). Assembly methods based on Type IIS restriction enzymes, known widely as
118 Golden Gate cloning, are founded on standard parts that can be characterised,
119 exchanged and assembled cheaply, easily, and in an automatable way without
120 proprietary tools and reagents (Engler *et al.*, 2009, 2014; Sarrion-Perdigones *et al.*,
121 2011; Werner *et al.*, 2012).
122
123 Type IIS assembly methods have been widely adopted in plant research laboratories
124 with many commonly used sequences being adapted for Type IIS assembly and
125 subsequently published and shared through public plasmid repositories such as
126 AddGene (Sarrion-Perdigones *et al.*, 2011; Weber *et al.*, 2011; Emami *et al.*, 2013;
127 Lampropoulos *et al.*, 2013; Binder *et al.*, 2014; Engler *et al.*, 2014; Vafaee *et al.*,
128 2014). Type IIS assembly systems have also been adopted for the engineering of
129 fungi (Terfrüchte *et al.*, 2014) and 'IP-Free' host expression systems have been
130 developed for bacteria, mammals and yeast (Whitman *et al.*, 2013).
131

132 To reap the benefits of the exponential increase in genomic information, DNA
133 assembly and bioengineering technologies, biological assembly standards must be
134 agreed for multicellular eukaryotes. A standard for plants must be applicable to the
135 diverse taxa that comprise Archaeplastida and also be capable of retaining the features
136 that minimize the need to re-invent common steps such as transferring genetic
137 material into plant genomes. In this letter, the authors of which include inventors,
138 developers and adopters of Golden Gate cloning methods from multiple international
139 institutions, we define a Type IIS genetic grammar for plants, extendible to all
140 eukaryotes. This sets a consensus for establishing a common language across the plant
141 field, putting in place the framework for a sequence and data repository for plant
142 parts.
143
144 **Golden Gate Cloning**
145 Golden Gate cloning is based on Type IIS restriction enzymes and enables parallel
146 assembly of multiple DNA parts in a one-pot, one-step reaction. Contrary to Type II
147 restriction enzymes, Type IIS restriction enzymes recognise non-palindromic
148 sequence motifs and cleave outside of their recognition site (**Figure 1A**). These
149 features enable the production of user-defined overhangs on either strand, which in
150 turn allow multiple parts to be assembled in a pre-determined order and orientation
151 using only one restriction enzyme. Parts are released from their original plasmids and
152 assembled into a new plasmid backbone in the same reaction, bypassing time-
153 consuming steps such as custom primer design, PCR amplification and gel
154 purification (**Figure 1B**).
155
156 The one-step digestion-ligation reaction can be performed with any collection of
157 plasmid vectors and parts providing that:
158
159 (a) Parts are housed in plasmids flanked by a convergent pair of Type IIS recognition
160 sequences
161 (b) The accepting plasmid has a divergent pair of recognition sequences for the same
162 enzyme, between which the part or parts will be assembled
163 (c) The parts themselves, and all plasmid backbones, are otherwise free of recognition
164 sites for this enzyme

165 (d) None of the parts are housed in a plasmid backbone with the same antibiotic
166 resistance as the accepting plasmid into which parts will be assembled
167 (e) The overhangs created by digestion with the Type IIS restriction enzymes are
168 unique and non-palindromic
169
170 To date, several laboratories have converted ‘in-house’ and previously published
171 plasmids for use with Golden Gate cloning and have assigned compatible overhangs
172 to standard elements such as promoters, coding sequences and terminators found in
173 eukaryotic genes (Sarrion-Perdigones *et al.*, 2011; Weber *et al.*, 2011; Emami *et al.*,
174 2013; Lampropoulos *et al.*, 2013; Binder *et al.*, 2014; Engler *et al.*, 2014). The
175 GoldenBraid2.0 (GB2.0) and Golden Gate Modular Cloning (MoClo) assembly
176 standards, the main features of which are described below, are both widely used
177 having been adopted by large communities of plant research laboratories such as the
178 European Cooperation in Science and Technology (COST) network for plant
179 metabolic engineering, the Engineering Nitrogen Symbiosis for Africa (ENSA)
180 project, the C4 Rice project and the Realizing Increased Photosynthetic Activity
181 (RIPE) project. MoClo and GB2.0 are largely, though not entirely, compatible. Other
182 standards have been developed independently resulting in parts that are non-
183 interchangeable with laboratories using MoClo or GB2.0. Even small variations
184 prevent the exchange of parts and hinder the creation of a registry of standard,
185 characterised, exchangeable parts for plants. The standard syntax defined below
186 addresses these points, establishing a common grammar to enable the sharing of parts
187 throughout the plant science community, whilst maintaining substantial compatibility
188 with the most widely adopted Type IIS-based standards.
189
190 **A Standard Type IIS Syntax for Plants**
191
192 *Plasmid backbones of standard parts*
193 For sequences to be assembled reliably in a desired order and in a single step, all
194 internal instances of the Type IIS restriction enzyme recognition sequence must be
195 removed. The removal of such sites and the cloning into a compatible backbone,
196 flanked by a convergent pair of Type IIS restriction enzyme recognition sequences, is
197 described as ‘domestication’. Assembly of standard parts into a complete
198 transcriptional unit uses the enzyme *Bsa*I. Standard parts for plants must minimally,

199 therefore, be domesticated for *Bsa*I (Figure 2). Parts must also be housed in plasmid
200 backbones that, apart from the convergent pair of *Bsa*I recognition sites flanking the
201 part, are otherwise free from this motif. The plasmid backbone should also not contain
202 bacterial resistance to ampicillin/carbenicillin or kanamycin as these are commonly
203 utilised in the plasmids in which standard parts will be assembled into complete
204 transcriptional units (e.g. Sarrion-Perdigones *et al.*, 2013; Engler *et al.*, 2014) (Figure
205 2). When released from its plasmid backbone by *Bsa*I, each part will contain specific,
206 four-base-pair, 5' overhangs, known as fusion sites (Figure 2).
207
208 For assembly of transcriptional units into multi-gene constructs MoClo and GB2.0
209 require that parts are free of at least one other enzyme. MoClo uses *Bpi*I to assemble
210 multiple transcriptional units in a single step, with subsequent assembly of larger
211 constructs using *Bsa*I or *Bsm*BI. GB2.0 uses *Bsa*I and *Bsm*BI for iterative assembly of
212 transcriptional units into multigene constructs (Figure 2). All three enzymes recognise
213 six base-pair sequences (and produce four-base-pair 5' overhangs) therefore
214 recognition sites are relatively rare. Compatibility with MoClo and GB2.0 multi-gene
215 assemble plasmid systems can be obtained by domesticating *Bpi*I and *Bsm*BI as well
216 as *Bsa*I recognition sequences (Figure 2).
217
218 *Standard parts*
219 A standard syntax for eukaryotic genes has been defined and twelve fusion points
220 assigned (Figure 3). Such complexity allows for the complex and precise engineering
221 of genes that is becoming increasingly important for plant synthetic biology.
222 Standard parts are sequences that have been cloned into a compatible backbone
223 (described above) and are flanked by a convergent pair of *Bsa*I recognition sequences
224 and two of the defined fusion sites. The sequence can comprise just one of the ten
225 defined parts of genetic syntax bounded by an adjacent pair of adjacent fusion sites.
226 However, when the full level of complexity is unnecessary, or if particular functional
227 elements such as N- or C-terminal tags are not required, standard parts can comprise
228 sequences that span multiple fusion sites (Figure 3).
229
230 The sequences that comprise the fusion sites have been selected both for maximum
231 compatibility in the one-step digestion-ligation reaction and to maximise biological
232 functionality. The 5' non-transcribed region is separated into core, proximal and distal

233 promoter sequences, with the core region containing the transcriptional start site
234 (TSS). The transcribed region is separated into coding parts and 5' and 3' untranslated
235 parts. For maximum flexibility, an ATG codon for methionine is wholly or partially
236 encoded into two fusion sites. The translated region, therefore, may be divided into
237 three or four parts. The 3' non-translated region is followed by the 3' non-transcribed
238 region, which contains the polyadenylation sequence (PAS). Amino acids coded by
239 fusion sites within the coding region have been rationally selected: Neutral, non-polar
240 amino acids, methionine and alanine, are encoded in the 3' overhangs of parts that
241 may be used to house signal and transit peptides in order to prevent interference with
242 recognition and cleavage. An alternative overhang, encoding a glycine, is also
243 included to give greater flexibility for the fusion of non-cleaved coding parts. Serine,
244 a small amino acid commonly used to link peptide and reporter tags, is encoded in the
245 overhang that will fuse C terminal tag parts to coding sequences.
246
247 *Universal acceptor plasmids*
248 Universal acceptor plasmids (UAP) allow the conversion of any sequence to a
249 standard part in a single step (Figure 4). This is achieved by polymerase chain
250 reaction amplification of desired sequences as a single fragment or, if restriction sites
251 need to be domesticated, as multiple fragments (Figure 4). The oligonucleotide
252 primers used for amplification add 5' sequences to allow cloning into the UAP, add
253 the standard fusion sites that the sequence will be flanked with when released from
254 the UAP as a standard part with *Bsa*I and can also introduce mutations (Figure 4).
255 Two UAPs, pUPD2 (https://gbcloning.org/feature/GB0307/) and pUAP1 (AddGene
256 #71721) can be used to create new standard parts in the chloramphenicol resistant
257 pSB1C3 backbone, in which the majority of BioBricks housed at the Registry of
258 Standard Parts are cloned. A spectinomycin resistant UAP, pAGM9121 has been
259 published previously (AddGene #52833 (Engler *et al.*, 2014)).
260
261 *Compatibility with multigene assembly systems*
262 Standard parts are assembled into transcriptional units in plasmid vectors that contain
263 the features and sequences required for delivery to the cell, for example Left (LB) and
264 Right Border (RB) sequences and an origin of replication for *Agrobacterium*-
265 mediated delivery. Subsequently, transcriptional units can be assembled into
266 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.

Summary

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 standardisation. 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 characterised standard parts for plants.

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 (AO and AL), BBSRC grant no. BB/L02182X/1(AW), the Spanish MINECO grant no. BIO2013-42193-R (DO), the BBSRC Institute Strategic Programme Grants 'Understanding and Exploiting Plant and Microbial Metabolism' and 'Biotic Interactions for Crop Productivity', the John Innes Foundation and the Gatsby Foundation. The authors also acknowledge the support of COST Action FA1006, PlantEngine.

300 **Anderson JC, Dueber JE, Leguia M, Wu GC, Goler JA, Arkin AP, Keasling JD**.
301 **2010**. BglBricks: A flexible standard for biological part assembly. *Journal of*
302 *Biological Engineering* **4**: 1.

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

306 **Cook C, Bastow R, Martin L**. **2014**. *Developing plant synthetic biology in the UK -*
307 *Challenges and opportunities*. Report from the 2013 GARNet meeting 'An
308 Introduction to Opportunities in Plant Synthetic Biology'.

309 **Ellis T, Adie T, Baldwin GS**. **2011**. DNA assembly for synthetic biology: from parts
310 to pathways and beyond. *Integrative biology: quantitative biosciences from nano to*
311 *macro* **3**: 109–18.

312 **Emami S, Yee M-C, Dinneny JR**. **2013**. A robust family of Golden Gate
313 Agrobacterium vectors for plant synthetic biology. *Frontiers in plant science* **4**: 339.

314 **Engler C, Gruetzner R, Kandzia R, Marillonnet S**. **2009**. Golden Gate shuffling: A
315 one-pot DNA shuffling method based on type IIs restriction enzymes. *PloS one* **4**:
316 e5553.

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

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

322 **Fesenko E, Edwards R**. **2014**. Plant synthetic biology: A new platform for industrial
323 biotechnology. *Journal of Experimental Botany* **65**: 1927–1937.

324 **Gibson D, Young L, Chuang R, Venter C, Hutchison III CA, Smith H**. **2009**.
325 Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature*
326 *Methods* **6**: 343–345.

327 **Knight T**. **2003**. Idempotent vector design for standard assembly of BioBricks. *MIT*
328 *Artificial Intelligence Laboratory; MIT Synthetic Biology Working Group*.

329 **Kok S, Stanton L, Slaby T**. **2014**. Rapid and reliable DNA assembly via ligase
330 cycling reaction. *ACS Synthetic Biology* **3**: 97–106.

331 **Lampropoulos A, Sutikovic Z, Wenzl C, Maegele I, Lohmann JU, Forner J**.
332 **2013**. GreenGate - A novel, versatile, and efficient cloning system for plant
333 transgenesis. *PloS one* **8**: e83043.

334 **Li MZ, Elledge SJ**. **2012**. SLIC: A method for sequence- and ligation-independent
335 cloning. In: Peccoud J, ed. Methods in Molecular Biology. Totowa, NJ: Humana
336 Press, 51–59.

337 **Liu W, Yuan JS, Stewart CN**. **2013**. Advanced genetic tools for plant
338 biotechnology. *Nature reviews. Genetics* **14**: 781–93.

339 **Patron NJ**. **2014**. DNA assembly for plant biology: techniques and tools. *Current*
340 *Opinion in Plant Biology* **19C**: 14–19.

341 **Phillips I, Silver P**. **2006**. *BBF RFC 23: A new BioBrick assembly strategy designed*
342 *for facile protein engineering*. http://hdl.handle.net/1721.1/32535

343 **Quan J, Tian J**. **2009**. Circular polymerase extension cloning of complex gene
344 libraries and pathways. *PloS one* **4**: e6441.

345 **Sarrion-Perdigones A, Falconi EE, Zandalinas SI, Juárez P, Fernández-del-**
346 **Carmen A, Granell A, Orzaez D**. **2011**. GoldenBraid: An iterative cloning system
347 for standardized assembly of reusable genetic modules. *PloS one* **6**: e21622.

348 **Rutten V, Munabi A, Riche F, Lewy G, Wilson H, Pipan M; Bhate S; Nghiem T-**
349 **A, Kaufhold W; Haseloff J, Rubert A, González A, Quijano A, Llopis I, Gavaldá**
350 **J, Estellés L, Vásquez M, Orzáez D, Deal C; Gray J, Spiegel M, Monsey S,**
351 **Middlemiss A, Day J, Patron NJ. 2015.** *BBF RFC 106: A Standard Type IIS Syntax*
352 *for Plants.* http://hdl.handle.net/1721.1/96069

353 **Sarrion-Perdigones A, Vazquez-Vilar M, Palací J, Castelijns B, Forment J,**
354 **Ziarsolo P, Blanca J, Granell A, Orzaez D**. **2013**. GoldenBraid2.0: A
355 comprehensive DNA assembly framework for plant synthetic biology. *Plant*
356 *Physiology*.

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

359 **Terfrüchte M, Joehnk B, Fajardo-Somera R, Braus GH, Riquelme M, Schipper**
360 **K, Feldbrügge M**. **2014**. Establishing a versatile Golden Gate cloning system for
361 genetic engineering in fungi. *Fungal Genetics and Biology* **62**: 1–10.

362 **Vafaee Y, Staniek A, Mancheno-Solano M, Warzecha H**. **2014**. A modular cloning
363 toolbox for the generation of chloroplast transformation vectors. *PloS one* **9**: e110222.

364 **Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S**. **2011**. A modular
365 cloning system for standardized assembly of multigene constructs. *PloS one* **6**:
366 e16765.

367 **Werner S, Engler C, Weber E, Gruetzner R, Marillonnet S**. **2012**. Fast track
368 assembly of multigene constructs using Golden Gate cloning and the MoClo system.
369 *Bioengineered Bugs* **3**: 38–43.

370 **Whitman L, Gore M, Ness J, Theotdorou E, Gustafsson C, Minshull J**. **2013**.
371 Rapid, scarless cloning of gene fragments using the Electra Vector System™. *Genetic*
372 *Engineering and Biotechnology News* **76**: 42.

373

374

375 **Figure 1** **A** Type IIS restriction enzymes such as *Bsa*I are directional, cleaving outside
376 of their non-palindromic recognition sequences. **B** Providing compatible overhangs
377 are produced on digestion, standard parts cloned in plasmid backbones flanked by a
378 pair of convergent Type IIS restriction enzyme recognition sites can be assembled in a
379 single digestion-ligation reaction into an acceptor plasmid with divergent Type IIS
380 restriction enzyme recognition sites and a unique bacterial selection cassette.

381 **Figure 2** **A** Standard parts for plants are free from *Bsa*I recognition sequences. To be
382 compatible with Golden Gate Modular Cloning (MoClo) and GoldenBraid2.0 (GB2.0)
383 they must also be free from *Bpi*I and *Bsm*BI recognition sequences. **B** Standard parts
384 are housed in plasmid backbones flanked by convergent *Bsa*I recognition sequences.
385 The plasmid backbones are otherwise free from *Bsa*I recognition sites. The plasmid
386 backbone should not confer bacterial resistance to ampicillin, carbenicillin or
387 kanamycin. When released from their backbone by *Bsa*I, parts are flanked by four-
388 base-pair 5' overhangs, known as fusion sites.

389 **Figure 3** Twelve fusion sites have been defined. These sites allow a multitude of
390 standard parts to be generated. Standard parts comprise any portion of a gene cloned
391 into a plasmid flanked by a convergent pair of *Bsa*I recognition sequences. Parts can
392 comprise the region between an adjacent pair of adjacent fusion sites. Alternatively, to
393 reduce complexity or when a particular functional element is not required, parts can
394 span multiple fusion sites (examples in pink boxes).

395 **Figure 4** **A** Universal acceptor plasmids (UAPs) comprise a small plasmid backbone
396 conferring resistance to spectinomycin or chloramphenicol in bacteria. They contain a
397 cloning site consisting of a pair of divergent Type IIS recognition sequences (e.g.
398 *Bpi*I*,* as depicted, or *Bsm*BI) flanked by overlapping convergent *Bsa*I recognition
399 sequences. **B** A sequence containing an illegal *Bsa*I recognition sequence can be
400 amplified in two fragments using oligonucleotide primers with 5' overhangs (red
401 dashed lines) that (i) introduce a mutation to destroy the illegal site (ii) add TypeIIS
402 recognition sequences (e.g. *Bpi*I*,* as depicted, or *Bsm*BI) and fusion sites to allow one
403 step digestion-ligation into the universal acceptor and (iii) add the desired fusion sites
404 (green numerals) that will define the type of standard part and that will flank the part
405 when re-released from the backbone with *Bsa*I.

A

BsaI

GGTCTCNNNNN
CCAGAGNNNNN

BsaI

NNNNNGAGACC
NNNNNCTCTGG

B

BsaI NNNN Part 1 NNNN BsaI
NNNN NNNN
Bacterial Selection A

BsaI NNNN Part 2 NNNN BsaI
NNNN NNNN
Bacterial Selection A

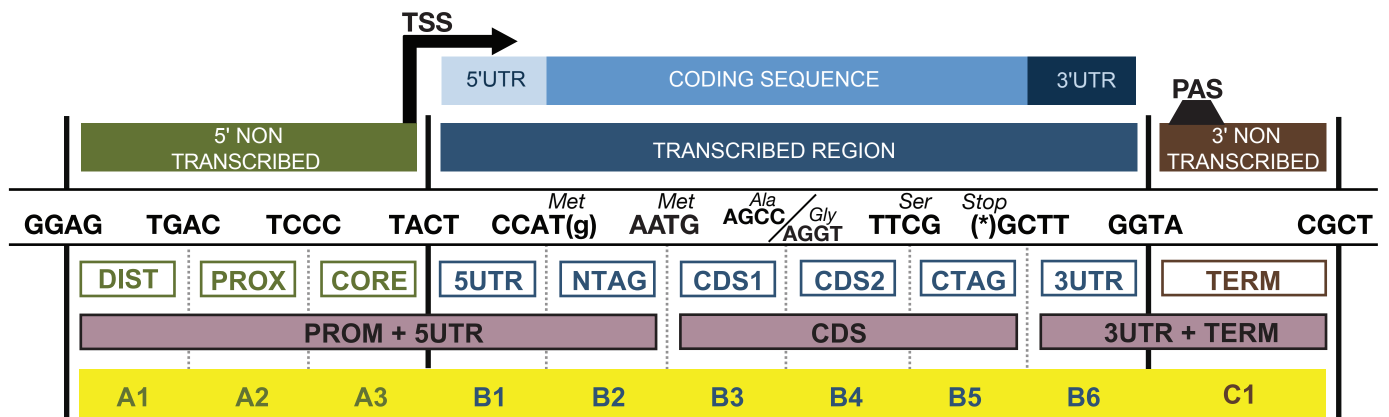
BsaI NNNN Part 3 NNNN BsaI
NNNN NNNN
Bacterial Selection A

BsaI BsaI
NNNN Acceptor NNNN
NNNN NNNN
Bacterial Selection B

One step digestion-ligation reaction with *BsaI* and T4 ligase.
Selection for colonies carrying plasmids with Bacterial selection B.

NNNN Part 1 NNNN Part 2 NNNN Part 3 NNNN
NNNN NNNN NNNN NNNN
Bacterial Selection B

TYPE	ENZYME	SEQUENCE	COMPATIBILITY
ILLEGAL	*BsaI*	GGTCTCNNNNN CCACGANNNNN	Plant Standard MoClo, GB2.0
AVOID	*BpiI*	GAAGACNNNNNN CTTCTGNNNNNN	MoClo
AVOID	*BsmBI*	CGTCTCNNNNN GCAGAGNNNNN	GB2.0 MoClo (Level 2i+)



POSITION	NAME	FUNCTION	5' OVERHANG	3' OVERHANG
A1	DIST	Distal promoter region, *cis* regulator or transcriptional enhancer	GGAG	TGAC
A2	PROX	Proximal promoter region, *cis* regulator or transcriptional enhancer	TGAC	TCCC
A3	CORE	Minimal promoter region, including transcription start site (TSS)	TCCC	TACT
A4	5UTR	5' untranslated region	TACT	CCAT
B2	NTAG	N terminal coding region	CCAT	AATG
B3	CDS1	Coding region - optional N terminal coding region	AATG	AGCC /AGGT
B4	CDS2	Coding region - no start or stop codon	AGCC /AGGT	TTCG
B5	CTAG	C terminal coding region	TTCG	GCTT
B6	3UTR	3' untranslated region	GCTT	GGTA
C1	TERM	Transcription terminator including polyadenylation signal (PAS)	GGTA	CGCT

A

BsaI BpiI BpiI BsaI

GGTCTCATTGTCTTC..........GAAGACAACGAGACC

CCAGAGTAACAGAAG..........CTTCTGTTGCTCTGG

BpiI BpiI cloning selection BsaI BsaI

Universal Acceptor Plasmid

SpecR/CamR

B

BsaI

nnnnnnnnnnnnnnGGTCTCnnnnnnnnnnnn

nnnnnnnnnnnnnnCCAGACnnnnnnnnnnnn

BpiI BpiI

nnGAAGACnnCTCA1234nnnnnnnnnnnnnnGGTCACnnGTCTTCnn

nnCTTCTGnnGAGT1234nnnnnnnnnnnnnnCCAGTGnnCAGAAGnn

BpiI BpiI

nnGAAGACnnGGTCACnnnnnnnnnnnn5678CGAGnnGTCTTCnn

nnCTTCTGnnCCAGTGnnnnnnnnnnnn5678GCTCnnCAGAAGnn

C

BsaI BsaI

GGTCTCA1234nnnnnnnnnnnnnnGGTCACnnnnnnnnnnnn5678CGAGACC

CCAGAGT1234nnnnnnnnnnnnnnCCAGTCnnnnnnnnnnnn5678GCTCTGG

BsaI part BsaI

Standard Part

SpecR/CamR