

Bricks and blueprints: methods and standards for DNA assembly

Arturo Casini¹, Marko Storch¹, Geoffrey S Baldwin^{1*}, Tom Ellis^{2*}

1. Centre for Synthetic Biology and Innovation, Department of Life Sciences, Imperial College London, SW7 2AZ, UK

2. Centre for Synthetic Biology and Innovation, Department of Bioengineering, Imperial College London, SW7 2AZ, UK

* Co-corresponding authors

Correspondence: Tom Ellis (E-mail: t.ellis@imperial.ac.uk; Phone: +44 (0)20 7594 7615) and Geoffrey S Baldwin (E-mail: g.baldwin@imperial.ac.uk; Phone: +44 (0)20 7594 5288)

Preface

DNA assembly is a key part of constructing gene expression systems and even whole chromosomes. In the past decade a plethora of powerful new DNA assembly methods including Gibson assembly, Golden Gate and LCR have been developed. In this Innovation article we discuss these methods and standards such as MoClo, GoldenBraid, MODAL and PaperClip, which have been developed to facilitate a streamlined assembly workflow, aid material exchange, and the creation of modular, reusable DNA parts.

Introduction

Our capacity to cut and paste DNA from different sources and to assemble it into gene constructs has been one of the key drivers of biological research and biotechnology over the past four decades. However, despite countless advances in molecular biology, the assembly of DNA parts into new constructs remains a craft that is both time consuming and unpredictable. The decreasing costs of gene synthesis promises to alleviate these limitations by providing custom-made double-stranded DNA fragments typically between 200 and 2000 bp in length¹. Nonetheless, gene synthesis does not eliminate the need for DNA assembly, which remains necessary for the production of constructs beyond one kilobase in size, both in research labs and at gene synthesis companies. DNA assembly also enables carrying out projects with more complex experimental needs and is especially valuable for building diverse plasmid libraries and creating multi-component systems, and has even been used to construct synthetic cells².

Addressing the limitations of DNA assembly methods has been one of the key goals of synthetic biology, a scientific discipline focused on the construction and testing of new or redesigned versions of genes, gene networks, pathways and cells^{3,4}. In order to tackle projects of increasing scale and complexity, researchers have invested significant effort into developing new tools for DNA assembly and into matching them with improved, lower-cost gene synthesis (for reviews on gene synthesis see REFS. 1,5^{1,5}), as well as a suite of important new tools for genome editing (**Box 1**). With these combined advances the field is now at a point where gene synthesis and DNA assembly can empower even undergraduate students to construct entire eukaryotic chromosomes⁶.

This acceleration in the scale of DNA assembly enables construction projects too complex to be drawn out on the back of an envelope, which instead require an engineering approach. In the past decade important

assembly methods such as Gibson assembly and Golden Gate have been developed^{7,8}, which define new protocols for joining together DNA parts. Alongside these methods, researchers have also developed various physical standards such as MODAL (Modular Overlap-Directed Assembly with Linkers)⁹ and MoClo (Modular Cloning system)¹² that define rules for the format of DNA parts that can be used with them. These physical standards facilitate the re-use of parts between experiments, exchange of parts between research groups and importantly provide modularity in construction. In this Innovation article we focus on the recent advances in DNA assembly methods and describe their matching physical standard. A summary of key DNA assembly methods and physical standards is presented in **Table 1**.

Endonuclease-mediated assembly

Gene cloning using restriction endonucleases and DNA ligases has been employed successfully for over 40 years¹⁰, but in the age of cheap DNA synthesis its limitations are becoming evident. It typically only joins two DNA parts at a time and the parts of interest should be compatible with restriction sites found in a target multiple cloning site (MCS) within the target vector. The restriction sites chosen directly affect cloning efficiency and can even lead to significant changes in plasmid function^{11,12}. To resolve such experimental inconsistencies, standards for DNA assembly using restriction endonucleases began emerging early this century. BioBricks was one of the first and was developed to address the *ad hoc* nature of cloning¹³. BioBrick DNA parts are flanked by standard prefix and suffix sequences that contain four defined restriction sites. These sequences enable a repeatable, idempotent assembly process: the ligation of two BioBrick parts produces a new, larger BioBrick part with the same physical format. As the BioBrick standard became adopted, various improvements (for example BglBricks) were developed that improved its flexibility and efficiency^{14–16}.

Although standards such as BioBricks allow parts to be rationally assembled into desired constructs, it is often quicker to modify existing constructs. New standard plasmid formats have thus been developed to facilitate swapping of parts between constructs. For rapid prototyping of *Escherichia coli* gene network constructs the Breadboard standard has been developed¹⁷. For working across a broad range of bacterial species, Standard European Vector Architecture (SEVA) plasmids are a standard that gives plasmids a modular structure with unique restriction sites that flank fundamental parts such as selection markers and origins of replication, which often need to be exchanged in order to efficiently operate in different organisms^{18,19}.

The drawback of all digestion and ligation methods is the need to remove forbidden digestion sites within DNA parts prior to cloning them. Homing endonucleases such as I-SceI have been proposed as a way to overcome this as they are equivalent to restriction endonucleases but only cut at long recognition sequences, which are unlikely to be found in cloning parts²⁰. The iBrick standard closely resembles BioBricks, but uses prefix and suffix sequences containing homing endonuclease recognition sites²¹. The HomeRun Vector Assembly System (HVAS) also exploits homing endonucleases but within a tiered framework that allows multi-part construction²². Although both iBrick and HVAS tackle the issue of forbidden restriction sites, they result in the formation of large ‘scar’ sequences when parts are recombined, as the long restriction enzyme-recognition sites remain in the final construct. Scar sequences can also pose a problem in BioBrick assembly, as the 8 bp sequence generated between two fused BioBricks can act as a destabilising sequence in *E. coli*²³.

Type IIS endonuclease assembly. Type IIS restriction endonucleases differ from traditional type II endonucleases because they cut DNA a few bases away from their recognition site, thereby allowing

freedom of choice of the overhang sequences they generate. The use of Type IIS endonucleases adds a significant advantage to DNA assembly because it allows scar-free, 'seamless' assembly, and just one enzyme can digest parts with different overhangs so that many of these can be ligated together in a defined order in one go. Golden Gate cloning is based on these principles, and has gained significant popularity especially in publically-available kits where many parts and/or highly repetitive sequences need to be assembled in one go²⁴⁻²⁶. One of the main challenges in assembly with Golden Gate is defining the position of the DNA parts within the final construct, which depends on the sequences of the short overhangs generated by digestion.

Golden Gate parts can be generated by PCR or by gene synthesis^{27,28}, but in order to minimize errors in assembly the basic assembly parts are usually first placed in 'entry vectors' and then digested and ligated into 'destination vectors' (**Figure 1**). To enable using Golden Gate for hierarchical assembly, a variety of physical standards have been developed that adopt a tiered assembly approach: in the first tier genes are assembled from their elementary parts (promoters, open reading frames, terminators, etc.; **Figure 1a**)²⁹⁻³² and then in the second tier, these genes are combined to form multi-gene systems (**Figure 1b**). The MoClo standard, first developed for plant molecular biology, uses a parallel approach for all tiers but requires the use of a large number of entry and destination vectors²⁹. The GoldenBraid standard reduces the number of required vectors by applying a pairwise approach for assembly beyond the gene-level, but at the cost of requiring more steps for larger constructs^{30,33}.

The power of Golden Gate to perform parallel assembly of multiple parts without PCR has made it very popular and recently a mammalian MoClo (mMoClo) standard was introduced³⁴. mMoClo enables the stable integration into defined sites in mammalian chromosomes of very large multigene constructs such as genetic switch programs that encode the regulated expression of counteracting transcription factors that establish heritable memory in cells³⁴. Golden Gate is best suited for applications where multiple genes must be expressed together, and consequently it is also used in kits for multiplex CRISPR-Cas9 genome editing³⁵. The large number of plasmids required for Golden Gate is a drawback, but in return researchers can format their DNA to physical standards that facilitate part re-use and sharing. However, as with classic restriction endonuclease-based cloning, Golden Gate still requires that endonuclease recognition sequences are absent within the DNA parts to enable their cloning. To overcome this limitation, programmed DNA methylation (which can prevent digestion by certain endonucleases) has been utilized in some standards such as GreenGate and MASTER in order to guide digestion only to desired sites^{32,36}. Impressively, methylation was used also in the related Pairwise Selection Assembly method to construct a complete 91 kb synthetic yeast chromosome arm without the need to consider forbidden sites³⁷.

Site-specific recombination

Site-specific recombination omits any need for restriction endonucleases and instead employs phage integrases, which are site-specific recombinases that recognise versions of *att* sequence motifs and catalyse DNA rearrangement between them. Integrases are employed by the popular commercial Gateway cloning method³⁸, one of the earliest DNA assembly standard, which uses the lambda integrase *in vitro* to catalyse directional cloning of DNA parts that are flanked by orthogonal versions of the lambda *attB* and *attP* sites (**Figure 2a**). This method is simple, efficient and reliable, and is widely used for the generation of clone libraries and for expression analysis in eukaryotic systems^{39,40}. By synthetically generating four additional orthogonal *att* recombination sequences, Gateway has also recently been expanded to enable the cloning of multiple parts simultaneously⁴¹. Similar non-commercial systems have also been developed that use alternative phage integrases, including phiBT1 and phiC31^{42,43}. For all these methods, reactions at

recombined *att* sites can be reversed by additionally providing either an excisionase (a bacteriophage excision protein) or a recombination directionality factor (RDF), which is an accessory protein that in combination with the integrase reverses the reaction and leads to excision rather than integration. For phiC31 integrase-based assembly, excision allows quick replacement of a single part within an already assembled construct, enabling the insertion of an alternative part or of a multi-part fragment that further expands the construct⁴³.

Due to the nature of integrase sites, recombinase-based assembly leaves repeated scar sequences between all assembled parts and these can be problematic for maintaining DNA integrity or for mRNA folding as they are typically long and palindromic. Creating the entry vectors for recombinase cloning is also not trivial, as the various *att* sites must each be matched with their specific partner sequence, whilst also maintaining the correct orientation on the DNA strand (**Figure 2b**).

Long overlap-based assembly

Long overlap-based assembly methods are *in vitro* and *in vivo* techniques that work by joining DNA fragments that share homologous sequences — usually between 20 bp and 40 bp in length — at their ends (**Figure 3a**). The long sequence homology ensures high efficiency and specificity of DNA assembly, meaning that most long overlap methods can easily assemble five or more DNA parts together in one step⁷. These methods are particularly popular for manipulating larger DNA fragments (>1 kb) as there is usually no requirement for removing restriction sites from within parts. Another reason for their popularity is that overlap regions can easily be added by PCR. The mechanism of action of these methods varies greatly: Circular Polymerase Extension Cloning (CPEC) is an evolution of Overlap Extension PCR (OE-PCR⁴⁴) and is essentially a high-fidelity PCR amplification, where template and primers are replaced by the DNA fragments to be assembled into a plasmid⁴⁵. As these are designed to share homology at their ends, the parts anneal to each other during PCR and act as primers for one another's amplification, until eventually a nicked circular molecule is generated. Other methods use enzymatic digestions to convert the homologous regions of different parts into single-strand overhangs that can directly anneal to each other. Uracil-Specific Excision Reagent (USER) assembly does so by producing nicks in the homology regions^{46,47}, whereas Gibson assembly instead employs a "chew-back" mechanism on one strand, improving on the previously described Sequence and Ligation-Independent Cloning (SLIC) method by the addition of *in vitro* DNA repair^{2,48}. In these methods, purified repair enzymes can also be replaced by cellular DNA repair systems; for example, the Seamless Ligation Cloning Extract (SLiCE) method can assemble DNA parts *in vitro* at very low cost by using *E. coli* cell extracts instead of an enzyme cocktail like that used by Gibson assembly⁴⁹. Other methods take a step further, relying on the repair machinery of live organisms like *Bacillus subtilis* and *Saccharomyces cerevisiae*, which are able to take up linear DNA parts with overlaps and assemble them spontaneously *in vivo* into the desired constructs via native homologous recombination⁵⁰⁻⁵².

Long overlap assembly methods have very few sequence restrictions but design requirements for the overlap sequences need to be considered if assembly is to be efficient, if modularity is required and especially if repetitive DNA parts are present. Such requirements define the many standards that have emerged for these methods. The most straightforward standards are developed for a specific task. For example, DNA Assembler performs the assembly of metabolic pathways in *S. cerevisiae* by initially building individual genes from elementary parts using OE-PCR. These genes are built with end-homology so that upon their transformation into yeast cells they recombine to assemble the desired metabolic pathway cluster⁵³. In contrast, the MODAL standard (**Figure 3b**) for modular construction is more flexible and works with methods as diverse as *S. cerevisiae in vivo* assembly, Gibson assembly and CPEC⁹. In MODAL, parts are

first standardised into modules by flanking each with universal prefix and suffix sequences. PCR with a set of orthogonal primers then amplifies the parts to add designed homology region 'linkers' that define the position and orientation of any part in the final construct, which can then be assembled by any of the above mentioned techniques (**Figure 3b**). The linkers are computationally designed⁵⁴ to ensure that they are orthogonal and that they comply with homology requirements that ensure efficient assembly. An alternative approach avoids PCR by instead adding the linkers through cloning the parts into a plasmid where the MCS is already flanked by the appropriate linker sequences. The linker-flanked parts can then be excised using restriction endonucleases that cut outside the linkers, and the linear DNA can then be assembled by the Gibson method⁵⁵.

Although linkers provide many benefits they also end up as foreign sequences between parts. This can be overcome by instead including 'bridges' into assembly reactions, which are oligonucleotides that overlap with the ends of two parts to be joined (**Figure 3c**). PaperClip is a standard that works with Gibson assembly which adopts a bridge-based approach⁵⁶, where bridges with a short 3 bp scar are created from linking-together short oligonucleotides that have homology to the end of each part. It is also the only standard currently compatible with Ligase Cycling Reaction (LCR) assembly, a unique overlap method that uses bridges for its mechanism (**Figure 3c**). In LCR, parts to be joined are mixed with bridge oligonucleotides in the presence of a thermostable DNA ligase and subjected to denaturing and annealing cycles. This brings consecutive parts together, fusing them at high temperature to ensure base-pair specificity. As fused parts act as the template for the annealing and ligation of more parts, the reaction is efficient and a successful one-step assembly of as many as 20 parts was reported⁵⁷.

Recent developments are now beginning to combine principles from different assembly strategies. BASIC (Biopart Assembly Standard for Idempotent Cloning) is an evolution of MODAL using orthogonal linkers to guide overlap assembly, but employing type II_s restriction digestion instead of PCR to attach these to DNA parts⁵⁸. This allows the efficiency of long overlap methods to be combined with the simplicity of type II_s cloning and maintains a simple single-tier structure without intermediate cloning steps. Methylation can also be used to protect linkers from digestion, enabling an idempotent format for hierarchical construction through further rounds of assembly.

Integration of multiple methods is also key to the Modular Mammalian Standard for construction of large multi-gene plasmids for mammalian cell experiments⁵⁹. This begins with an approach similar to the HVAS standard where Gateway is used to assemble elementary parts into genes on plasmids, which in this case have designed linker sequences flanking the integration sites. Assembled genes flanked by linkers are then excised by homing endonuclease digestion and further assembled together using Gibson assembly to place multiple genes into a destination plasmid that can be transfected into cells (**Figure 3d**). The assembled constructs maintain their linker sequences and homing endonuclease sites, which upon excision leave behind large scars but also enable further rounds of assembly without any need for PCR or consideration of forbidden restriction sites. This standard has already enabled the assembly of constructs bigger than 60 kb with more than 30 parts, allowing researchers to quickly build novel genetic programs of more than 10 interacting genes and test them in mammalian cells⁵⁹.

Conclusion and future perspectives

The variety of new assembly methods and standards described here are already being put to use around the world for impressively diverse tasks. Golden Gate assembly is at the heart of many of the kits for genome-editing (**Box 1**), so numerous labs will already be familiar with its ability to assemble many

different parts in a defined order in a single reaction. Likewise, Gibson assembly has become the method of choice in synthetic biology⁶⁰ and has even been used to assemble an entire mouse mitochondrial genome from hundreds of oligonucleotides⁶¹. Implementing standards for large DNA assembly projects is also beginning to bear fruit. By using a variation of the MoClo Golden Gate standard, researchers were recently able to automate the design and construction of 122 different versions of a cluster of 16 genes for nitrogen fixation, building from a starting library of 103 parts⁶². At the largest scale of DNA assembly, the landmark genome synthesis projects for *Mycoplasma genitalium*⁶³ and *Mycoplasma mycoides*⁶⁴ have shown that different scales of assembly require different methods: Gibson assembly can be used to join gene-size DNA fragments in a scale of up to a hundred kilobases, but beyond that *in vivo* recombination assembly in *S. cerevisiae* becomes necessary. The global project to construct a synthetic version of the yeast genome also recognises the need for different methods at different scales and utilises combinations of Gibson assembly, USER cloning, traditional digestion and ligation and *in vivo* recombination to hierarchically combine short DNA fragments into 50 kb synthetic ‘megachunks’ that replace their equivalent endogenous regions in the genome^{65,66}. Given that the work of this project is shared between teams around the world, it is not surprising that standardisation is required to ensure efficient progress.

Recognising that different methods work best at different scales is already leading to a new generation of assembly approaches where standard workflows integrate multiple methods together. By formatting parts and protocols according to a standard, future workflows will be set so that parts can be combined efficiently over sequential tiers and exchanged between projects around the world. As these standards are implemented the next logical steps for DNA assembly should recapitulate mature engineering disciplines: enabling software will be developed, automation of the labour will be introduced and ultimately the work of cloning will be commercially outsourced just as it is already done for gene synthesis and DNA sequencing.

Software tools will increasingly play an important role in DNA assembly, and are already required for the design of parts for modular DNA assembly^{54,55,59} and used in the experimental planning and quality control of large and complex projects^{67–69}. Many of these software tools can also control liquid handling robots and their ability to automate hundreds of complex modular assemblies has been recently demonstrated⁶². Not surprisingly, companies have emerged from these efforts that now sell advanced DNA assembly software to clients. These have quickly been followed by a second set of companies that offer to perform outsourced DNA assembly projects and sell lab automation equipment specifically designed to run new DNA assembly methods. As the cloning of everything from plasmids to whole chromosomes by DNA assembly methods moves towards becoming a commercial service, it is likely that further advances will specifically focus on decreasing costs and time while increasing fidelity. Dramatic improvements on all of these fronts may be achieved by incorporating microfluidics into workflows or by replacing enzymatic steps with physical and chemical routes for DNA assembly. However, it is more likely that the keys to success for future DNA assembly foundries will be efficient exploitation of cheap gene and oligo synthesis combined with smart use of low-cost next generation sequencing.

Acknowledgements

We wish to thank Max Jamilly for contributing to initial discussions. We also apologise to all whose work was not discussed due to space limitations. Research on DNA assembly in the groups of T.E. and G.S.B. is supported by the UK Engineering and Physical Research Council (EPSRC) grant EP/J02175X/1 and EU FP7 gra

Competing Interests Statement

The authors declare no competing interests.

Box 1: Genome Editing

Genome editing technologies rewrite DNA sequences in a site-specific manner within cells and can be used alongside DNA assembly methods in the construction and engineering of genomes. Multiplex Automatable Genome Engineering (MAGE) entails the delivery of synthetic DNA oligonucleotides into growing cells to mutate specific genomic sequences during DNA replication⁷⁰. It works efficiently in *E. coli* and was used impressively to recode all 321 TAG stop codons in the *E. coli* genome to TAA stop codons in order to provide a genomically-recoded organism capable of utilising non-standard amino acids⁷¹. Other genome editing methods rely on homologous recombination (*recombineering*), and, in *E. coli*, this can be used to insert or delete DNA at almost any desired locus via the lambda Red method which utilises a phage recombinase to recombine long regions of homology⁷². This has already enabled researchers to construct a size-reduced *E. coli* genome by performing large-scale deletions of unwanted elements⁷³. In other organisms, the emergence of programmable nucleases such as Zinc Finger Nucleases (ZFN), Tal-Effector Nucleases (TALEN) and, most recently, the CRISPR–Cas9 genome editing tool, means that site-specific cutting of chromosomes is also now possible, paving the way for recombineering in these cells via Homology-Directed Repair (HDR)⁷⁴. Genome editing technologies differ from DNA assembly methods as they edit existing sequence rather than combine DNA parts together, and thus combining the two approaches can be particularly powerful, as was recently demonstrated in the assembly and placement of pathway of 14 genes into multiple loci of the yeast genome in one go by combining standardised DNA assembly methods with site-specific integration using CRISPR–Cas9⁷⁵.

FIGURE LEGENDS

Figure 1: Examples of type IIS endonuclease-based methods and standards. (a) Diagram of the Golden Gate method, an endonuclease digestion and ligation-based technique that exploits the ability of the type IISs endonuclease BsaI to cut DNA a few bases away from its recognition site thereby leaving various overhangs that enable one-pot multipart assembly with high efficiency²⁴. Shown is an example of a first tier Golden Gate assembly workflow, in which genes (dark green, bottom) are constructed from a series of genetic elements such as Promoters, ORFs and Terminators (DNA parts, light green). Multiple parts are cloned into vectors that carry inward-facing BsaI restriction sites (Entry vectors, blue) and are then assembled in a simultaneous digestion and ligation reaction into a destination plasmid (Dest vector, orange) carrying outward-facing BsaI restriction sites and a drop-out screening cassette (here LacZ). Recognition sites for BsaI generate different 4 bp overhangs (differently coloured triangles and bars) that define the position of the various parts in the final construct. Simultaneous digestion and ligation forces incorrectly assembled plasmids to be recut by BsaI while correct ones are protected. (b) Diagram of the MoClo and GoldenBraid 2.0 standards^{29,30}. These provide two different approaches for the second tier of Golden Gate assembly, where the genes built in the first tier reaction are combined to form multi-gene constructs. MoClo (left) employs a parallel approach, in which in tier one each gene is assembled into destination plasmids flanked by digestion sites with different cut sequences (BsmBI, colors) that in tier two are digested to leave overhangs that define the final construct layout. GoldenBraid employs a pairwise approach, in which in tier one genes are cloned into pairs of destination plasmids (each with the LacZ screening cassette), and these are then assembled two-at-a-time in successive tiers by performing assembly alternately with BsmBI and BsaI type IIS endonucleases.

Figure 2: Site-specific recombination methods and standards. (a) General diagram for site-specific recombination-based cloning⁷⁶. An entry plasmid (blue) containing a gene (green) flanked by two orthogonal *attB* sites (*attB1* and *attB2*, red) is mixed with a destination plasmid (orange) containing a bacterial suicide gene (for example the toxin *ccdB*; black), flanked by two related *attP* sites (*attP1* and *attP2*, purple), that forces selection of recombination between the two plasmids. In the presence of the appropriate integrase enzymes, the two sites recombine to swap genes and form hybrid sites *attR* and *attL* (*attL1* and *attL2*, orange). Undesired constructs (shown as faded) are selected against by antibiotic selection and suicide gene expression. (b) Diagram of the HomeRun Vector Assembly System (HVAS) standard²². The first tier of assembly employs a Gateway reaction to build genes from elementary parts stored in Gateway entry plasmids (blue), which carry a specific pair of orthogonal recombination sites assigned to each part type (Promoter, ORF, Terminator; black). Genes are assembled in one of four special Gateway destination plasmids carrying matching recombination sites (orange), where they are flanked by two specific homing endonuclease sites (from a selection of four homing endonuclease (HE) possibilities, HE1-HE4). In the second tier of assembly (HomeRun Cloning), genes are excised from these sites by homing endonuclease digestion and sequentially ligated into a final destination plasmid (yellow) containing all four endonuclease sites.

Figure 3: Examples of long overlap-based assembly methods and standards. (a) Schematic of long overlap assembly. Parts for assembly have flanking overlap (OL) regions typically 40 nucleotides in length that define the final arrangement of the construct. Parts such as genes (green and brown) and the linearized plasmid (orange) are joined either by *in vitro* enzymatic reactions or *in vivo* recombination. The schematic depicts *in vitro* methods such as in Gibson assembly² or SLIC⁴⁸, where exonuclease digestion creates single-stranded, overhang overlaps that anneal to each other. (b) In the MODAL (Modular Overlap-Directed Assembly with Linkers) standard⁹ both genes (green) and backbone parts like the ampicillin resistance gene (AmpR, orange) and an origin of replication (Ori, yellow) are held in in entry plasmids (blue) and flanked by prefix and suffix sequences (P and S). These are PCR amplified with standard primers fused to linkers, so that the linker-flanked parts can be assembled to become a plasmid by a variety of long overlap assembly methods, e.g. CPEC (Circular Polymerase Extension Cloning). (c) Ligase Chain Reaction (LCR) assembly using a bridge-based approach⁵⁷. This method employs single-stranded bridge oligonucleotides complementary to the ends of two DNA parts (grey and blue) being assembled. Temperature cycling denatures parts and anneals them to the provided bridge oligonucleotides. A thermostable ligase then joins the DNA (*) where the bridge is. (d) **1.** The modular mammalian standard⁵⁹ uses Gateway recombination to build genes from elementary parts (for example Promoter and ORF, black) held in entry vectors (blue), with assembly-placing them into destination vectors (orange) that carry overlap sequences (OL) that guide the 2nd tier of assembly. **2.** Genes (green) with overlap sequences are then excised from these plasmids using homing endonuclease digestion (HE, brown), and using Gibson assembly are combined in a one-pot reaction into a final plasmid by joining the genes as well as two linearised plasmid fragments (yellow) that provide gentamycin selection (KanR, purple) and bacterial antibiotic selection and plasmid replication (Ori-AmpR, purple).

Table 1: Summary of physical standards in DNA assembly

Physical Standards	Underlying Methodology					Limitations		Workflow		
	Restriction & ligation	Homing endonuclease	Type IIS enzyme	Site-specific recombinase	Long overlap	PCR required ^a	Forbidden restriction sites ^b	Assembly tiers	Multipart assembly ^c	Hierarchical assembly
BioBrick ¹³ and BglBrick ¹⁵	X					No	4	1	No	Yes
iBrick ²¹		X				No	-	1	No	Yes
HVAS (HomeRun Vector Assembly System) ²²		X		X		No	-	2	Yes/No	No ^d
MoClo (Modular Cloning system) ²⁹			X			No	3	2	Yes/Yes	Yes
GoldenBraid 2.0 ³⁰			X			No	3	2+	Yes/No	Yes
GreenGate ³²			X			No	1	2	Yes/No	Yes
Binder <i>et al.</i> ³¹			X			No	3	2	Yes/Yes	Yes
PSA (Pairwise-specific assembly) ³⁷			X			No	-	1	No	Yes
DNA Assembler ⁵³					X	Yes	-	2	Yes/Yes	No
MODAL (Modular Overlap-Directed Assembly with Linkers) ⁹					X	Yes	-	1	Yes	No
BASIC (Biopart Assembly Standard for Idempotent Cloning) ⁵⁸			X		X	No	1	1	Yes	Yes
Torella <i>et al.</i> ^{55 e}	X				X	No	4 ^f	2	No/Yes	No
Guye <i>et al.</i> ⁵⁹		X		X	X	No	-	2	Yes/Yes	Yes
PaperClip ⁵⁶					X	No	-	1	Yes	No

^a Not considered here are single PCR amplifications that format DNA parts according to standard requirements (such as adding a prefix and a suffix) .

^b The number of different restriction sites that need to be absent from all DNA parts.

^c Denotes whether the standard allows parallel assembly of more than two DNA parts at a time. If a standard employs two tiers, the first tier (left) and the second tier (right) are evaluated separately. GoldenBraid can do more than 2 tiers of assembly, but beyond tier 1, all assemblies are two parts at a time.

^d HVAS does not allow hierarchical rounds of assembly but allows some degree of post-assembly modification of the constructs.

^e The standard developed by Torella *et al.* in its latest version provides two options to attach linkers to parts. The table shows the option presented in the original paper. The alternative route employs a PCR amplification to attach the linkers, thereby removing any first tier assembly⁷⁷.

^f The BioBrick or BglBrick standard is used for first tier assembly (with four forbidden sites) and a different digest is also required to begin second tier assembly.

AUTHOR BIOGRAPHIES

Arturo Casini

Dr Casini is a research associate at the Centre for Synthetic Biology and Innovation at Imperial College London. He obtained an undergraduate degree in Biology and an MRes degree in Systems and Synthetic Biology followed by a PhD developing new methods, physical standards and software tools for DNA assembly.

Marko Storch

Marko Storch earned his PhD in biophysics studying cytoskeletal proteins with Joe Howard at the Max Planck Institute CBG in Dresden. He followed his interest in synthetic biology joining Geoff Baldwin's group at Imperial College London as Marie Curie Fellow and is currently working on DNA assembly and directed evolution in vivo.

Geoff Baldwin

Dr Baldwin is a Reader in Biochemistry. He has spent his career at the interface of the physical and life sciences, with an undergraduate in Chemistry and a PhD in biochemistry. He has now moved into engineering via synthetic biology, with an active research programme in DNA assembly, part characterisation and in vivo approaches to directed evolution.

Tom Ellis

Tom Ellis is a senior lecturer in synthetic biology with both PhD and commercial experience in designing and testing DNA-binding. His postdoctoral expertise was in constructing novel gene networks using synthetic biology tools and his research group now focuses on genome synthesis and predictable engineering of cells with new functions.

References

1. Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **11**, 499–507 (2014).
2. Daniel G Gibson, Lei Young, Ray-Yuan Chuang, J Craig Venter, Clyde A Hutchison, H. O. S. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
3. Keasling, J. D. Synthetic biology for synthetic chemistry. *ACS Chem. Biol.* **3**, 64–76 (2008).
4. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12**, 381–90 (2014).
5. Czar, M. J., Anderson, J. C., Bader, J. S. & Peccoud, J. Gene synthesis demystified. *Trends in Biotechnology* **27**, 63–72 (2009).
6. Annaluru, N. *et al.* Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55–8 (2014).

7. Ellis, T., Adie, T. & Baldwin, G. S. DNA assembly for synthetic biology: from parts to pathways and beyond. *Integr. Biol. (Camb)*. **3**, 109–18 (2011).
8. Chao, R., Yuan, Y. & Zhao, H. Recent advances in DNA assembly technologies. *FEMS Yeast Res.* (2014). doi:10.1111/1567-1364.12171
9. Casini, A. *et al.* One-pot DNA construction for synthetic biology: the Modular Overlap-Directed Assembly with Linkers (MODAL) strategy. *Nucleic Acids Res.* **42**, e7 (2014).
10. Cohen, S. N., Chang, A. C., Boyer, H. W. & Helling, R. B. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 3240–4 (1973).
11. Crook, N. C., Freeman, E. S. & Alper, H. S. Re-engineering multicloning sites for function and convenience. *Nucleic Acids Res.* **39**, e92 (2011).
12. Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology--identifying causes of failure of synthetic biological systems. *Biotechnol. J.* **7**, 856–66 (2012).
13. Shetty, R. P., Endy, D. & Knight, T. F. Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.* **2**, 5 (2008).
14. Norville, J. E. *et al.* Introduction of customized inserts for s-treamlined assembly and optimization of BioBrick synthetic genetic circuits. *J. Biol. Eng.* **4**, 17 (2010).
15. Anderson, J. C. *et al.* BglBricks: A flexible standard for biological part assembly. *J. Biol. Eng.* **4**, 1 (2010).
16. Leguia, M., Brophy, J. A., Densmore, D., Asante, A. & Anderson, J. C. 2ab assembly: a methodology for automatable, high-throughput assembly of standard biological parts. *J. Biol. Eng.* **7**, 2 (2013).
17. Litcofsky, K. D., Afeyan, R. B., Krom, R. J., Khalil, A. S. & Collins, J. J. Iterative plug-and-play methodology for constructing and modifying synthetic gene networks. *Nat. Methods* **9**, 1077–80 (2012).
18. Silva-Rocha, R. *et al.* The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* 1–10 (2012). doi:10.1093/nar/gks1119
19. Martínez-García, E., Aparicio, T., Goñi-Moreno, A., Fraile, S. & de Lorenzo, V. SEVA 2.0: an update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Res.* (2014). doi:10.1093/nar/gku1114
20. Marcaida, M. J., Muñoz, I. G., Blanco, F. J., Prieto, J. & Montoya, G. Homing endonucleases: from basics to therapeutic applications. *Cell. Mol. Life Sci.* **67**, 727–48 (2010).
21. Liu, J.-K., Chen, W.-H., Ren, S.-X., Zhao, G.-P. & Wang, J. iBrick: A New Standard for Iterative Assembly of Biological Parts with Homing Endonucleases. *PLoS One* **9**, e110852 (2014).
22. Li, M. V *et al.* HomeRun Vector Assembly System: a flexible and standardized cloning system for assembly of multi-modular DNA constructs. *PLoS One* **9**, e100948 (2014).

23. Sleight, S. C. & Sauro, H. M. The visualization of evolutionary stability dynamics and competitive fitness of *Escherichia coli* engineered with randomized multi-gene circuits. *ACS Synth. Biol.* 130816142503008 (2013). doi:10.1021/sb400055h
24. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647 (2008).
25. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82 (2011).
26. Kamens, J. The Addgene repository: an international nonprofit plasmid and data resource. *Nucleic Acids Res.* (2014). doi:10.1093/nar/gku893
27. Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–60 (2013).
28. Linshiz, G. *et al.* PR-PR: Cross-platform laboratory automation system. *ACS Synth. Biol.* **3**, 515–524 (2014).
29. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. A modular cloning system for standardized assembly of multigene constructs. *PLoS One* **6**, e16765 (2011).
30. Sarrion-Perdigones, A. *et al.* GoldenBraid2.0: A comprehensive DNA assembly framework for Plant Synthetic Biology. *Plant Physiol.* (2013). doi:10.1104/pp.113.217661
31. Binder, A. *et al.* A modular plasmid assembly kit for multigene expression, gene silencing and silencing rescue in plants. *PLoS One* **9**, e88218 (2014).
32. Lampropoulos, A. *et al.* GreenGate---a novel, versatile, and efficient cloning system for plant transgenesis. *PLoS One* **8**, e83043 (2013).
33. Sarrion-Perdigones, A. *et al.* GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One* **6**, e21622 (2011).
34. Duportet, X. *et al.* A platform for rapid prototyping of synthetic gene networks in mammalian cells. *Nucleic Acids Res.* (2014). doi:10.1093/nar/gku1082
35. Sakuma, T., Nishikawa, A., Kume, S., Chayama, K. & Yamamoto, T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep.* **4**, 5400 (2014).
36. Chen, W.-H., Qin, Z.-J., Wang, J. & Zhao, G.-P. The MASTER (methylation-assisted tailorable ends rational) ligation method for seamless DNA assembly. *Nucleic Acids Res.* **41**, e93 (2013).
37. Blake, W. J. *et al.* Pairwise selection assembly for sequence-independent construction of long-length DNA. *Nucleic Acids Res.* **38**, 2594–602 (2010).
38. Hartley, J. L., Temple, G. F. & Brasch, M. A. DNA cloning using in vitro site-specific recombination. *Genome Res.* **10**, 1788–95 (2000).
39. Alberti, S., Gitler, A. D. & Lindquist, S. A suite of Gateway cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. *Yeast* **24**, 913–9 (2007).

40. Marsischky, G. & LaBaer, J. Many paths to many clones: a comparative look at high-throughput cloning methods. *Genome Res.* **14**, 2020–8 (2004).
41. Sasaki, Y. *et al.* Evidence for high specificity and efficiency of multiple recombination signals in mixed DNA cloning by the Multisite Gateway system. *J. Biotechnol.* **107**, 233–43 (2004).
42. Zhang, L., Zhao, G. & Ding, X. Tandem assembly of the epothilone biosynthetic gene cluster by in vitro site-specific recombination. *Sci. Rep.* **1**, 141 (2011).
43. Colloms, S. D. *et al.* Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination. *Nucleic Acids Res.* 1–10 (2013). doi:10.1093/nar/gkt1101
44. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61–8 (1989).
45. Quan, J. & Tian, J. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One* **4**, e6441 (2009).
46. Bitinaite, J. *et al.* USER friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Res.* **35**, 1992–2002 (2007).
47. Annaluru, N. *et al.* Assembling DNA fragments by USER fusion. *Methods Mol. Biol.* **852**, 77–95 (2012).
48. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods* **4**, 251–6 (2007).
49. Zhang, Y., Werling, U. & Edlmann, W. SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res.* **40**, e55 (2012).
50. Itaya, M., Fujita, K., Kuroki, A. & Tsuge, K. Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat. Methods* **5**, 41–3 (2008).
51. Raymond, C. K., Pownder, T. A. & Sexson, S. L. General method for plasmid construction using homologous recombination. *Biotechniques* **26**, 134–8, 140–1 (1999).
52. Gibson, D. G. *et al.* One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 20404–9 (2008).
53. Shao, Z., Zhao, H. & Zhao, H. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* **37**, e16 (2009).
54. Casini, A. *et al.* R2oDNA designer: computational design of biologically neutral synthetic DNA sequences. *ACS Synth. Biol.* **3**, 525–8 (2014).
55. Torella, J. P. *et al.* Rapid construction of insulated genetic circuits via synthetic sequence-guided isothermal assembly. *Nucleic Acids Res.* 1–9 (2013). doi:10.1093/nar/gkt860
56. Trubitsyna, M., Michlewski, G., Cai, Y., Elfick, A. & French, C. E. PaperClip: rapid multi-part DNA assembly from existing libraries. *Nucleic Acids Res.* 1–8 (2014). doi:10.1093/nar/gku829
57. De Kok, S. *et al.* Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synth. Biol.* **3**, 97–106 (2014).

58. Storch, M. *et al.* BAS-IC: a new Biopart Assembly Standard provides accurate, single-tier DNA assembly for synthetic biology. *Unpublished* (2015). doi:10.1021/sb500356d
59. Guye, P., Li, Y., Wroblewska, L., Duportet, X. & Weiss, R. Rapid, modular and reliable construction of complex mammalian gene circuits. *Nucleic Acids Res.* **41**, e156 (2013).
60. Kahl, L. J. & Endy, D. A survey of enabling technologies in synthetic biology. *J. Biol. Eng.* **7**, 13 (2013).
61. Gibson, D. G., Smith, H. O., Hutchison, C. A., Venter, J. C. & Merryman, C. Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods* **7**, 901–3 (2010).
62. Smanski, M. J. *et al.* Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* (2014). doi:10.1038/nbt.3063
63. Gibson, D. G. *et al.* Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319**, 1215–20 (2008).
64. Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–6 (2010).
65. Jovicevic, D., Blount, B. A. & Ellis, T. Total synthesis of a eukaryotic chromosome: Redesigning and SCRaMBLE-ing yeast. *Bioessays* **36**, 855–60 (2014).
66. Dymond, J. S. *et al.* Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* **477**, 471–6 (2011).
67. Hillson, N. J., Rosengarten, R. D. & Keasling, J. D. j5 DNA Assembly Design Automation Software. *ACS Synth. Biol.* **1**, 14–21 (2012).
68. Appleton, E., Tao, J., Haddock, T. & Densmore, D. Interactive assembly algorithms for molecular cloning. *Nat. Methods* **11**, 657–62 (2014).
69. Densmore, D. *et al.* Algorithms for automated DNA assembly. *Nucleic Acids Res.* **38**, 2607–16 (2010).
70. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).
71. Lajoie, M. J. *et al.* Genomically recoded organisms expand biological functions. *Science* **342**, 357–60 (2013).
72. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640–6645 (2000).
73. Pósfai, G. *et al.* Emergent properties of reduced-genome *Escherichia coli*. *Science* **312**, 1044–1046 (2006).
74. Kim, H. & Kim, J.-S. A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.* **15**, 321–34 (2014).
75. Horwitz, A. A. *et al.* Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas. *Cell Syst.* (2015). doi:10.1016/j.cels.2015.02.001

76. Groth, A. C. & Calos, M. P. Phage Integrases: Biology and Applications. *J. Mol. Biol.* **335**, 667–678 (2004).
77. Torella, J. P. *et al.* Unique nucleotide sequence-guided assembly of repetitive DNA parts for synthetic biology applications. *Nat. Protoc.* **9**, 2075–2089 (2014).