BASIC: a new Biopart Assembly Standard for Idempotent Cloning provides accurate, single-tier DNA assembly for synthetic biology

Marko Storch[1†], Arturo Casini[1†], Ben Mackrow[1], Toni Fleming[2], Harry Trewhitt[1], Tom Ellis[3], Geoff S. Baldwin*[1]

1. Department of Life Sciences & Centre for Synthetic Biology & Innovation, Imperial College, South Kensington Campus, London, SW7 2AZ, UK
2. Dr Reddy's Chirotech Centre, Milton Road, Cambridge, CB4 0PE, UK
3. Department of Bioengineering & Centre for Synthetic Biology & Innovation, Imperial College, South Kensington Campus, London, SW7 2AZ, UK

ABSTRACT

The ability to quickly and reliably assemble DNA constructs is one of the key enabling technologies for synthetic biology. Here we define a new Biopart Assembly Standard for Idempotent Cloning (BASIC), which exploits the principle of orthogonal linker based DNA assembly to define a new physical standard for DNA parts. Further, we demonstrate a new robust method for assembly, based on type IIs restriction cleavage and ligation of oligonucleotides with single stranded overhangs that determine the assembly order. It allows for efficient, parallel assembly with great accuracy: 4 part assemblies achieved 93% accuracy with single antibiotic selection and 99.7% accuracy with double antibiotic selection, while 7 part assemblies achieved 90% accuracy with double antibiotic selection. The linkers themselves may also be used as composable parts for RBS tuning or the creation of fusion proteins. The standard has one forbidden restriction site and provides for an idempotent, single tier organisation, allowing all parts and composite constructs to be maintained in the same format. This makes the BASIC standard conceptually simple at both the design and experimental levels.

INTRODUCTION

The ability to build newly-designed DNA constructs easily, quickly and with high accuracy is one of the key enabling technologies of Synthetic Biology(1, 2) and the adoption of a standard format for the assembly of genetic components is part of this vision(1, 3). The BioBrick standard(4) is a restriction-ligation-based format and its usefulness stems from the principle of idempotency, where assembled parts retain the prefix and suffix of the original, enabling successive rounds of hierarchical cloning. However, a recent survey(5) has highlighted that most synthetic biology researchers now use Gibson's isothermal method for their DNA assembly(6). This suggests that the advantages of being able to assemble five or more fragments of DNA in parallel and having no forbidden sequences or scars outweighs the usefulness of a widely-adopted standard in the eyes of many researchers.

Although the Gibson method can be adapted to a physical standard framework using synthetic sequences to guide assembly(3, 7, 8), it is mostly used ‘ad hoc’, with customised parts that are generally prepared via PCR amplifications. This has led to a return to bespoke assembly, where each reaction requires design, optimisation and verification. Furthermore, reliance on PCR can compromise fidelity through errors in amplification and is inefficient for very long sequences or those containing high GC content and repeat sequences. PCR is also difficult to implement in an automated workflow because reactions for individual parts have to be optimised and verified. A recent approach excludes PCR(7), but requires upstream cloning to define downstream assembly order, thus extending the workflow.

Aside from BioBricks, alternative restriction-based standards have been developed, including GoldenBraid(9, 10) and MoClo(11), which are based on the Golden-Gate (12) protocol that employs Type IIs restriction enzymes. A common feature of these approaches is that the entry vector of a part defines its position in the final destination vector, so that changing the order of the parts requires an additional round of cloning. Both MoClo and GoldenBraid adopt a tiered approach, which takes advantage of the consistent layout of transcription units: in the first tier of assembly, where elementary parts such as promoters, ORFs and terminators are assembled into transcription units, a fixed and predefined part order is adopted. This way this first round of assembly never requires changing entry vectors. The same strategy cannot be used in the second tier of assembly, where transcription units are assembled into multigene constructs, since it is usually necessary to retain complete freedom of design. Here MoClo adopts a parallel approach, which requires cloning in a different vector for each possible position, while GoldenBraid adopts a sequential approach that minimises the number of vectors necessary but only allows pairwise assembly. It was previously suggested that the relative advantages of MoClo and GoldenBraid were mutually exclusive*(9)*.

To address the limitations of current assembly technologies, we have developed BASIC (Biopart Assembly Standard for Idempotent Cloning), to bring together six key concepts: standard reusable parts; single-tier format (all parts are in the same format and are assembled using the same process); idempotent cloning; parallel (multipart) DNA assembly; size independence; automatability. Our previous assembly strategy was based on Modular Overlap Directed Assembly with Linkers (MODAL)*(3)*, which introduced the concept of computationally derived orthogonal linkers*(13)*. To address the key concepts we have developed a new method based on robust restriction/ligation reactions to ligate orthogonal oligonucleotide linkers with single stranded overhangs that define the assembly order. To further address many of the requirements of assembling DNA parts and biological pathways*(1)* we have enabled hierarchical cloning within a single-tier format and demonstrated that the linkers themselves can be used as composable parts encoding RBS sequences or peptide linkers for fusion proteins. This has been achieved within a standard format that facilitates re-use of both linkers and parts.

RESULTS & DISCUSSION

Specification and Design. The core of the BASIC physical DNA standard is constituted by the integrated prefix and suffix sequences (*i*P/*i*S), which were designed to be back compatible with our previous MODAL strategy*(3)* where they can act as PCR priming sites; alternatively we here define the BASIC assembly method, based on simple robust reactions. The *i*P and *i*S sequences were also designed to ensure compatibility with the creation of fusion proteins, either by BASIC or MODAL, by optimising the amino acid coding of both the short BASIC scars and the full *i*P/*i*S sequences (Figure 1a).

To avoid PCR, yet retain the advantages of overlap directed DNA assembly, we have revisited pre-PCR methods, where oligonucleotides were routinely ligated onto DNA ends to either provide restriction enzyme sites, or compatible sticky ends to direct molecular cloning*(14)*. The BASIC standard defines two inward-facing BsaI recognition sites to release the parts from a storage vector, leaving a 4 bp scar on the prefix end and a 6 bp scar on the suffix (Figure 1b). Digestion yields different 4 bp overhangs at the prefix and suffix, enabling end-specific ligation. Ligation of partially double-stranded oligonucleotide DNA linkers is performed simultaneously with BsaI digestion (Figure 1C). Non-ligated oligonucleotide linkers are then removed by a purification step to yield linker-adapted parts. Final assembly is achieved by annealing the linker-adapted parts in an ionic buffer at elevated temperature. No ligase is required in the final step and the nicked plasmid generated is readily repaired *in vivo* following transformation. Full details of the protocol and optimisation of the method are provided in Online Supplementary Information.

Figure 1

Linker sequences to guide assembly were an expanded set of 7 linkers based on our previously-used 40% GC content linkers designed by R2oDNA Designer*(3, 13)* (Supplementary Tables 3&4). We split each 45 bp linker sequence across 2 parts, with each containing a 12 bp double stranded region on the outer side and sharing the central 21 base single stranded overlap region as top and bottom strands (Figure 1b). Double-stranded regions at the sites of linker ligation are necessary for efficient activity of T4 DNA ligase(15). The use of 21 bp overlaps enables the use of elevated temperatures during final assembly by complementary annealing, facilitating the kinetics, thermodynamics and specificity of the homology-search process.

Evaluation of efficiency. To evaluate the efficiency and accuracy of our assembly method, and benchmark against our previous Gibson-based work, we generated a number of parts in BASIC format with *i*P and *i*S sequences flanking the part of interest (Supplementary Table 2). This formatting step only ever has to be performed once for any part since the storage plasmid carries no positional information for the DNA assembly process, which is directed through the subsequent choice of linkers. Parts prepared include those that are essential for cell survival (origin of replication (MB1), kanamycin resistance (Kan) chloramphenicol resistance (Cm), combined origin and kanamycin (Kan-MB1), and others that produce fluorescent proteins (GFP and RFP; both as expression cassettes and as separate open reading frame parts).

Benchmarking DNA assembly reactions were performed by creating plasmids in a modular format from this parts library. Constructs comprising 2 to 6 component parts with a single antibiotic marker, and 2 to 7 parts with double antibiotic selection were chosen for evaluation (Figure 2a). All final constructs (apart from D2) contain a fluorescent reporter, and accuracy of assembly was thus evaluated by observing the correct expression of reporters whilst assessing each construct's ability to replicate and confer the appropriate antibiotic resistance.

Figure 2

BASIC assembly reactions were performed four times for each of the 11 designated test constructs following an optimised protocol (Online Supplementary Information). Assembly efficiency was determined from the number of colonies and accuracy as the % of colonies with the correct antibiotic resistance expressing the correct fluorescence reporters (Figure 2).

The results of the DNA assembly benchmarking reveal that the efficiency of assembly decreases exponentially with the number of parts involved (Figure 2b). However, even with 6 or 7 parts, reactions routinely returned between 40 and 150 colonies, while 3-4 part assembly routinely returned more than 1000 colonies, demonstrating the overall efficiency of the process.

The more critical measure of DNA assembly is accuracy. With single antibiotic selection there is the possibility that the storage plasmid that carries either the Kan-MB1 composite part or the Kan cassette can return a viable non-fluorescent colony if it is not completely digested in the first step of the protocol. The assembly efficiency decreased exponentially with increasing number of parts, but the number of incorrect assemblies (which includes both white background colonies and colonies with the wrong fluorescent reporters) remained relatively constant. The incorrect assemblies thus became a larger proportion of the colony count, decreasing accuracy (Figure 2c). To address this we included a second antibiotic resistance cassette, chosen so that the final construct could be selected using double antibiotic selection without any of the starting constructs conferring resistance. This significantly reduces the proportion of incorrect assemblies, indicating that these arise largely through carryover of storage plasmids when only a single antibiotic marker is used (Figure 2c). The double antibiotic selection strategy thus provides a significant improvement in the accuracy of BASIC assembly and was therefore adopted as the standard method in subsequent assemblies.

Since the orthogonal linker sequences provide positional watermarks in the final assembly, they may be used to validate assembly since they act as ideal PCR primer sites. This strategy was used to evaluate the 5-part assembly, demonstrating the flexibility in re-ordering parts simply by changing the linker combinations ligated to each part. We assessed the assembly order of these reactions as well as the seven part construct by performing PCR reactions with a forward primer for the first linker and reverse primers for each of the other 4 or 6 linkers. The PCR products exhibit the anticipated ladder of increasing size demonstrating the correct order and presence of each part in the assembly (Supplementary Figure 1). This provides a useful screening method for DNA assembly verification and because the linkers are standardised, the PCR verification primers are also standardised (Supplementary Table 13). Because the DNA assembly workflow starts with plasmid DNA and does not involve PCR amplification, there is less of an imperative to sequence the final construct following positional verification of the parts, which is especially useful when constructing pathways and libraries.

Hierarchical assembly. In many cases it is advantageous to assemble a limited number of parts together in a module and then combine different modules to create more complex systems or to re-use modules in different assemblies. The single-tier approach of BASIC therefore requires an idempotent method by which the *i*P and *i*S sequences can be recapitulated during DNA assembly. The objective therefore was to encode *i*P and *i*S on linkers attached during DNA assembly but avoiding any modification to the protocol. To achieve this we investigated DNA methylation as a strategy to protect the BsaI site from digestion during the assembly process.

The cognate DNA methyltransferase of the BsaI restriction modification system is a C-5 methyltransferase, but its target within the BsaI recognition sequence is not known(16). We have therefore determined the pattern of methylation protection through *in vitro* digestion of fluorescently labelled oligonucleotides, with each of the 4 cytosine residues within the recognition site methylated in turn. The restriction digests clearly reveal that methylation of the bottom strand only partially protects the DNA from digestion, while methylation of either cytosine in the top strand effectively protects the DNA from digestion by BsaI (Supplementary Figure 2). We therefore propose a general single-tier workflow for BASIC, where *i*P and *i*S are recapitulated around the constructed cargo during assembly by methylation of specified linker oligonucleotides to avoid cleavage during the combined digestion/ligation step (Figure 3a).

Figure 3

To demonstrate this approach we separately constructed GFP and RFP expression cassettes from individual parts encoding a constitutive promoter (J23102) and RBS-ORFs for GFP and RFP: these cassettes were then used in a second round of assembly to construct a dual fluorescence plasmid (Figure 3b). Parallel reactions were also performed with non-methylated linkers to benchmark the efficiency of the idempotent assembly compared to standard linkers [for a detailed list of assembly order see Supplementary Table 7]. The 4-part first round of assembly proceeded with 99% accuracy and an efficiency that was only 10% lower than that with standard linkers (Figure 3c). The expression cassette constructs were then successfully used for construction of the dual reporter plasmid. This demonstrates that methylation of a single cytosine in the BsaI recognition sequence provides sufficient protection against BsaI digestion to enable an idempotent strategy without modification of the protocol. Maintaining the same protocol for all stages of assembly and for all parts creates an easy workflow for either bench-scale work or automation.

Linkers as composable parts encoding RBS sequences. One feature of synthetic biology is the ability to rationally compose parts to provide either tuneable or predictable behaviour. Using custom RBS sequences to regulate protein translation has become increasingly common(17-20). The use of synthetic linker sequences provides the opportunity to encode small parts within the linker, such as RBS sequences. In line with our modular standardised approach to DNA assembly, we chose to tune the output of fluorescent reporters by encoding known RBS sequences of different strengths with the expectation that local sequence context would provide additional variability(19). Four RBS sequences were selected from the iGEM Parts Registry, and encoded onto the double stranded portion of *i*P (Figure 4a; Supplementary Table 8). Two linker overhang sequences were designed using R2oDNA Designer software(13) that are suitable for assembly with the 4 RBS sequences used. These two linker overhangs are orthogonal to the other linkers used in this paper and thus it was possible to generate a library of RBS sequences that can be incorporated in two different locations within a single assembly.

Figure 4

To evaluate the tunability of protein expression using RBS-linkers, 4-part assemblies were performed with the four different strength RBS linkers to join a constitutive promoter to a GFP ORF part without an RBS, but with a start codon adjacent to its *i*P (Figure 4b). Two sets of assemblies were performed to evaluate the degree of variation caused by the minor context change produced by changing the overhang sequence in the two sets of RBS linkers. Additionally we evaluated the potential to perform combinatorial library assembly by including multiple RBS linkers for a single part using a combination of RBS1 and RBS3 in one instance and of all four RBS linkers in the other. The four RBS sequences clearly give distinct levels of GFP expression, while there is no significant difference due to the overhang sequence context of linker 1 vs. linker 2 (Figure 4c).

To evaluate the combinatorial RBS assemblies, a number of individual colonies were randomly selected from a quadrant of the plate and grown out in culture. Comparison of expression levels for assemblies with a single RBS linker demonstrate that each colony tested exhibited a fluorescence expression within the expected range for the RBS sequences used. An even distribution of all RBS sequences used was also seen, demonstrating that there was no obvious bias between the RBS sequences chosen. All possibilities of RBS variants in the library construction were found within a relatively small number of colonies analysed. Furthermore none of the randomly selected colonies for either the specific or library constructions were incorrect, again demonstrating the overall accuracy of the assembly process.

The constructed sequences were computationally evaluated for predicted expression strength using the reverse mode of the RBS Calculator(18) (Figure 4c; Supplementary Table 15). It is interesting to note that the expected levels of expression follow the anticipated order for RBS 1-3, while RBS 4 gives significantly lower than expected output. For the different linker contexts, the computational prediction was that there would be a significant difference in protein output when the RBS sequences were combined with the different linker sequences. However, experimental results demonstrate that there was minimal variation in the protein output when the RBS sequences were placed in different linker contexts (Figure 4; Supplementary Table 15). While RBS calculator tools provide a reasonable correlation between prediction and output on a larger sample size(21), our results demonstrate that with a small population accurate prediction remains difficult.

Creation of fusion proteins. To further develop the BASIC approach we designed linkers that enable the fusion of protein parts during the assembly process. The *i*P and *i*S sequences have been designed to be compatible with fusion proteins and we have designed linkers that provide complete read through of coding sequence to generate peptide sequences that can join two in frame protein ORFs. A GFP ORF part was generated omitting the stop codon and with the final codon in frame to *i*S, while an RFP ORF was generated without an RBS and with the Met start codon in frame to *i*P (Supplementary Table 2). Three fusion linkers have been designed to encode peptide fusions with different properties including both flexible and alpha-helical sequences (Supplementary Table 11). These were designed with different codons to avoid nucleotide repeats and the sequences were validated with R2oDNA Designer software to ensure compatibility with BASIC.

Figure 5

Constructs expressing GFP and RFP cassettes singly and on the same plasmid were then constructed in addition to test constructs with GFP fused in frame to RFP using the linkers that encode peptide fusions. To demonstrate the functionality of the linkers, cells expressing the protein fusions were grown to mid-log phase and their protein expression analysed by SDS-PAGE, which revealed that all three constructs containing fusion linkers expressed stable GFP-RFP fusion (Figure 5).

Conclusion. BASIC comprises both a standard format for DNA parts and a new method for efficient parallel assembly. Our standardised assembly reactions can be benchmarked against our previous 4 part assemblies performed using Gibson reactions with the same orthogonal linkers defining the junctions(3). Our previously published Gibson 4-part assembly gave 75% accuracy(3), while the similar 4-part BASIC assembly reported here gave 93% accuracy with single antitibiotic selection and 99.4% accuracy with double antitibiotic selection, and 7-part assembly gave 90% accuracy with double antitibiotic selection.

The single-tier format retains the greatest degree of flexibility and simplicity and the presence of only one forbidden restriction sequence minimises adoption requirements. Operations such as changing the position of a part, or even reversing the direction of a promoter or ORF can easily be accomplished by simply changing the linkers. Additionally the assembly workflow is completely PCR-free, which greatly enhances its reliability, reduces the chances of introducing sequence errors and avoids the limitations of PCR.

While double antibiotic selection provides a significant improvement in accuracy for larger assemblies, high accuracy and efficiency can be maintained for smaller assemblies of up to four parts with only single antibiotic selection. The mode of implementation can therefore be chosen by the user based on their specific requirements. Alternative strategies to reduce background may also be employed, such as PCR amplification of the part containing the selectable marker, followed by DpnI digestion. We have avoided this to maintain a uniform workflow for all parts.

The use of orthogonal sequences to direct assembly, together with the BASIC protocol offers significant advantages over existing DNA assembly technologies. We have demonstrated that it is possible to position the same promoter part in different locations with great accuracy and no loss of efficiency. This would not be possible with a scarless method, such as the original Gibson protocol(6), the recently reported ligase cycling reaction method or paperclip (22, 23), because the repeated DNA sequence homology would misdirect parts in the final assembly.

The ability to assemble parts as small as 153 bp is also of significance and utility: small parts are known to be problematic with assembly methods that rely on exonuclease digestion as they can readily be digested. However small parts are frequently required for important functions such as promoters and in BASIC or other restriction-ligation based methods(1, 24) these pose no problems.

BASIC linkers also provide a means to encode biological functions for even smaller parts: we have demonstrated here that RBS parts can be composed on the adapter regions of the oligonucleotide linkers, or the whole linker can be used to code for peptide sequences that generate fusion proteins.

Verification of the final construct in DNA assembly is a critical component of the workflow. The orthogonal linkers employed in BASIC provide effective watermark sequences for this purpose with ideal PCR primer properties. The standardisation of these components means that assembly verification can utilise a limited set of standardised primers and a standard protocol. Despite reductions in sequencing costs at the genome scale, sequence verification of whole plasmid constructs remains costly in terms of both time and money. But because PCR is not used in BASIC, there is less of an imperative to sequence the final construct following positional verification of the DNA parts.

In common with other standards, adoption of BASIC can facilitate the sharing and re-use of parts and this is enhanced here by the single-tier format of the approach. It would also be possible to reuse parts designed for other type IIs methods like Golden Gate simply by changing the linker ligation overhang sequence. While a significant number of oligonucleotides are required for this method, their standardisation and long-term viability means that economies of scale rapidly accumulate as more people within a single laboratory or group of laboratories adopt the methodology (an evaluation of cost is provided in Supplementary Table 16). Furthermore, the robustness and predictability of all the steps in the BASIC protocol will facilitate the translation of the workflow to an automated liquid handling platform.

METHODS

BASIC assembly protocol. A full protocol for laboratory use is provided as online supplementary material.

Methods. Full details of all other materials and methods are provided in online supplementary information.

REFERENCES

1. **Ellis, T., Adie, T., and Baldwin, G. S. (2011) DNA assembly for synthetic biology: from parts to pathways and beyond, *Integrative Biology 3*, 109-118.**
2. **Arkin, A. (2008) Setting the standard in synthetic biology, *Nat Biotechnol 26*, 771-774.**
3. **Casini, A., MacDonald, J. T., De Jonghe, J., Christodoulou, G., Freemont, P. S., Baldwin, G. S., and Ellis, T. (2014) One-pot DNA construction for synthetic biology: the Modular Overlap-Directed Assembly with Linkers (MODAL) strategy, *Nucleic Acids Res 42*, e7.**
4. **Knight, T. Draft Standard for Biobrick Biological Parts.**
5. **Kahl, L. J., and Endy, D. (2013) A survey of enabling technologies in synthetic biology, *J Biol Eng 7*, 13.**
6. **Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nature methods 6*, 343-345.**
7. **Torella, J. P., Boehm, C. R., Lienert, F., Chen, J. H., Way, J. C., and Silver, P. A. (2013) Rapid construction of insulated genetic circuits via synthetic sequence-guided isothermal assembly, *Nucleic Acids Res*.**
8. **Guye, P., Li, Y., Wroblewska, L., Duportet, X., and Weiss, R. (2013) Rapid, modular and reliable construction of complex mammalian gene circuits, *Nucleic Acids Res 41*, e156.**

9. Sarrion-Perdigones, A., Falconi, E. E., Zandalinas, S. I., Juarez, P., Fernandez-del-Carmen, A., Granell, A., and Orzaez, D. (2011) GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules, *PLoS One 6*, e21622.
10. Sarrion-Perdigones, A., Vazquez-Vilar, M., Palaci, J., Castelijns, B., Forment, J., Ziarsolo, P., Blanca, J., Granell, A., and Orzaez, D. (2013) GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology, *Plant physiology 162*, 1618-1631.
11. Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011) A modular cloning system for standardized assembly of multigene constructs, *PLoS One 6*, e16765.
12. Engler, C., Gruetzner, R., Kandzia, R., and Marillonnet, S. (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes, *PLoS One 4*, e5553.
13. Casini, A., Christodoulou, G., Freemont, P. S., Baldwin, G. S., Ellis, T., and Macdonald, J. T. (2014) R2oDNA Designer: Computational Design of Biologically Neutral Synthetic DNA Sequences, *ACS synthetic biology 10.1021/sb4001323*.
14. Sambrook, J., E.F.Fritsch, and Maniatis, T. (1989) *Molecular Cloning A Laboratory Manual*, 2nd Edition ed., Cold Spring Harbor Laboratory Press, New York.
15. Horspool, D. R., Coope, R. J., and Holt, R. A. (2010) Efficient assembly of very short oligonucleotides using T4 DNA Ligase, *BMC research notes 3*, 291.
16. Roberts, R. J., Vincze, T., Posfai, J., and Macelis, D. (2010) REBASE--a database for DNA restriction and modification: enzymes, genes and genomes, *Nucleic Acids Res 38*, D234-236.
17. Barrick, D., Villanueba, K., Childs, J., Kalil, R., Schneider, T. D., Lawrence, C. E., Gold, L., and Stormo, G. D. (1994) Quantitative analysis of ribosome binding sites in E.coli, *Nucleic Acids Res 22*, 1287-1295.
18. Salis, H. M. (2011) The ribosome binding site calculator, *Methods Enzymol 498*, 19-42.
19. Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009) Automated design of synthetic ribosome binding sites to control protein expression, *Nat Biotechnol 27*, 946-950.
20. Na, D., and Lee, D. (2010) RBSDesigner: software for designing synthetic ribosome binding sites that yields a desired level of protein expression, *Bioinformatics 26*, 2633-2634.
21. Reeve, B., Hargest, T., Gilbert, C., and Ellis, T. (2014) Predicting translation initiation rates for designing synthetic biology, *Front Bioeng Biotechnol 2*, 1.
22. deKok, S., Stanton, L. H., Slaby, T., Durot, M., Holmes, V. F., Patel, K. G., Platt, D., Shapland, E. B., Serber, Z., Dean, J., Newman, J. D., and Chandran, S. S. (2014) Rapid and reliable DNA assembly via ligase cycling reaction, *ACS Synth. Biol. 3*, 97-106.
23. Trubitsyna, M., Michlewski, G., Cai, Y., Elfick, A., and French, C. E. (2014) PaperClip: rapid multi-part DNA assembly from existing libraries, *Nucleic Acids Res 42*, e154.
24. Chao, R., Yuan, Y., and Zhao, H. (2014) Recent advances in DNA assembly technologies, *FEMS yeast research*.

ASSOCIATED CONTENT

Supporting Information

Supplementary information, figures and tables are provided to give further details on protocol development and methods. In addition a stand-alone lab protocol is also provided. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

*To whom correspondence should be addressed:
g.baldwin@imperial.ac.uk
†These authors contributed equally to this work

ACKNOWLEDGEMENT

The authors would like to thank the following organisations for funding: EU (KBBE.2011.5-289326; IEF Fellowship 628019); BBSRC (BB/J020036/1); EPSRC (EP/J02175X/1). The authors do not declare any conflict of interest.

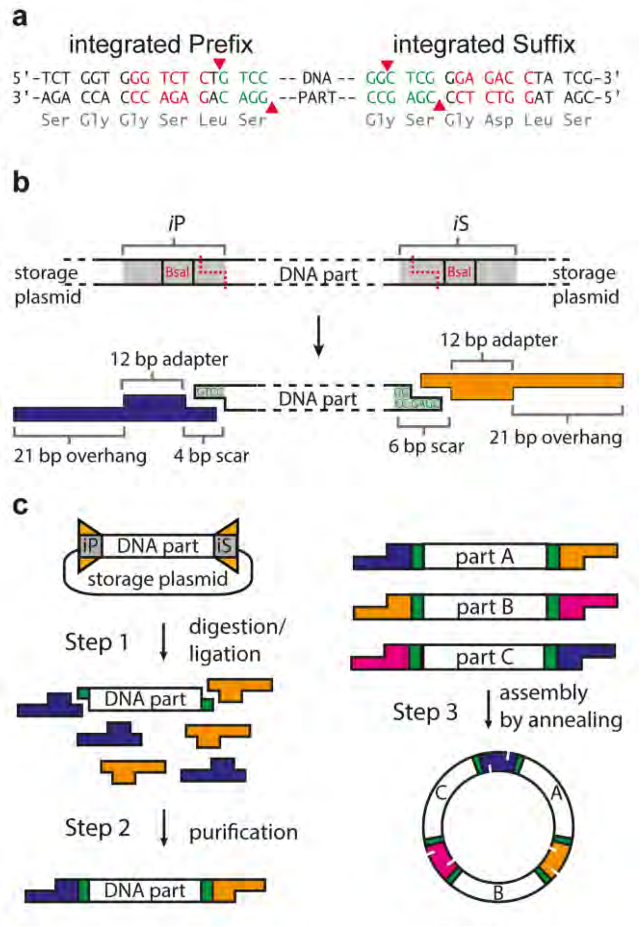


Figure 1. BASIC standard and method**. a**) Sequence of the integrated prefix (*i*P) and suffix (*i*S); the BsaI recognition sequence is shown in red with the cut sites marked (red triangles); the amino acid translation for each codon in the iP/iS are shown. **b**) During the assembly process *i*P and *i*S are cut to produce different sticky ends that enable differential ligation of linkers onto each end. **c**) The BASIC assembly workflow: Step 1: linkers are attached by simultaneous digestion and ligation. Step 2: unligated excess linkers are removed via magnetic bead purification. Step 3: purified linker-adapted parts are mixed and annealed in an ionic buffer to generate the desired final construct. [Protocol provided in Online Supplementary Information.]

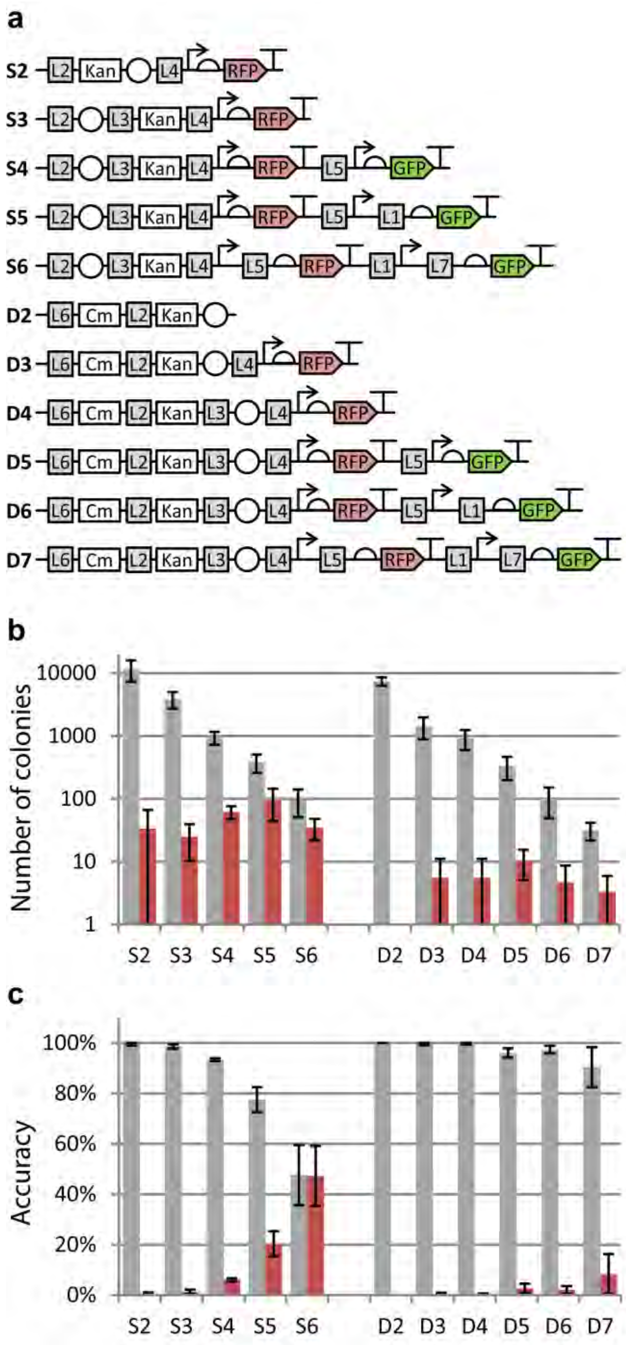


Figure 2. BASIC allows for highly efficient multi part assembly. **a**) Benchmarking DNA assembly reactions were performed creating constructs with 2 to 6 parts using single antibiotic selection (S2-S6 and 2 to 7 parts with double antibiotic selection (D2-D7) [Supplementary Table 1]. **b**) The number of colonies returned from each assembly is shown as the average of 4 repeat reactions with standard error of the mean (SEM; grey bars); the total number of incorrect assemblies that either had no fluorescence or incorrect fluorescence profiles are also shown (red bars). c) The accuracy of each assembly reaction was assessed as the % of colonies with the correct fluorescence profile for the designed assembly (grey bars); % incorrect assemblies are also shown (red bars). All data is shown as the average of 4 repeat reactions with SEM.

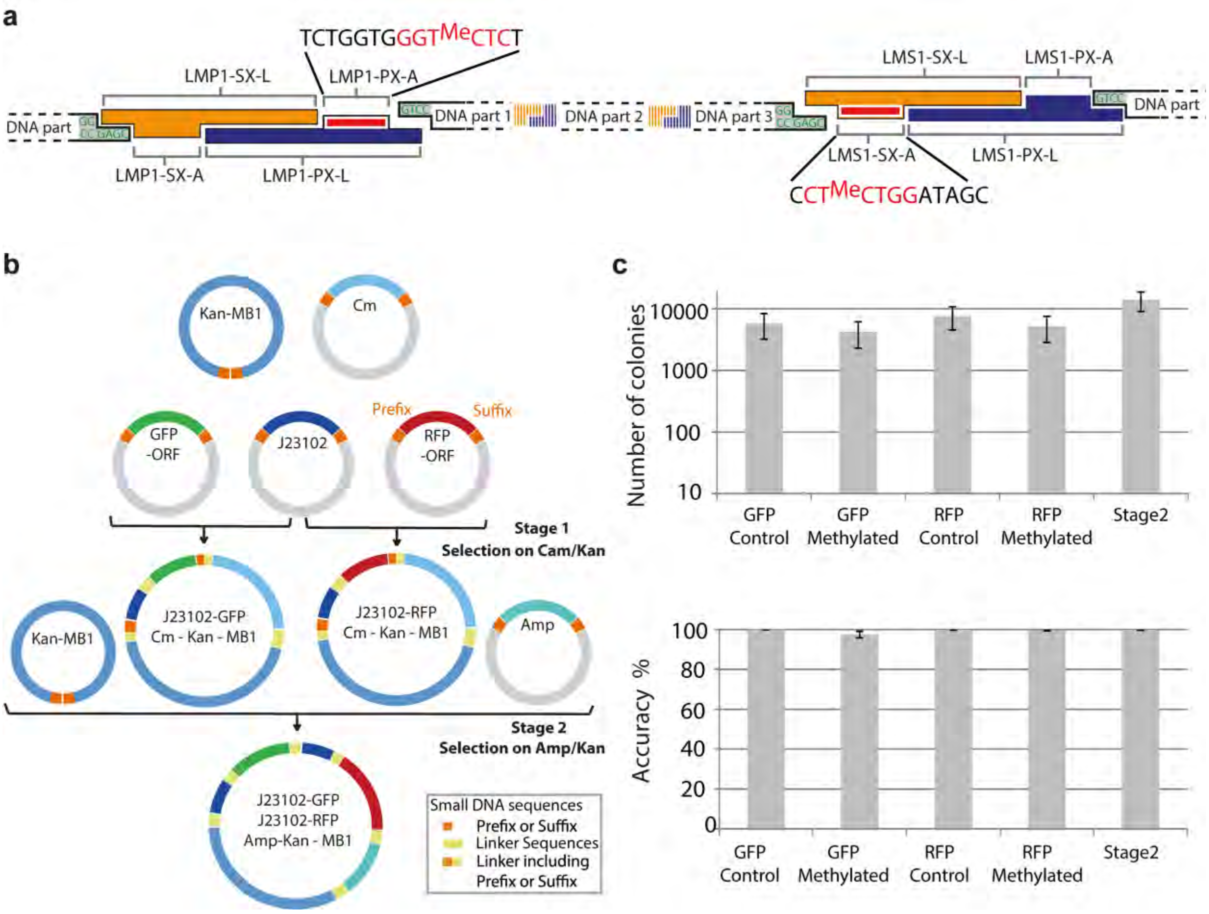


Figure 3. Hierarchical assembly using methylated linkers. **a**) Linker structure to recapitulate *i*P and *i*S adjacent to the parts being assembled. The methylated cytosine is located on the adapter oligonucleotide, which prevents digestion of the linker during the assembly process. **b**) Workflow to test idempotent DNA assembly using methylated linkers: in stage 1 GFP and RFP expressing cassettes are assembled flanked by *i*P and *i*S, backbone Kan-MB1 and Cm parts are located outside of *i*P and *i*S and so are not carried through in subsequent rounds. In stage 2 the previously assembled expression cassettes are used to assemble a double fluorescence reporter. **c**) Data from assembly reactions is shown for reactions with methylated linkers, control reactions with non-methylated linkers and stage 2 reactions. Data shown is the average of 4 repeat reactions with SEM for the number of colonies returned and the accuracy, determined as the % colonies with the correct fluorescence profile.

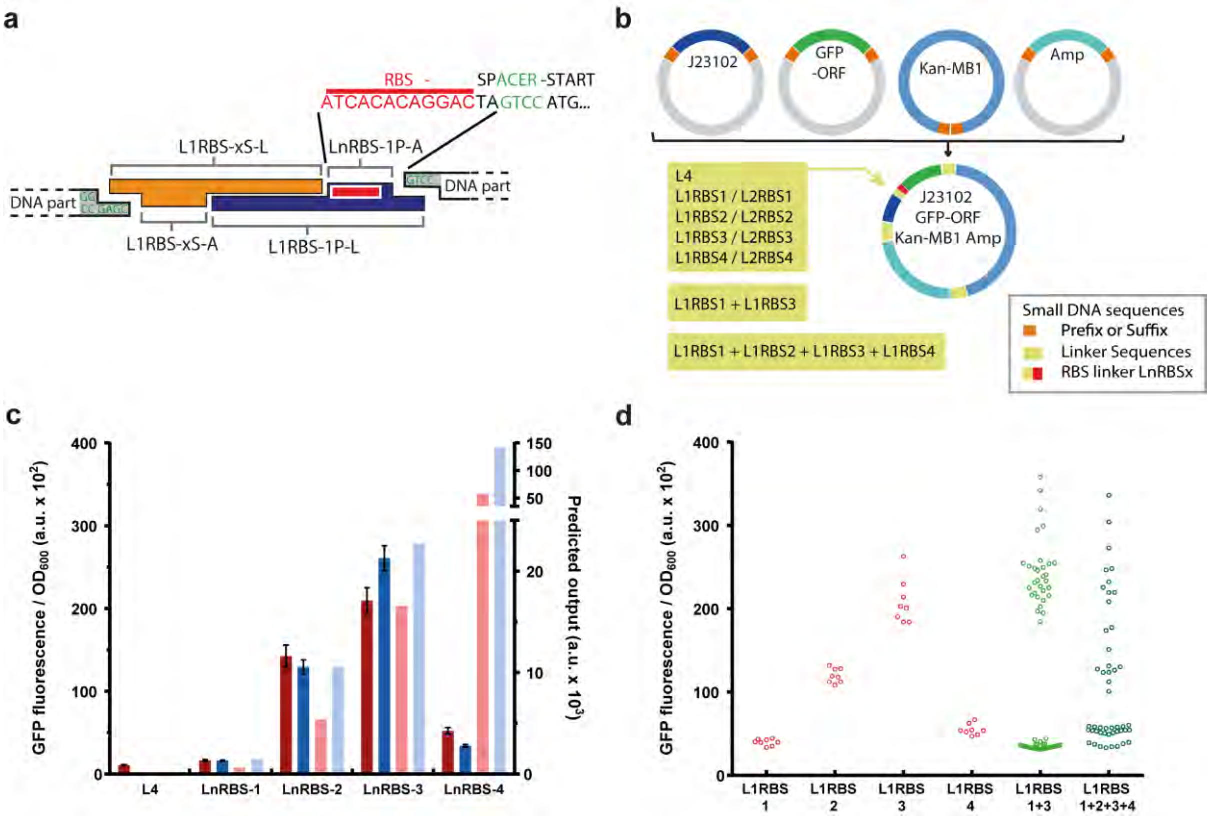


Figure 4. Tuning translation with RBS linkers. **a**) RBS sequences were encoded on the double stranded portion of the ligated linker oligonucleotide with a spacing region to ensure efficient translation efficiency [Supplementary Table 8]; the single stranded overlap of the linker does not encode the RBS and multiple RBS sequences were encoded with the same linker homology (LnRBSx, where n denotes the homology type within a series of x RBS sequences). **b**) Assembly strategy for constructs to test 4 RBS sequences within two different linker contexts and a control linker that does not encode an RBS (L4). **c**) GFP expression was evaluated after 6h growth and is shown normalised to OD_{600} for no RBS control (L4) and RBS1 to RBS4 with linker 1 (dark red bars) and linker 2 (dark blue bars). Predicted expression levels were calculated for all 4 RBS sequences in both linker contexts using the RBS calculator*(18)* and these are plotted for linker 1 (light red bars) and linker 2 (light blue bars). **d**) Assembly reactions were performed with single RBS linkers and also combinations of both two (RBS1&3) and four (RBS1-4) linkers to create a library of expression variants. Expression analysis of randomly selected colonies for these assemblies are shown as a dot plot.

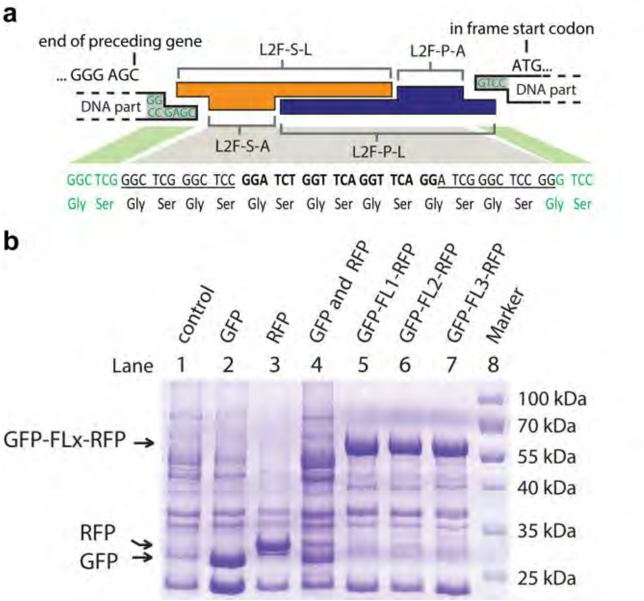


Figure 5. Creating fusion proteins with fusion linkers. **a**) Linkers were designed to provide an in frame polypeptide sequence to fuse two protein sequences, where the upstream gene had no stop codon and the downstream gene was in frame with *i*S. **b**) Constructs were created using 3 different fusion linkers [Supplementary Table 10] between GFP- and RFP-ORFs. SDS PAGE shows the expression of GFP and RFP in separate cells (lanes 2 and 3) and separately in the same cells (lane 4); the 3 fusion constructs of GFP and RFP are shown in lanes 5-7.

BASIC Protocol

0. Linker preparation

	volume to add	stock concentration	final concentration
Annealing buffer	49 µl		
Linker oligo	0.5 µl	100 µM	1 µM
Adapter oligo	0.5 µl	100 µM	1 µM
TOTAL Vol	50 µl		

Heat the mixed oligonucleotides at 95°C for 1 min and allow to cool slowly to room temperature, store at -20°C.

Annealing buffer: 10mM TRIS-HCl buffer pH7.9, 100mM NaCl, 10mM $MgCl_2$

1. Digestion/Ligation

	volume to add	stock concentration	final concentration
Water	8.5		
ATP	3 µl	10 mM	1 mM
NEBuffer 4	3 µl	10x	1x
BSA	3 µl	10x	1x
iP Linker	5 µl	1 µM	166.6 nM
iS Linker	5 µl	1 µM	166.6 nM
DNA Part[a]	1 µl	76 nM (~200 ng/µl)	2.5 nM (~6.7 ng/µl)
BsaI-HF[b]	1 µl	20 U/µl	0.66 U/µl
T4 Ligase[b]	0.5 µl	400 U/µl	6.6 U/µl
TOTAL Vol	30 µl		

Incubate using the following cycle: 37°C for 1 hour; 20°C for 20 minutes; 65°C for 20 minutes. [Extending the 37°C incubation time to up to 4 hours can enhance cleavage and overall efficiency.]

[a] Prepare parts at a concentration of 50 ng/µl per kb of the whole storage plasmid. i.e. for a 4 kb storage plasmid prepare a 200 ng/µl solution.

[b] It may be necessary to titrate enzyme levels to achieve complete digestion.

2. Purification

Use an AMPure XP magnetic DNA purification kit according to manufacturer's recommendations.
Use 54 µl beads per reaction and elute in 40 µl water.

3. Assembly

	volume to add	stock concentration	final concentration
DNA parts (each)	1 µl	1.5 nM[b]	0.15 nM[b]
BSA	1 µl	10x	1x
NEB4	1 µl	10x	1x
Water	up to 10 µl		
TOTAL Vol	10 µl		

Incubate at 50°C for 45 minutes (extending this time can enhance efficiency).

[b] Estimated assuming 80% purification efficiency.

4. Transformation

Transform 40 µl of chemically competent *E. coli* DH5α cells using 5 µl of assembly mix. [Competent cells with an efficiency of 10^9 cfu / µg pUC19 DNA are recommended, particularly for assembly reactions with large numbers of parts.]

Supporting Information:
BASIC: a new Biopart Assembly Standard for Idempotent Cloning provides accurate, single-tier DNA assembly for synthetic biology

Marko Storch[1†], Arturo Casini[1†], Ben Mackrow[1], Toni Fleming[2], Harry Trewhitt[1], Tom Ellis[3], Geoff S. Baldwin*[1]

1. Department of Life Sciences & Centre for Synthetic Biology & Innovation, Imperial College, South Kensington Campus, London, SW7 2AZ, UK
2. Dr Reddy's Chirotech Centre, Milton Road, Cambridge, CB4 0PE, UK
3. Department of Bioengineering & Centre for Synthetic Biology & Innovation, Imperial College, South Kensington Campus, London, SW7 2AZ, UK

*To whom correspondence should be addressed:
g.baldwin@imperial.ac.uk
[†]These authors contributed equally to this work

Supplementary Information

Table of contents

Assembly protocol optimisation

Oligonucleotide linkers. It was determined that prior annealing of the linker and adapter oligonucleotides was necessary for high efficiency assembly (*i.e.* prior formation of *e.g.* L1S and L1P, Supplementary Table 3). Tests demonstrated that a salt annealing buffer containing both $MgCl_2$ and NaCl provided optimal annealing conditions (Supplementary Online Material: BASIC protocol). In principle non-phosphorylated oligonucleotides could be purchased and polynucleotide kinase (PNK) included in the reaction as ATP is present in the ligation buffer. We have not tested this, but we have experienced small increases in efficiency when adding PNK even with phosphorylated oligonucleotides (PNK is not included as standard).

We have evaluated the efficiency of BASIC assembly using oligonucleotides stored for >12 months (oligonucleotide concentration 100 μM, 10 mM Tris pH 7.9, -20°C) against newly synthesised oligonucleotides and did not observed any appreciable loss of efficiency. Annealed linker and adapter oligonucleotides have been stored for up to 1 month in annealing buffer at 4°C without significant loss of efficiency in the BASIC reaction.

Linker ligation: The first step of the BASIC assembly protocol consists of the simultaneous digestion/ligation reaction. Optimisation of this step determined that complete digestion of the plasmid was essential for minimising background, especially if only a single antibiotic marker was used. The critical factors in this were the concentration of DNA and the quantity and ratio of *Bsa*I and T4 DNA ligase. We have experienced batch-to-batch variations and users may wish to perform test ligations on a two part assembly to determine optimal enzyme concentrations. Gel electrophoresis of digestion/ligation mixtures can provide a useful evaluation of digestion efficiency, but was not performed as standard. Temperature cycling between 16°C and 37°C gave no noticeable improvement in efficiency or accuracy compared to 37°C incubation for 1 h, 20°C for 20 mins., followed by heat inactivation at 65C for 20 mins.

Many of the DNA parts are stored pJET1 (ThermoFisher), which contains a BsaI site within its Amp gene. This may enhance the efficiency of the assembly process, but it has not been explicitly evaluated.

Purification: Purification was required to remove unincorporated linkers prior to annealing of the parts to create the final construct; magnetic beads provided a simple, cost-effective method for oligonucleotide removal. Evaluation of spin columns indicated a much lower efficiency of assembly indicating possible carry-over of unincorporated oligonucleotides.

Assembly: Following purification, adapted parts were mixed in an ionic buffer at 50°C for 45 min before transformation into chemically competent cells. Elevated temperature for annealing gave a significant improvement in efficiency and accuracy. Reactions at 37°C with T4 DNA ligase gave significantly higher error rates that we ascribe to inaccurate ligation or annealing of junctions caught in thermodynamic traps that may be stabilised by non-productive binding of T4 ligase. Such non-productive junctions are repaired *in vivo* following transformation and sequencing of incorrect junctions demonstrated that a single linker sequence was present, but joining incorrect parts. Assembly reactions at elevated temperature with Taq ligase gave no improvement over assembly reactions without ligase; exclusion of an enzyme in the second stage provided an improvement in both simplicity and cost.

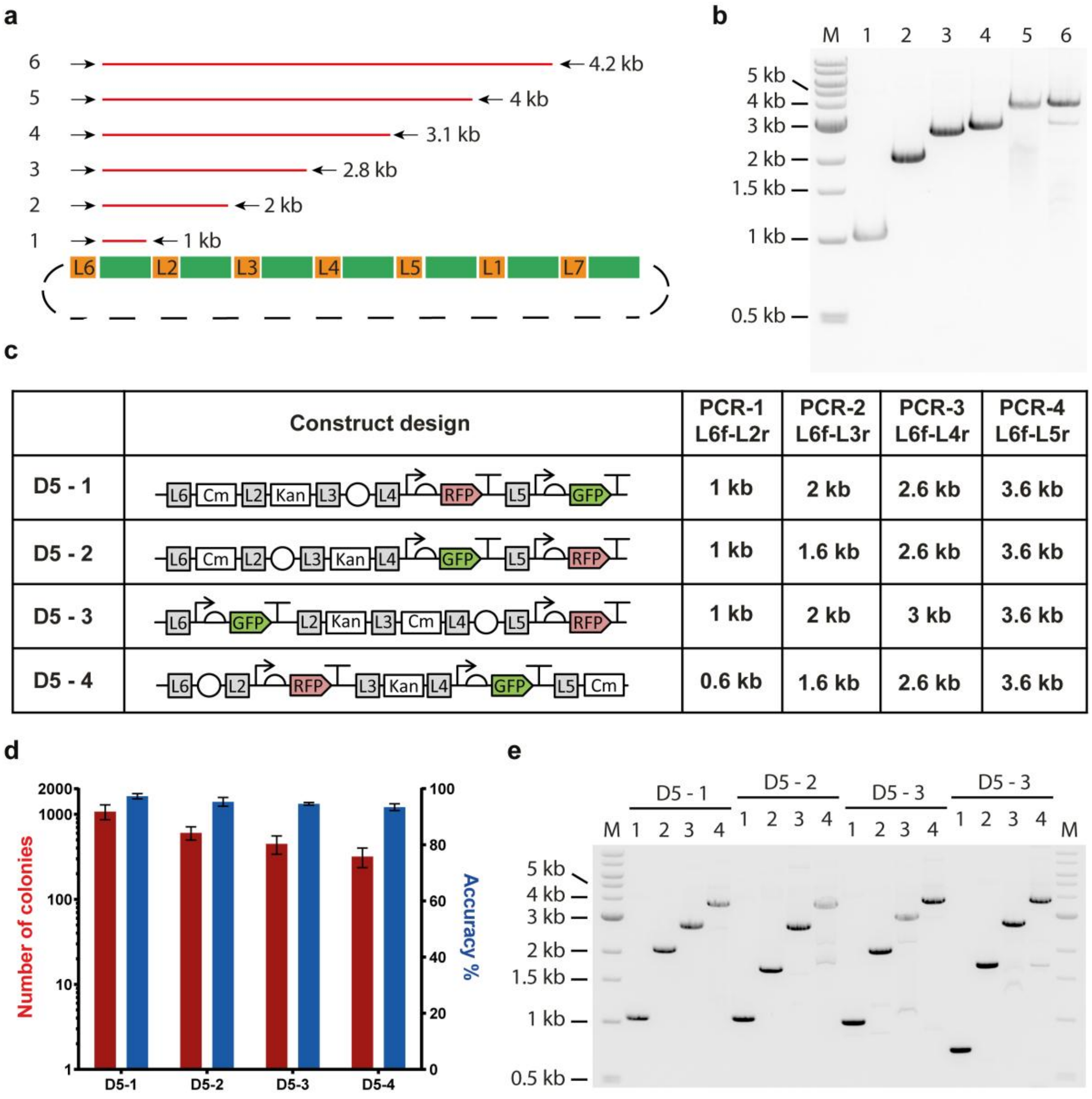
BASIC benchmarking results

Construct name	# of parts	Part composition	Number of colonies	Accuracy
S2	2	L2 Kan-MB1 L4 RFP	11658.3 ± 4281.8	99.5% ±0.5
S3	3	L2 Kan L3 MB1 L4 RFP	3875.0 ±1145.8	98.7% ±0.8
S4	4	L2 Kan L3 MB1 L4 RFP L5 GFP	944.4 ±219.7	93.3% ±0.6
S5	5	L2 Kan L3 MB1 L4 RFP L5 J23102 L1 GFP-ORF	382.5 ±124.2	77.5% ±5.0
S6	6	L2 Kan L3 MB1 L4 J23102 L5 RFP-ORF L1 J23102 L7 GFP-ORF	96.1 ±45.2	47.6% ±11.9
D2	2	L6 Cm L2 Kan-MB1	7416.6 ±1120.4	100% ±0.0
D3	3	L6 Cm L2 Kan-MB1 L4 RFP	1433.3 ±546.2	99.5% ±0.5
D4	4	L6 Cm L2 Kan L3 MB1 L4 RFP	920.0 ±324.3	99.7% ±0.3
D5	5	L6 Cm L2 Kan L3 MB1 L4 RFP L5 GFP	332.7 ±133.6	96.0% ±1.8
D6	6	L6 Cm L2 Kan L3 MB1 L4 RFP L5 J23102 L1 GFP-ORF	100.2 ±50.9	97.4% ±1.5
D7	7	L6 Cm L2 Kan L3 MB1 L4 J23102 L5 RFP-ORF L1 J23102 L7 GFP-ORF	31.6 ±10.0	90.4% ±7.9

Supplementary Table 1: Colony number and accuracy values for the constructs assembled to benchmark BASIC's efficiency, as presented in Figure 2. Constructs S2-S6 contain a single antibiotic marker, while constructs D2 to D7 have two. The table reports the total number of colonies obtained and the accuracy of each assembly reaction, assessed as the % of colonies with the correct fluorescence profile for the designed assembly. All data is shown as the average of 4 repeat reactions and SEM.

Assembly validation by PCR

Since BASIC assembled constructs contain standardized linker sequences as watermarks, these linker sequences lend themselves for deriving standard PCR primers to verify the desired construct architecture via colony PCR or PCR on mini-prepped DNA. We used this approach to here verify the 7 part construct assembled during benchmarking (D7; main paper Figure 2); PCR reactions were performed on DNA using a single forward primer in each reaction and with a different reverse primer from each of the other linker sequences. Resolution of reactions by agarose gel verified the presence of each of the linker sequences and the expected size according to the correct assembly (Supplementary Figure 1). It is in principle possible to use any combination of primer pairs derived from the linker sequences (Supplementary Table 3 & 13) to verify constructs as required.



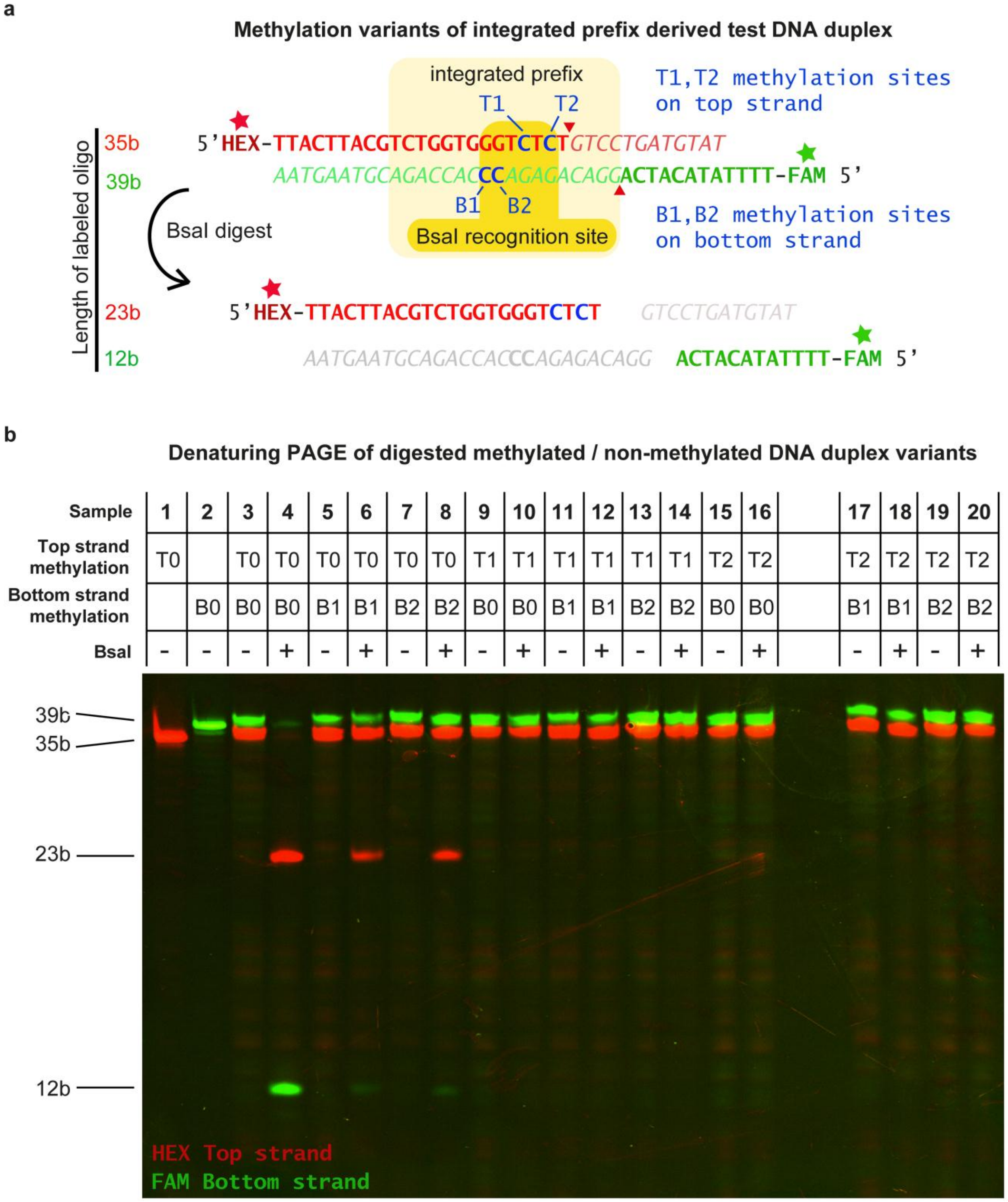
	Construct design	PCR-1 L6f-L2r	PCR-2 L6f-L3r	PCR-3 L6f-L4r	PCR-4 L6f-L5r
D5 - 1	L6 Cm L2 Kan L3 L4 RFP L5 GFP	1 kb	2 kb	2.6 kb	3.6 kb
D5 - 2	L6 Cm L2 L3 Kan L4 GFP L5 RFP	1 kb	1.6 kb	2.6 kb	3.6 kb
D5 - 3	L6 GFP L2 Kan L3 Cm L4 L5 RFP	1 kb	2 kb	3 kb	3.6 kb
D5 - 4	L6 L2 RFP L3 Kan L4 GFP L5 Cm	0.6 kb	1.6 kb	2.6 kb	3.6 kb

Supplementary Figure 1. Validation PCRs were performed for the 7 part construct D7 (main paper Figure 2). **a**) D7 is abstracted in linker (orange) and part (green) sequence. Primer pairs were chosen using a single forward primer (V-L6-F; Supplementary Table 13) and reverse primers that walked around each of the linker sequences. **b**) Reactions were analysed on a 0.8% agarose gel and showed the anticipated sizes for the correctly assembled construct for the different reverse primers used (lane 1: V-L2-R, 1 kb; lane 2: V-L3-R, 2 kb; lane 3: V-L4-R, 2.8 kb; lane 4: V-L5-R, 3.1 kb; lane 5: V-L1-R, 4 kb; lane 6: V-L7-R, 4.2 kb). **c**) The ability to assemble parts in alternative configurations was verified with 5-part constructs as shown; predicted sizes for validation PCRs are given for different primer combinations. **d**) Assembly efficiency and accuracy are shown for the different 5-part assemblies. **e**) Verification PCR demonstrates the anticipated sizes for the correct order of assembly from randomly selected colonies.

Protection of BsaI recognition site by methylation

In order to re-use BASIC assembled sequences in subsequent assemblies as parts, Prefix and Suffix regions have to be reconstituted by flanking linkers. These linkers need to contain BsaI sites, which have to be protected from digestion during the linker ligation step of the BASIC protocol. Methylation of enzyme recognition sites provides an opportunity for protecting restriction sites from digestion by their respective enzyme. We devised an assay to identify 5'methylation of cytosines within the BsaI recognition site that would provide protection against BsaI restriction.

To test BsaI restriction sensitivity towards 5'methylation we designed DNA duplex containing the prefix sequence with its BsaI restriction recognition site and labelled the Top strand with HEX and Bottom strand with FAM dye on their 5' sites. For each strand we ordered a control oligonucleotide without 5'methylation, and oligonucleotides with a 5'methylation on the first or second cytosine within the BsaI recognition side (Supplementary Table 14). All duplex combinations of the 3 Top and 3 Bottom strands were annealed and subjected to BsaI digestion. Denaturing PAGE was used to separate the resulting oligonucleotides. Depending on restriction efficiency, either the full length or restricted oligonucleotides carrying HEX or FAM dyes were visualized on a gel scanner (Supplementary Figure 2). Both T1 and T2 methylation variants provided strong protection against BsaI restriction and the T1 version was chosen as default methylation position for protecting BsaI restriction sites in methylated BASIC linkers (Supplementary Table 5).



Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Top strand methylation	T0		T0	T0	T0	T0	T0	T0	T1	T1	T1	T1	T1	T1	T2	T2	T2	T2	T2	T2
Bottom strand methylation		B0	B0	B0	B1	B1	B2	B2	B0	B0	B1	B1	B2	B2	B0	B0	B1	B1	B2	B2
BsaI	-	-	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+

Supplementary Figure 2. a) Oligonucleotides encoding *i*P and *i*S were either un-methylated, or methylated at the first or second cytosine on the Top or Bottom strand (T0, T1, T2, B0, B1, B2). All oligonucleotide combinations were annealed as for normal linker oligonucleotides. **b**) Oligonucleotides were digested with BsaI and separated by denaturing PAGE. Un-methylated oligonucleotides were fully digested (lane 4, T0 B0), while those methylated on either of the bottom positions (lane 6, T0 B1; lane 8, T0 B2) were partially protected from digestion. Methylation of either of the cytosines on the Top strands resulted in complete protection against digestion (lane 10, T1 B0; lane 16, T2 B0).

Materials & Methods

Part library construction

BASIC_1_pMB1 was created by combining pMB1 and the AMP gene (containing a BsaI site) via Gibson assembly including prefix and suffix flanking the pMB1 origin. BASIC_2_Kan-MB1 was created by combining the MB1 part (BASIC_1_MB1) with the Kan part (BASIC_3_Kan) with L2 between the parts and blunt ligation of the PCR product including the *i*P and *i*S primer extensions. For all other parts (BASIC_3-13), PCR (Phusion Polymerase, NEB) was used to attach *i*P and *i*S sequences via primer extension. The resulting PCR product was blunt-ligated into pJET1.2 (CloneJET, Fermentas) using T4 DNA Ligase (NEB). All parts were sequence verified and sequences can be accessed via Genbank accession numbers given in Supplementary Table 2.

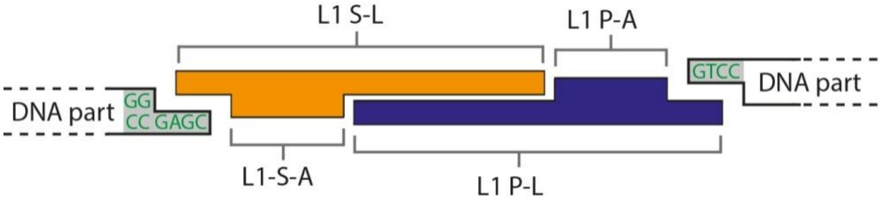
Storage plasmid name	**Part abbreviation**	**Size (bp)**	**Genbank accession number**
BASIC_1_MB1	MB1	1591	KP223695
BASIC_2_Kan-MB1	Kan-MB1	1683	KP223696
BASIC_3_Kan	Kan	3987	KP223697
BASIC_4_Cm	Cm	3970	KP223698
BASIC_5_Amp	Amp	3951	KP223699
BASIC_6_J23102	J23102	3163	KP223700
BASIC_7_J23101-RBS34-mCherry-B0015	RFP	3922	KP223706
BASIC_8_RBS34- mCherry -B0015	RFP-ORF	3881	KP223701
BASIC_9_ATG-mCherry -B0015	ATG-RFP-ORF	3861	KP223702
BASIC_10_J23101-RBS32-GFP-B0015	GFP	3935	KP223703
BASIC_11_RBS34-GFP-B0015	GFP-ORF	3891	KP223704
BASIC_12_ATG-GFP-B0015	ATG-GFP-ORF	3873	KP223705
BASIC_13_ATG-GFP-nostop	ATG-GFP-ORF-nostop	3730	KP223694

Supplementary Table 2: Storage plasmids containing BASIC parts. The first two parts (BASIC_1_pMB1 and BASIC_2_KAN-pMB1) are not stored in a pJET plasmid. MB1 is the MB1 origin of replication; Kan-MB1 is a composite part containing both Kanamycin resistance cassette and MB1; Kan is Kanamycin resistance cassette; Cam is Chloramphenicol resistance cassette; Amp is Ampicillin resistance cassette mutated to remove the BsaI site; J23101 and J23102 are promoters. Constructs _7 to _12 contain composites of parts coded in the order listed, where: RBS34 is the ribosome binding site from Bba00034 and similarly RBS32 is from Bba00032; mCherry is the mCherry ORF; B0015 is a terminator; GFP is superfolder GFP ORF; ATG means the part is truncated to remove RBS and starts in frame; nostop is where the stop codon has been removed and is in frame to *i*S to enable read-through of fusion proteins.

BASIC linker oligonucleotides

Neutral Linkers

All linker oligonucleotides (given 5’->3’) were ordered 5’ phosphorylated and HPLC purified from IDT. BASIC linkers are 45 bp DNA sequences used to connect 2 parts in BASIC format. Every linker (Ln; where n denotes the identifying linker number) is split in a suffix linker part (LnS) and a prefix linker part (LnP) each made up from one 37 base long linker oligonucleotide (LnS-L and LnP-L) and their respective 12 base adapter oligonucleotides (LnS-A and LnP-A).



Supplementary Figure 3. Structure of BASIC linkers.

Name			Sequence 5’->3’
L1	**L1S**	**L1S-L**	PO_4-CTCGttacttacgacactccgagacagtcagagggta
		L1S-A	PO_4-tgtcgtaagtaa
	L1P	**L1P-L**	PO_4-GGACtagttcaataaataccctctgactgtctcggag
		L1P-A	PO_4-tttattgaacta
L2	**L2S**	**L2S-L**	PO_4-CTCGatcggtgtgaaaagtcagtatccagtcgtgtag
		L2S-A	PO_4-tttcacaccgat
	L2P	**L2P-L**	PO_4-GGACaggtaataagaactacacgactggatactgact
		L2P-A	PO_4-ttcttattacct
L3	**L3S**	**L3S-L**	PO_4-CTCGatcacggcactacactcgttgctttatcggtat
		L3S-A	PO_4-tagtgccgtgat
	L3P	**L3P-L**	PO_4-GGACtctgtaataacaataccgataaagcaacgagtg
		L3P-A	PO_4-tgttattacaga
L4	**L4S**	**L4S-L**	PO_4-CTCGacccacgactattgactgctctgagaaagttga
		L4S-A	PO_4-atagtcgtgggt
	L4P	**L4P-L**	PO_4-GGACtaatcgtaacaatcaactttctcagagcagtca
		L4P-A	PO_4-ttgttacgatta
L5	**L5S**	**L5S-L**	PO_4-CTCGagaagtagtgccacagacagtattgcttacgag
		L5S-A	PO_4-ggcactacttct
	L5P	**L5P-L**	PO_4-GGACaggataaatcaactcgtaagcaatactgtctgt
		L5P-A	PO_4-ttgatttatcct
L6	**L6S**	**L6S-L**	PO_4-CTCGgtattgtaaagcacgaaacctacgataagagtg
		L6S-A	PO_4-gctttacaatac
	L6P	**L6P-L**	PO_4-GGACaaggagaactgacactcttatcgtaggtttcgt
		L6P-A	PO_4-tcagttctcctt
L7	**L7S**	**L7S-L**	PO_4-CTCGaacttttacgggtgccgactcactattacagac
		L7S-A	PO_4-cccgtaaaagtt
	L7P	**L7P-L**	PO_4-GGACagattgtagtaagtctgtaatagtgagtcggca
		L7P-A	PO_4-ttactacaatct

Supplementary Table 3: Oligonucleotides constituting the standard BASIC linkers. BsaI cut site annealing regions are shown in green; linker annealing regions are shown in blue.

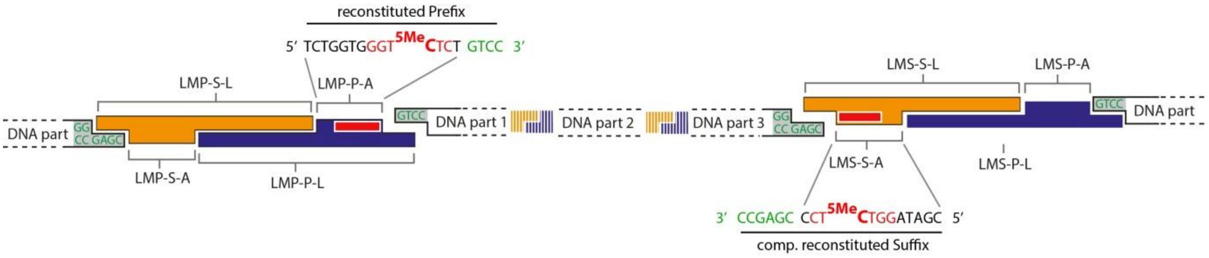
Resulting Neutral Linker sequences:

Name	Sequence 5'->3'
L1	CCGAGCTTACTTACGACACTCCGAGACAGTCAGAGGGTATTTATTGAACTAGTCC
L2	CCGAGCATCGGTGTGAAAAGTCAGTATCCAGTCGTGTAGTTCTTATTACCTGTCC
L3	CCGAGCATCACGGCACTACACTCGTTGCTTTATCGGTATTGTTATTACAGAGTCC
L4	CCGAGCACCCACGACTATTGACTGCTCTGAGAAAGTTGATTGTTACGATTAGTCC
L5	CCGAGCAGAAGTAGTGCCACAGACAGTATTGCTTACGAGTTGATTTATCCTGTCC
L6	CCGAGCGTATTGTAAAGCACGAAACCTACGATAAGAGTGTCAGTTCTCCTTGTCC
L7	CCGAGCAACTTTTACGGGTGCCGACTCACTATTACAGACTTACTACAATCTGTCC

Supplementary Table 4: Full linker sequences resulting from BASIC assembly with standard linkers: linker annealing regions are shown in blue; *i*S and *i*P scars are shown in green.

Methylated Linkers

Methylated linkers contain the full *i*P or *i*S, including the BsaI restriction site, to recapitulate the complete *i*P and *i*S around the assembly they flank, thus defining a new part constructed from a number of different input parts. They carry a 5'methylation in their BsaI recognition site protecting them from digestion during the combined digestion-ligation step. After amplification in *E.coli* the methylation is lost and the reconstituted prefix and suffix become functional. The linkers for the methylated prefix (LMiP) and suffix (LMiS) each consist of a suffix (LMP-S; LMS-S) and prefix (LMP-P; LMS-P) part since they have to link the newly assembled part with the backbone (Supplementary Figure 4). Both are made up from a linker oligonucleotide (LMP-S-L/ LMP-P-L; LMS-S-L/ LMS-P-L) and an adapter oligonucleotide (LMP-S-A/ LMP-P-A; LMS-S-A/ LMS-P-A). The methylated cytosines are encoded in the adapter oligonucleotides for both linkers (LMP-P-A; LMS-S-A).



Supplementary Figure 4. Structure of methylated BASIC linkers.

Name			Sequence 5'->3'
LMiP	**LMP-S**	**LMP-S-L**	PO_4-CTCGGGTAAGAACTCGCACTTCGTGGAAACACTATTA
		LMP-S-A	PO_4-CGAGTTCTTACC
	LMP-P	**LMP-P-L**	PO_4-GGACAGAGACCCACCAGATAATAGTGTTTCCACGAAGTG
		LMP-P-A	PO_4-TCTGGTGGGT/iMe-dC/TCT
LMiS	**LMS-S**	**LMS-S-L**	PO_4-CTCGGGAGACCTATCGGTAATAACAGTCCAATCTGGTGT
		LMS-S-A	PO_4-CGATAGGT/iMe-dC/TCC
	LMS-P	**LMS-P-L**	PO_4-GGACGATTCCGAAGTTACACCAGATTGGACTGTTATTAC
		LMS-P-A	PO_4-AACTTCGGAATC

Supplementary Table 5: Oligonucleotides constituting the methylated BASIC linkers. BsaI cut site annealing regions are shown in green; *i*P and *i*S sequences are shown in red; linker annealing regions are shown in blue. Methylation positions are shown as /iMe-dC/.

Resulting Methylated linker sequences

Name	Sequence 5’->3’
LMiP	CCGAGCGGTAAGAACTCGCACTTCGTGGAAACACTATTATCTGGTGGGTCTCTGTCC
LMiS	CCGAGCGGAGACCTATCGGTAATAACAGTCCAATCTGGTGTAACTTCGGAATCGTCC

Supplementary Table 6: Full linker sequences resulting from BASIC assembly with methylated linkers: *i*P and *i*S sequences are shown in red; linker annealing regions are shown in blue; *i*S and *i*P scars are shown in green.

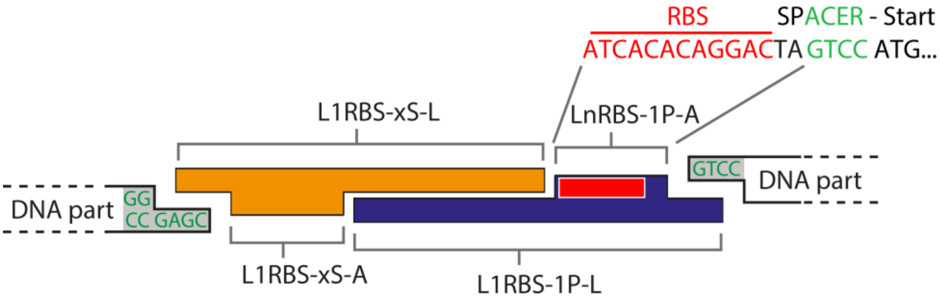
Construct name	# of parts	Part composition	New Part Name
GFP Control	4	L1 J23102 L2 GFP-ORF L3 Cm L4 Kan-MB1	n/a
GFP Methylated	4	LMiP J23102 L2 GFP-ORF LMiS Cm L4 Kan-MB1	J23102-GFP
RFP Control	4	L1 J23102 L2 RFP-ORF L3 Cm L4 Kan-MB1	n/a
RFP Methylated	4	LMiP J23102 L2 RFP-ORF LMiS Cm L4 Kan-MB1	J23102-RFP
Stage 2	4	L1 J23102-GFP L2 J23102-RFP L3 Amp L4 Kan-MB1	n/a

Supplementary Table 7: Constructs made to evaluate idempotent cloning with methylated oligonucleotides. GFP Methylated and RFP Methylated parts created in Stage 1 were used to create the dual reporter in Stage 2.

RBS Linkers

Functional RBS sequences were encoded within the adapter region of the linkers so that they were independent of the annealing region (Supplementary Figure 5). The length of the adapter and linker oligonucleotide for the Prefix site has been extended by 2 bases to accommodate a standard RBS of 12 bases plus 2 spacing bases. For each linker, versions with different RBS were derived forming RBS linker families. Since the RBS variation only changes the sequence of the two oligonucleotides forming the Prefix part of the linker, the 2 respective Suffix related oligonucleotides can be re-used (They therefore carry a x after the number indicating the RBSlinker family). Also the Prefix adapter oligonucleotide coding for the RBS can be re-used for different RBSlinker families for this specific RBS. RBS linkers (LnRBS) were designed with the use of R2oDNA software*(1)*.

Notation: LnRBS-xP/S-L/A with first number (n) indicating RBSLinker family and the second (x) the RBS in the resulting linker. The second position is not specified in the Suffix (S) linker and is denoted xS. Similarly the Prefix (P) adapter is not specific for any RBSLinker family and is denoted Ln.



Supplementary Figure 5. Structure of BASIC RBS linkers.

Name			Sequence 5′-3′
L1RBS1	**L1RBS-xS[1]**	**L1RBS-xS-L**	PO$_4$-CTCGTTGAACACCGTCTCAGGTAAGTATCAGTTGTAA
		L1RBS-xS-A	PO$_4$-GACGGTGTTCAA
	L1RBS-1P	**L1RBS-1P-L**	PO$_4$-GGACTAGTCCTGTGTGATTTACAACTGATACTTACCTGA
		LnRBS-1P-A[3]	PO$_4$-ATCACACAGGACTA
L1RBS2	**L1RBS-xS[1]**	**L1RBS-xS-L**	PO$_4$-CTCGTTGAACACCGTCTCAGGTAAGTATCAGTTGTAA
		L1RBS-xS-A	PO$_4$-GACGGTGTTCAA
	L1RBS-2P	**L1RBS-2P-L**	PO$_4$-GGACTATTTCCCCTCTTTTTACAACTGATACTTACCTGA
		LnRBS-2P-A[4]	PO$_4$-AAAGAGGGGAAATA
L1RBS3	**L1RBS-xS[1]**	**L1RBS-xS-L**	PO$_4$-CTCGTTGAACACCGTCTCAGGTAAGTATCAGTTGTAA
		L1RBS-xS-A	PO$_4$-GACGGTGTTCAA
	L1RBS-3P	**L1RBS-3P-L**	PO$_4$-GGACTATTTCTCCTCTTTTTACAACTGATACTTACCTGA
		LnRBS-3P-A[5]	PO$_4$-AAAGAGGAGAAATA
L1RBS4	**L1RBS-xS[1]**	**L1RBS-xS-L**	PO$_4$-CTCGTTGAACACCGTCTCAGGTAAGTATCAGTTGTAA
		L1RBS-xS-A	PO$_4$-GACGGTGTTCAA
	L1RBS-4P	**L1RBS-4P-L**	PO$_4$-GGACTACCTCCTTGTGATTTACAACTGATACTTACCTGA
		LnRBS-4P-A[6]	PO$_4$-ATCACAAGGAGGTA
L2RBS1	**L2RBS-xS[2]**	**L2RBS-xS-L**	PO$_4$-CTCGTGTTACTATTGGCTGAGATAAGGGTAGCAGAAA
		L2RBS-xS-A	PO$_4$-CCAATAGTAACA
	L2RBS-1P	**L2RBS-1P-L**	PO$_4$-GGACTAGTCCTGTGTGATTTTCTGCTACCCTTATCTCAG
		LnRBS-1P-A[3]	PO$_4$-ATCACACAGGACTA
L2RBS2	**L2RBS-xS[2]**	**L2RBS-xS-L**	PO$_4$-CTCGTGTTACTATTGGCTGAGATAAGGGTAGCAGAAA
		L2RBS-xS-A	PO$_4$-CCAATAGTAACA
	L2RBS-2P	**L2RBS-2P-L**	PO$_4$-GGACTATTTCCCCTCTTTTTTCTGCTACCCTTATCTCAG
		LnRBS-2P-A[4]	PO$_4$-AAAGAGGGGAAATA
L2RBS3	**L2RBS-xS[2]**	**L2RBS-xS-L**	PO$_4$-CTCGTGTTACTATTGGCTGAGATAAGGGTAGCAGAAA
		L2RBS-xS-A	PO$_4$-CCAATAGTAACA
	L2RBS-3P	**L2RBS-3P-L**	PO$_4$-GGACTATTTCTCCTCTTTTTTCTGCTACCCTTATCTCAG
		LnRBS-3P-A[5]	PO$_4$-AAAGAGGAGAAATA
L2RBS4	**L2RBS-xS[2]**	**L2RBS-xS-L**	PO$_4$-CTCGTGTTACTATTGGCTGAGATAAGGGTAGCAGAAA
		L2RBS-xS-A	PO$_4$-CCAATAGTAACA
	L2RBS-4P	**L2RBS-4P-L**	PO$_4$-GGACTACCTCCTTGTGATTTTCTGCTACCCTTATCTCAG
		LnRBS-4P-A[6]	PO$_4$-ATCACAAGGAGGTA

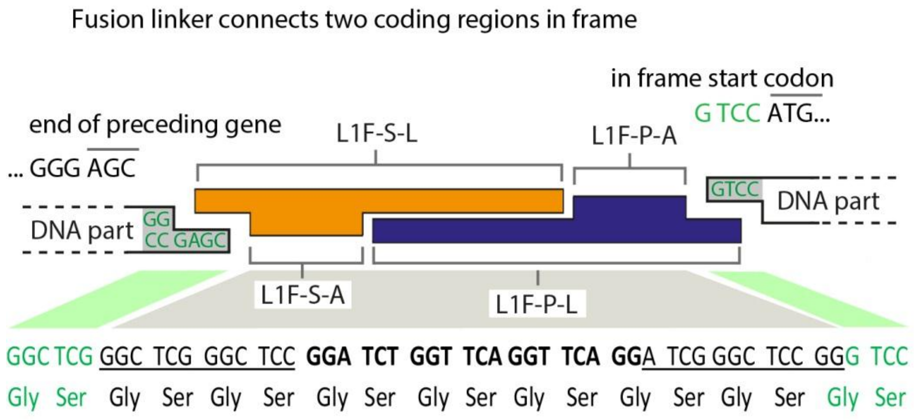
Supplementary Table 8: Oligonucleotides constituting the BASIC RBS linkers. RBS regions are shown in red; BsaI cut site annealing regions are shown in green; linker annealing regions are shown in blue. 1) and 2) suffix linker were used with different RBS specifying linker prefix parts ; 3),4),5),6) prefix adapter oligos encoding specific RBS sequence were reused with different linker suffix parts.

Resulting RBS linker sequences

Name	Sequence
L1RBS1	CCGAGCTTGAACACCGTCTCAGGTAAGTATCAGTTGTAAATCACACAGGACTAGTCC
L1RBS2	CCGAGCTTGAACACCGTCTCAGGTAAGTATCAGTTGTAAAAAGAGGGGAAATAGTCC
L1RBS3	CCGAGCTTGAACACCGTCTCAGGTAAGTATCAGTTGTAAAAAGAGGAGAAATAGTCC
L1RBS4	CCGAGCTTGAACACCGTCTCAGGTAAGTATCAGTTGTAAATCACAAGGAGGTAGTCC
L2RBS1	CCGAGCTGTTACTATTGGCTGAGATAAGGGTAGCAGAAAATCACACAGGACTAGTCC
L2RBS2	CCGAGCTGTTACTATTGGCTGAGATAAGGGTAGCAGAAAAAAGAGGGGAAATAGTCC
L2RBS3	CCGAGCTGTTACTATTGGCTGAGATAAGGGTAGCAGAAAAAAGAGGAGAAATAGTCC
L2RBS4	CCGAGCTGTTACTATTGGCTGAGATAAGGGTAGCAGAAAATCACAAGGAGGTAGTCC

Supplementary Table 9: Full linker sequences resulting from BASIC assembly with RBS linkers: RBS regions are shown in red; spacer regions are shown in orange; linker overlaps are shown in blue; *i*S and *i*P scars are shown in green.

Fusion Linkers



Supplementary Figure 6. Structure of BASIC Fusion linkers.

Name			Sequence 5'-3'
LF1	**LF1-S**	**LF1-S-L**	PO_4-CTCGGCCGAAGCGGCTGCTAAAGAAGCAGCTGCTAAAGAGGCGGC
		LF1-S-A	PO_4- AGCCGCTTCGGC
	LF1-P	**LF1-P-L**	PO_4- GGACCCTGCCTTGGCGGCCGCCTCTTTAGCAGCTGCTTCTTTAGC
		LF1-P-A	PO_4- CGCCAAGGCAGG
LF2	**LF2-S**	**LF2-S-L**	PO_4- CTCGGGCTCGGGCTCCGGATCTGGTTCAGGTTCAGG
		LF2-S-A	PO_4- GGAGCCCGAGCC
	LF2-P	**LF2-P-L**	PO_4- GGACCCGGAGCCCGATCCTGAACCTGAACCAGATCC
		LF2-P-A	PO_4- ATCGGGCTCCGG
LF3	**LF3-S**	**LF3-S-L**	PO_4- CTCGCTGCTTGAGAGCCCTAAAGCATTAGAAGAAGCACCTTGGCC
		LF3-S-A	PO_4- GCTCTCAAGCAG
	LF3-P	**LF3-P-L**	PO_4- GGACCCTTCTGGTGGAGGCCAAGGTGCTTCTTCTAATGCTTTAGG
		LF3-P-A	PO_4- TCCACCAGAAGG

Supplementary Table 10: Oligonucleotides constituting BASIC fusion linkers LF1-3. BsaI cut site annealing regions are shown in green; linker annealing regions are shown in blue.

Resulting fusion linker sequences and translations

Name	Sequence
LF1	GGCTCGGCCGAAGCGGCTGCTAAAGAAGCAGCTGCTAAAGAGGCGGCCGCCAAGGCAGGGTCC
LF1-translation	G S A E A A A K E A A A K E A A A K A G S
LF2	GGCTCGGGCTCGGGCTCCGGATCTGGTTCAGGTTCAGGATCGGGCTCCGGGTCC
LF2-translation	G S G S G S G S G S G S G S G S G S
LF3	GGCTCGCTGCTTGAGAGCCCTAAAGCATTAGAAGAAGCACCTTGGCCTCCACCAGAAGGGTCC
LF3-translation	G S L L E S P K A L E E A P W P P P E G S

Supplementary Table 11: Full linker sequences resulting from BASIC assembly with fusion linkers: linker overlaps are shown in blue; *i*S and *i*P scars are shown in green. LF1 encodes an alpha helical forming peptide; LF2 is a flexible peptide; LF3 is a naturally occurring flexible peptide.

Construct name	**Part composition**
Control	L1 Kan-MB1 L2 Cm
GFP	L1 Kan-MB1 L2 Cm L3 GFP
RFP	L1 Kan-MB1 L2 Cm L3 RFP
GFP & RFP	L1 Kan-MB1 L2 Cm L3 GFP L5 RFP
GFP-FL1-RFP	L1 Kan-MB1 L2 Cm L3 J23102 L1RBS3 ATG-GFP-ORF-nostop FL1 ATG-RFP
GFP-FL2-RFP	L1 Kan-MB1 L2 Cm L3 J23102 L1RBS3 ATG-GFP-ORF-nostop FL2 ATG-RFP
GFP-FL3-RFP	L1 Kan-MB1 L2 Cm L3 J23102 L1RBS3 ATG-GFP-ORF-nostop FL3 ATG-RFP

Supplementary Table 12: Constructs made to evaluate fusion linkers.

Linker design

All standard BASIC linkers were designed with the R2oDNA Designer software (R2oDNA Designer: Computational Design of Biologically Neutral Synthetic DNA Sequences, Casini *et al.*, 2014, www.r2odna.com). All other linkers were derived from R2oDNA Designer, rationally edited to accommodate specific features and verified again with R2oDNA Designer using the reverse mode.

Bacterial strains, reagents and protocols

Escherichia coli DH5alpha strain was used as the host to clone both storage plasmids and those assembled with BASIC. It was grown at 37°C in Luria-Bertani (LB) medium, with the appropriate antibiotic to select for cells transformed with the plasmid of interest at the following concentrations: kanamycin (50 µg/ml), ampicillin (50 µg/ml), chloramphenicol (25 µg/ml). Chemically competent cells were prepared as described in "High efficiency transformation of Escherichia coli with plasmids" (Inoue *et al*., 1990), except that cells are grown at 18°C in SOC medium. Competent cells were evaluated for efficiency to give at least 10^9 CFU/ug of pUC19. Cell transformation was performed as follows: 5 µl of solution containing DNA was mixed with 40 µl of competent cells in a 1.5 Eppendorf tube and incubated on ice for 20 mins. A heat shock was performed at 42°C for 45 seconds, after which the tubes were moved back to ice for 2 minutes. 670 µl of SOC medium was added to the tubes and they were placed in a 37°C shaking incubator for 1 hour. The tubes were spun at 13k rpm for 1 minute, and 600 µl of supernatant were discarded. The remaining liquid was either all plated, or divided in 10, 30 and 60 µl plating fractions as necessary.

All enzymes used were purchased from New England Biolabs (NEB) unless otherwise specified. BsaI-HF was used in all BASIC reactions. Sigma-Aldrich GenElute PCR Clean-Up Kit and EZNA Plasmid Mini Kit I were used, respectively, for PCR product and plasmid purifications unless otherwise specified. All oligonucleotides were synthesised by IDT. Oligonucleotides used for linkers in BASIC assembly were supplied 5' phosphorylated and HPLC purified.

Colony counting and visualization

All assembly reactions and cell transformations for each single assembly experiment repeat were performed on the same day, using the same batch of competent cells and plates to ensure comparability. Following colony growth, plates were scanned using a Fujifilm LAS5000 scanner: GFP expressing colonies were visualized by scanning with a blue (473 nm) laser and Fluorescein isothiocyanate (FITC) filter; RFP expressing colonies were visualized by scanning with a green (532 nm) laser and a long pass green (LPG) filter. Images were overlaid and aligned to correct for chromatic aberration using ImageJ software (NIH). The total numbers of colonies and colonies with different or no fluorescence profiles were counted manually. The assembly accuracy of each reaction was calculated as the percentage of total colonies for a reaction that showed the correct fluorescent expression profile.

PCR assembly validation

PCR reactions to validate DNA assembly were performed with plasmid DNA, primers as specified (Supplementary Figure 1; Supplementary Table 13) and Phusion Polymerase (NEB) in standard reaction conditions. The program used was: 3 minutes at 98°C, 30 cycles of 10 seconds at 98°C, 30 seconds at 65°C, 2 minutes 30s at 72°C, a final elongation step of 10 minutes at 72°C. Reactions were analysed on a 0.8% agarose gel. A complete list of oligonucleotides for use in assembly validation is provided in Supplementary Table 13.

Oligonucleotide name	**Sequence 5' -> 3'**
V-L1-F	GACACTCCGAGACAGTCAGAGGGTA
V-L1-R	TACCCTCTGACTGTCTCGGAGTGTC
V-L2-F	GTGTGAAAAGTCAGTATCCAGTCGTGTAGTTC
V-L2-R	GAACTACACGACTGGATACTGACTTTTCACAC
V-L3-F	GCACTACACTCGTTGCTTTATCGGTATTG
V-L3-R	CAATACCGATAAAGCAACGAGTGTAGTGC
V-L4-F	CGACTATTGACTGCTCTGAGAAAGTTGATTG
V-L4-R	CAATCAACTTTCTCAGAGCAGTCAATAGTCG
V-L5-F	GCCACAGACAGTATTGCTTACGAGTTG
V-L5-R	CAACTCGTAAGCAATACTGTCTGTGGC
V-L6-F	GCACGAAACCTACGATAAGAGTGTCAG
V-L6-R	CTGACACTCTTATCGTAGGTTTCGTGC
V-L7-F	GGGTGCCGACTCACTATTACAGACTTAC
V-L7-R	GTAAGTCTGTAATAGTGAGTCGGCACCC

Supplementary Table 13: List of primers that anneal to linker sequences for verification of assembled constructs.

Methylation sensitivity assay

100uM oligonucleotide stock solution (10mM Tris pH7.9) was diluted to 200nM working concentration (10mM Tris pH 7.9). Final concentration in restriction reaction was 100nM. DNA duplex was formed by mixing equal amounts of complementary oligonucleotides (200nM) and heating mixture in a heating block @ 100°C. After 5 min heat block was switched off and oligonucleotide mix was allowed to cool to room temperature.

5ul of duplex (200nM) was mixed with 2 µl H_20, 1 µl NEB4 buffer, 1 µl BSA10x (1mg/ml), 1 µl BsaI (20U) and incubation for 2h @ 37°C and heat inactivation at 65°C for 20min.

Samples were boiled for 2 min to denature DNA duplex and placed straight away on ice. 5 µl 3x denaturing loading buffer (95% formamide, 12mM EDTA) were added and 10ul were loaded onto 18% denaturing gel, which was run at 50°C. Gels were scanned on a Fujifilm LAS5000 laser scanner using 800V gain and 50 µm resolution. Lasers and filters used were 473nm/FITC for FAM and 532nm/Cy3 for HEX imaging.

Oligonucleotide name	Sequence (5'-3')
T0	**5HEX**-TT ACT TAC GTC TGG TGG GTC TCT GTC CTG ATG TAT
T1	**5HEX**-TT ACT TAC GTC TGG TGG GT**/iMe-dC/** TCT GTC CTG ATG TAT
T2	**5HEX**-TT ACT TAC GTC TGG TGG GTC T**/iMe-dC/**T GTC CTG ATG TAT
B0	**FAM**-TT TTA TAC ATC AGG ACA GAG ACC CAC CAG ACG TAA GTA A
B1	**FAM**-TT TTA TAC ATC AGG ACA GAG AC**/iMe-dC/** CAC CAG ACG TAA GTA A
B2	**FAM**-TT TTA TAC ATC AGG ACA GAG A**/iMe-dC/**C CAC CAG ACG TAA GTA A

Supplementary Table 14: Oligonucleotides used in the methylation sensitivity assay. 5' hexachlorofluorescein (HEX), 5' carboxyfluorescein (FAM), and 5-methyl deoxyCytosine (/iMe-dC/) were added during oligonucleotide synthesis.

RBS Linker experiments for GFP expression

8 colonies were picked from each Linker-RBS construct plate and used to inoculate 200µl LB medium (kanamycin 50 µg/ml) in a 96 well plate (Greiner) each. After overnight growth at 37°C shaking (600 rpm; Mikura). Cultures were normalized to OD 0.2 into a fresh 96 well plate in LB medium (total volume 200ul, kanamycin 50 µg/ml) and incubated as before. Measurements of OD_{600} and GFP fluorescence where taken after 3h and 6h in plate reader (Synergy HT, 485/528nm filters). Normalized GFP fluorescence was calculated by dividing the fluorescence value by the OD_{600} value for each well. In the negative control construct L4, neutral linker 4 was used instead of an RBS linker to join the Promoter part with the GPF ORF.

RBS Calculator was used to predict the output of the constructs created. Sequences from the start of the RBS linker up to the preceding *i*S scar and downstream 150 bases into the GFP ORF were submitted. *E. coli* MG 1655 was specified as the host. The expected start codon of the GFP ORF was always predicted to have the highest translation efficiency, and these values are plotted in Figure 4 (main paper) and Supplementary Table 15.

NAME	# of parts	Order	Predicted expression	Observed expression
L4	4	L6 Kan-MB1 L2 Amp L3 J23102 L4 ATG-GFP-B0015	5	1087 ±43
L1RBS1	4	L6 Kan-MB1 L2 Amp L3 J23102 L1RBS1 ATG-GFP-ORF	639	1633 ±108
L1RBS2	4	L6 Kan-MB1 L2 Amp L3 J23102 L1RBS2 ATG-GFP-ORF	5380	14247 ±1345
L1RBS3	4	L6 Kan-MB1 L2 Amp L3 J23102 L1RBS3 ATG-GFP-ORF	16573	20939 ±1575
L1RBS4	4	L6 Kan-MB1 L2 Amp L3 J23102 L1RBS4 ATG-GFP-ORF	57705	5230 ±387
L2RBS1	4	L6 Kan-MB1 L2 Amp L3 J23102 L2RBS1 ATG-GFP-ORF	1437	1622 ±78
L2RBS2	4	L6 Kan-MB1 L2 Amp L3 J23102 L2RBS2 ATG-GFP-ORF	10567	12953 ±842
L2RBS3	4	L6 Kan-MB1 L2 Amp L3 J23102 L2RBS3 ATG-GFP-ORF	22711	26074 ±1506
L2RBS4	4	L6 Kan-MB1 L2 Amp L3 J23102 L2RBS4 ATG-GFP-ORF	141945	3424 ±152
L1RBS1&3	4	L6 Kan-MB1 L2 Amp L3 J23102 L1RBS1 or L1RBS3 ATG-GFP-ORF	n/a	n/a
L1RBS1-4	4	L6 Kan-MB1 L2 Amp L3 J23102 L1RBS1 or L1RBS2 or L1RBS3 or L1RBS4 ATG-GFP-ORF	n/a	n/a

Supplementary Table 15: Constructs assembled for RBS linker assay. Linkers are shown in blue and are as described in Supplementary Tables 3 and 8; parts are as described in Supplementary Table 2. Predicted expression levels were obtained from the RBS calculator and observed expression were from fluorescence/OD_{600} after 3 hours from 8 samples with SEM.

Cost analysis

BASIC assembly only requires commonly used laboratory reagents. Most of the cost comes from the enzymes used in the digestion/ligation reaction that attaches linkers to DNA parts and from the magnetic beads used during purification (Supplementary Table 16).

To ensure that oligonucleotide quality did not affect assembly efficiency, all the linker and adaptor oligonucleotides used for BASIC assembly were synthesised by IDT at 100 nM scale, and supplied HPLC purified and phosphorylated at the 5' end. C-5 methylation was also incorporated during oligonucleotide synthesis but adds only a minor cost. This routinely provided enough DNA for well in excess of 2000 reactions. While the upfront cost of high quality synthesis is high it scales considerably so that the cost per reaction is very low (Supplementary Table 16).

Reagent	Cost (GBP)	Volume (µl)	# reactions	Cost per part
BsaI-HF	169.60	250	250	0.68
T4 DNA ligase	166.40	250	500	0.33
AMPure XP beads	689.67	60000	1111	0.62
Oligonucleotides	182.42	28000	2800	0.07
			Total per part:	**1.70**

Supplementary Table 16: Summary of BASIC's cost for each part being assembled.

1. **Casini, A., Christodoulou, G., Freemont, P. S., Baldwin, G. S., Ellis, T., and Macdonald, J. T. (2014) R2oDNA Designer: Computational Design of Biologically Neutral Synthetic DNA Sequences, *ACS synthetic biology 10.1021/sb4001323*.**