SYNTHETIC BIOLOGY

Synthetic Biology, 2022, **7(1)**, 1–8
DOI: https://doi.org/10.1093/synbio/ysac023
Advance access publication date: 11 October 2022
Software

basicsynbio and the BASIC SEVA collection: software and vectors for an established DNA assembly method

Matthew C. Haines[1,2,*], Benedict Carling[3], James Marshall[3], Vasily A. Shenshin[4], Geoff S. Baldwin[4,5], Paul Freemont[1,2,6,*], and Marko Storch[1,2,*]

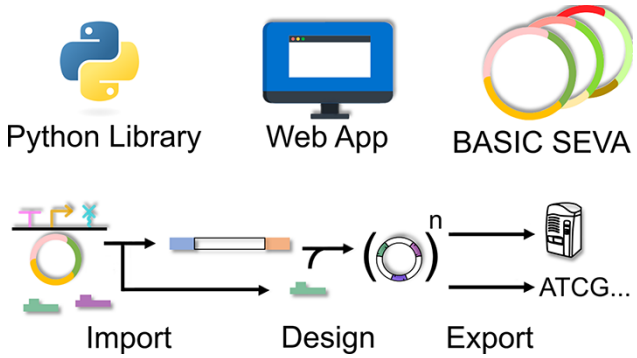
[1]Department of Infectious Disease, Sir Alexander Fleming Building, South Kensington Campus, Imperial College London, London SW7 2AZ, UK
[2]London Biofoundry, Imperial College Translation and Innovation Hub, London W12 0BZ, UK
[3]Department of Bioengineering, Imperial College London, London, Westminster SW7 2AZ, UK
[4]Department of Life Sciences, Imperial College London, London, Westminster SW7 2AZ, UK
[5]Imperial College Centre for Synthetic Biology, Imperial College London, London SW7 2AZ, UK
[6]UK DRI Care Research and Technology Centre, Imperial College London, Hammersmith Campus, Du Cane Road, London W12 0NN, UK
*Corresponding authors: E-mails: matthew.haines10@imperial.ac.uk; p.freemont@imperial.ac.uk and m.storch@imperial.ac.uk

Abstract

Standardized deoxyribonucleic acid (DNA) assembly methods utilizing modular components provide a powerful framework to explore designs and iterate through Design–Build–Test–Learn cycles. Biopart Assembly Standard for Idempotent Cloning (BASIC) DNA assembly uses modular parts and linkers, is highly accurate, easy to automate, free for academic and commercial use and enables hierarchical assemblies through an idempotent format. These features enable applications including pathway engineering, ribosome binding site (RBS) tuning, fusion protein engineering and multiplexed guide ribonucleic acid (RNA) expression. In this work, we present basicsynbio, open-source software encompassing a Web App (https://basicsynbio.web.app/) and Python Package (https://github.com/LondonBiofoundry/basicsynbio), enabling BASIC construct design via simple drag-and-drop operations or programmatically. With basicsynbio, users can access commonly used BASIC parts and linkers while designing new parts and assemblies with exception handling for common errors. Users can export sequence data and create instructions for manual or acoustic liquid-handling platforms. Instruction generation relies on the BasicBuild Open Standard, which is parsed for bespoke workflows and is serializable in JavaScript Object Notation for transfer and storage. We demonstrate basicsynbio, assembling 30 vectors using sequences including modules from the Standard European Vector Architecture (SEVA). The BASIC SEVA vector collection is compatible with BASIC and Golden Gate using BsaI. Vectors contain one of six antibiotic resistance markers and five origins of replication from different compatibility groups. The collection is available via Addgene under an OpenMTA agreement. Furthermore, vector sequences are available from within the basicsynbio application programming interface with other collections of parts and linkers, providing a powerful environment for designing assemblies for bioengineering applications.

Key words: bioinformatics; DNA assembly; synthetic biology; vectors; cloning

Graphical Abstract



1. Introduction

DNA assembly is an essential tool in Synthetic Biology and Life Sciences, required for building genetic designs and iterating through the Design-Build-Test-Learn cycle (1, 2). A large repertoire of DNA assembly methods are available to researchers, and the choice of a suitable method will depend on factors such as the freedom to

Submitted: 3 April 2022; **Received (in revised form):** 6 September 2022; **Accepted:** 10 October 2022

include forbidden restriction sites, the availability of part libraries or a need for high accuracy (2, 3). Standardized and modular DNA assembly methods are ideal for high-throughput and hierarchical assemblies enabling the cost-effective generation of large numbers of constructs with high accuracy while encouraging the reuse of parts across designs (2, 4–7).

Biopart Assembly Standard for Idempotent Cloning (BASIC) DNA assembly is a standardized DNA assembly method, which utilizes modular parts and linkers as functional units (7–10). The method benefits from several desirable attributes including a single-part storage format and assembling up to 14 parts and linkers per round with >90% accuracy (7). Given that linkers can encode functional sequences such as RBSs and fusion protein linkers, diverse constructs are feasible within a single round of assembly. It is also free for academic and commercial use and only requires the absence of one restriction enzyme site (BsaI). It is easy to automate the physical workflow (10) and conduct hierarchical assemblies since parts are stored in a single format and assemblies are ubiquitously returned with flanking sequences reconstituting this format, enabled by the underlying single-tier, idempotent architecture.

BASIC compares favorably with modular methods based on Golden Gate assembly (5, 6), where multiple restriction enzymes are utilized and assemblies not conforming to standard transcriptional units, e.g. operons, are not supported. Notably, BASIC DNA assembly was successfully applied to several areas of Synthetic Biology and Life Sciences research including combinatorial pathway engineering (3, 11, 12), synthetic operon (10) and small non-coding RNA circuit design (13), combinatorial guide RNA expression for gene editing (14), ribosome binding site (RBS) tuning and fusion protein engineering (7).

In this work, we developed basicsynbio design software with several aims. First, we make DNA sequences for commonly used parts and linkers accessible from a single source, making them easier to access and maintain. Second, we introduce exception handling, reducing failure rates caused by design errors. Third, we enable users to export a variety of data types for downstream building, validating and sharing of assemblies. In particular, we provide users with a data standard describing multiple assemblies, which is easily parsed into custom workflows, enabling the automation of BASIC DNA assembly on further liquid-handling platforms and the generation of instructions for manual workflows. These aims are achieved in the context of the basicsynbio Python Package or Web App, which facilitate the programmatic generation of large numbers of constructs and their sequence data or provide a user-friendly drag-and-drop interface, respectively. This extends our previous work DNA-BOT (10), which automated BASIC DNA assembly specifically for the Opentrons platform. We demonstrate basicsynbio by designing and exporting data for a collection of 30 vectors containing several modules from the Standard European Vector Architecture (SEVA) database (15, 16). We subsequently build and deposit the collection on Addgene and make the sequences accessible via the basicsynbio application programming interface (API), enabling access for BASIC DNA assembly users and the community.

2. Materials and methods

2.1 Preparation of BASIC linkers and parts

Apart from BSEVA_L1, all BASIC linkers were acquired from Biolegio (BBT-18500) and prepped according to the manufacturer's instructions. Oligos for BSEVA_L1 (Supplementary Table S1) were ordered from Integrated DNA Technologies, Inc., and linker halves were prepared as previously described (9).

Unless specified, all plasmid DNA was prepped using Omega BIO-TEK E.Z.N.A.® Plasmid Mini Kit II according to the manufacturer's instructions. All plasmid DNA was quantified using a Qubit™ dsDNA BR Assay Kit (Thermo Scientific™ Q32850).

Each BASIC SEVA vector is composed of three parts: T0 + marker part, ori + T1 part and mScarlet counter-selection cassette. Initially, each was either chemically synthesized or amplified from SEVA vectors (16) with primers incorporating *iP* and *iS* sequences upstream and downstream, respectively. The resulting linear sequences were cloned into an appropriate vector, prior to prepping plasmid DNA. Specifically, T0 + marker parts were blunt cloned into pJET1.2 (Thermo Scientific™ K1231) according to the manufacturer's instructions. ori + T1 and mScarlet counter-selection cassette parts were assembled as described in the Supplementary Data (oris.gb and addgene_submission_notebook.pdf) using BASIC DNA assembly (8). Constructs were plated on LB-agar (Formedium) supplemented with 100 μg/ml carbenicillin and incubated at 30 or 37°C prior to picking colonies and prepping plasmid DNA. All parts were sequence verified via Sanger Sequencing prior to assembly.

2.2 Assembly and validation of the BASIC SEVA collection

All assemblies were designed *in silico* using basicsynbio (Supplementary Data: addgene_submission_notebook.pdf). Echo instructions for the 'Assembly reaction' step of the workflow and manual instructions for the entire workflow were exported (see the Supplementary Data).

Clip reaction and MagBead purification steps were implemented as described in the manual instructions (Supplementary Data: BASIC_SEVA_collection_v10_manual.pdf), transferring purified clip reactions to an Echo® Qualified 384-Well Polypropylene Microplate (Beckman 001-14555). Purified clip reactions were mixed by executing the echo_clips_1.csv script (Supplementary Data) on a Beckman Echo 525 Acoustic Liquid Handler, using a 96-well destination plate (Azenta Life Sciences 4ti-0960). Double distilled water (ddH_2O) and 10× assembly buffer solutions were transferred to the same destination plate by executing the echo_water_buffer_1.csv script with both solutions transferred from an Echo® Qualified Reservoir, 2 × 3 Well, Polypropylene Microplate (Beckman 001-11101). The destination plate containing assemblies was sealed with a polymerase chain reaction (PCR) foil seal (Azenta Life Sciences 4ti-0550), vortexed and centrifuged prior to incubating at 50°C for 45 min. Twenty-five microliters of NEB® 5-alpha Competent *Escherichia coli* (*E. coli*) cells (C2987) were added to each assembly reaction on ice. Transformation reaction mixtures were incubated for 20 min on ice, followed by heat shock at 42°C for 15 sec, recovery on ice for 2 min, the addition of 150 μl of SOC media (Formedium) and outgrowth at 30°C for 2 hr. Cells were plated on LB agar containing antibiotics concentrations illustrated in Figure 4b. Plasmid DNA from pink colonies was prepped as described above for corresponding ori + T1 parts.

Prior to Addgene submission, we verified the presence of the correct ori + T1 part, sequencing each vector using the BSEVA_L1_overhang sequencing primer (Supplementary Table S1). The resulting data were analyzed using cMatch (17) to verify sequence identity.

3. Results and discussion

3.1 basicsynbio workflow

We conceived a typical workflow for users implementing basicsynbio (Figure 1a). Initially, users would access collections of parts

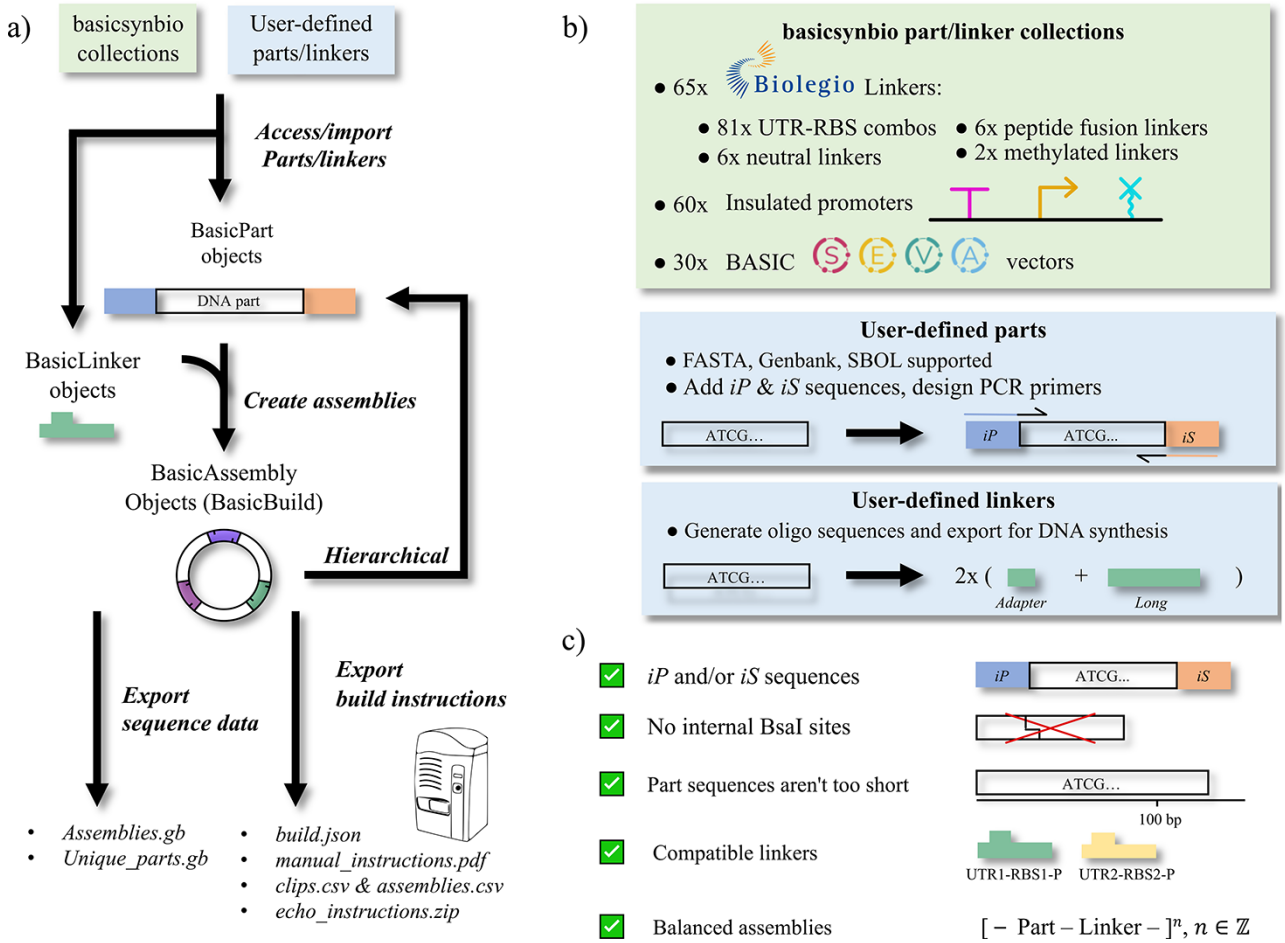


Figure 1. basicsynbio workflow and functionality: (a) typical basicsynbio workflow: BasicPart and BasicLinker objects are imported from internal collections or user-defined sources. Combinations of BasicParts and BasicLinkers initiate BasicAssembly objects, describing assemblies. Sequence data and build instructions are exported for reference and to aid downstream workflows, respectively. Multiple data types are exportable, including those for DNA assembly using an acoustic liquid handler. (b) Options for importing parts and linkers. basicsynbio part and linker collections contain >150 sequences compatible with BASIC DNA assembly, accessible from within the Python Package and Web App. Numerous biological sequence formats are supported for user-defined parts. When creating parts, users can add *iP* and *iS* sequences and/or design necessary PCR primers. Users can define linkers using the API and export required adapter and long oligonucleotide sequences for prefix and suffix linker halves. (c) Error handling in basicsynbio. Objects are checked for common errors that could lead to failure during assembly. For instance, using linker halves incompatible with each other in the same assembly raises an exception, as do unbalanced assemblies.

and linkers available from the basicsynbio API, in addition to importing their own. These BasicPart and BasicLinker objects are combined initiating BasicAssembly objects representing assembled constructs. A key advantage of BASIC DNA assembly is its idempotency, meaning that sequences in assemblies flanked by LMP and LMS linkers are themselves BasicParts and can function in subsequently larger constructs. basicsynbio facilitates this, enabling users to convert BasicAssembly objects into BasicParts, ready to initiate next-tier, larger BasicAssembly objects. Once the user has specified all the desired BasicAssembly objects, various data types are available to export (Figure 1a and Supplementary Figure S1). Users can export sequence data representing BasicAssembly and BasicPart objects in GenBank via the Web App or in formats supported by Biopython (18) via the Python Package. Notably, all features are preserved, maintaining annotations in the resulting assemblies. In addition to exporting sequence data, users can export build instructions, for instance, instructions for manual or automated workflows, e.g. pdf instructions for manual workflows or csv files to program a Beckman Echo robot.

To aid accessibility of existing core BASIC DNA assembly part and linker sequences, we include PartLinkerCollection objects, accessible from the API and containing instances of commonly used BasicParts and BasicLinkers. Notable collections are illustrated in Figure 1b, including BASIC_BIOLEGIO_LINKERS, BASIC_PROMOTER_PARTS and BASIC_SEVA_PARTS, which contain all 65 commercially available Biolegio linkers, including linkers for 81 different untranslated region (UTR)/RBS combinations, a collection of 60 inducible and constitutive promoters, insulated by different combinations of upstream terminators and downstream RiboJ sequences (19) and a collection of 30 vectors containing several SEVA modules (16), respectively. To aid the exploration of PartLinkerCollections, users can visualize individual parts and linkers via the Web App using SeqViz or DNA Features Viewer (20) (Supplementary Figures S2 and S3). Different versions of a given PartLinkerCollection are supported, enabling future updates where required. Furthermore, users can contribute new PartLinkerCollections as described in the online documentation (https://londonbiofoundry.github.io/basicsynbio/). We hope that this will

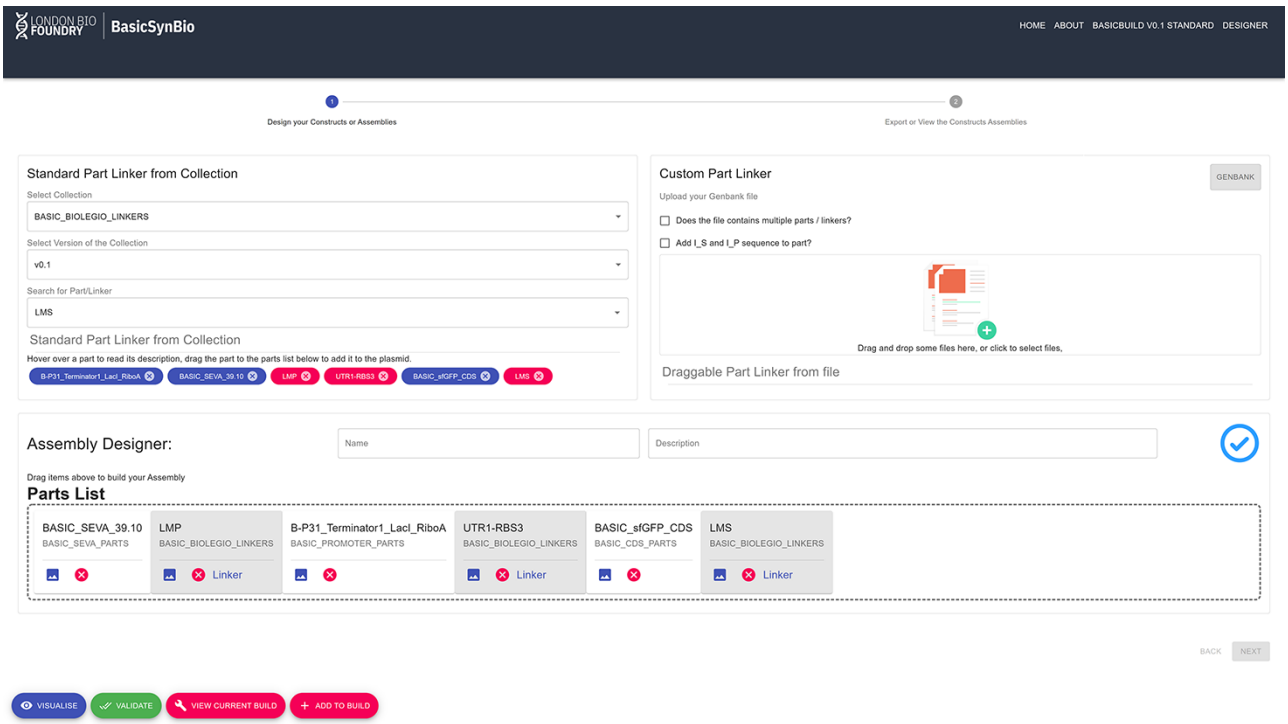


Figure 2. Screenshot of basicsynbio Web App designer. In this example, a construct expressing sfGFP under an Isopropyl β-d-1-thiogalactopyranoside-inducible, insulated promoter is created in the Assembly Designer. BASIC_SEVA_39.10, BP31_Terminator_LacI_RiboA and BASIC_sfGFP_CDS BasicPart objects were selected from BASIC_SEVA_PARTS, BASIC_PROMOTER_PARTS and BASIC_CDS_PARTS collections, respectively. These are combined with LMP, UTR1-RBS3 and LMS linkers from the BASIC_BIOLEGIO_LINKERS collection. The assembly is checked against errors (check mark) and can be visualized prior to adding to a BasicBuild (bottom left).

encourage the BASIC DNA assembly user community to share collections of new part and linker sequences for different applications between laboratories and institutions.

In addition to the above PartLinkerCollections, users can import parts from local files or external sources and/or create new linkers using the API, greatly expanding the number of possible assemblies. Users may import parts specified in commonly used file formats such as FASTA, GenBank and SBOL (Figures 1b and 2). Furthermore, to aid the generation of new parts, users can automatically add required *iP* and *iS* sequences (7) to the 5′ and 3′ ends of input DNA sequences, respectively. It is also possible to design primers to add *iP* and *iS* sequences to parts via overhang PCR with the aid of Primer3 (21). This enables cost-effective conversion of existing DNA sequences into a physical part, avoiding the need for *de novo* DNA synthesis. For a given linker, users can calculate the four oligonucleotide sequences required to generate linker halves, an adapter and long oligonucleotide for each linker half (7, 9) (Figure 1b). This feature aids the generation of custom linkers required for specific applications or organisms.

For the successful implementation of BASIC DNA assembly, imported parts and designed assemblies must satisfy several conditions (Figure 1c). For instance, if the length of a part is significantly shorter than 100 bp, the linker-ligated part would be lost during the purification step of assembly. Additionally, internal BsaI sites are not allowed in BASIC parts and specific linkers can only be used once per assembly round, while BasicPart and BasicLinker objects must alternate, with equal numbers of each. Where a user designs an assembly that does not satisfy the above conditions, basicsynbio raises exceptions preventing subsequent experimental failure, increasing robustness.

To implement the basicsynbio workflow illustrated in Figure 1a, users can utilize the open-source Python Package or Web App. Python iterator patterns combined with the basicsynbio package allow users to initiate large numbers of BasicAssembly objects programmatically, facilitating the exploration of large design spaces with 100s of constructs feasible with BASIC DNA assembly (10). Meanwhile, the designer interface of the Web App (Figure 2) offers users an intuitive way to create BasicAssembly objects by dragging and dropping selected BasicPart and BasicLinker objects. In addition to visualizing parts using the Web App (Supplementary Figures S2 and S3), users can dynamically visualize assemblies to ensure that they contain the desired sequence prior to implementing the checks illustrated in Figure 1c.

3.2 BasicBuild Open Standard

Following design, a user builds their collection of assemblies. To determine build instructions, a user has to make multiple calculations (10). First, the user calculates the unique set of clip reactions required by all assemblies. Each clip reaction is defined by a BasicPart in combination with BasicLinker prefix and suffix halves. Second, the user needs to associate each unique clip reaction with the assemblies requiring it. From this, the user calculates the absolute number of each clip given that each can support 15–30 assemblies, depending on the workflow. Third, the user makes calculations ensuring a final part concentration of 2.5 nM following the clip reaction setup, maximizing efficiency. These three parameters guide liquid-handling operations during clip reaction setup and assembly stages of the BASIC workflow. Previously, we implemented this for a specific liquid-handling platform (10), and in

```
"unique_parts": {
    "UP0": {
        "sequence": "GGTAAGAACTCGCACTTCGTGGAAACACTATTATCTG…",
        "id": "bb90c9b9713d85ff695493fde3f0051a",
        "name": "BASIC_SEVA_19",
        "description": "BASIC SEVA vector containing Ampic…",
        "part mass per 30 µL clip reaction (ng)": 155,
        "clip reactions": [
            "CR0"
        ]
    },…
}

"unique_linkers": {
    "UL0": {
        "id": "1b62864ef7e8ad38d30e92a2dbc17670",
        "linker_class": "<class 'basicsynbio.main.BasicLin…",
        "sequence": "GGCTCGGGAGACCTATCGGTAATAACAGTCCAATCTG…",
        "prefix_id": "LMS-P",
        "suffix_id": "LMS-S",
        "overhang_slice_params": null,
        "prefix clip reactions": [
            "CR0"
        ],
        "suffix clip reactions": [
            "CR1",
            "CR2",
            "CR3"
        ]
    },…
}

"clips_data": {
    "CR0": {
        "prefix": {
            "key": "UL0",
            "prefix_id": "LMS-P"
        },
        "part": {
            "key": "UP0",
            "id": "bb90c9b9713d85ff695493fde3f0051a",
            "name": "BASIC_SEVA_18"
        },
        "suffix": {
            "key": "UL1",
            "suffix_id": "LMP-S"
        },
        "total assemblies": 3,
        "assembly keys": [
            "A0",
            "A1",
            "A2"
        ]
    },…
}

"assembly_data": {
    "A0": {
        "id": "sfGFP",
        "clip reactions": [
            "CR0",
            "CR1"
        ]
    },…
},
```


Figure 3. BasicBuild Open Standard: key objects from a BasicBuild example, serialized in JSON: 'unique_parts', 'unique_linkers', 'clips_data' and 'assembly_data'. For clarity, values have been shortened and similar objects replaced with '…'. Using the BasicBuild Open Standard, the generation of functions for custom workflows is simplified as demonstrated in this work for manual and acoustic liquid-handler instructions. A specification of the BasicBuild Open Standard v0.1 is available online (https://basicsynbio.web.app/basicbuild-standard).

this work, we describe a standard enabling bespoke manual and automated workflows.

The BasicBuild Open Standard is a data structure given in JavaScript Object Notation (JSON), which contains the information outlined above. The most uptodate definition for this standard is available online at https://basicsynbio.web.app/basicbuild-standard. Figure 3 illustrates the four nested objects from an example. The clips_data object contains data on each clip reaction required for the build. Further information on clip reaction parts and linkers is available within unique_parts and unique_linkers objects, where the corresponding key can be used to access this information, e.g. 'UP0' to access the first part within unique_parts. A link between each clip reaction and the assemblies using it is established by both the 'assembly keys' attribute in the clips_data object and the 'clips reactions' attribute of the assembly data object. The absolute number of each clip reaction can be calculated using the clips_data 'total assemblies' attribute, considering the number of assemblies supplied by each purified clip reaction (15–30 depending on the method). To aid the addition of parts to a final concentration of 2.5 nM in clip reactions, a 'part mass for 30 µl clip reaction (ng)' attribute is provided, where the addition of the associated mass to a 30 µl clip reaction results in the desired final concentration. Further to providing the parameters required to build assemblies, the BasicBuild Open Standard includes additional data, e.g. a unique ID for each assembly. We implement the BasicBuild Open Standard within the basicsynbio API as a python class. To demonstrate the standard's flexibility, we have written functions that parse BasicBuild class objects into manual instructions in pdf format and instructions for a Beckman Echo liquid-handling platform.

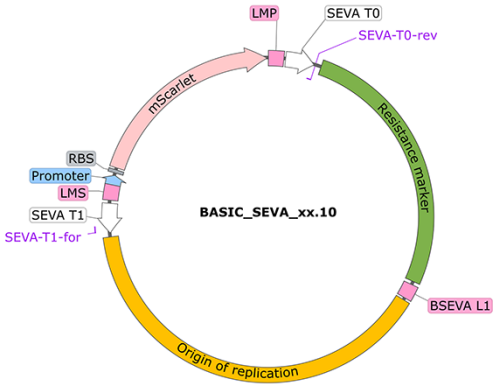
It is also worth noting that the basicsynbio API can decode BasicBuilds serialized in JSON creating a class object. As such, designs serialized in one location can be transferred securely to the location of manufacturing, decoded and processed into build instructions specific to the facility. We envision that this will allow designers to work agnostically of the protocol or facility used for building, freeing them to focus on other steps of the Design-Build-Test-Learn cycle.

3.3 BASIC SEVA collection

To demonstrate basicsynbio, we designed a collection of vectors illustrated in Figure 4a. Each contains a specific combination of antibiotic resistance marker and origin of replication (ori) flanked by SEVA T0 and T1 terminators. The terminators prevent transcriptional readthrough, maintaining plasmid stability in an equivalent manner to vectors from the SEVA database (15). To enable these vectors to function in BASIC DNA assembly and Golden Gate workflows using BsaI, we flank the terminator, resistance marker and ori components with LMP and LMS linkers. In contrast to vectors from the SEVA database, we omit the origin of transfer (OriT) sequence from our collection. This reduces the risk of unintended transfer to natural microorganisms, and if desired, users can incorporate this sequence when generating constructs for specific applications.

The marker and ori are joined with a neutral BASIC linker sequence (BSEVA L1), enabling combinatorial assembly of the collection using BASIC. This linker sequence was designed using DNA Chisel (22) to retain maximum plasmid stability. Specifically, sequences that could lead to expressions such as promoters and RBSs were avoided and complementarity to the *E. coli* MG1655 genome or existing BASIC parts and linkers was minimized, reducing the propensity for recombination in strains with active homologous recombination machinery (23). Additional bespoke linkers would benefit diverse applications in the future. As such, future work could focus on incorporating relevant

a)



b)

BASIC_SEVA_ x x .<version>

ID	Marker	µg/mL
1	Ampicillin	100[1]
2	Kanamycin	50
3	Chloramphenicol	25
4	Spectinomycin/ Streptomycin	100[2]
5a	Tetracycline	5
6	Gentamicin	10

ID	Ori
5a	RSF1010
6	p15A
7	pSC101
7_pKD46	pSC101 (pKD46)
9	pBR322-ROP

Figure 4. BASIC SEVA vector collection: (a) plasmid map of the collection. BASIC linkers (LMP, LMS & BSEVA_L1) and SEVA terminators (T0 & T1) are indicated. Origins of replication and antibiotic resistance markers are joined via a neutral BSEVA L1 linker. An mScarlet counter-selection cassette, compromising promoter, RBS and mScarlet CDS, is flanked by methylated LMP and LMS linkers, resulting in drop-out during assembly. SEVA-T0-rev and SEVA-T1-for sequencing primer binding sites are indicated. (b) Nomenclature of the BASIC SEVA collection. Vectors are named through the combination of two strings that reflect the origin of replication and resistance marker identities. Where an integer value is given, the associated module is identical to that provided by SEVA (16). The remaining modules are discussed in the text. The nomenclature also provides a versioning system, equivalent to that used by GenBank (https://www.ncbi.nlm.nih.gov/genbank/release/). Antibiotic concentrations used to isolate plasmids in this study are given ([1]carbenicillin was used instead of ampicillin; [2]streptomycin was not tested in this study but has previously been demonstrated (31)).

DNA Chisel functionality into the basicsynbio framework, making linker design more seamless.

Since the sequence outside the LMP/S flanked region is lost during the assembly of downstream constructs using these vectors, we incorporated an mScarlet counter-selection cassette within this region. Thus, colonies containing these vectors display a pink phenotype allowing for visual counterselection against assembly background from the undigested, original vector.

The BASIC SEVA collection generated in this work encompasses 30 vectors, and every combination of six markers and five oris is described in Figure 4b. The six markers correspond with the first six markers used by the SEVA collection. Apart from the tetracycline module, all marker modules are identical in sequence to SEVA modules. For reasons outlined in the Supplementary Data (BASIC SEVA modules, Supplementary Figures S4 and S5), we use an alternative tetracycline sequence but note its difference from that used by SEVA by assigning a '5a' ID as opposed to '5'. We selected four oris from the SEVA database to include in our collection with three identicals in sequence (6, 7, 9). The fourth (ori 5a) is homologous to the SEVA RSF1010 module with the sequence previously reported in the literature (24).

To assemble the collection, we split the design in Figure 4a into three BASIC parts: T0 + marker part, ori + T1 part and mScarlet counter-selection cassette. Prior to assembling the collection, we cloned, prepped and sequence verified each part (Section 2). We subsequently assembled each vector of the collection *in silico* using basicsynbio (Supplementary Data: addgene_submission_notebook.pdf), naming each vector according to the BASIC SEVA nomenclature (Figure 4b). Using the resulting BasicBuild, we exported manual instructions for the entire assembly and Echo liquid-handling instructions for the 'Assembly reaction' step of the workflow, aiding building (Section 2). Following assembly and transformation, we selected colonies containing the antibiotic and counter-selection cassette using relevant antibiotics (Figure 4b) and by picking pink colonies, respectively.

Sequencing the collection, we identified deletion or insertion mutations at the border of the gentamicin coding sequence (CDS) N-terminus and 5′-UTR for plasmids containing p15A and pBR322 oris (BASIC_SEVA_66.10 and 69.10). We isolated plasmid DNA from a further three colonies for each design and observed similar mutations (Supplementary Figure S6). In contrast to pSC101 and RSF1010, both p15A and pBR322 belong to the ColE1 class of plasmids (25) where the copy number is regulated by RNA transcripts (26). It is conceivable that the SEVA gentamicin cassette sequence interferes in this process although this remains to be determined.

To provide users with plasmids conferring gentamicin resistance and containing p15A and pBR322 oris, we selected constructs from those sequenced having identical mutations and designate them BASIC_SEVA_66.11 and 69.11 (Supplementary Figure S6). A search revealed that this mutation in the gentamicin cassette has previously been reported (AJ247370.1). Sequence data describing the entire collection were exported with the assistance of basicsynbio and used to generate a basicsynbio PartLinkerCollection, making the BASIC SEVA collection accessible to other basicsynbio users via the API.

The five ori modules used to build the collection enable a variety of applications. Notably, we include a temperature-sensitive ori (7_pKD46), not present in the SEVA database. This ori is identical in sequence to that previously described (27) and is unstable at 37°C, requiring growth at 30°C. Plasmids harboring this ori are ideal for applications where plasmid curing is a requirement, e.g. strain engineering (27–29). The three SEVA oris used (6, 7 and 9) are from three different incompatibility groups (30), enabling applications requiring multiple plasmid types in the same host. Furthermore, these three oris provide a range of copy numbers to tune gene expression. As previously discussed (15), the remaining ori (RSF1010) is compatible with a broad range of hosts/chasses, enabling applications suited to non-model organisms. While a high copy number ori was not included in the collection (Supplementary Data—BASIC SEVA modules), plasmids containing pBR322 can be amplified with the addition of chloramphenicol, increasing plasmid yield (30). Additionally, we observed a relatively higher yield for vector BASIC_SEVA_39.10, which contains both pBR322 ori and a chloramphenicol marker (data not shown), suggesting that this vector is suitable for applications requiring higher yields of plasmid DNA.

In conclusion, with the basicsynbio Python Package and Web App, users can access commonly used parts and linkers, robustly design new parts, linkers and assemblies while exporting sequence data and build instructions. We also outline

the BasicBuild Open Standard, enabling the facile generation of custom build instructions as demonstrated in this work for manual workflows and workflows using acoustic liquid handlers. To demonstrate basicsynbio, we design and assemble a collection of 30 vectors using modules from the SEVA database. Sequence data for this collection are available for users via the basicsynbio API, while plasmids were deposited on Addgene. In combination with other accessible parts and linkers, users can easily and robustly design a large repertoire of assemblies, enabling applications in Synthetic Biology and the Life Sciences.

Supplementary data

Supplementary data are available at SYNBIO Online.

Data availability

Additional data for the basicsynbio Python Package and Web App are available at https://github.com/LondonBiofoundry/basicsynbio and https://basicsynbio.web.app/, respectively.

Material availability statement

BASIC SEVA plasmids were deposited on Addgene (Deposit 80391, Addgene ID 178913–178942 inc.) and made available under Addgene's OpenMTA agreement.

Funding

P.F. and M.C.H. were supported by the UKRI Engineering and Physical Sciences Research Council (EP/T013788/1).

Acknowledgments

We would like to thank Alexis Casas for discussion on software and the BasicBuild Open Standard.

Author contributions

M.C.H. participated in conceptualization, investigation, software, visualization, writing—original draft, writing—review and editing and project administration. B.C. contributed to software, writing—review and editing and visualization. J.M. participated in investigation and writing—review and editing. V.A.S. participated in investigation and writing—review and editing. G.S.B. participated in conceptualization, supervision and writing—review and editing. M.S. participated in conceptualization, investigation, supervision, project administration and writing—review and editing. P.F. participated in funding acquisition, supervision and writing—review and editing.

Conflict of interest statement. P.F. sits on the SAB of Tierra Biosciences and is Chair at Solena Materials.

References

1. Freemont,P.S. (2019) Synthetic biology industry: data-driven design is creating new opportunities in biotechnology. *Emerg. Top Life Sci.*, **3**, 651–657.
2. Casini,A., Storch,M., Baldwin,G.S. and Ellis,T. (2015) Bricks and blueprints: methods and standards for DNA assembly. *Nat. Rev. Mol. Cell Biol.*, **16**, 568–576.
3. Young,R., Haines,M., Storch,M. and Freemont,P.S. (2021) Combinatorial metabolic pathway assembly approaches and toolkits for modular assembly. *Metab. Eng.*, **63**, 81–101.
4. Deb,S.S. and Reshamwala,S.M.S. (2020) Recent progress in DNA parts standardization and characterization. In Singh Vijai (ed). *Advances in Synthetic Biology* . Springer, Singapore, pp. 43–69.
5. 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.
6. Sarrion-Perdigones,A., Falconi,E.E., Zandalinas,S.I., Juárez,P., Fernández-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.
7. Storch,M., Casini,A., Mackrow,B., Fleming,T., Trewhitt,H., Ellis,T. and Baldwin,G.S. (2015) BASIC: a new biopart assembly standard for idempotent cloning provides accurate, single-tier DNA assembly for synthetic biology. *ACS Synth. Biol.*, **4**, 781–787.
8. Storch,M., Dwijayanti,A., Mallick,H., Haines,M.C. and Baldwin,G.S. (2020) BASIC: a simple and accurate modular DNA assembly method. *Methods Mol. Biol.*, **2205**, 239–253.
9. Storch,M., Casini,A., Mackrow,B., Ellis,T. and Baldwin,G.S. (2017) BASIC: A simple and accurate modular DNA assembly method. In: Hughes Randall A (ed). *Methods in Molecular Biology*, Vol. 1472, Humana Press Inc., London, UK, pp. 79–91.
10. Storch,M., Haines,M.C. and Baldwin,G.S. (2020) DNA-BOT: a low-cost, automated DNA assembly platform for synthetic biology. *Synth. Biol.*, **5**, ysaa010.
11. Bartasun,P., Prandi,N., Storch,M., Aknin,Y., Bennett,M., Palma,A., Baldwin,G., Sakuragi,Y., Jones,P.R. and Rowland,J. (2019) The effect of modulating the quantity of enzymes in a model ethanol pathway on metabolic flux in *Synechocystis* sp. PCC 6803. *Peer J.*, **7**, e7529.
12. Casas,A., Bultelle,M., Motraghi,C. and Kitney,R. (2022) PASIV: A pooled approach-based workflow to overcome toxicity-induced design of experiments failures and inefficiencies. *ACS Synth. Biol.*, 11, 1272–1291.
13. Dwijayanti,A., Storch,M., Stan,G.-B. and Baldwin,G.S. (2022) A modular RNA interference system for multiplexed gene regulation. *Nucleic Acids Res.*, **50**, 1783–1793.
14. Benns,H.J., Storch,M., Fisher,F.R., Alves,E., Wincott,C.J., Baum,J., Baldwin,G.S., Weerapana,E., Tate,E.W. and Child,M.A. (2021) Prioritization of antimicrobial targets by CRISPR-based oligo recombineering 1 2. *bioRxiv*.
15. Silva-Rocha,R., Martínez-García,E., Calles,B., Chavarría,M., Arce-Rodríguez,A., de Las Heras,A., Páez-Espino,A.D., Durante-Rodríguez,G., Kim,J., Nikel,P.I. *et al.* (2013) The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **41**, D666–D675.
16. Martínez-García,E., Goñi-Moreno,A., Bartley,B., McLaughlin,J., Sánchez-Sampedro,L., Pascual Del Pozo,H., Prieto Hernández,C., Marletta,A.S., De Lucrezia,D., Sánchez-Fernández,G. *et al.* (2020) SEVA 3.0: an update of the Standard European Vector Architecture for enabling portability of genetic constructs among diverse bacterial hosts. *Nucleic Acids Res.* **48**, D1164–D1170.
17. Casas,A., Bultelle,M., Motraghi,C. and Kitney,R.I. (2022) Removing the Bottleneck: Introducing cMatch - A Lightweight Tool for Construct-Matching in Synthetic Biology. *Front. Bioeng. Biotechnol.*, **9**, 785131. 10.3389/fbioe.2021.785131.
18. Cock,P.J.A., Antao,T., Chang,J.T., Chapman,B.A., Cox,C.J., Dalke,A., Friedberg,I., Hamelryck,T., Kauff,F., Wilczynski,B. *et al.* (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, **25**, 1422–1423.
19. Dwijayanti,A. (2019) Engineering standardised and modular biological controllers for efficient design and easy implementation

in synthetic genetic circuits. *PhD thesis*, Imperial College London. 10.25560/82476.
20. Zulkower,V. and Rosser,S. (2020) DNA Features Viewer: a sequence annotation formatting and plotting library for Python. *Bioinformatics*, **36**, 4350–4352.
21. Untergasser,A., Cutcutache,I., Koressaar,T., Ye,J., Faircloth,B.C., Remm,M. and Rozen,S.G. (2012) Primer3—new capabilities and interfaces. *Nucleic Acids Res.*, **40**, e115.
22. Zulkower,V. and Rosser,S. (2020) DNA Chisel, a versatile sequence optimizer. *Bioinformatics*, **36**, 4508–4509.
23. Sleight,S.C., Bartley,B.A., Lieviant,J.A. and Sauro,H.M. (2010) Designing and engineering evolutionary robust genetic circuits. *J. Biol. Eng.*, **4**, 12.
24. Cook,T.B., Rand,J.M., Nurani,W., Courtney,D.K., Liu,S.A. and Pfleger,B.F. (2018) Genetic tools for reliable gene expression and recombineering in *Pseudomonas putida*. *J. Ind. Microbiol. Biotechnol.*, **45**, 517–527.
25. Selzer,G., Som,T., Itoh,T. and Tomizawa,J. (1983) The origin of replication of plasmid p15A and comparative studies on the nucleotide sequences around the origin of related plasmids. *Cell*, **32**, 119–129.
26. Del Solar,G., Giraldo,R., Ruiz-Echevarría,M.J., Espinosa,M. and Díaz-Orejas,R. (1998) Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.*, **62**, 434–464.
27. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 6640–6645.
28. Jensen,S.I., Lennen,R.M., Herrgård,M.J. and Nielsen,A.T. (2016) Seven gene deletions in seven days: fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress. *Sci. Rep.*, **5**, 17874.
29. Vo,P.L.H., Ronda,C., Klompe,S.E., Chen,E.E., Acree,C., Wang,H.H. and Sternberg,S.H. (2021) CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nat. Biotechnol.*, **39**, 480–489.
30. Sambrook,J. and Russell,D. (2001) *Molecular Cloning a Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory Press: New York.
31. de Lorenzo,V., Eltis,L., Kessler,B. and Timmis,K.N. (1993) Analysis of *Pseudomonas* gene products using lacIq/Ptrp-lac plasmids and transposons that confer conditional phenotypes. *Gene*, **123**, 17–24.