PlasmidMaker is a versatile, automated, and high throughput end-to-end platform for plasmid construction

Behnam Enghiad1,2,3, Pu Xue1,2,3, Nilmani Singh1, Aashutosh Girish Boob1,2, Chengyou Shi1,2, Vassily Andrew Petrov1, Roy Liu1, Siddhartha Suryanarayana Peri1, Stephan Thomas Lane1, Emily Danielle Gaither1 & Huimin Zhao1,2

Plasmids are used extensively in basic and applied biology. However, design and construction of plasmids, specifically the ones carrying complex genetic information, remains one of the most time-consuming, labor-intensive, and rate-limiting steps in performing sophisticated biological experiments. Here, we report the development of a versatile, robust, automated end-to-end platform named PlasmidMaker that allows error-free construction of plasmids with virtually any sequences in a high throughput manner. This platform consists of a most versatile DNA assembly method using Pyrococcus furiosus Argonaute (PfAgo)-based artificial restriction enzymes, a user-friendly frontend for plasmid design, and a backend that streamlines the workflow and integration with a robotic system. As a proof of concept, we used this platform to generate 101 plasmids from six different species ranging from 5 to 18 kb in size from up to 11 DNA fragments. PlasmidMaker should greatly expand the potential of synthetic biology.

1 Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL, USA. 2 Department of Chemical and Biomolecular Engineering, University of Illinois Urbana-Champaign, Urbana, IL, USA. 3 These authors contributed equally: Behnam Enghiad, Pu Xue. ✉email: zhao5@illinois.edu
Plasmids are one of the most fundamental tools for recombinant DNA technologies. Design and construction of plasmids from smaller DNA parts to form complex functional DNA molecules such as biochemical pathways and genetic circuits is essential in molecular biology and represents one of the key steps enabling the design, build, test, and learn (DBTL) cycle in synthetic biology. During the past two decades, a variety of innovative methods for plasmid construction have been developed. Some of the most notable examples include sequence homology based methods such as Sequence and Ligation Independent Cloning (SLIC), isothermal Gibson assembly, uracil-excision based cloning, and yeast homologous recombination, or restriction digestion based methods such as MASTER ligation and Golden Gate assembly. Despite the development of numerous DNA assembly techniques, due to different capabilities and limitations of each approach, assembly of complex plasmid DNA molecules often requires trial of multiple techniques to find the suitable assembly approach for the DNA of interest or utilization of a multi-step hierarchical assembly scheme. As a result, construction of plasmid DNA remains one of the most time-consuming, labor-intensive, and inflexible steps in the DBTL cycle of synthetic biology, hampering the speed and scale of performing complex biological experiments.

Several strategies have been developed to address this limitation by automating the construction of plasmid DNA. For example, high throughput synthesis of transcription activator-like effector nucleases (TALENs) was achieved using automated construction of plasmids via Golden Gate method. In addition, a web-based software tool combined with a DNA assembly protocol using the Type-IIS restriction endonuclease based Modular Cloning technique was automated for efficient production of DNA fragments. Although these automated plasmid construction methods have shown high productivity with accuracy, due to utilization of restriction digestion based methods such as Golden Gate assembly or addition of computationally designed linkers to join fragments, these platforms still lack flexibility in construction of plasmids with virtually any DNA sequence. This lack of flexibility is mainly due to inherent limitations of restriction digestion based assembly methods such as presence of restriction enzymes’ recognition sequences on the DNA of interest or addition of scar sequences to improve the efficiency of DNA assembly. Moreover, most of the progress on automated DNA assembly has focused on the “build” part, where the initial design and final confirmation of plasmids are discrete. Therefore, it is highly desirable to develop a DNA assembly strategy that can assemble any DNA sequence with high fidelity and robustness and create an end-to-end pipeline for automated plasmid construction.

In this work, we report a robust, versatile, and automated end-to-end platform for plasmid construction named PlasmidMaker that enables scalable construction of virtually any plasmids in a high throughput manner. To implement this platform, we first developed a versatile, scalable, parallel, robust, and accurate method for assembly of multiple DNA fragments using *Pyrococcus furiosus* Argonaute (PfAgo) based artificial restriction enzymes (AREs). We then designed both frontand backend software for customers to build their specific DNA fragments using a user-friendly web interface and for our technicians to collect essential information that is required for DNA assembly, respectively. Finally, we integrated the DNA assembly method and the software with a robotic system named Illinois Biological Foundry for Advanced Biomanufacturing (iBioFAB) to create a nearly fully automated workflow for PlasmidMaker. As a proof of concept, we constructed 101 plasmids across six different species (bacteria, yeast, plants, and mammals) which involves ~2 × 10^4 pipetting steps using this PlasmidMaker platform. Plasmids with sizes ranging from 5 to 18 kb were assembled from up to 11 DNA fragments with limited human intervention. Our method allows assembly of fragments with GC content as high as 77% and error-free assembly of plasmids as large as 27 kb including the ones containing multiple repeats from up to 10 DNA fragments.

**Results**

**PlasmidMaker overview.** The overall end-to-end pipeline for PlasmidMaker is shown in Fig. 1. There are four major steps involved in the automated workflow for plasmid construction: (1) In the Design part, users can design plasmids by arranging DNA fragments in their preferred manner as well as search for common plasmids from a database to serve as their templates by using the frontend software. After an order is received, a technician can perform quality check by using the backend software. The sequences which pass the checking criteria can then be sorted into picklists and sent for construction. (2) In the Build part, according to the worklists generated, PCR plates for amplification of fragments are prepared. The Tecan FluentControl liquid handler is controlled by Momentum Work Scheduling software and F5 robotic arm to dilute and mix primers, templates, guides, and master mix into every single well of the 96-well plate. (3) Also in the Build part, one-pot automated PfAgo digestion, purification, ligation, and transformation of different plasmids are performed on the iBioFAB. (4) In the Test part, constructed plasmids purified by automated minipreps are verified through restriction digestion and gel electrophoresis. Correctly assembled plasmids are then re-cultured for making frozen stocks by the liquid handler. This high throughput pipeline allows completion of multiple plates of DNA assembly via standardized and integrated unit operations. Researchers only need to supervise the whole system while the programmed software and robotic arms implement the complicated experiments.

**PlasmidMaker development: 1. DNA assembly method.** The DNA assembly method using PfAgo-based AREs is shown in the bottom panel of Fig. 1. In the first step, linear DNA molecules including fragments created by polymerase chain reaction (PCR), linearized circular plasmids, or synthesized dsDNA are mixed in a one-pot reaction containing wild type (WT) PfAgo enzyme, a mutant PfAgo enzyme (PfAgo H745D, a.k.a. PfAgo*), and single-stranded DNA guides. The linear DNA fragments are designed such that they each share 24 bp sequence homology to their neighboring assembly fragments. The ssDNA guides are also designed such that after digestion, 5’ sticky ends of 12 nt length are created (the combination of PfAgo enzyme and DNA guides will be referred to as PfAgo/ARE hereafter). After digestion of all linear DNA molecules in a one-pot reaction, the digestion products are purified and assembled using a high-fidelity DNA ligase in the second and final step. The assembly products are then transformed into *E. coli* cells and the resulting colonies are checked for correct clones.

Previously we were able to demonstrate PfAgo’s capability to create small size user defined sticky ends (i.e., 2–5 nt) on dsDNA molecules. Although small size sticky ends like the ends created by most type II restriction enzymes are routinely used in DNA ligation/assembly applications, due to their low levels of base pair hybridization as well as limitations in the number of orthogonal overhangs, small size sticky ends are not ideal for assembly of multiple large DNA molecules. As PfAgo/AREs can theoretically create any size of sticky ends, we decided to test this capability by creating 5–12 nt sticky ends on two linear DNA molecules ends (Supplementary Fig. 2a). To determine the efficiency of PfAgo
cleavage on DNA ends for different sticky end sizes, we placed 24 bp of DNA homology at the ends of PCR reactions to make specific DNA fragments via liquid handlers. Next, one-pot digestion, ligation, and transformation is performed inside iBioFAB to assemble DNA fragments. The received oligos in either 96-well format or 384-well format are transferred into corresponding destination wells for PCR reactions to make specific primers and guides. The received oligos are picked from the robotic system for DNA fragments via liquid handlers. Next, one-pot digestion, ligation, and transformation is performed inside iBioFAB to assemble DNA fragments. In the first step, linear DNA molecules ends are digested with WT and engineered PfAgo/AREs in a one-pot reaction. The AREs generate 5′ sticky ends of 12 nt length. After purification, the digested DNA molecules are assembled in vitro using a high-fidelity DNA ligase and assembly products are transformed into E. coli cells for screening. Finally, constructed plasmids (input for “Test” part) are checked and the correct plasmids are stocked using the robotic system.

Other than advantages such as higher levels of base pair hybridization and increased probability for creation of orthogonal recognition sequences, for both 9 and 12 nt AREs with different GC contents as a cofactor for efficient digestion of DNA molecules. One of the factors affecting strand dissociation is DNA GC content. To test whether the GC content of PfAgo/AREs recognition sequence can affect their cleavage efficiencies, we placed 10 randomly generated recognition sequences for both 9 and 12 nt AREs with different GC-contents and GC-distributions on the ends of two linear DNA fragments and analyzed DNA cleavage efficiencies by PfAgo/AREs for each recognition sequence (Fig. 2a). We speculated that other than PfAgo/AREs recognition sequence, the overall GC content of the whole DNA fragment might also play a role in PfAgo/AREs cleavage efficiency. Therefore, we performed the same experiments on three sets of linear DNA molecules with three different GC content ranges. Previously, we created a mutant PfAgo* which can utilize either Mg2+ or Mn2+ ions as a cofactor for efficient digestion of DNA molecules. In our initial screenings, PfAgo* demonstrated different sequence affinities compared to the WT PfAgo enzyme. Therefore, we decided to also test the effects of GC content on PfAgo*/AREs.
As shown in Fig. 2b and Supplementary Figs. 4a-c, except for cases where the AREs recognition sequence GC content was ~80%, both WT PfAgo/AREs and PfAgo*/AREs were able to cleave their target DNA molecules. As expected, WT PfAgo and PfAgo* demonstrate different sequence affinities. When the GC content of the recognition sequence is ~10%, WT PfAgo demonstrates higher cleavage efficiency compared to PfAgo*. On the other hand, if the GC content of the recognition sequence is higher than 50%, PfAgo* cleaves its target at higher efficiency. While for DNA fragments with overall GC-contents of ~36% (i.e., Pyk1p-2μ) and ~51% (i.e., amp-CrtI), no significant difference in DNA cleavage/assembly efficiency was observed, both PfAgo/AREs and PfAgo*/AREs exhibited lower cleavage efficiencies when the overall DNA GC content was increased to ~73% (i.e., Sco5072-Sco5076). In addition, creation of different length of sticky ends (i.e., 9 or 12 nt) and the recognition sequence GC-distribution did not have major effects on AREs' cleavage efficiencies. Taken together, these results indicate that PfAgo/AREs can be programmed to cleave linear DNA ends with a wide range of GC-contents (0-75%) (Fig. 2b) and used in the assembly of multiple linear DNA fragments.

After testing the effects of both 9 and 12 nt sticky ends as well as a variety of DNA ligases on assembly of multiple DNA fragments (see Supplementary Text and Supplementary Fig. 5), we decided to use 12 nt sticky ends and HiFi Taq DNA ligase in the final version of our method and evaluated its performance.

To characterize the method’s error rate, we designed a five-fragment assembly for a 7.2 kb plasmid (pAmp-ZeaX) harboring a functional zeaxanthin pathway in E. coli (Supplementary Fig. 5a). As the resulting colonies harboring the full pathway produce yellow zeaxanthin, this plasmid allows quick assessment of DNA assembly fidelity based on colonies’ colors. However, production of yellow zeaxanthin does not necessarily guarantee an entirely correct assembly without any sequence errors. To evaluate whether DNA assembly by PfAgo/AREs would result in sequence errors between assembly junctions, we sequenced the five assembly junctions targeted by PfAgo/AREs on pAmp-ZeaX plasmid for 14 randomly selected yellow colonies. Out of the 70 sequenced junctions, all junctions showed correct assembly without any sequence errors. To find out whether any of the white colonies harbor the full pathway but with sequence mutations, we analyzed purified plasmids from three white colonies by restriction digestion. Interestingly, one of the three plasmids showed correct digestion pattern for the complete 7.2 kb plasmid. After full plasmid sequencing, we found out that the lack of zeaxanthin production is due to four base pair deletions at the CrtZ gene stop codon. This section however is not targeted by any of the PfAgo/AREs and all the five junctions targeted by PfAgo/AREs showed correct assembly without any sequence errors. Since in plasmid pAmp-ZeaX, CrtZ and CrtB genes share 62 bp of sequence homology on their ends and they are amplified in one fragment, we speculate that this deletion occurs during the PCR reaction and due to the distance of mutation to the PfAgo/AREs targets, it cannot be correlated to PfAgo/AREs’ cleavage. Consequently, these results indicate that
the resulting cleavage products can be accurately assembled using at least 24 different DNA guides in a one-pot reaction and as well as gfp included four 500 bp repeated TEF1 promoter sequences. To minimize the effect of DNA size on transformation efficiency and hence the number of acquired colonies, we sought to assemble this plasmid with up to 10 DNA fragments. Taken together, these results suggest that plasmid molecules with sizes up to 27 kb including the ones with sequence repeats can be efficiently assembled by P/Ago/AREs generating 12 nt sticky ends from up to 10 DNA fragments. All experiments were performed in three biological replicates. For all graphs, error bars for colony number and assembly fidelity, show standard deviation (s.d.) and standard error (s.e.m) respectively. Source data are provided as a Source Data file.

PlasmidMaker development: II. Software. To empower a broader set of users to use PlasmidMaker, we developed a web-based frontend software that is used to place orders to iBioFAB and a complete backend pipeline for automated assembly (Fig. 4a). We used industry-standard tools and practices for all phases of the software lifecycle. To facilitate rapid development, we selected the ReactJS framework for its widespread support for common open-source libraries and existing design templates. We selected the Django framework for backend processing for interoperability with existing laboratory software written in Python. The frontend is a fast and reactive website to design plasmids that can be sent through the automated workflow pipeline. Users must register to create an account that keeps track of their plasmids. In the main view, users can create new projects, which starts the process for plasmid construction. Once a project is selected, the project editor is opened, where the users either select from an existing library of plasmid fragments or upload custom fragments. We provide a text-based search and a sequence-based search to explore available fragments in the library. The main view includes a plasmid visualizer to display the final output, and a sequence visualizer to view which components can be placed in the project (Supplementary Fig. 7). Once the user has created an order of their liking, the fragments are annotated each assembly and analyzed them by restriction digestion. As shown in Fig. 3b, both the assembly fidelity and the number of acquired colonies are inversely affected by the size of final product. However, despite the effects of size on DNA assembly, we were able to assemble the 26.8 kb plasmid using seven DNA fragments with ~63% fidelity and acquire ~4200 colonies.

To analyze the effect of the final assembly product size on the assembly fidelity and the number of acquired colonies, we sought to assemble three different plasmids with varying sizes ranging from 12.3 to 26.8 kb using seven DNA fragments. These plasmids included a 12.3 kb plasmid harboring an (R,R)-2,3-butanediol (BDO) pathway and a gfp gene, a 17.5 kb plasmid harboring the BDO and zeaxanthin pathways and a gfp gene, and a 26.8 kb plasmid harboring an n-butanol pathway as well as gfp and mCherry genes. The 26.8 kb plasmid also included four 500 bp repeated TEF1 promoter sequences. To analyze the effect of DNA size on transformation efficiencies and hence the number of acquired colonies, the assembly mixtures were introduced into E. coli cells by electroporation. For evaluation of the assembly fidelity, we picked 10 colonies from each assembly and analyzed them by restriction digestion. As shown in Fig. 3b, both the assembly fidelity and the number of acquired colonies are inversely affected by the size of final product. However, despite the effects of size on DNA assembly, we were able to assemble the 26.8 kb plasmid using seven DNA fragments with ~63% fidelity and acquire ~4200 colonies.

To analyze the effect of the final assembly product size on the assembly fidelity and the number of acquired colonies, we sought to assemble three different plasmids with varying sizes ranging from 12.3 to 26.8 kb using seven DNA fragments. These plasmids included a 12.3 kb plasmid harboring an (R,R)-2,3-butanediol (BDO) pathway and a gfp gene, a 17.5 kb plasmid harboring the BDO and zeaxanthin pathways and a gfp gene, and a 26.8 kb plasmid harboring an n-butanol pathway as well as gfp and mCherry genes. The 26.8 kb plasmid also included four 500 bp repeated TEF1 promoter sequences. To analyze the effect of DNA size on transformation efficiencies and hence the number of acquired colonies, the assembly mixtures were introduced into E. coli cells by electroporation. For evaluation of the assembly fidelity, we picked 10 colonies from each assembly and analyzed them by restriction digestion. As shown in Fig. 3b, both the assembly fidelity and the number of acquired colonies are inversely affected by the size of final product. However, despite the effects of size on DNA assembly, we were able to assemble the 26.8 kb plasmid using seven DNA fragments with ~63% fidelity and acquire ~4200 colonies.

To analyze the effect of the final assembly product size on the assembly fidelity and the number of acquired colonies, we sought to assemble three different plasmids with varying sizes ranging from 12.3 to 26.8 kb using seven DNA fragments. These plasmids included a 12.3 kb plasmid harboring an (R,R)-2,3-butanediol (BDO) pathway and a gfp gene, a 17.5 kb plasmid harboring the BDO and zeaxanthin pathways and a gfp gene, and a 26.8 kb plasmid harboring an n-butanol pathway as well as gfp and mCherry genes. The 26.8 kb plasmid also included four 500 bp repeated TEF1 promoter sequences. To analyze the effect of DNA size on transformation efficiencies and hence the number of acquired colonies, the assembly mixtures were introduced into E. coli cells by electroporation. For evaluation of the assembly fidelity, we picked 10 colonies from each assembly and analyzed them by restriction digestion. As shown in Fig. 3b, both the assembly fidelity and the number of acquired colonies are inversely affected by the size of final product. However, despite the effects of size on DNA assembly, we were able to assemble the 26.8 kb plasmid using seven DNA fragments with ~63% fidelity and acquire ~4200 colonies.
to obtain guides and primers. The order can also be submitted to iBioFAB for synthesis of the customized plasmid. Updates to the order after submission are tracked in real time through the projects tab. At each phase of an order, users receive updates as to the progress of their order. We use the React frontend library to serve performant webview renders to the client, and to ensure protection from cyber-attacks, an authentication system is implemented which obfuscates it from Cross-Site Scripting attacks through expirable cookies.

One of the challenges in designing a user interface that is intuitive to use is the layout of the controls and system state. To validate our design decisions, we incorporate a rolling-feedback method for collecting sentiment and experience data. After each iteration of the interface, we send requests to perform challenges through expirable cookies. The recognition sequence is finalized as soon as the guides satisfying the design rules for PfAgo digestion and Hifi Taq ligation are obtained. To simplify and improve the fidelity of PfAgo-based assembly, we developed a computational framework to generate design components given the annotated ‘DNA’ file (Supplementary Fig. 8). We first obtain the plasmid sequence, guide search space and the off-target library from the nucleotide sequence of the annotated fragments. The guide search space is defined as the region obtained by joining 40 bp from the end of the adjoining fragments for locating the guides. Based on the GC content of the fragments and the guide search space, the code provides suggestions on which PfAgo enzymes to use for the assembly.

The guide library is created using the guide search space of 80 bp and therefore, contains 57 recognition sequences of 24 bp to screen for. For each junction, a 24 bp recognition sequence (containing two guides of 16 bp each) is selected based on the digestion and ligation criteria (Fig. 4b). The digestion criteria include constraints based on the GC content, presence of G-quadruplex and more than 4 consecutive A’s or T’s and off-targeting to maximize the efficiency of creating desired sticky ends using PfAgo and PfAgo6. The ligation criteria are based on avoiding palindromic sticky ends and minimizing the mismatch ligation below a threshold score based on the mismatch fidelity profile of Hifi Taq ligase22. Once the viable guides are obtained, primers are then designed based on the position of the recognition sequence to include overhangs in the PCR amplicons for generating complementary 12 bp sticky ends (Supplementary Fig. 9). The length of the primer binding site is decided based on the GC content of the 18 bp binding region to ensure...
amplification works across fragments of varying GC content. The primer binding site is further modified such that forward and reverse primers have $T_m$ values within 1 °C. The $T_m$ values of the primers are calculated using the `calcTm` function of Primer3. For assembling smaller fragments such as linkers and E. coli promoters (<80 bp), the workflow was modified to design primers for combining smaller fragments with the adjacent fragments via PCR overhangs and afterward, guides and primers are generated for assembly of the modified fragments.

To verify whether the plasmid assembly is correct, the script also generates the combination of restriction enzymes to use such that bands formed while performing gel electrophoresis are well separated and observable. The restriction enzymes can be selected from a list of commonly used restriction enzymes provided in Supplementary Table 1. We first check for the recognition sequence of the restriction enzymes and locate them in the plasmid sequence. Then, we look at the combinations of the restriction enzymes and discard the ones that do not satisfy the following criteria so as to observe well separated bands during gel electrophoresis: (1) Smallest band size $\geq$100 bp; (2) Largest band size $\leq$9500 bp; (3) If the difference between subsequent bands $<A \times$ band size, combination discarded where $A$ is the accuracy of Fragment Analyzer and is set to 0.1 (Supplementary Fig. 10). If multiple `.DNA` files are provided in a folder, the script generates files for guides and primers for each plasmid assembly as well as a combined file for a list of restriction enzymes for correct assembly verification and repeated fragments across plasmids. We also address the problems in the assembly of the plasmid library through quality control scripts. All the details of the algorithm are explained in brief in the Supplementary Information.

We employed the Django Python framework to develop the backend server of the web application. Such a framework allows for seamless development of API endpoints, database management, and background tasks through a model-view-controller framework (Supplementary Fig. 11). The backend has a set of API calls that the frontend could call for necessary stored information from user data to custom plasmid information (Supplementary Table 2).

We used a PostgreSQL database to store experimental plasmids from a set of template plasmids, and important parts (promoters, genes, terminators) from the templates. With such storage of data, users can search genetic parts that they would like to add to their desired custom plasmids using the web interface. After an order is submitted, a background job managed by a Celery queue is started. Upon the order fulfillment, a notification is sent to the user when their custom plasmids are ready for download.

The automated DNA assembly requires picklist/worklist generation to smoothly integrate the liquid handling on different instruments on the iBioFAB (Fig. 4c). Using custom Python scripts and the output of the order queue, we generate picklists for various assembly steps. A custom Python script combines and removes duplicated DNA sequences from the primers/guides file generated from the plasmidMaker order queue for one DNA assembly batch and formats it for ordering. Using the information sheet from the manufacturer, another Python script generates picklist to add water to the lyophilized DNA. A subsequent picklist dilutes and transfers primers to a separate 384-well plate, which will work as ECHO liquid handler source plate for PCR set up. The locations of primers/guides in 96-well/384-well plate for each plasmid is written in a csv file using a Python script and this file is used as reference file for creating liquid handling picklists during DNA assembly.

For setting up PCR, a Python script reads the reference csv files along with the location of PCR templates and creates picklist for mixing of primers and templates from 384-well plate to 96-well PCR plate in ECHO liquid handler. The Python script generates multiple PCR picklists for dividing the PCRs into multiple plates if the total number of PCRs exceeds 92. Each 96-well PCR plate has 92 PCRs and empty wells are used for DNA ladder for Fragment Analyzer run. Some DNA assemblies require insertions of $<80$ bp DNA fragments, which is achieved by designing extended PCR primers to add the inserts through PCR. A separate Python script creates the picklists for first round of PCR for adding $<80$ bp inserts. The PCR from first round is used as template for second PCR to generate the PCR fragments for assembly. The PCR products are analyzed using Fragment Analyzer and a Python script compares the expected and observed DNA sizes. When PCRs fail, a picklist is created combining all failed PCRs, followed by optimizing the PCRs. After all PCRs are successful, another Python script generates the picklist for guide mixing and phosphorylation. For digestion of PCR fragments with P/Ago/AreS, a Python script generates picklist for equimolar mixing of purified PCR fragments and phosphorylated DNA guides. After the digestion, PCR fragments are purified, ligated, and transformed in E. coli cells for colony and verification of correct assembly.
Fig. 5 Overview of the automated plasmid construction workflow. a A flowchart showing different instruments across iBioFAB connected through Momentum™ to orchestrate the plasmid construction process. The whole process starts from design of oligos and ends at preparation of frozen stocks of correctly assembled plasmids. The steps of experiments are listed in black character and the corresponding instrument are listed in blue. Rectangle shape: automated steps without incubation and shaking. Oval shape: automated steps with incubation and shaking. The detailed protocols are available in Supplementary Fig. 12. b Schematic overview for a closed automated plasmid construction workflow inside iBioFAB with all the detailed procedures and equipment listed.
in the confirmation module: one for quick identification of common restriction enzymes among different constructed plasmids, and the other for matching the predicted and actual band sizes efficiently.

**PlasmidMaker validation: automated construction of 101 plasmids.** To validate the PlasmidMaker platform, we sought to construct 101 plasmids for a wide variety of commonly used model organisms including *Escherichia coli*, *Saccharomyces cerevisiae*, *Issatchenkia orientalis*, *Arabidopsis thaliana*, Human embryonic kidney (HEK) 293 T cells, and *Streptomyces griseoflavus*. The summary of all the constructed plasmids including their assembly fidelity, plasmids size, number of fragments in total, fragments size range, and primers/guides involved is listed in Table 1. This table demonstrates the robustness and capability of PlasmidMaker in construction of a variety of plasmid with different sizes and number of fragments. Restriction digestion checks by Fragment Analyzer for the 101 constructed plasmids is available in Supplementary Fig. 14. In overall, we were able to check by Fragment Analyzer for the 101 constructed plasmids is different sizes and number of fragments. Restriction digestion of PlasmidMaker in construction of a variety of plasmid with different sizes and number of fragments (11 fragments) showing the lowest fidelity. We were also able to construct plasmids containing multiple fragments with GC content as high as 77%. Because of the uniqueness of PfAgo/AREs, our DNA assembly method provides the most versatile approach for construction of plasmids containing multiple repeats and secondary structures without any hindrance from "forbidden" DNA sequences. The detailed information and features of each individual plasmid are shown in Supplementary Table 3.

**Discussion**

In this study, we developed a versatile and automated end-to-end platform for design and construction of virtually any plasmids in a high throughput manner. This platform is built upon an innovative PfAgo/AREs-based DNA assembly strategy and seamless integration of a computational infrastructure (i.e., forefront and backend) with a fully-integrated robotic system (i.e., iBioFAB).

An ideal method for assembly of DNA molecules should allow quick, scarless, error-free, sequence independent, and parallel assembly of multiple DNA molecules in a predetermined order and with high efficiency. While DNA assembly strategies based on native type IIIs restriction enzymes such as Golden Gate assembly allow parallel assembly of multiple DNA molecules in a predetermined order with high efficiency and low error rate, they generally suffer from two major limitations: (1) All DNA fragments used in assembly are required to be free of the restriction enzyme recognition sequences. As restriction enzyme recognition sequences are typically very short (~6 bp), this requirement limits their application in assembly of large DNA molecules due to low probabilities of finding restriction enzymes which do not cleave the DNA molecules in the middle; (2) restriction enzymes only generate short (~4 nt) sticky ends. Due to the small space of possible orthogonal 4 nt sticky ends that do not result in mismatch ligation and at the same time demonstrate high ligation efficiencies, assembly of multiple DNA fragments is generally performed using sets of standardized, predefined sticky ends, which lead to assemblies with scar sequences. Unlike restriction enzyme based DNA assembly methods, long overlap homology based methods such as SLIC, SLICE, or isothermal Gibson assembly are not limited by the presence of "forbidden" DNA sequences in the middle of DNA fragments. However, homology-based DNA assembly methods are not truly sequence independent. As most homology-based methods rely on creation of long ssDNA ends through exonuclease treatment, the success rate and efficiency of these methods drops significantly if the generated long ssDNA ends can form stable secondary structures or contain sequence repeats. In addition to exonuclease treatment, some of the homology-based methods such as isothermal Gibson assembly also rely on end-filling using DNA polymerase enzymes which can result in creation of junctions with sequence errors. As described elsewhere, in some cases, the error rate for junctions assembled by Gibson assembly can be as high as 100%.

Compared to native restriction enzymes, PfAgo/AREs offer programmability as well as much longer recognition sequences. As a result, assembly of DNA molecules using PfAgo/AREs is not limited by the presence of “forbidden” sequences on the DNA of interest. Additionally, PfAgo/AREs can also be programmed to create varying sticky end sizes. Generation of longer sticky ends compared to native restriction enzymes (i.e., 12 nt in this study) results in substantially higher probabilities for acquisition of orthogonal sticky ends as well as possibility of utilization of thermophilic DNA ligases which offer remarkable ligation specificities. Unlike exonuclease enzymes, PfAgo/AREs can create defined sticky end sizes. Compared to long ssDNA ends generated by exonuclease treatment (20–40 nt), creation of defined sticky ends with moderate size (i.e., 12 nt) significantly reduces the probability of DNA secondary structure formation. PfAgo/AREs also offer capabilities in assembly of fragments containing sequence repeats at their ends. As the location for cleavage of DNA molecules by PfAgo/AREs is predefined, if this location (i.e., 24 bp) does not lie within the repeated sequence, presence of the repeated sequence would not interfere with assembly by PfAgo/AREs. Lastly, as presented in this study, due to use of defined sticky ends and DNA ligation, DNA assembly with PfAgo/AREs can be essentially performed with 0% error rates.

Traditionally, multiple cloning procedures are performed manually, which is labor-intensive and error-prone. Specifically, the possibility of human induced errors becomes very high when one trying to clone a large batch of plasmids, where numerous pipetting steps with tiny volumes of materials are required. With

![Table 1 Summary of results for construction of 101 plasmids across bacteria, yeast, plant, and mammalian cells. Source data are provided as a Source Data file.](https://doi.org/10.1038/s41467-022-30355-y)
automation of the cloning process, the assembly time is decreased tremendously as well as errors introduced by researchers can be avoided. In this work, ~2 × 10^6 pipetting steps and transfer volume down to 2.5 mL are included to complete the automated plasmid construction. It will be very challenging to perform all these repetitive procedures manually with both precision and accuracy.

To establish an automated end-to-end platform for plasmid construction using PfAgo/AREs-based DNA assembly, we developed a computational infrastructure consisting of a user-friendly frontend for plasmid design and a backend that streamlines the workflow and integration with our iBioFAB. The computational framework enables robust and high throughput construction of plasmids with high fidelity by incorporating complex criteria such as minimizing off-target sites and reducing the chances of mismatch ligation between the fragments which are strenuous to optimize manually. It also features the first bioinformatics algorithm for guide DNA selection of Argonauta proteins. The digestion efficiency of PfAgo/AREs can be further increased by incorporating design rules for position-dependent base preference of PfAgo enzyme or its variants. Insights can be drawn from the computational model for predicting the sequence-specific activity of Cas9 variant and applied to PfAgo/AREs. Moreover, the ligation criteria can be improved by profiling mismatch ligation specifically for 12 nt sticky ends for HiFi Taq ligase.

One key challenge of building a fully automated DBTL cycle is smooth integration of the user frontend, the technician frontend, the software backend, and the robotic hardware-end. Users of the web interface are likely to be familiar with plasmid design and assembly technique in general but need to be able to design plasmids without knowledge of innovative PfAgo/AREs. To facilitate a simple design process, the frontend uses validation of all user input to ensure compatibility with the assembly process. After we receive the requests for plasmid assembly from researchers on the PlasmidMaker user frontend, the backend software will perform BLAST to search for sequences and annotate them. Meanwhile, the technician will generate the picklist for ordering according to the quality check results from software-end. If no primers and guides are found, a redesign of the assembly and/or an iterative check will be required. After we get the parts for assembly, the robotic hardware-end will perform the experiments following the inputs from the Design part. Ideally, no human intervention is needed in these steps, but if anything goes wrong in the middle of this cycle, the error will be recorded in our database and the continued run must be stopped. The role of the researcher is to monitor the whole workflow and respond quickly once an error occurs. Therefore, rational decision makings and expertise in automation are preferred for technicians before running the whole system.

One of the major challenges for high throughput assembly is robust integration of liquid handling steps over different instruments for different molecular biology assays. While worklists can be manually generated for each instrument, it is highly error-prone and labor-intensive. Here, we have developed an end-to-end Python module to generate all required worklists for DNA assembly on iBioFAB. The Python module takes the primer/guide information sheet as input and generates liquid handling worklists for various assays on iBioFAB such as PCR, guide mixing, phosphorylation, equimolar mixing of PCR fragments and PfAgo/AREs digestion. The versatility of the worklist generation lies in the fact that it creates all necessary worklists for different reactions before we begin the assembly and works under many different assembly conditions.

Finally, to limit the human intervention in the automated workflow for plasmid construction, it is important to confirm the correct PCR amplification, the number of colonies being generated, and the digestion pattern at the end of each module, respectively (Supplementary Fig. 14 and Supplementary Table 3). Based on the results from the construction of 101 plasmids, the rate-limiting step is the PCR amplification of the fragments to be assembled. The success rate of PCR depends on many factors such as GC content, primer dimerization, primer multiple binding sites, secondary structures, length of fragment, annealing temperature, and DNA polymerase. In general, we could obtain more than 90% correctly amplified fragments after the PCR reaction in module 1. For the incorrect ones, we had to set up different conditions for troubleshooting, such as trying different DNA polymerases, using gradient annealing temperatures, adding GC enhancers, and adding dimethyl sulfoxide (DMSO) to the PCR master mix. If the PCR still fails, new primers need to be designed and new quality check procedures need to be conducted. Therefore, it is useful to develop a standardized, repeatable protocol for troubleshooting steps as to improve the overall assembly efficiency. The other problem we faced during this workflow is the generation of false-positive colonies with only circular backbone template plasmids. We solved this by adding DpnI and incubating for 2 h after the PCR reaction and before purification in module 1 to get rid of the template structures. We also performed the Plasmid-Safe protocol after the ligation step in module 2 to improve the chance of getting true positive colonies.

In conclusion, PlasmidMaker is a powerful platform consisting of advanced software and hardware for rapid construction of virtually any plasmids in a high throughput manner. It addresses a critical bottleneck in basic and applied biology and should greatly accelerate the development of synthetic biology for biotechnological and biomedical applications.

**Methods**

**Bacterial strains and reagents.** *E. coli* strain NER108 (New England Biolabs, MA) was used for all cloning experiments. Restriction enzymes, T4 polymerase kinase, T4 DNA ligase, T7 DNA ligase, *E. coli* DNA ligase, Taq DNA ligase, HiFi Taq DNA ligase, ATP, and Q5 DNA polymerase were purchased from New England Biolabs, MA. Plasmid-safe ATP-dependent DNase was obtained from Lucigen, WI. DNA purification buffers PB and PE were purchased from Qiagen, Germany. All DNA oligonucleotides were ordered from Integrated DNA Technologies (Coraville, IA). All DNA guides were ordered unphosphorylated and the phosphorylation reaction was performed by T4 PNK enzyme prior to DNA cleavage.

**Expression and purification of WT PfAgo and PfAgo* proteins.** The expression plasmids for WT PfAgo (pET28a-PfAgo) and PfAgo* (pET28a-PfAgo4745D) were transformed into *E. coli* KRX (Promega, WI) following manufacturer’s protocol. The plasmids were cultivated overnight at 37 °C in 5 mL of LB medium supplemented with 50 μg/mL kanamycin. Following overnight incubation, 2 mL of culture was transferred into 400 mL of Terrific Broth containing 50 μg/mL kanamycin in a 2 L baffled flask and incubated at 37 °C until the OD_{600} of 1.2–1.5 was reached. The cultures were then cold shocked by incubation in an ice bath for 15 min and protein expression was induced by addition of isopropyl β-D-thiogalactopyranoside (IPTG) and L-rhamnose to final concentrations of 1 mM and 0.1% (w/v), respectively. Expression was continued by incubation at 30 °C for 20 h. Cells were harvested by centrifugation at 5,000 × g for 10 min and after removing the supernatant, the cell paste was stored at –80 °C until purification.

For the purification step, frozen cells were thawed at room temperature and then fully resuspended in 30 mL of buffer 1 (20 mM Tris-HCl pH 8.0, 1 M NaCl). The resuspended cells were lysed by sonication for 10 min (30% amplitude, 5 on, 5 off). The solution was then centrifuged 3 times at 20,000 g at 4 °C for 10 min intervals and the supernatant was used for purification using fast protein liquid chromatography (FPLC) and 1 mL Strept-Tactin Superflow high capacity column (IBA Lifesciences, Germany). After equilibration with five column volume (CV) of buffer I, the column was loaded with the supernatant and washed with 10 CV of buffer I. The proteins were eluted with 10 CV of buffer II (20 mM Tris-HCl pH 8.0, 1 M NaCl, 2.5 mM d-desthiobiotin) and concentrated with Amicon Ultra-15 50 kDa filtration units (MilliporeSigma, MA). The purified proteins were diluted to ~1 mg/mL concentration using storage buffer (20 mM HEPEs pH 7.0, 300 mM NaCl, 15% (v/v) glycerol) and the aliquots were stored at ~80 °C.
Ligation of cleavage products by HiFi Taq DNA ligase. The DNA ligase was carried out in a 20 µL reaction containing 5 µL of purified cleavage products, 2 µL (part number: DPN100500). The total running time of automated gel purification was incubated at 37 °C for 2 h and cooled to 10 °C until transformation. Based on our experiments, 0.5 µL of DpnI restriction enzyme can also optionally be added to the ligation reaction to help with lowering the number of false positive colonies formed by carryover of PCR templates during DNA purification. However, we only used DpnI in assembly of the 26.8 kb plasmid using 8, 9, or 10 DNA fragments.

Plasmid-safe treatment of the assembly products (optional). Following ligation, 2.3 µL of 10 mM ATP and 1 µL of plasmid-safe DNase was added to the ligation mixture and the sample was incubated at 37 °C for 90 min, 70 °C for 30 min followed by cooling to 10 °C until transformation. This plasmid-safe treatment step was only used in assembly of the 26.8 kb plasmid using 8, 9, or 10 DNA fragments.

**E. coli transformation.** For chemical transformation, 5 µL of the ligation mixture was added to 50 µL of chemically competent NEB10b cells (New England Biolabs) and transformation was performed following the manufacturer’s protocol. For electro- transformation, the ligation mixture was first dialyzed against ddH2O using membrane filters (MilliporeSigma, MA, cat. no. VSWP02500) for 30 min at room temperature. 5 µL of the dialyzed mixture was then added to 25 µL of electrocompelement NEB10b. E. coli cells (New England Biolabs), mixed gently and transferred to a 2 mm electroporation cuvette. Electroporation was performed using Gene Pulser XCell Electroporation system (Bio-Rad, CA) with 2500 V, 200 Ω, 25 µF condition.

Analysis of plasmid assembly. Following transformation, 1 mL of SOC medium was added to the cells and the suspended cells were transferred into 14 ml round-bottom Falcon tubes. The culture incubated at 37 °C for 60 min with shaking at 250 rpm. The total concentration of different colonies was then plated on LB agar plates containing appropriate antibiotics (100 µg/mL carbenicillin) using cell spreaders and incubated at 37 °C until colonies appeared. To increase the intensity of the yellow color for colonies harboring the zea1nthyaxin pathway, after ~16 h of incubation at 37 °C, the plates were placed overnight at room temperature and the colonies with the numbers calculated from the following day. To recover the correct colonies, at least 10 single colonies were picked and grown at 37 °C in LB medium supplemented with appropriate antibiotics. The plasmid DNA was purified from the cultures using Qiagen Spin Miniprep Kit (Qiagen, Germany) following the manufacturer’s protocol. The purified plasmids were then digested by appropriate restriction enzymes and the samples were analyzed by agarose gel electrophoresis.

**Automation of plasmid construction.** BioFAB was used to automate the assembly of DNA parts for the target 101 plasmids, including dilution of primers and guides, PCR reaction, gel electrophoresis, purification, guides phosphorylation, digestion, ligation, transformation, cell cultivation, plasmid extraction, and so on. SnapGene viewer software was used to analyze DNA files. SnapGene-reader was used to read all the 101 annotated plasmids files. The overall and detailed workflows of the experiments are shown in Fig. 5 and Supplementary Fig. 12, respectively. First, the primers and guides were diluted into 100 µM using pure water after the orders were received from IDT. The primers were transferred into 384-well plate from the 96-well IDT plate for module 1 sample preparation of PCR reaction. The guides were stored in 96-well plate for guide phosphorylation in module 2. The DNA plate samples were diluted to 500 ng and added into 384-well plates as well. The primers and templates were then transferred from 384-well source plate into 96-well destination plate using Echo 550 liquid handler. In each well of the destination plate, it contained the specific primers and templates for one PCR reaction according to the input CherryPick csv file. Meanwhile, the PCR reaction master mix was prepared by Tecan FluentControl liquid handler in a 96-well format. Since the volume required to dilute every sample into same final concentration was different, Flexible-Channel Arm (FCA) tips were used to perform the liquid transfer. This allowed efficient liquid handling with different volumes being transferred in each well. We also created a worklist in Fluent script for FCA that kept us on track. After another incubation at 37 °C for 2 hr, the plate was sent to Fluent again to prepare for gel electrophoresis check of the fragment sizes. We used the Fragment Analyzer dsDNA 920 Reagent Kit (Agnel Technologies, Santa Clara, CA) to perform the automated gel electrophoresis analysis. Premade marker plate (30 µL/well + mineral oil overlay), inset buffer plate (1000 µL/well), and 40 µl gel with 4µL PonceauRating dye were put into the corresponding trays in Fragment Analyzer. For preparation of the sample plate, 22 µL of 1x TE buffer was added to each well first by MCA, followed by 2 µL of the PCR product from the 96-well plate after DpnI digestion. Ladders were added to wells A1 and H12. For these colonies, we chose the ladder with a calibrated range from 75 bp to 15,000 bp (part number: ONP-9229-KIT). After another incubation at 37 °C for 1 h, the electrophoresis took 1.5 h, then we could decide whether we obtained the correct band patterns or not using the ProSize data analysis software. After verification of the
sizes, we performed the PCR purification using the Zymo-96 DNA clean-up column-based kit (D4018) on Fluent liquid handler. The DNA quantification step was done independently using the high throughput Luminar Microfluidic system (Unchained Labs, Pleasanton, CA). Only 2 pl of eluted samples were needed to determine DNA concentrations.

The output csv file of the DNA amounts was used to calculate the volume needed for each fragment in every assembly in order to obtain equimolar (1 μg of fragments in total). The fragments were mixed by Echo 550 and the guides were mixed and phosphorylated by Fluent and thermal cyclers, respectively (Fig. 4). The mixed PCR fragments and guides were then digested using PflAgo and purified using the Zymo-96 clean-up kit. Next, ligation of purified digested fragments was performed, followed by 1.5 hr incubation with plasmid-safe nuclease. The treated product was then transformed in NEB-10E E. coli competent cells and plated on LB agar plates and moved to Cytomat_2C2 by F5 robotic arm for overnight culture at 37 °C. On the second day, four colonies were picked from each of the plates using Pickolo colony-picker (SciRobotics, Israel) and inoculated in 96-deepwell plate with 1.4 mL of LB + antibiotic media in each well (more colonies were picked if we obtained false positive results after gel checking). The seed culture was grown overnight at 390 rpm and 37 °C. On the following day, 50 μL of culture was aspirated out first for preparation of frozen glycerol stocks. The rest of cultures were spun down at 15 min and 3,900 rpm using the Agilent centrifuge. Supernatants were removed and the cell pellets were kept for plasmid extraction. For miniprep, we first tried magnetic bead-based method using the MagiET plasmid DNA Kit (K2792) (ThermoFisher Scientific, Waltham, MA) on the Fluent. However, the concentration obtained was generally low with ~20 ng/μL in 30 μL elution volume. Then, we developed a vacuum-based plasmid extraction protocol named “TeVacs” on Fluent using the PureLink Pro Quick96 plasmid purification Kit (K211004A) (Invitrogen, Carlsbad, CA). Our customized protocol features the use of vacuum manifold for efficient removal of washing buffer and final elution of plasmid DNA into a 96-well plate placed on a 3D-printed holder. This alternative method took less than 90 min and yielded a DNA concentration of ~80 ng/μL in 50 μL elution volume. The purified plasmids were then digested using the selected common restriction enzymes listed in Supplementary Table 1 and analyzed by Fragment Analyzer. The correctly digested patterns of these plasmids were shown in Supplementary Fig. 14. Lastly, frozen stocks were prepared on Fluent using 40% glycerol and the overnight culture of correct clone in a 1:1 volume volume condition. Either FCA tips or MCA tips were used depending on the number of plasmids created.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** The data supporting the findings of this study are available within the article and its Supplementary Information files or uploaded through public repositories. If specific data is believed to be missing, that data is available from the corresponding author upon request. The plasmids for expression of PflAgo enzyme and its mutant version are available through Addgene.

**Code availability.** The custom codes for the frontend and backend software used in this study are available at Zenodo: https://zenodo.org/record/5812313 (https://doi.org/10.5281/zenodo.5812313). The source code to design guides, primers for PflAgo-based plasmid assembly and select restriction enzymes for assembly verification on a local machine is available at Zenodo and on Github at https://github.com/ibiofab/PlasmidMaker_GuideDNA. The web application is available on the following web address: https://biofoundry.web.illinois.edu/. The datasets used to populate the PostgreSQL database are available in the Supplementary Information files in a SnapGene format. The sequences can be converted into a Postgres format using the database scripts found in our code repository.

Received: 5 January 2022; Accepted: 28 April 2022; Published online: 16 May 2022

**References**

1. Hughes, R. A. & Ellington, A. D. Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. Cold Spring Harb. Perspect. Biol. 9, a023812 (2017).

2. Ellis, T., Adie, T. & Baldwin, G. S. DNA assembly for synthetic biology: from parts to pathways and beyond. Integr. Biol. 3, 109–118 (2011).

3. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant mixed-strain DNA via SLIC Cameras and SLICE. Nucleic Acids Res. 35, 251–256 (2007).

4. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345 (2009).

5. Geu-Florès, F., Nour-Eldin, H. H., Nielsen, M. T. & Halkier, B. A. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res. 35, e57 (2007).

6. Nour-Eldin, H. H., Hansen, B. G., Norholm, M. H., Jensen, J. K. & Halkier, B. A. Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. Nucleic Acids Res. 34, e122–e122 (2006).

7. Xiao, Z., Zhao, H. & Zhao, H. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. Nucleic Acids Res. 37, e16–e16 (2009).

8. Gibson, D. G. et al. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome. Proc. Natl Acad. Sci. 105, 20404–20409 (2008).

9. Chen, W.-H., Qin, Z.-J., Wang., J. & Zhao, G.-P. The MASTER (methylation-assisted tailorable ends rational) ligation method for seamless DNA assembly. Nucleic Acids Res. 41, e93–e93 (2013).

10. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS one 4, e5553 (2009).

11. Ortiz, L., Pavan, M., McCarthy, L., Timmens, J. & Densmore, D. M. Automated robotic liquid handling assembly of modular DNA devices. J. Vis. Exp. 130, e54703 (2017).

12. Exley, K. et al. Utilising datasheets for the informed automated design and build of a synthetic metabolic pathway. J. Biol. Eng. 13, 1–10 (2019).

13. Chao, R. et al. Fully automated one-step synthesis of single-transcript TALEN biosensors using a biological foundry. ACS Synth. Biol. 6, 678–685 (2017).

14. Storck, M., Haines, M. C. & Baldwin, G. S. DNA-BOT: a low-cost, automated DNA assembly platform for synthetic biology. Synth. Biol. 5, ysaa010 (2020).

15. Storck, M. et al. BASIC: a new biopart assembly standard for idempotent cloning provides accurate, single-tier DNA assembly for synthetic biology. ACS Synth. Biol. 4, 781–787 (2015).

16. Engilad, B. & Zhao, H. Programmed DNA-guided artificial restriction enzymes. ACS Synth. Biol. 6, 752–757 (2017).

17. Marmur, J. & Doty, P. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. 5, 109–118 (1962).

18. Engilad, B. Development and applications of programmable DNA-guided Argonauta-based artificial restriction enzymes. Doctoral dissertation, University of Illinois at Urbana-Champaign (2020).

19. Shi, S., Liang, Y., Zhang, M. M., Ang, E. L. & Zhao, H. A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in Saccharomyces cerevisiae. Metab. Eng. 33, 19–27 (2016).

20. Shi, S. et al. Metabolic engineering of a synergetic pathway for n-butanol production in Saccharomyces cerevisiae. Sci. Rep. 6, 1–10 (2016).

21. Sheng, Y., Mancino, V. & Birren, B. Transformation of Escherichia coli with large DNA molecules by electroporation. Nucleic Acids Res. 23, 1990–1996 (1995).

22. Lehman, G. J. et al. A high-throughput assay for the comprehensive profiling of DNA ligation fidelity. Nucleic Acids Res. 44, e14–e14 (2016).

23. Untergasser, A. et al. Primer3—new capabilities and interfaces. Nucleic Acids Res. 40, e115–e115 (2012).

24. HamelRad, M., Weisberg, S., Chao, R., Lian, J. & Zhao, H. Highly efficient single-pot scalable Golden Gate assembly. ACS Synth. Biol. 8, 1047–1054 (2019).

25. Potapov, V. et al. Comprehensive profiling of four base overhang ligation fidelity by T4 DNA ligase and application to DNA assembly. ACS Synth. Biol. 7, 2665–2674 (2018).

26. Zhang, Y., Werling, U. & Edelmann, W. SLiCE: a novel bacterial cell extract-based DNA cloning method. Nucleic Acids Res. 40, e55–e55 (2012).

27. Ma, X., Xiang, X. & Huo, Y. X. Developing a transformation-independent and unbiased qPCR assay to rapidly evaluate the determinants of DNA assembly efficiency. Engineering 5, 803–810 (2019).

28. Islam, M. et al. Optimizing T4 DNA polymerase conditions enhances the efficiency of one-step sequence and ligation-independent cloning. BioTechniques 63, 125–130 (2017).

29. Bordat, A., Hovefengbas, M. C. & German-Retana, S. Gibson assembly: an efficient and accurate multi-part DNA assembly method. ACS Synth. Biol. 8, e93 (2013).

30. Boob, A. & Zhao, H. Can deep learning solve the Cas9 dilemma? CRISPR J. 4, 118–120 (2021).

31. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. J. Mol. Biol. 215, 403–410 (1990).
35. Enghiad, B., et al. Source code to design guides and primers for PlasmidMaker is a versatile, automated, and high throughput end-to-end platform for plasmid construction. https://doi.org/10.5281/zenodo.6467756.

Acknowledgements
This work was supported by U.S. Department of Energy award DE-SC0018420. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy. We thank Jonathan Chow Ning for providing the zeaxanthin plasmid template and helpful discussions about initial design of experiments. The online tool BioRender (biorender.com) was used to create Fig. 1’s upper panel, Fig. 4a, b, and c, Fig. 5b, and Supplementary Fig. 8, 9.

Author contributions
B.E. and H.Z. conceived and designed the study. B.E., P.X., N.S., A.B., V.A.P., and H.Z. wrote the manuscript. B.E. developed PfAgo-based DNA assembly method as well as PfAgo mutant, performed all the method development experiments, and devised assembly design rules. P.X., N.S., and C.S. designed and developed all the robotic automation workflow and performed the construction of all 101 plasmids using iBioFAB. E.D.G. and S.L. helped with initial robotic scripting on Tecan FluentControl liquid handler. N.S. created the Python modules for worklist generation for liquid handling on iBioFAB and helped develop quality control scripts. A.B. developed the computational model for guide DNA, primer design, and selection of restriction enzymes for plasmid assembly verification. V.A.P., R.L., S.P., and S.L. developed the web interface and software pipeline for enabling plasmid construction workflows.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-30355-y.

Correspondence and requests for materials should be addressed to Huimin Zhao.

Peer review information Nature Communications thanks Guo-Qiang Chen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2022