RSC Chemical Biology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d5cb00317b

Scarless one-tube genome assembly *via* computationally optimized uracil-DNA glycosylase reactions

Rachel M. Carson,† Patrick M. Needham,† Pedro J. Fernandez Mendoza and Sam R. Nugen*

Synthetic biology enables the creation of systems such as bacteriophage (phage)-based biosensors, leveraging the innate specificity and efficiency of phages to rapidly identify pathogens. However, the current genome assembly and editing methods, including Gibson Assembly, Golden Gate Assembly, and CRISPR-Cas systems, have limitations that can hinder speed and flexibility, especially when complex modifications are needed. This study introduces a novel means for generating engineered bacteriophages through a one-pot, modular *in vitro* genome assembly platform utilizing uracil-DNA glycosylase, which allows genome modification without requiring extended overlaps, the removal of restriction enzyme sites, a Cas system, or homologous recombination. The design also minimizes the risk of secondary structure formation (*e.g.*, hairpins), allowing for a more efficient assembly of fragments. To demonstrate functional genome engineering, we incorporated a NanoLuc luciferase reporter gene into the T7 genome, producing a recombinant phage capable of detecting *E. coli*, a strategy consistent with our previous work on waterborne pathogen detection. This platform enables rapid and flexible synthetic genome construction with high functional assembly efficiency, with broad applications in phage engineering, biosensing, and synthetic biology.

Received 8th December 2025,
Accepted 6th May 2026

DOI: 10.1039/d5cb00317b

rsc.li/rsc-chembio

Introduction

There are two main objectives of synthetic biology. The first is utilizing unnatural molecules to produce new behaviours in biology. The second is to interchange biological parts to modify the system to function unnaturally.[1] To accomplish these two goals, it is imperative to continue developing methodologies that assist in generating these augmented organisms or systems. Many novel technologies have emerged to minimize error rates, maximize construct size, and reduce the time required to successfully develop functional synthetic genomes. Through synthetic biology, many prominent and new technologies have emerged, including therapeutics, energy, food, and diagnostics. This endeavour has produced robust strategies, including gibson assembly (GA),[2,3] golden gate assembly (GGA),[4] and CRISPR-Cas effector systems.[5,6]

Developing novel approaches for generating synthetic genomes is imperative for understanding the fundamental nature of biological systems, developing therapeutics, generating novel organisms, and harnessing biological systems for scaling production of desired products. In 2003, it was reported that PhiX174 could be constructed using synthetic oligonucleotides.[7]

This significant achievement demonstrated that increased tool development enabled the generation of synthetic genomes. Since the development of GA and GGA, these techniques have proven to be instrumental for the innovation of molecular cloning techniques and the production of engineered viral particles.

GA and GGA are methods that rely upon the assembly of parts and have been used to construct circular genomes, plasmids, or vectors. GA joins adjacent fragments that overlap by 20–100 bases. First, a 3′ exonuclease generates 5′ ssDNA overhangs on each fragment, allowing adjacent fragments to hybridize at these designed junctions. The length of the individual ssDNA can vary, as it is determined by enzyme activity. Therefore, small fragments are incompatible with this method as they may be completely digested. A DNA polymerase fills any gaps that the exonuclease may have formed after annealing adjacent fragments. Finally, a ligase joins the fragments, producing a continuous DNA molecule.[3,8] This method has previously been used to assemble large DNA pieces, but has resulted in mutations in the final product.[8,9] Alternatively, GGA uses Type IIs restriction enzymes to cut the ends of the fragments, resulting in ssDNA overhangs of ~4 bases, which are complementary

Cornell University Ithaca, NY, USA. E-mail: snugen@cornell.edu
† These authors contributed equally to this work.

to the overhang of the adjacent fragment. This assembly method has become very reliable for plasmids but has some issues with phage-size genomes. In a process known as ''domestication'', endogenous restriction enzyme recognition sites must be removed from the engineered genome, meaning the method is not scarless.[10] Removal of endogenous restriction enzyme sites alters the native genome sequence, introducing permanent sequence changes ('scars') that remain after assembly. Additionally, the relatively small sticky ends (~4 bases) can result in some fragments being omitted due to mismatches and limitations on specificity. Although both GA and GGA have been used to modify and produce mutant T7 phage particles before, a method that does not require domestication (scarless) and allows predefined T_m complementarity of ssDNA sticky ends on adjacent fragments would significantly improve the efficiency of assembling larger genomes.[9–11]

Other eminent methods for the generation of engineered systems involve the utilization of CRISPR-Cas effector systems. CRISPR-Cas systems have been previously shown to be especially efficacious in promoting homologous recombination (HR) for the engineering of genomes.[6,12] However, this engineering approach requires downstream selection of the mutants of interest, expression of the engineering system in a viable host, and optimization of the cutting machinery. These obstacles cost time and resources.

To enhance the synthetic biology toolkit, we developed a method that leverages Uracil DNA Glycosylase (UDG) for the building of fully modular and customizable synthetic genomes, which can then be rebooted through transformation. Specifically, we use Thermolabile Uracil-Specific Excision Reagent II (USER II) from New England Biolabs (NEB). It contains a UDG and Endonuclease III, which has lyase activity to cleave the DNA backbone after the excision of uracil by the UDG and leaves a 5′ phosphate after cleavage.[13] USER II was designed for various applications, including the selective removal of the second strand for directional RNA library prep, cleavage of NEBNext hairpin loop adaptors, DNA sequencing, and cloning.[14–17] In terms of cloning applications, USER-related technologies have been targeted for vector construction and expression applications.[17–20] These works focused on the incorporation of amplicons of a range of sizes from hundreds to low thousands of base pairs. USER II has not been previously used to assemble and engineer phage genomes *in vitro*.

Herein, we have developed a computational design and accomplished USER II assembly of a T7 genome (~40 kb) using uracil-containing primers. The T7 genome was divided into seven overlapping fragments using computationally optimized primer sets. The overlaps were at the primer locations, each containing a 5′ adenine and a 3′ thymine. The primers were synthesized with one internal and the 3′ thymine replaced by uracil. Therefore, the primer sequences were removed, leaving the complementary ssDNA. By precisely designing the primer T_m, which is influenced by the PCR buffer and components, we can determine the T_m of the ssDNA overlaps for the fragment assembly in the final *Taq* ligase reaction buffer. Annealing and ligation temperatures were set ~8 °C below the least stable overhang T_m. This reduces the likelihood of secondary structure formation or incorrect assembly. While this will result in efficient and specific assembly, the method relies on precise T_m calculations in PCR and *Taq* ligase buffers. We use established T_m models to ensure calculated T_m precision and experimentally determine corrective offsets for the PCR and *Taq* ligase buffers using melting curve analysis.[21–26] Including an internal uracil in the primers prevents re-annealing due to a T_m drop of the shorter oligo lengths. Additionally, previous work in optimizing uracil excision in oligo sequences is incorporated into the coding to increase assembly efficiency.[27–29] This work serves to illustrate the greater applicability of USER-based technologies, beyond vector cloning, to generate infectious and engineered viral particles from large amplicon fragments. To serve as a proof-of-concept, the reporter gene NanoLuc luciferase, from a T7-Select generated T7 phage, is incorporated into a wild-type T7 bacteriophage using this USER-based strategy.

Results and discussion

Fragment design

To support the modular assembly of the ~40 kb circularized T7 genome, the genome sequence was divided into seven overlapping fragments to make set FX.1, including one that spans the concatemer junctions (Fig. 1a and b). FX.1 met the defined thermodynamic and structural criteria and was optimized for high predicted UDG efficiency, based on empirical data.

For each fragment junction, a 1 kbp window was analysed to identify suitable overlap sequences. Within each window, at least 21 candidate sequences (primer sets) were identified that satisfied the constraints for T_m, primer-dimer avoidance, and USER II cloning compatibility.

The thermal stability of the ssDNA overhangs generated by USER II digestion was evaluated to promote efficient annealing during the assembly process (Fig. 1c). In the FX.1 fragment set,

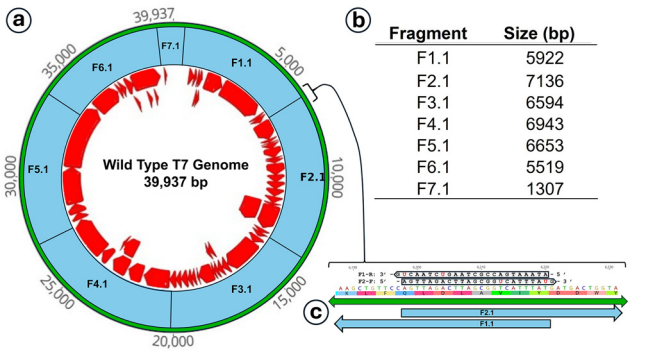


Fig. 1 Reassembly strategy for the T7 genome. (a) Circular map of the 39 937 bp T7 genome annotated with overlapping DNA fragments (blue) and gene locations (red). The reassembly scheme comprises seven fragments, including F7.1, which bridges the genomic termini to enable circularization. (b) Lengths of the individual genome fragments used in the assembly. (c) Enlarged view of the overlapping region, designed using computational primer design optimized for USER (Uracil-Specific Excision Reagent) assembly. Primers include uracil residues, indicated in red, to facilitate seamless fragment joining.

Table 1 Calculated melting temperatures of ssDNA fragment termini

Overlap	Overlap sequence	Calculated T_m [a]
F7.1–F1.1	ACACGGCGAATAGCCAT	59.9
F1.1–F2.1	ACAATGCTTAAGGTCGCT	57.6
F2.1–F3.1	AAACCAAGCTCAACAGCTTT	59.6
F3.1–F4.1	AACTAAGCAAGAAACTTACAGAGT	59.0
F4.1–F5.1	AAGTACACAATGGCTGGT	57.2
F5.1–F6.1	AGAAAGAAGGCTTCCTTCGT	60.1
F6.1–F7.1	ATGCACTATAGACCACGGAT	59.2

[a] T_m was calculated using the nearest Neighbor model and the reaction buffer composition.

the calculated average T_m of the ssDNA overlaps in the assembly buffer (see USER II optimization in Methods) is 58.9 °C ± 1.2, and a minimum T_m of 57.2 °C (Table 1). From these data, an annealing temperature of 50.0 °C was selected for the downstream assembly of digested fragments.

These values guided the selection of annealing-ligation temperatures, approximately 8 °C below the lowest T_m in each set, thereby optimizing hybridization efficiency while minimizing the formation of secondary structures or mismatches. Consistent overlap stability across fragments validates the computational primer design strategy and supports the feasibility of accurate, directional genome assembly using USER II cloning. The primers used for synthesizing the fragments are listed in Table S1.

Fragment synthesis

Initial validation of fragments generated through PCR from the phage DNA (NRG-P0097, rebooted T7 phage from Bocca Scientific, and NRG-P0098, T7 select phage containing the NanoLuc luciferase gene, respectively) was done using gel electrophoresis. NRG-P0097 was used to produce fragments F1.1–F3.1 and F5.1–F7.1, while NRG-P0098 was used for fragment F4.1. Incorporating F4.1 would provide the rebooted phage the ability to express NanoLuc luciferase upon infecting *E. coli.* This serves as an efficient experimental means for illustrating the efficacy of this assembly method for engineering phages because the NRG-P0098 was initially generated using the T7-Select system, which relies upon a modified T7 genome and restriction enzyme digestion for the incorporation of the intended modification.[30]

Gel electrophoresis was utilized to assess amplicon lengths of the intended fragments. The approximate sizes of the fragments were confirmed *via* sequencing (Fig. 2 and 3). This screening confirmed minimal off-target amplification, as there are no intense secondary bands that further complicate the purification of the amplicons of interest. Fig. 2 illustrates all seven fragments and their range of sizes. Notably, there is a secondary band in F1.1, which requires this fragment to be purified *via* gel extraction. While there are some other secondary band formations in F4.1 and F6.1, these bands were not a concern, as the maximum length for DNA purified by the Qiagen PCR Purification kit used was 10 kb (see methods).

Additionally, sequencing confirmed the fragments (see Fig. 1b for expected fragment sizes). Sequencing revealed that mutations were introduced through PCR compared to the wild-type genome (Fig. 3a–g). F6.1 has a single substitution at the

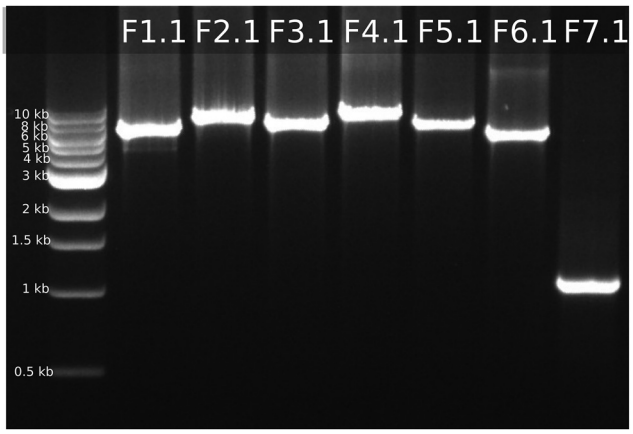


Fig. 2 Gel electrophoresis of individual fragments from the FX.1 assembly. Agarose gel showing DNA fragments F1.1 through F7.1 alongside a 1 kb DNA ladder. The distinct bands confirm successful amplification and expected fragment sizes for each component of the FX.1 assembly.

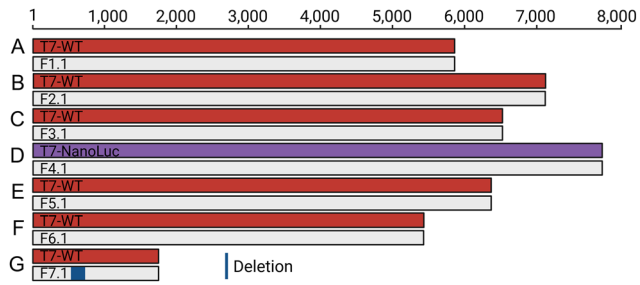


Fig. 3 Sequencing alignments of FX.1 amplicons compared to reference phage genomes. Pairwise alignments between FX.1 assembly fragments and their corresponding regions in the reference genomes. Black boxes indicate nucleotide mismatches. (A) FX.1 Fragment 1.1 (grey) aligned with the NRG-P0097 genome (red). (B) FX.1 Fragment 2.1 (grey) aligned with the NRG-P0097 genome (red). (C) FX.1 Fragment 3 (grey) aligned with the NRG-P0097 genome (red). (D) FX.1 Fragment 4.1 (yellow) aligned with the T7 NanoLuc genome (purple). (E) FX.1 Fragment 5.1 (grey) aligned with the NRG-P0097 genome (red). (F) FX.1 Fragment 6.1 (grey) aligned with the NRG-P0097 genome (red). (G) FX.1 Fragment 7.1 (grey) aligned with the NRG-P0097 genome (red). The blue line indicates deletions.

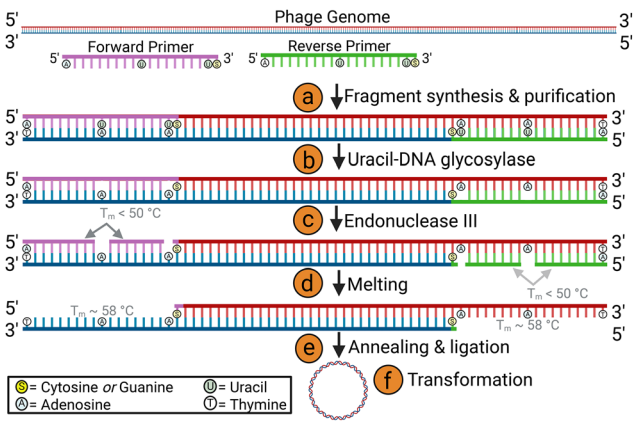
last base pair of the sequence, where NRG-P0097 contains an A and the fragment contains a G (Fig. 3f). Since it is not part of a coding region, the substitution does not affect the formation of the phage. It can also be explained as a sequencing error, as the primer region cannot be sequenced due to the presence of uracil in the amplicons. F7.1 shows a gap between 632–741 compared to the NRG-P0097 (Fig. 3g). F7.1 contains the concatemer junction, which is the region that contains the mature linkage between the right end and left end of the T7 DNA and is essential for phage genome processing and packaging.[31] Therefore, as long as the concatemer region has the correct 5′ and 3′ joining ends, the phage will be appropriately processed and packaged.[31] This confirms that the reassembled genome is capable of producing functional phage particles. As previously mentioned, the sequencing data could not confirm the presence of uracil in the amplicons; however, the integration can be verified through the assembly of corresponding

fragment junctions. Without digestion by UDG, ligation of the junctions cannot occur.

USER II optimization

To continue to examine this assembly method, an investigation into the activation of the 5′ and 3′ ends of the DNA fragments *via* USER II was essential. Without proper digestion of the ends, the full assembly would not be formed, as this method is designed to generate a circularized phage genome (Scheme 1). To accomplish this investigation, single-read qPCR analysis was conducted with a fluorescently tagged fragment on both the 5′ (SUN fluorophore) and 3′ (TYE665) ends (Fig. SI-1). As the ends of the fragments are digested, there is a decrease in fluorescence. This design helps to discern optimal conditions for activating the fragments for proper assembly.

The single-timepoint qPCR reads demonstrate a loss in relative fluorescence units (RFU) as the digestion time with UDG increases (Fig. SI-1a). This is in accordance with the expected outcome of digestion time because the UDG removes the uracil bases, which were inserted *via* PCR amplification. The results from Fig. SI-1b reveal a constant level of fluorescence for sample NC. Sample NC does not contain any uracil bases, meaning there is no digestion to cause a loss in fluorescence. Sample PC does not contain any fluorophores and, as a result, exhibits very minimal fluorescence signal. Sample F1.1 incorporates both uracil and respective fluorophore tags and illustrates that there is a general trend of a decrease in fluorescence as related to digestion time, with increasing deviation in signal as digestion time increases. See Table S2 for primer sequences. As a result of this data, longer digestion times were explored to further examine if the efficiency of the enzymatic digestion could be optimized.



Scheme 1 Assembly scheme for assembling a T7 genome from fragments. (a) Seven overlapping PCR fragments were generated using uracil-containing primers. (b) Fragments were combined in a single tube with UDG, endonuclease III, and Taq DNA ligase at 37 °C. (c) UDG excised uracil bases, and endonuclease III cleaved the resulting abasic sites. (d) The mixture was heated to 60 °C and cooled to 50 °C to denature secondary structures and promote specific annealing. (e) Complementary single-stranded overhangs annealed, and Taq ligase sealed the nicks to form a circular genome. (f) The assembled genome was introduced into host cells, producing infectious recombinant phages. *UDG and endonuclease III are contained in the USER II cocktail.

Fragment assembly

To better define the assembly of the fragments, sequential ligation junctions were examined for all fragments. This was done to optimize the ligation time and assess the "stickiness" of the USER II-generated overhangs. All fragments were UDG-digested for 45 minutes, then ligated at a 2× enzyme concentration. This information would inform the efficacy of the junction or overhang design and further illustrate the efficacy of USER II digestion. Success of the ligation was judged through the appearance of an upper band forming at approximately 15–20 kb. This indicates the joining of the two fragments of interest for F1.1 + F2.1, F2.1 + F3.1, F3.1 + F4.1, and F5.1 + F6.1 (Fig. 4). The band of interest, which shows ligation or joining of the fragments for F6.1 + F7.1 and F7.1 + F1.1, is the higher band, indicating a size difference of approximately 1 kb. This is the additive product of F7.1, which is 1307 bp, and F6.1 and F1.1, which are 5922 bp and 5519 bp, respectively (Fig. 4b). Based on this information, ligation is occurring for all fragment junctions.

Examining the results of this study, all fragments are joined *via* incubation with T4 ligase (NEB) (Fig. 4). ImageJ aided analysis of the ligated band formation relative to time illustrates that running the ligation for 30 minutes to 1 hour is more optimal than 15 minutes (Fig. SI-2). Given that 30 minutes and one hour ligation times were more optimal for band formation, one hour is used for all subsequent assemblies.

Genome reboot

USER II enzyme concentration was further optimized by comparing the phage titers (PFU mL^{-1}) obtained from rebooting fragment assemblies. Rebooting was performed using heat shock transformation to assess the feasibility of reviving phage from non-purified genome samples. Transformations were successful at USER II concentrations of 1× and 3.5× with Taq ligase present, and at 2× and 5× both with and without Taq ligase. 0× Taq replicates were examined to discern more information regarding the strength and specificity of the USER II generated overhangs. The recovery of infective phage in the absence of Taq ligase suggests that endogenous host repair mechanisms may seal remaining nicks after transformation, although this mechanism was not directly tested. The recombinant phages were subsequently named NRG-P0099 (5× USER II with 0× Taq), NRG-P0100 (2× USER II with 0× Taq), NRG-P0101 (1× USER II

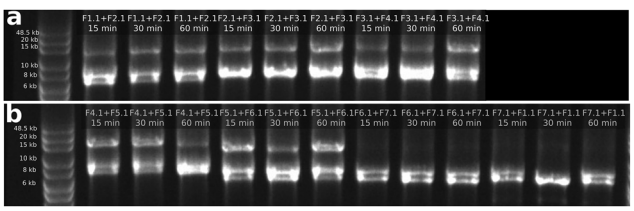


Fig. 4 Gel electrophoresis visualizing the annealing and ligation efficiency of individual fragment junctions from the FX.1 assembly run as separate reactions. The reaction was performed at varying times (15, 30, and 60 min). (a) Assembly of fragment pairs F1.1 + F2.1, F2.1 + F3.1, and F3.1 + F4.1. (b) Assembly of fragment pairs F4.1 + F5.1, F5.1 + F6.1, F6.1 + F7.1, and F7.1 + F1.1.

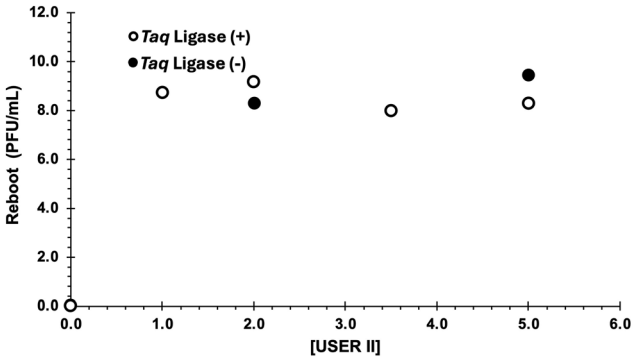


Fig. 5 Comparison of phage titers (PFU mL^{-1}) following transformation reboot between different USER II Concentrations and the presence of Taq ligase. The white circles represent the USER II Concentration (0×, 1×, 2×, 3.5×, and 5×) titers that included Taq ligase in their reaction mix. Black circles indicate the USER II Concentration (2× and 5×) titers without Taq ligase in their reaction mix.

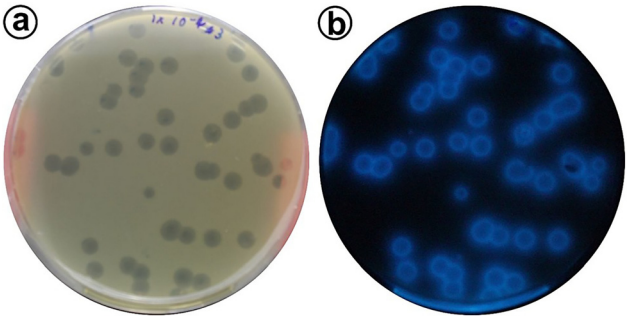


Fig. 6 Visual confirmation of recombinant phage production using NanoLuc reporter expression. (a) Representative plaque assay plate showing phage plaques formed following transformation of the assembled synthetic T7 genome into *E. coli*. (b) The same plate imaged after the addition of Nano-Glo® substrate, removal of ambient light, and extended exposure time, revealing luminescent plaques. The uniform luminescence across all plaques demonstrates functional expression of the NanoLuc reporter in recovered phage particles.

with 1× Taq), NRG-P0102 (2× USER II with 1× Taq), NRG-P0103 (3.5× USER II with 1× Taq), and NRG-P0104 (5× USER II with 1× Taq). In contrast, no viable phage was recovered at 0× USER II concentration, indicating that the enzyme is essential for successful genome assembly (Fig. 5 and Table S3). Differences in phage titer were observed across USER II concentrations and between conditions with or without Taq ligase. To statistically evaluate these effects, a two-way ANOVA was performed on log-transformed titer values (Fig. 5), testing two null hypotheses: first, that variations in USER II concentration do not affect phage titer, and second, that the presence of Taq ligase, in combination with USER II concentration, does not influence the titer. Both null hypotheses were rejected, with *p*-values below 0.05 (Table S4), indicating statistically significant effects.

To further investigate these differences, Tukey-adjusted pairwise comparisons of mean efficiencies were conducted. Three additional null hypotheses were tested: that USER II concentration does not significantly affect phage titer in the presence of Taq ligase; that USER II concentration has no effect in the absence of Taq ligase; and that the presence or absence of Taq ligase does not influence titer. All three hypotheses were rejected, with *p*-values below 0.05 (Table S4), confirming that both USER II concentration and Taq ligase significantly impact rebooting efficiency. As shown in Fig. 5, the highest mean titer (log-transformed) among samples with Taq ligase was observed at a 2× USER II concentration (NRG-P0102). In contrast, the highest titer without Taq ligase occurred at 5× USER II concentration (NRG-P0099). The difference in number of infective particles produced from USER II digested DNA is most likely an artifact of limiting reactants. Fragment samples are for USER II assembly are small volumes of DNA at high concentrations. Therefore, a small variance of one fragment can result in lower plaquing numbers as there are not as many complete assemblies between the samples.

The Nano-Glo® assay was used as a rapid screening to confirm integration and expression of the NanoLuc reporter gene. Notably, there are no apparent plaques that indicate the lack of NanoLuc expression (Fig. 6). All observed plaques (44/44) expressed NanoLuc, indicating successful recovery of functional recombinant phage particles, as is expected from a genome assembly-based method. Potential homologous recombination or expression of luciferase from *E. coli* alone is not evident in the results of these studies because viable *E. coli* cells do not generate luminescence. Rather, the luminescence is localized within the phage clearings, further confirming the generation of luciferase from the intended phages. To confirm that this was not a result of some other mode of contamination, 0× USER II was examined, this sample did not utilize USER II enzyme, so all DNA was theoretically undigested. This did not generate any infective phage particles. Plaquing results from 0× USER II are shown in Fig. SI-3 which indicate that there are no infective phages formed without the presence of USER II to catalyze the assembly.

Following the Nano-Glo® assay, single plaques were picked and analysed *via* sequencing. The alignment of the sequences indicates that the synthetic genomes are assembled as intended (Fig. 7). All fragments, except F4.1, are derived from NRG-P0097 (red sequence). Additionally, the insertion of the NanoLuc gene from NRG-P0098 was successful in all mutants in F4.1. The insertion of NanoLuc into our intended mutant was to serve as a proof-of-concept to produce an engineered infective phage particle. Based on this information, this method works effectively for incorporating desired insertions while maintaining the components of the initial genome used as the template.

Furthermore, upon analysing each sequence, the sequences with the single-nucleotide polymorphisms (SNPs) were 5× USER II (NRG-P0104) and 1× USER II (NRG-P0101), while 2× USER II (NRG-P0102) did not have any SNPs (Fig. 7 and Table S5). The exact substitution information suggests that there are likely sequencing errors that are inherently present as a result of the extended sequence length (Fig. 7). Oxford Nanopore states that the accuracy of their instrumentation is $>99.99\%$, leaving an error rate of $<0.01\%$.[32] The length of the

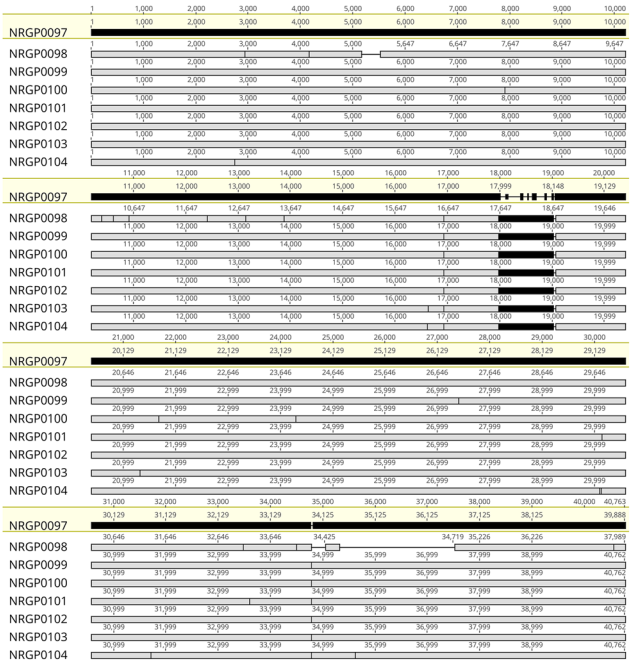


Fig. 7 Whole-genome alignment of USER-assembly T7 phages relative to the wild-type genome NRG-P0097. Each horizontal bar represents the full genome of a phage variant (NRG-P0097 through NRG-P0104). Vertical black bars denote single-nucleotide polymorphism (SNPs) relative to the wild-type T7genome (NRG-P0097), and horizontal thin black lines represent deletion events. Black boxes indicate insertions, the NanoLuc reporter cassette, originating from the NRG-P0098 template and successfully incorporated into all engineered genomes. Each genome is shown at an equal scale to allow direct visual comparison of mutation positions and overall genome integrity.

genomes is such that there will be 4.0763 errors, at most, per T7 phage genome analysed. This holds true for the mutants generated through USER II-based assembly, as the errors shown in Fig. 7 are all three SNPs or fewer. This further supports that junction assembly is effective with or without *Taq Ligase* and does not introduce further SNPs. Additionally, these SNPs, if not sequencing errors in their entirety, were not detrimental to the infectivity of the produced engineered T7 particles, as they are still infective (Fig. 7).

Materials & methods

Cultures

Bacterial cultures were propagated in Luria-Bertani (LB) broth from Thermofisher Scientific (Waltham, MA, USA) at 37 °C while shaking at 90 to 250 rpm overnight (~16 hours). The host used for phage propagation was NRG-234, originally called ATCC BAA-1025 (*E. coli* BL21), which was procured from the American Type Culture Collection (ATCC) (Manassas, VA, USA). The transformation mixture was grown on NEB 10-beta Competent *E. coli* (NEB, Ipswich, Massachusetts, USA), which was grown in an overnight culture. Wild-type phage T7 was procured from rebooting the phage DNA from Bocca Scientific, Dedham Massachusetts, USA, and subsequently renamed NRG-P0097. T7 phage containing the Nanoluc-CBM gene, NRG-P0098, from Hinkley *et al.* (2018), was also used.[30]

Chemicals

The chemicals used in various protocols came from different sources. Phenol-chloroform-isoamyl alcohol was obtained from Invitrogen, Waltham, Massachusetts, USA. Molecular biology grade chloroform was purchased from MilliporeSigma, Burlington, Massachusetts, USA. ThermoFisher, Waltham, Massachusetts, USA, supplied 100% molecular-grade ethanol and 100% molecular-grade isopropanol. MgCl2 came from Sigma-Aldrich, St. Louis, Missouri, USA, while CaCl2 was acquired from VWR, West Chester, PA, USA. The Nano-Glo® substrate was obtained from Promega, Madison, Wisconsin, USA, and 1,4-Dithiothreitol (DTT) was supplied by Roche, Indianapolis, IN, USA. NEB, Ipswich, Massachusetts, USA, provided DMSO, NAD^+, and dNTPs.

Sequencing

Fragment sequencing for PCR amplicons and Whole-genome sequencing (WGS) was performed by Plasmidsaurus using Oxford Nanopore Technologies. PCR amplicon samples were processed under the ''Standard Purified' service category, and WGS samples were processed under the ''Big Purified'' service category, utilizing linear/PCR sequencing protocols. Custom analysis and annotation were provided as part of the service.

Phage propagation

The Phage on Tap method, as described by Bonilla and Barr,[33] was used to propagate phages. Briefly, high-titer lysates were generated by infecting *E. coli* cultures in either liquid or plate-based formats. For liquid propagation, phages were added to log-phase bacterial cultures in LB broth supplemented with calcium and magnesium and then incubated until lysis occurred. For plate lysates, individual plaques were picked, amplified on agar overlays, and harvested with SM buffer. Lysates were clarified by centrifugation and stored at 4 °C for downstream purification.

Picked isolation

A pipette tip was used to pick a single plaque from an LB plate, and it was placed in an Eppendorf tube containing 1 mL of SM buffer to mix by pipetting up and down approximately five times. After removing the pipette tip, the tube was vortexed for five minutes. A double agar overlay plaque assay was utilized to propagate the picked plaques.

One milliliter of NRG-234 bacterial suspension and 1 mL of the picked phage lysate were added to 5 mL of melted LB top agar (0.8% w/v), which had cooled to 55 °C. After mixing the solution by inverting the tube several times, it was poured onto LB Miller bottom agar (1.4% w/v) in circular plates to solidify at room temperature. The plates were then inverted and incubated overnight at 37 °C.

After incubation, 5 mL of SM buffer was added to the top of the plate and placed on a tabletop shaker at 70 rpm for approximately 4 hours at room temperature. The buffer was collected from the plate and transferred to a 50 mL conical

tube. It was purified by centrifuging the cell debris at 3260× *g* for approximately 30 minutes. A 0.22 μm pore size PDVF vacuum filter from Corning (Corning, NY, USA) was then used to filter the supernatant. The phage lysate was stored at 4 °C until needed.

Liquid propagation

To propagate a large volume of phage lysate, 100 mL of LB broth was supplemented with 1 mM $CaCl_2$ and $MgCl_2$, then inoculated with 1 mL of an overnight culture of NRG-234. The culture was incubated at 37 °C with agitation (~90 rpm) until it reached an OD_{600} greater than 0.2. Next, 100 μL of high-titer phage lysate (>10^8 PFU mL^{-1}) was added to the liquid culture and incubated at 37 °C with agitation until the lysate cleared (~2 hours).

The phage supernatant was purified by first pelleting the cell debris through centrifugation at 3260× *g* for approximately 30 minutes. A 0.22 μm pore size PDVF vacuum filter from Corning (Corning, NY, USA) was then used to filter the supernatant. After filtration, an Amicon 100 kDa MWCO ultra-centrifugal filter device (MilliporeSigma, Burlington, MA, USA) was utilized to purify and concentrate the phage supernatant. To purify and concentrate the phage supernatant, 15 mLs were pipetted into the Amicon device and centrifuged at 3260× *g* for the time required for the liquid to pass through the filter, approximately 10–30 minutes, and this process was repeated until all the supernatant was filtered. SM buffer (5.8 g NaCl, 2.0 g $MgSO_4{\cdot}7H_2O$, 50 mL 1 M Tris-HCl pH 7.4, in 1 L dH_2O) was used to wash the filter three times, at 15 mL each time, to purify the phage lysate. At least 250 μL of liquid was retained in the filter to prevent it from drying out after each centrifugation during the purification and concentration steps. To remove the phages from the filter, 1 to 1.5 mL of phage lysate was left during the final wash. The phage lysate was extracted from the Amicon device after being vortexed for approximately 5 minutes to release any concentrated phages from the filter.

Double agar overlay plaque assays were employed to determine the titer of the purified phage lysate.[34] SM buffer was used to prepare a 1 : 10 dilution series of phage lysate. Two hundred microliters of NRG-234 bacterial suspension and 100 μL of phage lysate were added to 5 mL of melted LB top agar (0.8% w/v), which had been cooled to 55 °C. After mixing the solution by inverting the tube several times, it was poured onto LB Miller agar (1.4% w/v) in circular plates to solidify at room temperature. The plates were then inverted and incubated overnight at 37 °C.

Phage DNA extraction

Two DNA extraction protocols were used to isolate the phage DNA. The extended protocol of Norgen Biotek's Phage Genome Isolation Kit (Norgen Biotek, ON, Canada) was employed for the initial phage DNA extraction and sequencing of the optimized USER II concentration phages.

To extract phage DNA, phenol-chloroform-isoamyl alcohol precipitation was performed. First, 10 μL of DNase I was added to 1 mL of phage (~10^{10} PFU mL^{-1}) in a 1.5 mL Eppendorf tube and incubated at room temperature for 15 minutes to remove residual bacterial DNA. The mixture was then incubated for 5 minutes at 75 °C to inactivate DNase I. Next, 4 μL of proteinase *K* was added and incubated at 55 °C for 30 minutes to degrade contaminant proteins. An equal volume of phenol-chloroform-isoamyl alcohol (24 : 24 : 1) was added to the treated phage, mixed on a rotor mixer for 5 minutes at 40 rpm. This mixture was centrifuged for 5 minutes at 16 200× *g* to create a clear separation between the aqueous and organic layers. The aqueous (top) layer was removed and placed into another 1.5 mL microcentrifuge tube. An equal volume of phenol-chloroform-isoamyl alcohol was reapplied, and the mixing and centrifugation were repeated. From this point forward, the procedure is the same as traditional ethanol precipitation. The aqueous layer was again transferred to another 1.5 mL tube, to which an equal volume of chloroform was added and mixed on a rotor mixer for 5 minutes at 40 rpm. This mixture was centrifuged at 16 200× *g* for 5 minutes. This chloroform wash step is repeated. The top aqueous layer was then placed in a new 2 mL Eppendorf tube, and 0.1 volume of 3 M sodium acetate was incorporated. After adding 2.5 volumes of −20 °C chilled 100% ethanol, the solution was mixed for 5 minutes at 40 rpm. The sample was spun down at maximum speed in the centrifuge for 30 minutes at 4 °C to form a DNA pellet. The supernatant was removed, and the pellet was dislodged with 800 μL of 70% ethanol. Centrifugation was repeated at 16 200× *g* for 30 minutes at 4 °C, followed by another wash with 70% ethanol and centrifugation for 30 minutes at 4 °C. After removing the 70% ethanol supernatant, 1 mL of 100% ethanol was added to resuspend the DNA pellet, followed by centrifugation for 30 minutes at 4 °C. The supernatant was removed, and the tube was left open to allow the residual ethanol to evaporate. Once dry, 30 μL of nuclease-free water was used to resuspend the sample. The DNA concentration and absorbance 260/280 and 260/230 ratios were quantified using a Nanodrop.

Primer and fragment design

To facilitate the assembly of the approximately 40 kb T7 bacteriophage genome, the genome was segmented into seven overlapping fragments, including one that spans the concatemer junctions formed during replication. The design facilitates genome circularization during *in vitro* assembly, mimicking concatemeric replication intermediates.

An automated Excel VBA workflow was developed to support primer design and fragment segmentation. This workflow enumerates, scores, and ranks primer pairs around each intended junction using nearest neighbor thermodynamics, buffer-specific corrections, and an uracil DNA glycosylase (UDG) efficiency lookup trained on a comprehensive sequence library. These features ensure compatibility with USER II cloning and promote high excision efficiency and well-matched overhang stability during annealing and ligation.

Each junction was scanned within a 1000 bp window to identify candidate primer sequences. Forward primer candidates were generated by sliding start positions and testing lengths from 18 to 32 nucleotides. Candidates were required

to meet several sequence-level constraints: both ends must be clamped with G or C, the second base from the 5′ end must be A, and the penultimate base must be T. These positions are replaced with uracil during synthesis to create defined USER II cleavage sites. Homopolymer runs of five or more identical bases were excluded. The reverse complement of each candidate was also evaluated to ensure both members of the primer pair met the same constraints.

To ensure directional assembly and reduce mis-annealing, each reverse primer was designed so that its reverse complement, offset by a GC clamp, served as the forward primer for the adjacent fragment. This strategy created defined overlaps between fragments based on primer sequences. Melting temperatures (T_m) of the single-stranded DNA overhangs generated by USER II digestion were calculated for the annealing-ligation buffer. The total DNA concentration used for the calculation was 2 pmol, comprising 1 pmol of USER II-digested fragments and 1 pmol of released primer fragments.

Melting temperatures were calculated using the nearest-neighbor thermodynamic model (DNA_NN3) with salt corrections and adjusted to align with New England Biolabs (NEB) calculator values for Q5U buffer conditions. The NEB-corrected T_m was validated against the NEB online tool and constrained to fall between 60 and 65 °C. The annealing-ligation temperature was selected to be approximately 8 °C below the T_m of the least stable overhang to promote efficient hybridization while minimizing secondary structure formation and mismatched base pairing. Primer design followed the criteria summarized in Table 2, with a target T_m of approximately 62 °C under Q5U buffer conditions (New England Biolabs, Ipswich, MA).[35]

To enhance USER II cleavage efficiency, the macro prioritized seven-nucleotide motifs centred on each thymine to be converted to uracil, using excision efficiency data from Hölz *et al.* (2019) rescaled from 1 to 100. This lookup was performed for both the internal uracil within the primer body and the terminal uracil near each end. The internal uracil was selected to prevent re-annealing and to introduce a second programmable nicking site. The macro scanned the primer body from position 8 to the end minus 7, evaluated all thymine-centred motifs, and retained the highest efficiency motif. The flanking 5′ and 3′ segments were exported for both strands to ensure matched thermodynamic properties after USER II treatment.

Additional constraints included minimizing primer-dimer and hairpin formation (T_m < 50 °C), maximizing UDG efficiency based on empirical data from Lietard and Somoza,[36] and favouring high G/C content adjacent to uracil residues. Primers were synthesized with the 3′ T and the selected internal T replaced by U, creating two independent excision sites per primer. After USER II treatment, the primer regions are removed and complementary single-stranded overlaps are revealed at each junction, enabling directional and scarless assembly.

Fragment synthesis

PCR amplification using Q5U DNA polymerase. PCR amplification of DNA fragments was performed using Q5U Hot Start High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA). DNA templates were derived from the NRGP0097 and NRGP0098 strains. Specifically, fragments F1-F3, F1.1-F3.1, F5-F7, and F5.1-F7.1 were amplified from NRGP0097, while fragments F4 and F4.1, containing the NanoLuc-CBM gene, were amplified from NRGP0098. Template DNA concentrations were adjusted to 1–1000 ng per reaction.

PCR reactions (50 μL) were prepared using a modified NEB protocol and included the following components: 10 μL of 5× Q5U Reaction Buffer, 1 μL of 10 mM dNTPs, 1 μL of DMSO, 0.5 μL (1 unit) of Q5U polymerase, 2.5 μL each of 10 μM forward and reverse primers, and 2 μL of template DNA. Nuclease-free water was added to reach the final volume. Primer sequences are listed in Table S1.

Thermal cycling conditions were as follows: initial denaturation step at 98 °C for 2 minutes; 30 cycles of denaturation at 98 °C for 10 seconds, annealing at 65 °C for 20 seconds, extension at 72 °C for fragment-specific durations, followed by a final extension at 72 °C for 5 minutes. PCR products were validated by electrophoresis on a 1% TAE agarose gel. Fragments F3.1, F4, F4.1, F5, F5.1, F6, and F6.1 were purified using the QIAquick PCR Purification Kit (Qiagen, Venlo, Netherlands). Fragments F1, F1.1, F2, F2.1, and F3 were gel-purified and extracted using the QIAquick Gel Extraction Kit (Qiagen).

An overnight *E. coli* strain NRG-234 culture was grown at 37 °C for 16 hours with shaking at 250 rpm. A 1 : 100 dilution of this culture was inoculated into fresh media and grown to an OD_{600} of 0.4 under the same conditions. Once this optical density was achieved, 10 μL of NRGP0097 phage stock ($\geq 10^{10}$ PFU mL^{-1})

Table 2 Constraints used for primer selection and pairing

The initial search constraints:
- $T_m \sim 62$ °C in PCR master mix (Nearest Neighbor Model)
- 5′ **A**, 3′ [a]**TS** and an additional internal **T**: the two fragments created after removal of the **T**'s must have T_m < 50 °C
- Primer dimer/hairpin T_m < 50 °C.
- High predicted efficiency for uracil-DNA glycosylase based on experimental dataset (graciously provided by Lietard and Somoza).
- High G/C content directly adjacent to the 5′ **A**, 3′ **T**, and internal **T**. This is a prioritized favourability but not strictly required.
- 40–60% G/C content: This would mark as favourable but not required.

Primer selection constraints:
- Distance between forward and reverse primers pairs defines fragments (<7.5 kb).
- Reverse-complement of reverse primer serves as forward primer of adjacent fragment resulting in the fragment overlaps being the primer sequences.
- T_m of all post-USER II, non-complementary ssDNA) T_m < 50 °C, preventing incorrect assembly order.

[a] TS = TC or TG, providing a GC clamp.

was added and incubated for 30 minutes at 37 °C while shaking at 90 rpm.

Following the infection, the culture was centrifuged at 3260× *g* for 15 minutes. The supernatant was discarded, and the pellet was resuspended in the residual media by vortexing. Fragment 7 was amplified directly from 2 μL of resuspended cells using colony PCR with the Q5U protocol mentioned above. The primer sequences are listed in Table S1. PCR cycling conditions included an initial denaturation step at 98 °C for 2 minutes, followed by 30 cycles of denaturation at 98 °C for 10 seconds, annealing at 67 °C for 20 seconds, and extension at 72 °C for 33 seconds, with a final extension at 72 °C for 5 minutes. The product was validated on a 1% TAE agarose gel and purified using the QIAquick PCR Purification Kit.

PCR product purification

The Qiagen QIAquick PCR purification kit and modified protocol were utilized to purify PCR amplicons (Qiagen, Venlo, Netherlands). In brief, nuclease-free water was heated to 50 °C. Five volumes of PB buffer were combined with one volume of the PCR reaction and 10 μL of 3 M sodium acetate (pH 5.0). The mixture was vortexed and spun down. This solution was then added to a QIAquick column and centrifuged at 16 200× *g* for 60 seconds. Afterward, the flow-through was discarded, and the column was returned to its catch tube. A total of 750 μL of Buffer PE wash buffer was added to the column twice, and the column was centrifuged at 16 200× *g* for 60 seconds each time. The column was centrifuged at 16 200× *g* for 10 minutes to remove any residual wash buffer. Following this centrifugation, the column was removed from the catch tube and transferred to a clean 1.5 mL Eppendorf tube. To elute the DNA, 30 μL of nuclease-free water heated to 50 °C was added to the center of the column membrane. The column was allowed to stand for 10 minutes before centrifugation at 16 200× *g* for 3 minutes. The DNA concentration was analysed on a Nanodrop.

Gel purification

The Qiagen QIAquick Gel Extraction kit (Qiagen, Venlo, Netherlands) and a modified protocol were used to isolate and purify the desired DNA bands from the PCR amplicons. A 100 mL solution of 1% TAE ultra-pure LMP agarose containing ethidium bromide was run with 8 PCR amplicons of each fragment at 75–90 V for approximately 2 hours. After the DNA bands were completely separated, the desired band was excised using a sharp scalpel or a Promega x-tracta™ Gel Extractor (Promega, Madison, WI, United States) and placed into a 15 mL tube. The gel slice was weighed, and 3 volumes of Buffer QG were added to one gel volume in the tube. The solution was then vortexed for about 5 seconds and incubated at 50 °C for 10 minutes in a shaking multithermal (programmable thermal incubator). Once the gel slices fully dissolved, 10 μL of 3 M sodium acetate, pH 5.0, was added to the solution and vortexed. One gel volume of isopropanol was subsequently added to the sample and vortexed. To bind the DNA, up to 750 μL of the sample was added to a QIAquick column in a 2 mL collection tube and centrifuged for 60 seconds at 16 200× *g*. The flow-through was discarded, and the DNA binding step was repeated until all the sample was bound to the column, accommodating a maximum of 400 mg of gel per spin column. The column was washed with 500 μL of QG buffer and centrifuged at 16 200× *g* for 1 minute, discarding the flow-through. A total of 750 μL of Buffer PE wash buffer was added to the column and allowed to stand for 5 minutes, followed by centrifugation for 1 minute at 16 200× *g* with the flow-through discarded. A final centrifugation at 16 200× *g* for 10 minutes was performed to remove the residual wash buffer. The column was then transferred to a clean 1.5 mL Eppendorf tube, and 30 μL of 50 °C heated nuclease-free water was added to the center of the column membrane. The column was allowed to stand for 10 minutes to enhance the DNA concentration before being centrifuged at 16 200× *g* for 3 minutes. The DNA concentration was analyzed on a Nanodrop. To validate the PCR amplicons, gel electrophoresis using 0.8% TBE gels with ethidium bromide was performed to visualize the amplicon lengths, using the NEB 1 KB ladder as a reference.

USER II optimization

USER II enzyme digestion time was optimized by testing the removal of fluorescent tags from the 5′ and 3′ ends of F1.1. The uracil-containing primers for F1.1 were labeled with an IDT fluorophore, SUN, on the 5′ end and TYE665 on the 3′ end, using the forward and reverse primers during PCR. Positive controls (PC) were prepared by performing PCR with uracil-containing F1.1 primers that were not fluorescently labelled. Negative controls (NC) were prepared using fluorescently tagged F1.1 primers that did not contain uracil. The USER II enzyme reaction mix was then prepared by combining 1 pmol of PCR amplicon, 1 unit (1 μL) of USER II enzyme, and 10× (5 μL) of rCutSmart™ Buffer. Nuclease-free water was added to bring the total volume to 50 μL. Six samples of each control and F1.1 were prepared for each timeframe. The tubes were then placed in a thermocycler and incubated at 37 °C for 0, 15, 45, and 90 minutes. The enzyme was subsequently heat-inactivated for 10 minutes at 65 °C, followed by placement on ice for 5 minutes. PCR purification was performed using the Qiagen QIAquick PCR purification kit to remove any residual fluorescent signal. Nuclease-free water was used as the blank, and 20 μL of it was pipetted into an Applied Biosystems MicroAmp™ Optical 96-Well Reaction Plate (ThermoFisher Scientific, Waltham, Massachusetts, USA). 20 μL of each purified PCR amplicon was pipetted into the plate. The plate was then sealed with an Applied Biosystems MicroAmp™Optical Adhesive Film (ThermoFisher Scientific, Waltham, Massachusetts, USA).

The qPCR QuantStudio™ 6 Flex system (ThermoFisher Scientific, Waltham, Massachusetts, USA) was used to analyze the fluorescent signal. To set up the qPCR machine, a new experiment was selected. In the Experiment properties, the QuantStudio™ 6 Flex system was chosen for the type of instrument used in the experiment. The 96-well (0.2 mL) plate was selected for the block used in the experiment. The type of experiment chosen was Presence/Absence, and *Taq*Man® Reagents were selected to detect the target sequence. Standard was chosen for the desired instrument properties. Pre-PCR Read and

Amplification were deselected at the bottom of the screen. In the Define section, two targets were selected. For target 1, the VIC reporter was chosen for SUN, and CY5 was selected for TYE665, as they match the wavelengths of the fluorescent tags, which the qPCR system does not have as specific options. No quenchers needed to be selected. The passive reference was set to None. Samples were labelled with Control, NC, PC, and F1.1, along with the duration of USER II enzyme digestion. In the Assign section, the 96-well plate image was labelled with sample names in their respective positions. All wells containing samples were also labelled with both reporters, marked as U for the task. In the Run Method section, the amount of reagent per well was set at 20 μL. The run time was set for 60 seconds, with the temperature unchanged. Raw data was collected and analyzed. The designed primers for this method are listed in Table S2 and the coding in Data S2.

USER II Enzyme concentration was optimized by testing the addition of 0 units, 1 unit, 2 units, 3.5 units, and 5 units of the enzyme in the one-pot system. The different concentrations were joined by a total of 1 pmol of Uracil-containing DNA, 1× (5 μL) rCurtsmart buffer (10×), 80 units (2 μL) *Taq* DNA Ligase, 1× (5 μL) of *Taq* DNA Ligase Reaction Buffer, 1 μL NAD^+, and 0.5 μL DTT in a 200 μL tube. The total reaction volume was brought up to 50 μL with nuclease-free water. The tubes were placed in a thermocycler at 37 °C for 3 hours. The temperature then increased to 60 °C and slowly decreased to 50 °C at a rate of −1 °C per minute. The temperature was held at 50 °C for 1 hour, and then the tube was placed in −20 °C freezer for storage until the time for rebooting the assembled genomes.

To assess the necessity of *Taq* DNA ligase in the one-pot assembly system, reactions were prepared with and without the enzyme using USER II concentrations of 2× and 5×. Each reaction contained 1 pmol of uracil-containing DNA fragments and an assembly buffer composed of 5 μL 10× rCutSmart buffer, 5 μL *Taq* DNA Ligase Reaction Buffer, 1 μL NAD^+, and 0.5 μL DTT. The final volume was adjusted to 50 μL with nuclease-free water. Reactions were incubated in a thermocycler at 37 °C for 3 hours, followed by a gradual temperature increase to 60 °C and a controlled cooling to 50 °C at a rate of −1 °C per minute. The temperature was then held at 50 °C for 1 hour to facilitate annealing and ligation. Assembled genomes were stored at −20 °C until rebooting.

Genome reboot

Transformation. The NEB 10-beta Competent *E. coli* (High Efficiency) (C3019H) and modified protocol were utilized to transform the modified T7 genome plasmid to reboot the phage. The competent *E. coli* cells were thawed on ice for 10 minutes. Next, 1–5 μL of 1 pg–100 ng of assembled DNA was added to the cell mixture and mixed by carefully flicking the tube 4–5 times without vortexing. The mixture was then placed on ice for 30 minutes before being heat shocked at 42 °C for exactly 30 seconds in a multitherm. Subsequently, the mixture was placed on ice for 5 minutes, after which 950 μL of room temperature NEB 10-beta/Stable Outgrowth Medium was gently added. The mixture was then incubated in a shaking environment (90 rpm) at 37 °C for 1–2 hours. Dilution sets of 1 : 10 dilutions to 10^{-8} were made for each transformation using SM buffer in triplicate. 100 μL of each dilution, starting with 10^0, was added to 200 μL of NEB 10-beta overnight culture. The mixture was then added to 5 mL LB top agar, mixed by inverting, and poured onto warmed LB plates. The plates were inverted after solidifying and placed in the 37 °C incubator overnight. The titer (PFU mL^{-1}) of each transformation mixture was calculated using plates containing 30–300 plaques and the formula.

Nano-Glo® assay

The Promega Nano-Glo® Luciferase Assay System and modified protocol were used to detect the presence and expression of the NanoLuc reporter gene. The Nano-Glo® Luciferase Assay Buffer was thawed to room temperature, while the Nano-Glo® Luciferase Assay Substrate was always kept on ice. A substrate-to-buffer ratio of 1 : 50 was used to prepare 1 mL of working solution per plate. This 1 mL of working solution was then used to flood the plate containing plaques after transformation, followed by a 3-minute incubation before imaging in a dark box using a Canon EOS Rebel T6 Camera (EF-S 18-55 IS II Kit) (Canon Inc., Ōita, Japan). The camera was set to TV: 30'', +/−: 4, and ISO: 1600 for visualization of luminescence. The camera settings for the before pictures were TV: 1/6, +/−: 0, and ISO: Auto.

ImageJ gel analysis

Gel analysis was conducted with Fiji/ImageJ version 2.16.0. Analysis began by converting the gel image into an 8-bit image, followed by cropping the gel from the top wells to the last marker on the gel ladder. The image was then inverted. Straight lines were drawn for each well (−90°), and a plot profile was generated for each respective well. Following this, the peaks were picked using the multiselect tool to determine the x and y coordinates (x being the distance migrated in pixels and y being the grey value of the point). The gray value of the points was then subtracted from the maximum value of 250 for relative intensity analysis. Relative intensity analysis was conducted by assuming that all the bands occurring in the well sum to a total value of 1, so the ligation band is divided by the sum of the grey values of all components in the well.

Statistical analysis

All statistical analyses were conducted using R Studio (Version 2025.05.1 + 513)[37] and the R programming language (R version 4.5.1 (2025-06-13)).[38] Library packages utilized include *readxl*[39] for importing and reading Excel files; *multcomp*[40] for the multiple comparisons of group means; *multcompView*[41] for the visualization of pairwise comparison significance; *emmeans*[42] for the *post hoc* comparisons of estimated marginal means; *ggeffects*[43] for generating and converting the estimated marginal means and adjusted predictions from regression models into data frames for visualization with *ggplot2.*[44]

Conclusions

This study demonstrates the successful development and application of a scarless, modular genome assembly method based

on uracil-DNA glycosylase chemistry. This is a further development of USER-based DNA technologies, expanding its utility from cloning and expression vector assembly to modification and complete assembly of larger genomes. The method was used to construct a fully synthetic T7 bacteriophage genome, which was rebooted into viable, infective phage particles. All plaques produced from the transformation expressed the NanoLuc reporter gene, confirming high assembly fidelity (44/44) and functional genome reconstruction. Sequencing verified correct fragment junctions and reporter insertion, with no evidence of assembly-induced errors at designed overlaps. The absence of mutations at fragment junctions indicates that USER II–mediated assembly does not introduce sequence errors at designed overlaps; observed substitutions elsewhere are most consistent with sequencing or PCR-amplification artifacts. These results confirm that the method is both reliable and efficient for phage genome engineering.

This platform offers several advantages over existing genome editing and assembly techniques. Unlike CRISPR-based systems, which rely on host-mediated recombination and are limited by phage-encoded anti-CRISPR proteins, this method enables direct editing of phage genomes, including regions that are toxic to the host or otherwise inaccessible. Compared to GGA, this approach is scarless and uses longer, computationally optimized, thermodynamically favorable overhangs, which improve assembly specificity and efficiency. GGA leaves scars because it requires domestication of the genome of interest, while this method does not require the removal of native restriction enzyme sites throughout the genome for successful modification. It also avoids the limitations of GA, which can be hindered by the formation of long ssDNA prone to secondary structure and by the degradation of small fragments due to exonuclease activity.

Importantly, this method is compatible with both transformation-based and cell-free rebooting strategies. Its design is compatible with *in vitro* TXTL-based rebooting strategies, suggesting potential applicability to fully cell-free workflows. This opens the door to rapid, high-throughput prototyping of synthetic phages and other DNA constructs without the need for live bacterial hosts. For these approaches and more, the 2× USER II concentration with Taq Ligase will serve as a sufficient starting point for assembling the genomes of interest.

As an enabling technology, this system has broad potential applications. It can accelerate the development of engineered phages for diagnostics, therapeutics, and biocontrol, particularly in the context of antimicrobial resistance. Beyond phages, the method may be adapted for the assembly of other large DNA constructs, including synthetic chromosomes, biosensors, and gene circuits. Exploration into application of this method with larger genomes is underway to further characterize this method. Its precision, flexibility, and independence from host constraints make it a valuable addition to the synthetic biology toolkit.

Author contributions

Rachel M. Carson: methodology, validation, formal analysis, investigation, data curation, writing – original draft, writing – review & editing, visualization. Patrick M. Needham: methodology, validation, formal analysis, investigation, data curation, writing – original draft, writing – review & editing, visualization. Pedro J. Fernandez Mendoza: software. Sam R. Nugen: software, conceptualization, writing – review & editing, visualization, supervision, project administration, funding acquisition.

Conflicts of interest

The authors have no conflict of interest.

Abbreviations

ANOVA	Analysis of variance
DTT	Dithiothreitol
LB	Luria-Bertani (broth or agar)
MWCO	Molecular weight cut-off
NAD^+	Nicotinamide adenine dinucleotide (oxidized form)
NC	Negative control
PC	Positive control
PCR	Polymerase chain reaction
PFU	Plaque forming units
qPCR	Quantitative PCR
RFU	Relative fluorescence units
SM buffer	Standard phage buffer
ssDNA	Single-stranded DNA
TAE/TBE	Tris-Acetate-EDTA/Tris-Borate-EDTA(buffer systems)
T_m	Melting temperature
TXTL	Transcription-translation (cell-free expression system)
UDG	Uracil-DNA glycosylase
USER	Uracil-specific excision reagent

Data availability

All data supporting the findings of this study are available within the article and its supplementary information files. This includes: Primer sequences used for fragment synthesis (Supplemental Table S1 and S2). Calculated melting temperatures and overlap sequences for genome assembly (Table 1). Sequencing alignments of assembled genome fragments (Fig. 3 and Fig. 7). Quantitative PCR fluorescence data for USER II digestion optimization (Supplemental Fig. SI-1). Phage titer measurements and statistical analyses (Fig. 5 and Supplemental Table S4). Raw gel electrophoresis images for fragment validation and ligation efficiency (Fig. 2 and 4). Raw reads for the whole-genome sequencing for recombinant phages below will be made available upon reasonable request:

NRG-P0097 → Wild-type T7 phage used as the template for fragments F1.1–F3.1 and F5.1–F7.1

NRG-P0098 → Engineered T7 phage containing the NanoLuc-CBM gene, used as the template for fragment F4.1

NRG-P0100 → Engineered phage assembled with 2× [USER II enzyme], 0× [*Taq* ligase]

NRG-P0099 → Engineered phage assembled with 5× [USER II enzyme] and 0× [*Taq* ligase]

NRG-P0101 → Engineered phage assembled with 1× [USER II enzyme] and 1× [*Taq* ligase]

NRG-P0102 → Engineered phage assembled with 2× [USER II enzyme] and 1× [*Taq* ligase] (highest fidelity)

NRG-P0103 → Engineered phage assembled with 3.5× [USER II enzyme] and 1× [*Taq* ligase]

NRG-P0104 → Engineered phage assembled with 5× [USER II enzyme] and 1× [*Taq* ligase]

Custom Excel VBA macros used for primer design and thermodynamic scoring are available upon reasonable request to the corresponding author (snugen@cornell.edu).

Supplementaray information (SI) is available. See DOI: https://doi.org/10.1039/d5cb00317b.

Acknowledgements

This work was supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) under grant R21EB034941. Additional funding was provided by the United States Department of Agriculture National Institute of Food and Agriculture (USDA-NIFA) through a Predoctoral Fellowship (AFRI 2022-11321) and a National Needs Fellowship (2023-67011-40359). The TOC figure was created using BioRender.

References

1 S. A. Benner and A. M. Sismour, Synthetic biology, *Nat. Rev. Genet.*, 2005, **6**(7), 533–543, DOI: 10.1038/nrg1637.

2 D. G. Gibson, G. A. Benders, K. C. Axelrod, J. Zaveri, M. A. Algire and M. Moodie, *et al.*, One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic Mycoplasma genitalium genome, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**(51), 20404–20409, DOI: 10.1073/pnas.0811011106.

3 D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison and H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nat. Methods*, 2009, **6**(5), 343–345, DOI: 10.1038/nmeth.1318.

4 C. Engler, R. Kandzia and S. Marillonnet, A One Pot, One Step, Precision Cloning Method with High Throughput Capability, *PLoS One*, 2008, **3**(11), e3647, DOI: 10.1371/journal.pone.0003647.

5 B. Martel and S. Moineau, CRISPR-Cas: an efficient tool for genome engineering of virulent bacteriophages, *Nucleic Acids Res.*, 2014, **42**(14), 9504–9513, DOI: 10.1093/nar/gku628.

6 P. Tao, X. Wu, W. C. Tang, J. Zhu and V. Rao, Engineering of Bacteriophage T4 Genome Using CRISPR-Cas9, *ACS Synth. Biol.*, 2017, **6**(10), 1952–1961, DOI: 10.1021/acssynbio.7b00179.

7 H. O. Smith, C. A. Hutchison, C. Pfannkoch and J. C. Venter, Generating a synthetic genome by whole genome assembly: φX174 bacteriophage from synthetic oligonucleotides, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**(26), 15440–15445, DOI: 10.1073/pnas.2237126100.

8 D. G. Gibson, G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson and J. Zaveri, *et al.*, Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome, *Science*, 2008, **319**(5867), 1215–1220.

9 E. M. Pulkkinen, T. C. Hinkley and S. R. Nugen, Utilizing *in vitro* DNA assembly to engineer a synthetic T7 Nanoluc reporter phage for *Escherichia coli* detection, *Integr. Biol.*, 2019, DOI: 10.1093/intbio/zyz005.

10 J. M. Pryor, V. Potapov, K. Bilotti, N. Pokhrel and G. J. S. Lohman, Rapid 40 kb Genome Construction from 52 Parts through Data-optimized Assembly Design, *ACS Synth. Biol.*, 2022, **11**(6), 2036–2042, DOI: 10.1021/acssynbio.1c00525.

11 D. R. Parker, A. P. Sikkema, R. K. Anderson, G. J. S. Lohman and S. R. Nugen, Recoded Bacteriophage Genome for Bio-Orthogonal-Enabled Concentration and Detection of *E. coli* in Drinking Water, *ACS Synth. Biol.*, 2026, **15**(1), 233–242, DOI: 10.1021/acssynbio.5c00665.

12 M. M. Duong, C. M. Carmody, Q. Ma, J. E. Peters and S. R. Nugen, Optimization of T4 phage engineering via CRISPR/Cas9, *Sci. Rep.*, 2020, **10**(1), 18229, DOI: 10.1038/s41598-020-75426-6.

13 Thermolabile USER® II Enzyme | NEB, [cited 2025 Sep 11]. Available from: https://www.neb.com/en-us/products/m5508-thermolabile-user-ii-enzyme.

14 Applications of USER® and Thermolabile USER II Enzymes | NEB. [cited 2025 Sep 12]. Available from: https://www.neb.com/en-us/applications/cloning-and-synthetic-biology/user-cloning/applications-of-user-and-thermolabile-user-ii-enzymes?_srsltid=AfmBOor7I_nAAEhWkRkUO0DqMxh1dsbczWdKF14keVogMEBzJLBMDjMJ.

15 J. Bitinaite and N. M. Nichols, DNA cloning and engineering by uracil excision, *Curr. Protoc. Mol. Biol.*, 2009, **86**, 3.21.1–3.21.16, DOI: 10.1002/0471142727.mb0321s86.

16 R. Vaisvila and J. Bitinaite, Gene synthesis by assembly of deoxyuridine-containing oligonucleotides, *Methods Mol. Biol.*, 2013, **978**, 165–171, DOI: 10.1007/978-1-62703-293-3_12 PubMed PMID: 23423896.

17 J. Bitinaite, M. Rubino, K. H. Varma, I. Schildkraut, R. Vaisvila and R. Vaiskunaite, USER™ friendly DNA engineering and cloning method by uracil excision, *Nucleic Acids Res.*, 2007, **35**(6), 1992–2002, DOI: 10.1093/nar/gkm041.

18 A. Rashtchian, C. G. Thornton and G. Heidecker, A novel method for site-directed mutagenesis using PCR and uracil DNA glycosylase, *PCR Methods Appl.*, 1992, **2**(2), 124–130, DOI: 10.1101/gr.2.2.124 PubMed PMID: 1477668.

19 P. E. Nisson, A. Rashtchian and P. C. Watkins, Rapid and efficient cloning of Alu-PCR products using uracil DNA glycosylase, *PCR Methods Appl.*, 1991, **1**(2), 120–123, DOI: 10.1101/gr.1.2.120 PubMed PMID: 1726854.

20 C. Smith, P. J. Day and M. R. Walker, Generation of cohesive ends on PCR products by UDG-mediated excision of dU, and application for cloning into restriction digest-linearized vectors, *PCR Methods Appl.*, 1993, **2**(4), 328–332, DOI: 10.1101/gr.2.4.328 PubMed PMID: 8324507.

21 J. SantaLucia, A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**(4), 1460–1465.

22 T. Xia, J. SantaLucia, M. E. Burkard, S. J. Schroeder, X. Jiao and C. Cox, *et al.*, Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs, *Biochemistry*, 1998, **37**(42), 14719–14735, DOI: 10.1021/bi9809425 PubMed PMID: 9778347.

23 H. T. Allawi and J. SantaLucia, Thermodynamics and NMR of internal GT mismatches in DNA, *Biochemistry*, 1997, **36**(34), 10581–10594, DOI: 10.1021/bi962590c.

24 J. SantaLucia, H. T. Allawi and A. Seneviratne, Improved nearest-neighbor parameters for predicting DNA duplex stability, *Biochemistry*, 1996, **35**(11), 3555–3562.

25 N. Sugimoto, S. Nakano, M. Katoh, A. Matsumura, H. Nakamuta and T. Ohmichi, *et al.*, Thermodynamic Parameters to Predict Stability of Rna/DNA Hybrid Duplexes, *Biochemistry*, 1995, **34**(35), 11211–11216.

26 N. von Ahsen, C. T. Wittwer and E. Schütz, Oligonucleotide melting temperatures under PCR conditions: nearest-neighbor corrections for Mg(2+), deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas, *Clin. Chem.*, 2001, **47**(11), 1956–1961PubMed PMID: 11673362.

27 M. E. Jørgensen, N. Wulff, M. Nafisi, D. Xu, C. Wang and S. K. Lambertz, *et al.*, Design and direct assembly of synthesized uracil-containing non-clonal DNA fragments into vectors by USERTM cloning, *Bio-Protoc.*, 2017, **7**(22), e2615–e2615.

28 H. J. Genee, M. T. Bonde, F. O. Bagger, J. B. Jespersen, M. O. Sommer and R. Wernersson, *et al.*, Software-supported USER cloning strategies for site-directed mutagenesis and DNA assembly, *ACS Synth. Biol.*, 2015, **4**(3), 342–349.

29 K. Hölz, A. Pavlic, J. Lietard and M. M. Somoza, Specificity and efficiency of the uracil DNA glycosylase-mediated strand cleavage surveyed on large sequence libraries, *Sci. Rep.*, 2019, **9**(1), 17822.

30 T. C. Hinkley, S. Garing, S. Singh, A. L. M. Le Ny, K. P. Nichols and J. E. Peters, *et al.*, Reporter bacteriophage $T7_{NLC}$utilizes a novel NanoLuc::CBM fusion for the ultrasensitive detection of*Escherichia coli*in water, *Analyst*, 2018, **143**(17), 4074–4082, DOI: 10.1039/c8an00781k.

31 Y. B. Chung and D. C. Hinkle, Bacteriophage T7 DNA Packaging: I. Plasmids containing a T7 replication origin and the T7 concatemer junction are packaged into transducing particles during phage infection, *J. Mol. Biol.*, 1990, **216**(4), 911–926, DOI: 10.1016/S0022-2836(99)80010-2.

32 Oxford Nanopore Technologies. [cited 2025 Sep 12]. Nanopore sequencing accuracy. Available from: https://nanoporetech.com/platform/accuracy.

33 N. Bonilla, M. I. Rojas, G. Netto Flores Cruz, S. H. Hung, F. Rohwer and J. J. Barr, Phage on tap-a quick and efficient protocol for the preparation of bacteriophage laboratory stocks, *PeerJ*, 2016, **4**, e2261, DOI: 10.7717/peerj.2261 PubMed Central PMCID: PMC4975003.

34 A. M. Kropinski, A. Mazzocco, T. E. Waddell, E. Lingohr and R. P. Johnson, Enumeration of Bacteriophages by Double Agar Overlay Plaque Assay, in *Methods in Molecular Biology*, Humana Press, 2009, pp. 69–76.

35 R. Owczarzy, B. G. Moreira, Y. You, M. A. Behlke and J. A. Walder, Predicting Stability of DNA Duplexes in Solutions Containing Magnesium and Monovalent Cations, *Biochemistry*, 2008, **47**(19), 5336–5353, DOI: 10.1021/bi702363u.

36 J. Lietard, D. Ameur, M. J. Damha and M. M. Somoza, High-Density RNA Microarrays Synthesized In Situ by Photolithography, *Angew. Chem., Int. Ed.*, 2018, **57**(46), 15257–15261, DOI: 10.1002/anie.201806895.

37 Posit team, RStudio: Integrated Development Environment for R. Boston, MA: Posit Software, PBC; 2025. Available from: https://www.posit.co.

38 R Core Team. R: The R Project for Statistical Computing. 2025 [cited 2025 Jul 19]. Available from: https://www.r-project.org/.

39 H. Wickham and J. Bryan, Readxl: Read Excel Files. The R Foundation; 2015 [cited 2025 Jul 19]. (CRAN: Contributed Packages). Available from: https://CRAN.R-project.org/package = readxl 10.32614/cran.package.readxl.

40 T. Hothorn, F. Bretz and P. Westfall, Simultaneous Inference in General Parametric Models, *Biom. J.*, 2008, **50**(3), 346–363, DOI: 10.1002/bimj.200810425.

41 Spencer Graves, Hans-Peter Piepho and Luciano Selzer with help from Sundar Dorai-Raj. multcompView: Visualizations of Paired Comparisons. The R Foundation; 2006 [cited 2025 Jul 19]. (CRAN: Contributed Packages). Available from: https://CRAN.R-project.org/package = multcompView 10.32614/cran.package.multcompview.

42 Lenth RV. emmeans: Estimated Marginal Means, aka Least-Squares Means. The R Foundation; 2017 [cited 2025 Jul 19]. (CRAN: Contributed Packages). Available from: https://CRAN.R-project.org/package = emmeans 10.32614/cran.package.emmeans.

43 D. Lüdecke, ggeffects: Tidy Data Frames of Marginal Effects from Regression Models, *J. Open Source Software*, 2018, **3**(26), 772, DOI: 10.21105/joss.00772.

44 H. Wickham, Cham: Springer International Publishing, 2016 [cited 2025 Jul 19], Available from: https://link.springer.com/10.1007/978-3-319-24277-4, DOI: 10.1007/978-3-319-24277-4.