Plasmid Library Construction From Genomic DNA

Valeria Florez-Cardona,[1,2] Jessica Khani,[1] Emily McNutt,[1] Bruno Manta,[2,3] and Mehmet Berkmen[1,3]

[1]New England Biolabs, Ipswich, Massachusetts
[2]Institut Pasteur de Montevideo, Montevideo, Uruguay
[3]Corresponding authors: *bmanta@pasteur.edu.uy*, *berkmen@neb.com*

Published in the Molecular Biology section

Functional genomic approaches have been effective at uncovering the function of uncharacterized genes and identifying new functions for known genes. Often these approaches rely on an *in vivo* screen or selection to associate genes with a phenotype of interest. These selections and screens are dependent upon the expression of proteins encoded in genomic DNA from an expression vector, such as a plasmid. Despite the utility of genomic DNA plasmid libraries, the protocols for their construction have remained unchanged in the past 40 years. Here, we present a procedure for constructing plasmid libraries from genomic DNA. This procedure is scalable and relies on simple techniques and common laboratory equipment and reagents. Briefly, the genomic DNA is extracted and then physically fragmented with a g-TUBE, overhangs are repaired, and fragments are selectively purified with magnetic beads to obtain an average fragment size of 2.5 kb. Blunted fragments are ligated into a blunt-end-digested and dephosphorylated vector. Finally, the library is amplified by electroporating the ligation into a high-transformation-efficiency *Escherichia coli* strain and extracting the plasmid DNA from the transformants. As a proof of concept, we built and sequenced three genomic libraries from different genomes and calculated their coverage using a next-generation sequencing (NGS) workflow. © 2025 The Author(s). Current Protocols published by Wiley Periodicals LLC.

Basic Protocol: Plasmid library construction
Alternate Protocol: Selection of gDNA fragments using SageELF gel fractionator
Support Protocol 1: Extraction of gDNA with phenol/chloroform
Support Protocol 2: Vector preparation

Keywords: bacterial selection • genomic DNA • plasmid library construction

How to cite this article:
Florez-Cardona, V., Khani, J., McNutt, E., Manta, B., & Berkmen, M. (2025). Plasmid library construction from genomic DNA. *Current Protocols*, *5,* e70088. doi: 10.1002/cpz1.70088

INTRODUCTION

There are many different strategies for discovering gene-encoded functions, which can be broadly divided into two classes: functional (*in vitro* or *in vivo*) and *in silico* (computational) approaches. Computational approaches rely on large databases to identify gene functions encoded in DNA/RNA sequences (Bar-Joseph et al., 2003; Pascal et al., 2021;

CURRENT PROTOCOLS
A Wiley Brand

Current Protocols e70088, Volume 5
Published in Wiley Online Library (wileyonlinelibrary.com).
doi: 10.1002/cpz1.70088

Wei-Po & Wen-Shyong, 2009). Functional approaches use either *in vitro* screens or *in vivo* selections or screens, which involve testing a large number of candidate genes for a desired activity and usually require the efficient deployment of time- and resource-consuming high-throughput methodology (Chuzel et al., 2018; Sévin et al., 2017). *In vivo* genetic selection methods offer the advantage of rapid identification of a DNA-encoded activity in a low-throughput manner, and can yield unexpected biological insights (Walker et al., 2022; Yuang et al., 2015). Overall, functional genomics methods provide a powerful approach to interrogate the biochemical repertoire coded by uncultivable microorganisms.

Genetic selection methods rely on living cells that can grow only when the desired activity is functional under a given set of conditions. Strains used for genetic selection are usually genetically tractable laboratory model organisms, such as *Escherichia coli*, *Bacillus subtilis*, or *Saccharomyces cerevisiae* (Biver et al., 2013; Craig et al., 2013; Loock et al., 2023; Villamizar et al., 2017). Success in discovering the desired activity depends on the diversity of the DNA library being screened. DNA libraries can be made from genomic DNA (gDNA) from a single organism, environmental DNA (eDNA) from complex sources such as soil or the gut microbiome, or synthetic DNA that is manually curated with a library of specific genes or mutant variants of a singular gene. The isolated DNA must be of high quality, quantity, and purity to be efficiently ligated into an expression vector. Various expression vectors may be used, such as plasmids with 1- to 10-kb inserts (Benoit et al., 2016; Chow, 2005), fosmids with 10- to 50-kb inserts (Zhou et al., 2018), or bacterial artificial chromosomes (BACs) with 150- to 350-kb inserts (Farrar & Donnison, 2007). These libraries are not commercially available and must be constructed by each researcher.

Plasmid-based libraries offer the advantage of high transformation efficiency compared to fosmid or BAC libraries (Asakawa et al., 1997; Felczykowska et al., 2014; Wan et al., 2006), and the same library can be used in multiple selections in different genetic backgrounds. Additionally, unlike with large inserts containing fosmids or BACs, identification of the genetically encoded function of interest is straightforward because plasmids harbor small inserts containing, on average, fewer than three genes. The strategy for the construction of plasmid libraries has remained basically unchanged from that introduced in 1988 by Handelsmann and colleagues (Handelsman et al., 1998) and can be described in five steps: (i) isolation of genomic or environmental DNA, (ii) fragmentation of the DNA, (iii) isolation of DNA fragments of the desired size range, (iv) ligation between DNA fragments and plasmid expression vector, and (v) amplification of the ligated library (Dobrijevic et al., 2013; Escuder-Rodríguez et al., 2022; Nilewski et al., 2022; Villamizar et al., 2017). These library construction protocols have been predominantly focused on cloning eDNA into a vector. The eDNA is partially digested with a restriction enzyme (usually a frequent four-base cutter, such as Sau3AI). The digested eDNA fragments are separated by size using agarose gel electrophoresis and then purified through conventional methods such as commercial gel DNA extraction kits. Enzymes such as the 4-base recognition siteenzyme Sau3AI will generate cohesive ends compatible with BamHI-digested vectors. After digestion with BamHI, the vector is dephosphorylated to minimize self-ligation. Ligation is carried out by T4 DNA ligase with incubation at 16°C overnight. The ligation is transformed into a host strain, and finally cells harboring plasmids are frozen and stored at −80°C (Díaz-Rullo et al., 2021; González-Pastor & Mirete, 2010; Simon & Daniel, 2017). Although this method has been widely used, we propose three key changes to improve the feasibility and reproducibility of library production. First, we suggest avoiding agarose gel separation for fragment size selection. During DNA extraction from gel, a significant proportion of DNA is lost (Wagner et al., 2015), thus reducing genome representation in the final library. Instead, we used physical fragmentation using G-TUBEs. Second, we use a blunt-end ligation

method that can be completed in 15 min, much faster than the standard overnight T4 ligase ligation method. Third, we suggest storing the library as plasmids rather than as transformed cells. This approach has two major advantages: first, the library can be used with different *E. coli* host strains, allowing various genetic selections to be performed; and second, it reduces freeze/thaw cycles that can affect cell viability (Bircher et al., 2018), thus reducing library efficacy.

In developing the protocols provided here, we focused on building DNA libraries using genomic DNA instead of eDNA, for the following reasons. First, gDNA comes from known species, whose genome is known or could be sequenced to identify all genes/ORFs encoded. Second, gDNA extraction is easy compared with eDNA extraction, and is reproducible if the sample can be repeatedly grown. Finally, using a single genome with a defined size enables us to confirm that there is sufficient coverage of the entire genome represented in the final plasmid library. We are aware that the "microbial dark matter" (Osburn et al., 2024) encoded in complex microbial ecosystems is a rich source of new activities, many with potential technological or biomedical interest. "Cloning the metagenome," as originally described by Handelsmann and coworkers (Handelsmann et al., 1998), can also be achieved with our protocols but requires some special considerations, in particular in regard to DNA extraction and fragmentation. Although there are common steps for constructing a plasmid gDNA library, there is no standard methodology, as the procedure depends on the objective of the genetic selection and the availability of resources. For example, if the aim is to capture biochemical pathways in multi-genic operons, large gDNA inserts will be preferred. However, if there is a limited source of gDNA, samples could be used to optimize libraries with a high percentage of inserts.

We present a procedure for robust and reproducible production of high-quality plasmid libraries from genomic DNA. We rely on simple techniques, common laboratory equipment, and affordable reagents. We further improve on the standard method by incorporating modern molecular biology methods to isolate and shear gDNA and evaluate the quality of the final library using next-generation sequencing (NGS). The goal of these protocols is to construct plasmid libraries for future genetic selections aimed to discover functions encoded in genes. Although we designed the procedure to obtain $\sim 2\times$ coverage of an average prokaryotic genome, we are aware that not all gDNA sequences will be represented equally, for example DNA sequences whose products are toxic or those that may be cloned but fail to be expressed or lack cofactors required for activity in the recombinant host cells (Green & Rao, 1998).

STRATEGIC PLANNING

We have separated this procedure into four main steps that take 5 days in total to complete (Fig. 1): (1) gDNA extraction (Support Protocol 1), (2) vector preparation (Support Protocol 2), (3) library construction (Basic Protocol, steps 1-50), and (4) library amplification (Basic Protocol, steps 51-73).

The first step (Day 1) is gDNA extraction, which can take 5 hr or longer. Scaling up the gDNA isolation step to multiple samples may lengthen this step to a full day, meaning that vector preparation will occur the next day. With the gDNA and vector prepared (Day 2), the process continues with library construction (Basic Protocol), comprising seven stages. The first three stages are (i) gDNA fragmentation, (ii) fragment blunting, and (iii) fragment size selection. These steps can be accomplished in 2 hr. Once gDNA fragments are prepared, the process continues with ligation (iv) and library assembly, which take $\sim$3 hr to complete. This stage is scalable and can easily accommodate the construction of several libraries simultaneously. The next stage is small-scale transformation (v), in which an aliquot of the ligation reaction is electroporated into high-efficiency electrocompetent *E. coli* cells. This stage is for evaluating the efficiency of the ligation reaction(s)

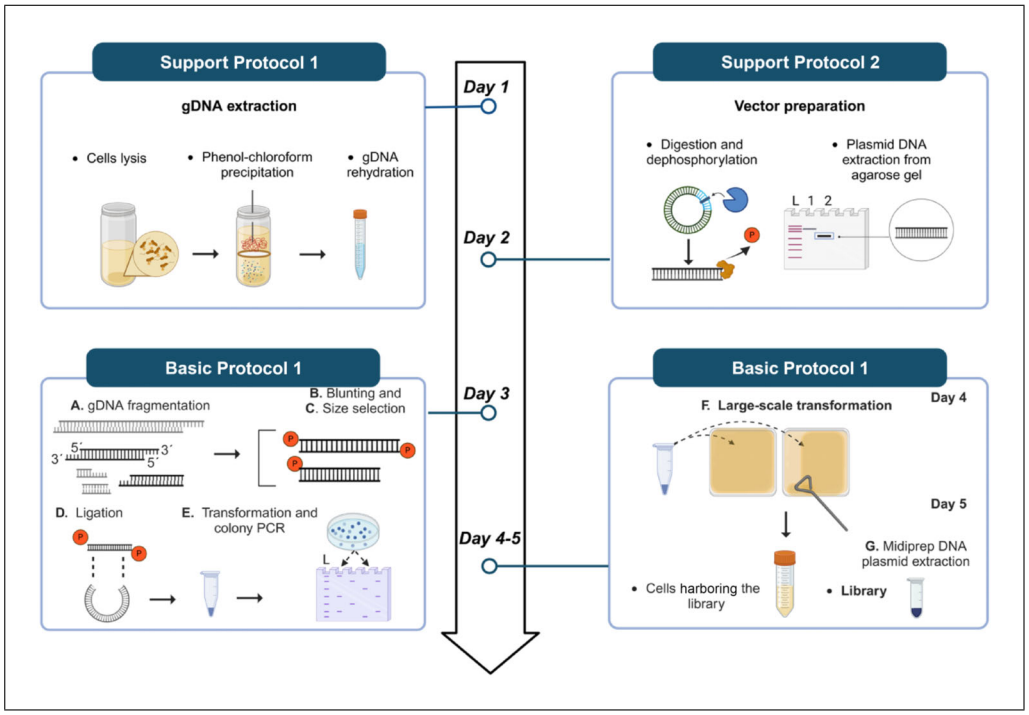

Figure 1 Strategic planning scheme for gDNA plasmid library construction protocol. The first step is to purify gDNA using Support Protocol 1 (Day 1). Three steps are undertaken to extract 1-5 mg of total gDNA. After gDNA is extracted, Support Protocol 2 (Day 2) is followed to prepare the expression plasmid for ligation. To ensure that only blunt-end-digested and dephosphorylated plasmid is utilized, linearized plasmid is purified from an agarose gel and its migration is compared (2) with a DNA size ladder (L) and an undigested aliquot (1). Basic Protocol (spanning days 3-5) begins with gDNA fragmentation (**A**), blunting of overhangs and phosphorylation of 5′ termini (**B**), and selection of sheared gDNA with the desired range (**C**). These are followed by the ligation of gDNA fragments with linearized expression plasmid (**D**), transformation of an aliquot of ligation reaction into a high-efficiency electrocompetent *E. coli* strain, and then testing of the ligation efficiency by amplifying inserts in the expression plasmid through colony PCR (**E**). If >50% of the colonies have inserted fragments of >1.5 kb, the procedure continues with large-scale transformation (**F**) starting with spreading 1 ml of transformants onto large, square plates, to scale up the number of colonies harboring plasmids. Finally, a midiprep plasmid DNA extraction is performed (**G**) to recover the concentrated gDNA plasmid library.

and takes at least 24 hr, as cells must grow to produce visible colonies. If the proportion of plasmids with inserts (assessed by colony PCR) is >50%, the process continues with the next stage (vi), large-scale library amplification, in which all remaining ligation reactions are electroporated into the same *E. coli* strain, this time with all transformants plated onto large agar plates to recover as many colonies as possible. The last stage (vii), library amplification, involves the extraction of the recombinant plasmids from the colonies obtained in stage (vi) to obtain highly concentrated amounts of the final plasmid gDNA library.

CAUTION: The phenol/chloroform mix utilized for gDNA extraction (Support Protocol 1) is classified as toxic by regulatory agencies. Strict adherence to biosafety protocols for use, disposal, and the appropriate safety equipment is required.

PLASMID LIBRARY CONSTRUCTION

BASIC PROTOCOL

To construct high-quality plasmid gDNA libraries, the genome needs to be sheared, and the ends of the DNA fragments must be repaired to prepare the DNA for ligation with the expression vector.

Fragmentation

There are many methods available for gDNA fragmentation. A partial enzymatic digestion with a four-base frequent cutter such as Sau3AI can result in semi-random DNA fragments. Physical fragmentation through shear force can be used to generate fragments without sequence bias. To generate randomly fragmented gDNA in a controlled reproducible manner, we use a g-TUBE (Covaris), a device that shears DNA by forcing the solution to pass through an orifice by centrifugation. The g-TUBE can shear genomic DNA into fragments ranging from 100 bp to 20 kb, depending on the centrifugation force and the number of cycles (Durin et al., 2012). Total recovery tends to be high. For example, following manufacturer instructions, we fragmented 7 µg of *E. coli* gDNA by centrifugation for 1 min at 16,000 × *g*, yielding 98% of the original gDNA. Unlike enzymatic shearing, physical shearing avoids restriction enzyme sequence bias, and it may yield more reproducible results across species. However, physical shredding will result in gDNA fragments with heterogeneous overhangs that prevent blunt-end ligation with the vector. We use the Quick Blunting Kit (New England Biolabs, NEB) to repair the 5′ and 3′ overhangs and add a 5′-phosphate group to make the DNA ends compatible with blunt-end ligation. Although fragmentation by g-TUBE is precise, it has been reported that fragments ranging from 100 bp to 20 kb can be present in the final elution (Covaris, 2020). Because small DNA fragments are preferentially ligated into the vector, they can "poison" the ligation reaction and become over-represented in the final library, despite being low in abundance in the sheared DNA (Díaz-Rullo et al., 2021; Hoorspool, et al., 2010; Rich et al., 2023). To remove unwanted small DNA fragments, we use magnetic bead purification to isolate DNA fragments ≥1.5 kb in size (AMpure Beads, Beckman Coulter) using a low bead-to-sample ratio.

Ligation

The enzyme most commonly used to ligate inserts into a plasmid is bacteriophage T4 ATP-dependent DNA ligase (Gansauge et al., 2017). T4 DNA ligase catalyzes the formation of phosphodiester bonds between adjacent 3′-hydroxyl and 5′-phosphate groups in double-stranded DNA for either blunt- or cohesive-end fragments (Pheiffer & Zimmerman, 1983; Rosi et al., 1997; Wilson & Murray, 1979). One of the most important parameters during ligation is the temperature. The optimal temperature for T4 DNA ligase activity is 37°C; however, the ideal temperature for ligation depends on the strand ends. For cohesive ends, the incubation temperature is generally between 12°C and 16°C, as at higher temperatures single strands of DNA can melt, reducing ligation efficiency. For blunt-ended fragments, ligations are carried out at room temperature with high concentration of T4 DNA ligase. Blunt-end ligation is naturally less efficient than cohesive-end ligation because there are no overhangs to hold the DNA molecules in position. However, efficiency can be enhanced 1000-fold by molecular crowding techniques, such as the addition of polyethylene glycol (PEG) to the reaction (Janner et al., 2013; Liu & Schwartz, 1992; Motohashi, 2019). It should be noted that most studies of DNA ligation parameters have been conducted with DNA inserts of ≤800 bp, a fact to consider when comparing results from prior publications with those obtained using this protocol optimized for an average insert size of 2.5 kb (Potapov et al., 2018; Wilson et al., 2013).

With this in mind, it is imperative that the quality of the library be evaluated before proceeding to use it for experimental purposes. In this protocol, we used the Blunt/TA Ligase Master Mix (NEB), which relies on a modified T4 ligase and a ligation enhancer for faster reaction times at room temperature. To evaluate the quality of the plasmid gDNA library, a small aliquot of the ligation is transformed into 10-β competent *E. coli* (NEB) engineered for high-efficiency transformation. *E. coli* 10-β is deficient in periplasmic

endonuclease I (*end*AI) and has high insert stability due to mutation in the recombinase gene *recA1*, making it ideal for plasmid recovery.

Once it is confirmed that the plasmid library has a low frequency of empty vectors and the average insert length is sufficient, the ligation reaction of the plasmid library assembly needs to be amplified. Multiple aliquots of electrocompetent *E. coli* 10-β are transformed with the remaining ligation reaction, and all transformants are plated on large, square bioassay plates to isolate as many transformants as possible. Transformants are scraped off the plate and pooled, and plasmids are purified from the harvested cells using a DNA midiprep protocol.

Materials

Purified gDNA (Support Protocol 1)
Nuclease-free water (NEB, cat. no. B1500S)
Bioanalyzer Genomic DNA ScreenTape assay kit (Agilent Technologies, cat. no. 5067-5365)
Bioanalyzer Genomic DNA ScreenTape reagents: Ladder and Loading Buffer (Agilent Technologies, cat. no. 5067-5366)
Qubit reagents: 1× dsDNA BR working solution (Invitrogen, cat. no. Q33262)
Qubit 1× dsDNA BR Standard #1 and #2 (Invitrogen, cat. no. Q33263)
Quick Blunting Kit (NEB, cat. no. E1201S)
95% ethanol
AMPure magnetic beads (Beckman Coulter, cat. no. A63880)
Blunt/TA Ligase Master Mix (NEB, cat. no. M0367S)
Linearized, purified vector (Support Protocol 2)
10-β competent *E. coli* (high efficiency; NEB, cat. no. C3020S), genotype *E. coli* DHB10 *Δ(ara-leu)* 7697 *araD139 fhu*A Δ*lacX74 galK16 galE15* e14-Φ80*dlacZ*Δ*M15 recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1* Δ(*mrr-hsdRMS-mcrBC*)
10-β/Stable Outgrowth Medium (NEB, cat. no. B9035S)
LB agar with appropriate antibiotic(s)

Q5 High Fidelity 2× Master Mix (NEB, cat. no. 0492S)
Appropriate primers
1% agarose gel (Lonza, cat. no. 54905)
5× Tris-borate-EDTA (TBE) buffer (Thermo Fisher, cat. no. J63487.3)
1-kb DNA ladder (NEB, cat. no. N3232L)
6× gel loading dye, purple (NEB, cat. no. B7024S)
Plasmid Midiprep Extraction Kit (100; Qiagen, cat. no. 12145)
g-TUBES (Covaris, cat. no. 520079)
Tabletop centrifuge (Eppendorf 5424, cat. no. 5404000014)
Bioanalyzer equipment (Agilent Technologies, cat. no. G2991BA)
Qubit 4 fluorometer (Invitrogen, cat. no. Q33238)
1.5-ml tubes (Eppendorf, cat. no. 0030123611)
NanoDrop spectrophotometer (Thermo Fisher, cat. no. 13-400-525)
Magnetic rack (NEB, cat. no. S1509S)
Minicentrifuge (Thermo Fisher, cat. no. 75004061)
PCR tubes (Eppendorf, cat. no. 951010006)
1-mm electroporation cuvettes (BTX, cat. no. 45-0124)
Gene Pulser Xcell Electroporator (Bio-Rad, cat. no. 165-2100)
Incubator (Eppendorf, cat. no. S44I300001)
Sterile cell spreader (VWR, cat. no. 612-5496)
90-mm petri plate (Corning, cat. no. CLS430599-60EA)
Thermocycler

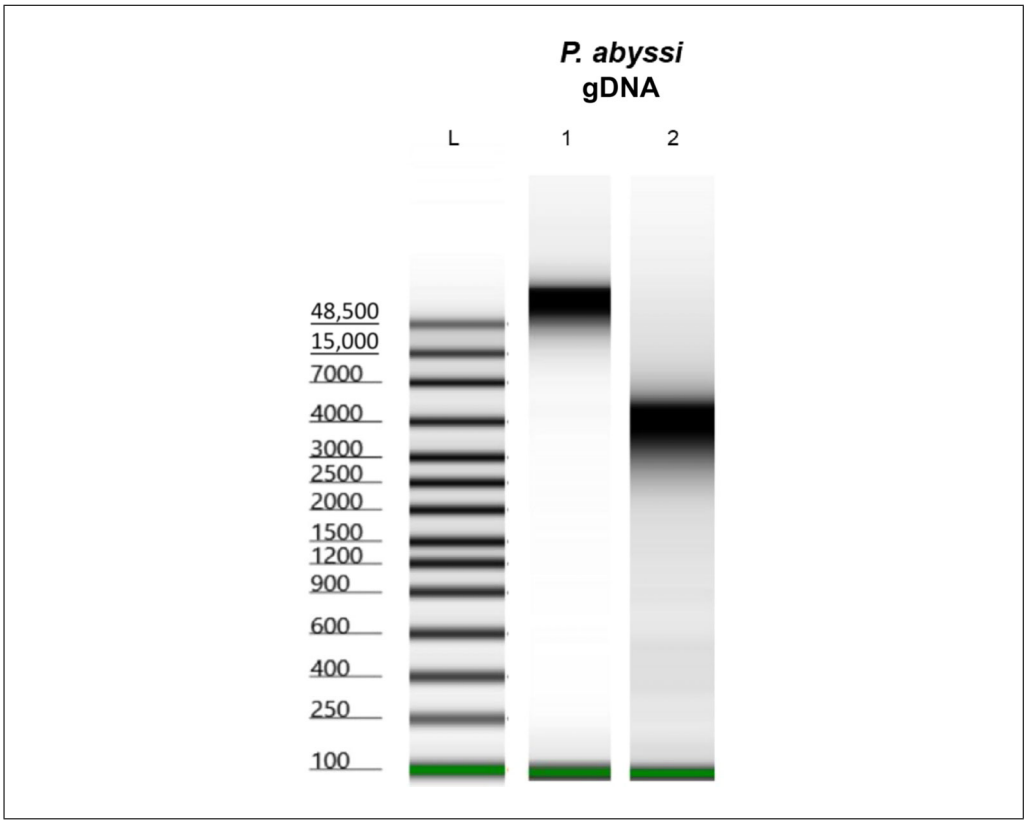

Figure 2 Comparison of extracted and fragmented gDNA. (1) *Pyrococcus abyssi* gDNA before and (2) after g-TUBE fragmentation. DNA ladder sizes are in bp.

Apparatus and equipment for agarose gel electrophoresis
245 × 245-mm square bioassay plate (Corning, cat. no. 431272)
50-ml Falcon tubes (Corning, cat. no. CLS430829-500EA)

gDNA fragmentation

1. Resuspend between 4 and 7 µg of gDNA from Support Protocol 1 (one aliquot) in nuclease-free water in a final volume between 110 and 150 µl.
2. Load sample into the g-TUBE.

   Avoid loading <100 µl or >10 µg of gDNA into the g-TUBES. In our experience, gDNA concentrations of >17 µg can result in rupture of the g-TUBE membrane. Always keep samples on ice.

3. Close the g-TUBE and screw the cap on firmly. The upper and bottom cap can be sealed with Parafilm to ensure samples will not leak during centrifugation.
4. Place the g-TUBE into the centrifuge with the blue screw-cap up. Use an extra g-TUBE with equal volume of water to balance the rotor if necessary.

   The g-TUBE protocol recommends using an Eppendorf 5424-R centrifuge.

5. Centrifuge 1 min at 4000 × *g*, room temperature. Flip the tube over and repeat centrifugation.
6. To check fragment size, mix 1 µl of sample with 9 µl of Genomic DNA TapeStation ScreenTape Loading Buffer and do the same with the ladder control.

   The average fragment size should be between 2.5 and 15 kb, as shown in Figure 2.

   Alternatively, gDNA fragmentation can be visualized using gel electrophoresis with a 1% agarose gel.

Table 1 Reaction Components for Generating Blunt-end gDNA

Component	Volume (µl)
Minimum 4-5 µg gDNA fragments (from step 7)	From 1 to 19
10× blunting buffer	2.5
1 mM dNTPs mix (250 µM each)	2.5
Blunt enzyme mix	1.0
Nuclease-free water	Variable
Total volume	25

7. Quantify gDNA using a Qubit 1× dsDNA Broad Range Assay Kit according to the manufacturer's instructions.

   You can use either NanoDrop or Qubit to quantify the gDNA.

8. Assess the purity of the gDNA using a NanoDrop spectrophotometer, checking both A_{260}/A_{280} and A_{260}/A_{230} ratios (as described in the introduction to Support Protocol 1) to assess gDNA quantity, purity, and integrity.

gDNA fragment blunting

9. Use a Quick Blunting Kit to repair 5′ and 3′ overhangs to generate blunt-end gDNA fragments, mixing the components listed in Table 1 in a 1.5-ml tube.

10. Incubate reaction at room temperature for 30 min. To avoid damaging DNA with long incubation periods at 70-80°C, skip the heat inactivation of the enzyme mix suggested by the manufacturer.

gDNA fragment size selection

Here, we describe an efficient method utilizing readily available reagents to select the size of gDNA fragments by varying the concentration of magnetic AMPure beads that bind the DNA specifically. AMPure beads enable cDNA fragments within a desired size range to be selectively purified by varying the bead-to-sample ratio. Larger DNA fragments (> 2000 bp) bind preferentially to the beads at lower bead-to-sample ratios of 1:2, while smaller fragments can be excluded at ratios of 3:1, providing a tunable size selection with bead ratios varying from 0.5× to 2× sample volume. However, other strategies may be used depending on the availability of equipment and materials, i.e., agarose gel electrophoresis or its variants (such as the lateral DNA fractionator described in the Alternate Protocol).

11. Thawing AMPure bead stock solution at room temperature 30 min before use.

12. Prepare fresh 80% ethanol solution with nuclease-free water.

13. Add 0.4 vol AMpure beads to the blunting reaction from step 8, and gently mix by pipetting until the solution becomes homogeneous.

    If the total blunting reaction volume is 25 µl, add 10 µl AMpure beads.

14. Incubate 10 min at room temperature to facilitate the binding of gDNA fragments to the beads.

15. Place the tube in the magnetic rack, let the beads migrate toward the magnet for 2 min, and remove and discard the supernatant using a pipet.

16. Wash the beads by adding 200 µl of 80% ethanol and immediately discarding it. Repeat this step.

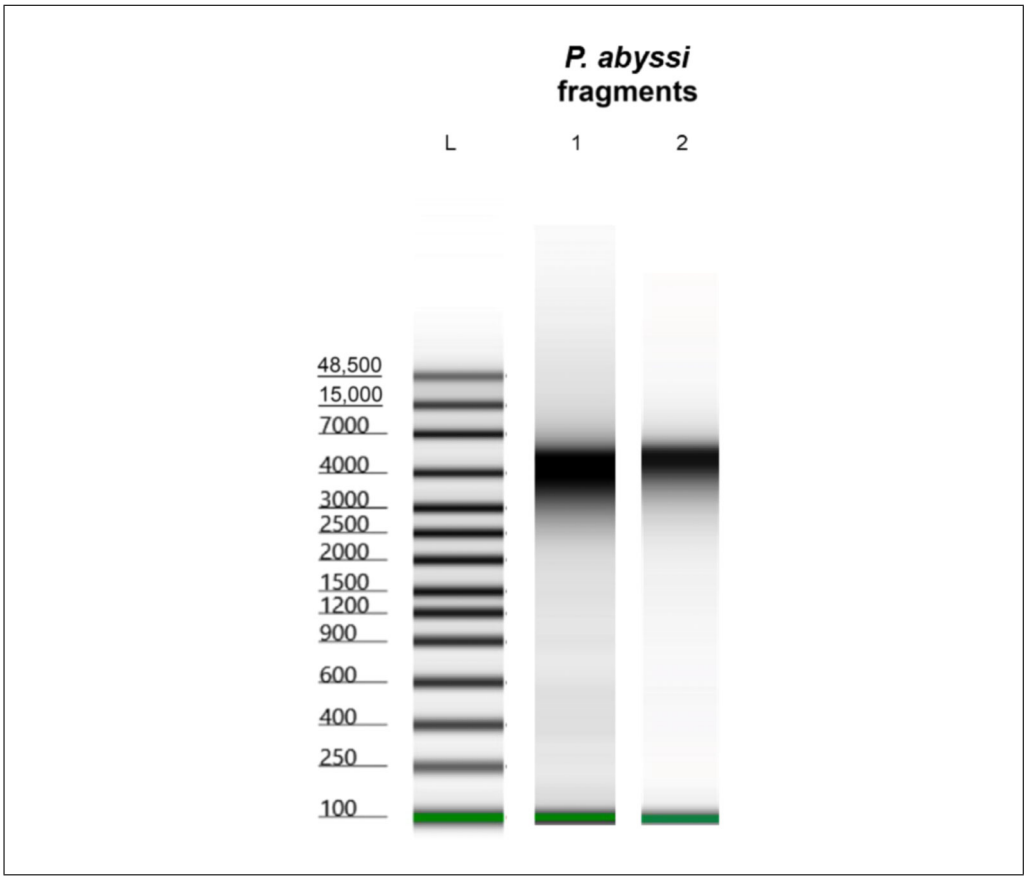

Figure 3 Fragment size selection with magnetic beads. As an example, a comparison of fragmented *Pyrococcus abyssi* gDNA before (1) and after (2) size selection with magnetic beads from tape station is shown. DNA ladder sizes are in bp.

17. Spin down the beads for 1 min on a tabletop centrifuge, remove remaining ethanol traces with a pipet, and let the tube air dry on a rack for few minutes.
18. Let ethanol evaporate by air drying for no more than 30 s with the lid open. Do not let the beads become dry.
19. Remove the tubes from the magnetic rack. Resuspend the beads in 20 µl nuclease-free water. Pipet up and down gently until the solution becomes homogeneous. Let stand for 5 min to allow the DNA to become resuspended.
20. Place the tubes back on the magnetic rack and incubate until the beads have all migrated toward the magnet. Without disturbing the beads, remove the supernatant.

    This is your eluted DNA.
21. Measure DNA quality parameters: (i) quantify using Qubit 1× dsDNA Broad Range Kit, (ii) assess purity using NanoDrop, and (iii) assess integrity using 1 µl of sample with the Bioanalyzer Genomic DNA ScreenTape reagents (Ladder and Loading Buffer) Kit, both according to manufacturer's instructions (see steps 6-8 and Fig. 3).

    A range of fragments between ∼2.5 kb and ∼10 kb should be observed. Fragment size ranges vary according to the integrity of gDNA from Support Protocol 1.
22. Store the blunt-ended repaired and purified gDNA samples at 4°C for immediate or next-day use. If the samples will not be used within that time frame, store them at −20°C.

    Avoid freeze/thaw cycles of the gDNA samples.

Table 2 Reaction Components for Ligating gDNA Fragments to Expression Plasmid Vector

Component	Vector control	Ligation
Nuclease-free water	[a]	[a]
gDNA sample	–	300 ng[b]
Vector sample	200 ng[b]	100 ng[b]
Blunt/TA ligase master mix	10 µl	10 µl
Total volume	20 µl	20 µl

[a] Adjust volumes depending on reaction.
[b] Adjust volumes according to the DNA concentration.

Ligation

23. Thaw an aliquot of Blunt/TA Ligase Master Mix on ice 20 min before use.

24. Mix gDNA fragments (from step 20) with linearized vector (from Support Protocol 2) in a PCR tube, along with the other components listed in Table 2, in the order given in the table.

    The manufacturer's instructions suggest performing a 10-µl reaction, with half the total volume composed of Blunt/TA Master Mix. However, for this protocol, the final reaction volume has been scaled up to 20 µl while preserving the component proportions. The other half of the reaction is composed of vector and gDNA, and volumes must be adjusted, always preserving a vector-to-insert molar ratio of 1:3.

25. Gently mix reaction components.

26. Spin down both reactions in a minicentrifuge and incubate them at room temperature for 20 min.

27. Purify reactions using 1.5× vol of AMpure beads (as compared to the of total ligation reaction) as described in steps 9-18. Elute DNA fragments in 20 µl nuclease-free water.

    If the total volume of the blunting reaction is 20 µl, add 30 µl AMPure beads.

28. Measure quantity and purity of ligated DNA as in step 19.

    Keep samples on ice for immediate use; otherwise, store at −20°C. Avoid storing both reactions at 4°C for >1 day.

Small-scale transformation to check ligation efficiency

29. Thaw three 50-µl aliquots of competent *E. coli* 10-β cells on ice for ∼30 min before use. If not using commercial electrocompetent cells, prepare your own and test for high-efficiency transformation using empty vector (Potapov et al., 2018).

30. Prechill 1-mm electroporation cuvettes on ice.

31. Prechill on ice an aliquot of prepared vector from Support Protocol 2, vector treated with ligase, and the ligation reaction from step 26.

32. Prewarm 10-β Stable Outgrowth Medium at 30°C. If using your own competent cells, prewarm SOC medium instead.

33. Mix 10-β competent cells with control and ligation reactions. Add ∼10 ng of each reaction; avoid adding >2 µl volume.

    These next steps are for evaluating the percentage of plasmids with inserts and their average size before proceeding with the rest of the protocol. In testing the quality of the vector prepared in Support Protocol 2, two controls are necessary, to evaluate efficiency of (i) digestion and (ii) dephosphorylation.

Table 3 Colony Counts Expected From the Small-scale Ligation Transformation

Reactions	Expected colonies	Description
Linearized vector	0%-5% of ligation	Measures digestion efficiency; a high number of colonies indicates high percentage of uncut vector
Ligated vector control	0%-30% of ligation	Measures dephosphorylation efficiency; a high number of colonies could indicate inefficient dephosphorylation or low quality of inserts due to remaining overhangs
Ligation	>5000	Measures ligation efficiency; a high number of colonies indicates efficient electroporation.

34. Transfer the cell and ligation mix into the electroporation cuvette, avoiding the formation of bubbles.
35. Set Bio-Rad GenePulser Xcell electroporator to preset protocol for *E. coli* at 1.8 kV, 25 µF.
36. Electroporate the cells, and check the value of current flow and the decay curve to confirm a successful electroporation. If arcing occurs, discard transformation and repeat with a smaller volume of ligation reaction.

    Current decay should be between 1.770 and 1.789 V.
37. Immediately add 950 µl of 10-β/Stable Outgrowth Medium for recovery of cells and transfer the cells into a fresh sterile 1.5-ml tube.
38. Incubate cells at 30°C with shaking at 200 rpm for 2 hr.
39. Equilibrate 90-mm LB agar plate(s) supplemented with corresponding antibiotic at room temperature.
40. After incubation, prepare 1:10 dilutions of control and ligation reactions by pipetting 100 µl cells into 900 µl of 10-β/Stable Outgrowth Medium. Mix by vortexing.

    This dilution is made to plate a countable number of transformants. Because electroporation can be very efficient, plating undiluted transformation can result in a lawn of cells unsuitable for counting. Several dilutions can be plated to find the ideal number of cells per plate. Agar plates with <20 and >200 colonies are not accurate. Accordingly, dilute cells to achieve 20-200 colonies per agar plate.
41. Transfer 100 µl of 1:10 dilution of cells to an agar plate from step 36. Use sterile glass beads or cell spreaders to evenly plate the transformation.
42. Incubate agar plate at 30°C overnight to avoid thermal stress and protein misfolding. Adjust incubation time and temperature if needed depending on the experimental requirements (e.g., when selecting for thermophilic or psychrophilic enzymes).
43. Count colonies and calculate the number of colony-forming units (cfu/ml) for each reaction.

    To calculate the efficiency of digestion and dephosphorylation of the plasmid vector, count the number of colonies from the two controls and divide by the total number of transformants. For example: if there are five colonies from the linearized vector control and 50 colonies from the total transformants, the percentage of undigested plasmid is 10%. This measurement is used to evaluate the quality of the ligation reaction. Ideally the number of colonies from the ligation transformation is greater than either of the controls. See Table 3 for expected number of colonies from the plasmid library ligations.

Colony PCR to assess the proportion of clones with insert

The goal of this part of the protocol is to determine the percentage of vectors that carry inserts. To control for primer binding, PCR amplification, and identification of PCR

Table 4 PCR Reaction Components to Check Number and Size of Plasmids Containing gDNA Inserts

Component	Volume (µl)
Q5 High-Fidelity 2× Master Mix	12.5
10 µM forward primer	1.25
10 µM reverse primer	1.25
NF water	9.0
Colony template DNA (from Basic Protocol, step 42)	1.0
Total volume	25

products, use a strain harboring the empty vector and another harboring the vector with an insert. Primers should be designed to anneal to the vector backbone and to amplify the vector's multiple cloning site region. Primers should be ≥500 bp apart to ensure a detectable band for the empty vector control.

44. Pick between 12 and 20 colonies from the ligation reaction plate from step 41 and resuspend each in 10 µl nuclease-free water in PCR tubes. Heat cells in the thermocycler for 5 min at 95°C for 5 min and then cool for 30 s at 10°C.

45. Set up PCR reaction using Q5 High Fidelity 2× Master Mix according to the manufacturer's instructions. A master mix for the reaction can be made with the Q5 High Fidelity 2× Master Mix, primers, and nuclease-free water, as described in Table 4. If using a master mix, aliquot 24 µl of mix per tube.

46. Mix PCR master mix reagents by vortexing and centrifuge mix to collect at the bottom of the tube. Fill the PCR tubes with 24 µl each of the PCR master mix.

47. Add 1 µl of each resuspended colony into one PCR tube from step 42. Mix each reaction by pipetting. Spin down all tubes.

48. Run PCR according to manufacturer's instructions, using calculated primer annealing temperatures.

    In this step, it is important to consider that gDNA inserts will vary in size. Even if the majority of the gDNA fragments were purified within a certain size range, smaller or larger fragments will be present. Thus, the PCR extension time should be set to 6 min to amplify fragments up to ∼10 kb.

49. Mix 5 µl of each PCR reaction mixed with 1 µl of 6× gel loading dye and load the 6 µl mixture into a 1% agarose gel. Run the gel at 100 V for ∼1 hr to allow adequate separation.

50. Image the gel. Check if the PCR control reactions have the expected size. Count the number of plasmids with amplicons larger than the empty vector control, and calculate the average size of the inserts. PCR reactions with no amplicons, including at the expected empty vector size, may indicate the presence of a gDNA insert that is too large to be amplified using 6-min extension time. Should there be a significant number of these non-amplified inserts, consider increasing the extension time.

    Ideally, >50% of the plasmids should contain an insert with an average size of >2 kb. If these criteria are met, continue with the protocol.

Large-scale transformation to amplify plasmid library

51. Fill 245 × 245-mm square bioassay plates with LB agar and corresponding antibiotic.

    Large plates require ∼200 ml of medium each.

52. Calculate the volume of ligation reactions that remained from step 31. Transform up to 20 ng of total DNA in each electroporation.
53. Thaw all necessary 50 µl aliquots of 10-β competent *E. coli* cells on ice.
54. Chill 1 mm electroporation cuvettes on ice.
55. Prewarm 10-β/Stable Outgrowth Medium at 30°C.
56. Mix 10-β competent *E. coli* with aliquots of ligation reaction.
57. Repeat electroporation as described in steps 32-35.
58. Incubate cells at 30°C shaking at 200 rpm for 2 hr.

    To count the cfu/ml for the library, 100 µl of transformants should be reserved from one of the transformations to dilute and plate as the colonies on the large bioassay plates will be too numerous to count.

59. Pipet 100 µl of transformants and dilute 1:10 into 900 µl of 10-β/Stable Outgrowth Medium.
60. Plate 100 µl diluted cells from step 52 in 90 mm LB agar supplemented with appropriate antibiotic. Additional dilutions can be made and plated as needed.
61. Plate 100 µl diluted cells from step 52 in 90 mm LB agar supplemented with corresponding antibiotic.
62. Transfer the remaining transformed cells onto a large bioassay LB agar plate (245 mm × 245 mm) from step 50 and spread cells all over the plate evenly, using sterile glass beads or a cell spreader. One plate is used per 1 ml transformation.
63. The next day, count colonies from dilution of 90-mm plate(s).

    On average, there should be 400,000 colonies per large bioassay plate (245 mm × 245 mm). Calculate total number of transformants by multiplying the number of cells with the dilution factor used. Generally, the expected number of transformants are between 3×10^5 and 8×10^5.

Library amplification

The plasmid library generated from the ligation should not be used directly in a genetic selection. Not only will the assembled libraries be consumed quickly, but the transformation efficiency of the ligation reaction is lower than transforming isolated plasmid. Also, some DNA inserts may be present in low frequency and can be lost in subsequent selections. It is therefore suggested to amplify the library using a highly efficient E. coli strain designed for plasmid DNA uptake. Purifying plasmid gDNA from 3-8 × 10^5 transformants to obtain >400 ng/µl plasmids is ideal.

64. Weigh two empty 50-ml Falcon tubes.
65. Add 5 ml sterile LB medium to the bioassay plates and spread by gently swirling the plate. Using a sterile scraper, scrape all cells from the surface to one of the lower corners of the plate. Be gentle; do not break the agar, as the pieces will contaminate downstream steps. Add more medium if necessary.
66. Carefully pour or aspirate the medium containing the cells into a sterile 50-ml Falcon tube on ice.

    If the volume of suspension is >20 ml, split the sample into two tubes.

67. Centrifuge the pooled colonies for 15 min at 10,000 × *g*, 4°C.
68. Discard supernatant and place the tube on ice.

69. Weigh each tube with cells and calculate how many grams of cells were harvested by comparing the extra weight measured in step 63. Expect between 2 and 3 g of cells in each tube.

70. Use Plasmid Midiprep Extraction Kit 100 (Qiagen) following manufacturer's protocol for low copy plasmid/fosmid, and immediately start purifying plasmid library.

71. Resuspend DNA pellet in 80-100 µl nuclease-free water by pipetting.

72. Measure the quantity and purity of the final library as described in step 17.

    The expected final yield of the plasmid library varies from 400 ng to 1.6 µg per µl.

73. Divide the plasmid library into small aliquots of ∼5 µl and store at −20°C for short-term or −80°C for long-term storage.

ALTERNATE PROTOCOL

SELECTION OF gDNA FRAGMENTS USING SageELF GEL FRACTIONATOR

The following alternative protocol uses the Sage Electrophoretic Lateral Fractionator (SageELF) gel electrophoresis system to more accurately select the fragmented gDNA (Sage Science, Beverly, MA). This protocol replaces steps 11-20 of the Basic Protocol; the procedure should then continue with the ligation protocol, beginning with step 21 of the Basic Protocol.

Materials

SageELF kit (0.75% precast agarose gel cassettes, buffer and fluorescent label; Sage Science, cat. no. ELD2010, ELD7510, ELD4010)
Fragmented gDNA (Basic Protocol, step 20)
Nuclease-free water (NEB, cat. no. B1500S)

SageELF instrument (Sage Science, cat. no. EL00038)
1.5-ml tubes (Eppendorf, cat. no. 0030123611)
Magnetic beads (Beckman Coulter, cat. no. A63880)
Magnetic rack (NEB, cat. no. S1509S)
Bioanalyzer ScreenTape (Agilent Technologies, cat. no. 5067-5365)
Bioanalyzer ScreenTape reagents: Ladder and Loading Buffer (Agilent Technologies, cat. no. 5067-5366)
Bioanalyzer equipment (Agilent Technologies, cat. no. G2991BA)

1. Using the SageELF software, set the running protocol for the 0.75% gel cassette and set the fragment separation for fragments from 1 kb to 18 kb, with an average size of 5000 bp.

2. Prepare precast gel and load fragmented gDNA sample according manufacturer's instructions.

3. Set the instruments according to the manufacturer's instructions and run the samples.

4. When the protocol is complete, the fragmented gDNA will be separated into thirteen wells containing buffers. Transfer each sample into a separate 1.5-ml tube.

5. Confirm gDNA fragment size by analyzing 1 µl of each separated fraction with a TapeStation as described in the Basic Protocol, step 5.

6. Select samples with desired fragment sizes and pool them in a 1.5-ml tube.

   Typically, we select fragments of 2.5-8 kb in size.

7. Measure the total volume of the pooled sample.

8. Clean up the fragments by adding 1× volume of magnetic beads to the pooled sample. Follow the magnetic bead cleanup protocol described in the Basic Protocol, steps 9-20.

It is important to elute fragments in nuclease-free water. After cleanup, either proceed immediately with the ligation reaction or store it at −20°C.

EXTRACTION of gDNA WITH PHENOL/CHLOROFORM

SUPPORT PROTOCOL 1

In this protocol, gDNA can be extracted either from a fresh culture or from frozen cells. Genomic DNA is isolated from cells using the phenol/chloroform extraction method. Briefly, the cells are lysed in a buffer containing proteinase K, Triton X-100 for lysis, and sucrose, which decreases physical damage that can affect gDNA integrity. Phenol/chloroform/isoamyl alcohol is added alongside sodium dodecyl sulfate (SDS) to denature proteins and to separate the gDNA from the rest of the cellular components. Upon centrifugation, the gDNA remains dissolved in the aqueous upper layer, the middle interphase is formed by insoluble membrane proteins, and the bottom layer is the phenolic phase containing insoluble proteins and lipids. gDNA is recovered from the aqueous phase by precipitating with isopropanol (Sambrook et al., 1989). It is then assessed for quantity, purity, and integrity.

The quantity of gDNA is measured in ng or µg/µl. To accurately assess the concentration of gDNA, we suggest using a Qubit instrument, which measures concentration via fluorescence using a DNA-specific probe (Nakayama et al., 2016), instead of a NanoDrop spectrophotometer, which simply measures the absorbance at 260 nm (Gallagher & Desjardins, 2006). The purity of the gDNA is determined by spectrophotometry (we suggest using SD'S a NanoDrop), with two absorbance (A) ratios measured: A_{260}/A_{280} nm and A_{260}/A_{230} nm. Nucleic acid absorbs at 260 nm, whereas proteins absorb at 280 nm; thus, the absorbance of the sample at 260 nm is compared to that at 280 nm to determine the purity of the sample. An A_{260}/A_{280} ratio of ∼1.8 is optimal; higher ratios indicate protein contamination of the sample. Proteins, such as nucleases and DNA-binding proteins, can interfere with downstream cloning procedures. Alcohols (ethanol, phenol, or isopropanol), as well as other contaminants such as HCl that may copurify with DNA during extraction, can poison downstream reactions and are detected at 230 nm. An A_{260}/A_{230} nm ratio of 2.0 is optimal, and lower ratios indicate the presence of contaminants. Finally, the integrity of the gDNA should be assessed either via gel electrophoresis or using a bioanalyzer system (Panaro et al., 2000). The gDNA should ideally be observed as a narrow band of high molecular weight, indicating that the gDNA is in large fragments. A "smear" or large range of DNA sizes indicates degradation of the DNA due to either physical or enzymatic causes. Large, homogenous fragments are key to downstream library construction steps.

NOTE: If starting with fresh cells, follow the protocol below. Use a rich medium and an incubation temperature appropriate for optimal growth of your strain of interest. If using previously frozen cells, start with at least 4 g of wet cell pellet. Thaw cell pellet on ice to prevent premature lysis and proceed to step 6 of the protocol.

Materials

Electrocompetent *E. coli* 10-β (NEB cat. no C3020K) from which to construct the gDNA library
Growth medium appropriate for bacterial strain
Buffers I, II, and III (see recipes)
Lysozyme powder (Thermo Scientific, cat. no. 89833)
0.25 M EDTA, pH 8, filtered solution (Invitrogen, cat. no. 15575-038)
Sucrose (Sigma-Aldrich, cat. no. S0389-500G)

10% Triton X-100 (Thermo Fisher, cat. no 85111)
Sodium dodecyl sulfate (SDS; Thermo Fisher, cat. no. AM9820)
25:24:1 (v/v/v/) phenol/chloroform/isoamyl alcohol saturated with 10 mM Tris buffer, pH 8/10 mM EDTA (Sigma-Aldrich, cat. no. P3803-400ML)
Fresh 70% ethanol (diluted from 95%, Reagent Alcohol, Pharmco, cat. no. 241000190CSGL)
5 M sodium chloride (Sigma-Aldrich cat. no. S6546-1litre)
Fresh isopropanol (99% Isopropyl Alcohol, Pharmco, cat. no. 231HPLC99)
1× dsDNA BR working solution (Invitrogen, cat. no. Q33262)
Qubit 1× dsDNA BR Standards #1 and #2 (Invitrogen, cat. no. Q33263)

250-ml plastic bottles, sterile (Nalgene, cat. no. 02-896-2F)
Tabletop centrifuge (Eppendorf 5424-R)
5-ml aspirating pipet (VWR, cat. no. 612-5885)
Pipettor (Eppendorf, cat. no. 3123000080)
Glass test tubes, sterile
Glass flask, sterile (Nalgene, cat. no. 342023-1000)
50-ml Falcon tubes
Glass rods (VWR, cat. no. 59060-105), sterile
Qubit 4 fluorometer (Invitrogen, cat. no. Q33238)
Bioanalyzer equipment (4200 TapeStation, Agilent Technologies, cat. no. G2991BA)
Genomic DNA ScreenTape (Agilent Technologies, cat. no. 5067-5365)
Genomic DNA reagents for tape station (Agilent Technologies, cat. no. 5067-5366)
NanoDrop spectrophotometer (Thermo Fisher, cat. no. 13-400-525)

1. Inoculate 5 ml of growth medium with the bacterial strain of interest.
2. Grow cells overnight at optimal growth temperature with shaking at 250 rpm.
3. Inoculate 500 ml of growth medium with 1 ml of overnight culture (1/500 final dilution).
4. Incubate culture overnight at optimal growth temperature shaking at 250 rpm.
5. Split the 500 ml of overnight culture in two sterile 250-ml centrifuge bottles. Spin 15 min at 10,000 × *g*, room temperature, and discard supernatant. Either extract gDNA from the cell pellet immediately or store cells at −80°C. To calculate wet weight of cell prior to DNA extraction, weigh two empty, dry sterile 250-ml centrifuge bottles. Split the 500 ml of overnight culture between the two sterile bottles. Spin 15 min at 10,000 × *g*, ??°C, and discard supernatant, being carefully to remove excess supernatant without discarding cells. Invert bottles onto absorbent paper and wait 3-5 min. Re-weigh the two 250-ml centrifuge bottles. The extra weight corresponds to the approximate wet weight of cells. Extract gDNA from the cell pellet immediately or store cells at −80°C.

   Freezing cells will increase lysis efficiency, but the water crystals formed may damage gDNA integrity. If using previously frozen cell pellets, thaw pellets on ice to prevent premature lysis.

6. If using previously frozen cell pellets, thaw pellets on ice to prevent premature lysis.
7. Resuspend cells in buffer I (resuspension solution) on ice. Use 5 ml buffer per gram of cell pellet: e.g., use 20 ml of buffer for the suggested 4 g of cell pellet. Mix well with a pipet until the suspension is homogeneous. Avoid vortexing the sample as this could shear the gDNA.
8. Prepare at least 12 ml of fresh lysozyme (20 mg/ml) in 0.25 M EDTA, pH 8.0, and add 6 ml to each bottle of resuspended cells.

9. Swirl for 30 s to mix and then incubate on ice for 2 hr.
10. Check lysis: Withdraw two 500-µl aliquots of the lysozyme-treated cell suspension and transfer into two sterile glass test tubes. Add 500 µl buffer II (lysis solution) and 100 µl of 10% SDS to one tube. Gently shake the tube and compare with the untreated sample in the other tube for viscosity and clarity. Cells are lysed when the mixture becomes transparent. If lysis has not occurred, continue incubating solution on ice for an additional hour or more, repeating the lysis check above every hour until cells have been sufficiently lysed.
11. When cell suspension turns transparent, add 25 ml of buffer II plus 5 ml of 10% SDS. Mix the suspension by gently swirling it to achieve complete lysis.

    At this point, the total volume in the bottles is ∼70 ml.

12. Add 70 ml (∼1 vol) of 25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol. Shake manually and vigorously for 2-3 min, until the solution becomes white and homogeneous.
13. Spin at 30 min ≥18,000 × *g*, 4°C.

    Poor mixing at step 11 can result in a failure of the layers to separate

14. Transfer the upper, aqueous layer (it should be clear and viscous; this layer contains gDNA) to a sterile glass flask using a 5-ml serological pipet. Avoid touching the bottom of the middle interphase layer, as phenol can compromise final gDNA quality. If residual phenol is visible at the bottom of the solution, if precipitates are visible, or if an odor is detected, spin again at 30 min ≥18,000 × *g*, 4°C, and carefully transfer the top layer to a new glass flask.
15. Prepare three 50-ml Falcon tubes each containing 20 ml of 70% ethanol.
16. Measure the volume of suspension recovered in step 13 with a pipet.
17. Add 5 M NaCl to the solution to obtain a final concentration of 0.4 M (for example, add 6 ml 5 M NaCl to 70 ml aqueous phase). Gently swirl mixture to precipitate proteins.
18. Gently add 0.7 volumes of 99% isopropanol to the top of the solution. For example, to the 76 ml of aqueous solution in the example in step 16, add 54 ml isopropanol.
19. Using a sterile glass rod, gently mix the isopropanol with the solution and let the solution separate again. Watch for the formation of fibrous, white precipitated gDNA.
20. Pick up the precipitated gDNA fiber with the same glass rod by rolling or "spooling" it. Let the liquid drain from the rod into the flask
21. Submerge and rinse the gDNA (attached to the glass rod) three times into the Falcon tubes containing ethanol from step 14 in each subsequent tube. This step helps to remove residual proteins, phenol, and isopropanol.
22. Air dry the gDNA stuck on glass rod on a rack for 5 min. Do not let the glass rod touch anything.
23. Immerse the glass rod with dried gDNA into a 50-ml Falcon tube containing 10 ml of buffer III (TE buffer, pH 8.0). Gently agitate the rod to help gDNA dissolve. If necessary, add an extra 5 ml buffer to help resuspend the gDNA.
24. To let the gDNA dissolve, incubate it at 4°C overnight.

    The next day, check to see if the gDNA is completely solubilized. gDNA that has not become fully resuspended would be a visible and transparent pellet floating in the solution.

Depending on the amount of gDNA purified, add between 1 or 2 ml of buffer III to completely dissolve it.

25. Assess gDNA quantity, purity, and integrity as described in the introduction to Support Protocol 1.

SUPPORT PROTOCOL 2

VECTOR PREPARATION

This protocol utilizes plasmids as expression vectors for the gDNA fragments. To construct high-quality gDNA libraries, it is important that most of the plasmids contain inserted genomic DNA ≥1 kb in size. The percentage of empty vector clones in the library should be as low as possible. In this protocol, the genome is physically fragmented, resulting in heterogeneous 5′ and 3′ ends. The Quick Blunting Kit is used to treat the gDNA and generate 5′-phosphorylated blunt-end DNA fragments that are suitable for blunt-end cloning. To prepare a compatible vector, the plasmid will be digested with a blunt-end-generating restriction enzyme such as SmaI. Following digestion, the vector must be dephosphorylated, as removal of the 5′ phosphates from the vector will prevent vector recircularization and enhance subsequent ligation efficiency. This can be done with an enzyme such as calf intestinal phosphatase (CIP), bacterial phosphatase, or shrimp alkaline phosphatase; we suggest using CIP because its protocol is relatively straightforward (Green & Sambrook, 2020a, 2020b). Finally, the digested, dephosphorylated vector is size selected through agarose gel electrophoresis followed by gel extraction and purification. This step is crucial for minimizing the amount of undigested vector. Reducing the amount of uncut vector and preventing self-ligation of cut vector via dephosphorylation will reduce the amount of empty plasmid present in the final gDNA library.

In selecting the vector, it is important that it contain a blunt-end restriction site at the multi-cloning site, to allow blunt-end ligation in Basic Protocol 1. We also recommend not using a vector for which the expression of the protein is driven by the T7 promoter, as this might lead to toxicity due to overexpression, and not all the strains harbor T7 RNA polymerase, limiting the use of the resulting plasmid library. In our work, we have used the vector pAL, for which the expression of the proteins is driven by the strong ribosomal RNA promoter (McNutt et al., 2024).

Materials

Protein expression plasmid vector of your choice containing a blunt-end restriction site at the multi-cloning site (see protocol introduction)
Midiprep Plasmid DNA kit (Qiagen Plasmid Midi Kit 100]; Qiagen, cat. no. 12145)
Nuclease free water (NF-water, NEB, cat. no B1500)
Blunt-end restriction enzyme: e.g., SmaI (NEB, cat. no. R0141S)
Sterile LB liquid medium (Thermo Fisher, cat. no. L3522-1KG)
10× CutSmart Buffer (NEB, cat. no B6004)
Alkaline phosphatase (Quick CIP or rSAP, NEB, cat. no. M0525S or M0371S)
0.7% fresh agarose gel (SeaKem LE Agarose; Lonza, cat. no. 50004)
AccuGENE 5× TBE Buffer (Lonza, cat. no. 50836)
Monarch DNA Gel Extraction Kit (NEB, cat. no. T1020S)
1× dsDNA BR working solution (Invitrogen, cat. no. Q33262)
Qubit 1× dsDNA BR Standards #1 and #2 (Invitrogen, cat. no. Q33263)

Gel box and gel electrophoresis apparatus for agarose gel
Razor blade (VWR, cat. no. 55411-050)
Thermal mixer (Eppendorf, cat. no. 5382000023)
1.5-ml tube (Eppendorf, cat. no. 022363301)
Qubit 4 Fluorometer (Invitrogen, cat. no. Q33238)
NanoDrop spectrophotometer (Thermo Fisher, cat. no. 13-400-525)

Table 5 Reaction Components for Digesting Expression Plasmid to Create Linearized Vector

Component	Volume
Nuclease-free water	[a]
Vector	[a]
10× CutSmart buffer	5 µl
SmaI	4 µl
Total volume	50 µl

[a] Nuclease-free water, vector, and CutSmart buffer are from the midiprep kit. Adjust volumes according to the vector DNA concentration. Use between 3 and 5 µg of total plasmid DNA.

1. In a 1.5-ml tube, mix the components listed in Table 5 in the order given.
2. Mix by pipetting gently. Avoid vortexing.
3. Incubate the reaction at 37°C for 2 hr.
4. Add 3 µl Quick CIP and incubate for 1 hr at 37°C.
5. Check if the vector is completely digested by analyzing 4 µl of the digestion reaction, alongside an aliquot of the undigested vector, by gel electrophoresis on a 1% agarose gel. Linearized vector must run according to its expected size.

   If the vector is not completely digested—i.e., bands representative of undigested nonlinear plasmid DNA are visible—add additional restriction enzyme (e.g., 1-2 µl SmaI) and continue incubating the restriction digestion reaction for another 1 hr at 37°C.
6. Once the vector is completely linearized, proceed with gel extraction.
7. Prepare 0.7% agarose gel in 1× TBE buffer.
8. Cast gel in the gel box. Do not cast gel >7 mm in depth.
9. Mix the digestion reaction with 1× gel loading dye and load each sample into a well. Load an undigested vector control. Run gel at 70 V for 50 min.
10. Extract linearized plasmid DNA from gel by cutting it out with a razor. Minimize the volume of gel by cutting as close to the DNA band as possible.
11. Purify plasmid DNA from the gel using the Monarch DNA Gel Extraction Kit according to manufacturer's instructions.
12. Elute DNA by adding 30 µl nuclease-free water.
13. Measure the quantity and purity of linearized vector DNA as detailed in the Basic Protocol, step 19.

    During DNA gel extraction, ∼50% of plasmid DNA can be lost.
14. Use immediately or freeze samples at −20°C. Proceed with ligation as described in the Basic Protocol, step 21.

    Avoid freeze/thaw cycles of vector samples. For ligations, use fresh (i.e., ≥48 hr old) digested vector.

REAGENTS AND SOLUTIONS

Resuspension solution, 1×

125 g sucrose (Sigma-Aldrich, cat. no. S0389-500G)
25 ml 1 M Tris•Cl, pH 8.0 (Invitrogen, cat. no. AM9855G)
Add deionized H_2O to 500 ml final volume
Store up to 6 months at 4°C.

Lysis solution, 1×

62.5 ml 0.25 M EDTA
12.5 ml Tris•Cl, pH 8.0 (Invitrogen, cat. no. AM9855G)
10 ml 25% Triton X-100
Add deionized H_2O to 500 ml final volume
Store up to 6 months at 4°C.

TE buffer

1 ml 1 M Tris•Cl, pH 8.0 (Invitrogen, cat. no. AM9855G)
0.2 ml 0.5 M EDTA, pH 8.0
Add deionized H_2O to 100 ml final volume
Store up to 6 months at 4°C.

COMMENTARY

Critical Parameters and Troubleshooting

Certain steps in this protocol are critical, and special attention should be paid to these points. When preparing the vector for cloning (Support Protocol 2), the linearized vector is gel extracted. Using DNA extracted from an agarose gel often will result in a low yield and quality of the DNA. At least 80 ng/µl of vector should remain after gel extraction (step 11). This quantity of vector is sufficient for all required ligations. When assessing the

Table 6 Summary of Common Problems Encountered During Plasmid Library Construction and Their Potential Solutions

Problem	Possible cause	Solution
Difficulty fragmenting gDNA with g-TUBE	Very long and compacted gDNA strands; high GC content; low gDNA purity	Repeat the fragmentation step until desired fragments are produced. If the problem does not resolve, repeat gDNA extraction.
Low yield of plasmid DNA after gel extraction	Large amounts of agarose from gel slice	Cut a small, precise gel band.
	When eluting with nuclease-free water, pH was not ideal for dissolving DNA	Use an elution buffer and then purify DNA with 1.5× magnetic beads.
Few transformants	Low yield of DNA into the ligation	If using T4 DNA ligase kit, use a ligation enhancer.
	Low ligation efficiency	Ensure that purity of DNA is within the acceptable ranges.
		Change vector/insert ratio.
High percentage of empty vector	Inefficient dephosphorylation	Use >80 U of phosphatase in an overnight reaction.
		Try different phosphatases.
	Inefficient restriction digestion	Increase incubation time of restriction digest.
	Failure to separate cut/uncut vector during gel electrophoresis	Increase time and/or voltage during gel electrophoresis, thus ensuring the linearized vector will separate from nonlinearized vector.
	Poor ligation efficiency	Change insert vector ratio.
		Increase ligase concentration or ligation incubation time.
Insert average size <1000 bp	Contamination with small fragments of DNA	Check integrity of final purified fragmented gDNA to ensure there is minimal degradation below the desired size.
		Purify and size select blunt-end fragmented gDNA again.
		Repeat gDNA extraction and remove small fragments.

purity of the vector, pay special attention to the $A_{260}/_2A_{30}$ ratio which measures alcohol contamination. The ratio should be in the range of 1.7-2.0, ideally 1.8. It is critical not to have any traces of alcohol during ligation, as concentrations >25% can denature nucleic acids and lead to lower ligation efficiency (Behera et al., 2013; Herskovits et al., 1961). If the purity of the vector is not in the acceptable range, re-purify plasmid DNA with magnetic beads, according to the manufacturer's protocol.

Another key step is the extraction of the gDNA from the source organism. If the extracted gDNA is highly fragmented (step 22 of Support Protocol 1) and is visualized as a smear on an agarose gel or tape station, it is not suitable for end-repair and cloning (Basic Protocol). Before starting the Basic Protocol, re-purify the gDNA with a 0.4× ratio of AMPure beads according to manufacturer's instructions. This step helps remove the small DNA fragments (<50 bp) that are preferentially ligated into the plasmid and can lower the quality of the final plasmid library. If a smear is still visible after bead purification, start Support Protocol 1 from the beginning and extract fresh gDNA. A summary of common problems encountered during plasmid gDNA library construction, and their potential solutions, is provided in Table 6.

Understanding Results

As a proof of concept and to validate the results of the protocols presented here, we constructed three libraries, for *Pyrococcus abyssi*, *Aeropyrum pernix*, and *Deinococcus radiophilus* (two thermophilic archaea and an extremophile bacterium). There are two general methods for quantifying the quality of the final gDNA plasmid library: (1) colony PCR of individual transformants and (2) next-generation sequencing (NGS) analysis of the entire library. PCR is sufficient to evaluate the percent of clones that contain an insert. However, colony PCR cannot be used to assess the quality of the final library in terms of sequence diversity and genome coverage. We used both approaches in these library construction protocols: Colony PCR was used to assess the percent of clones with inserts prior to the library amplification step, and NGS was used to analyze the quality of the

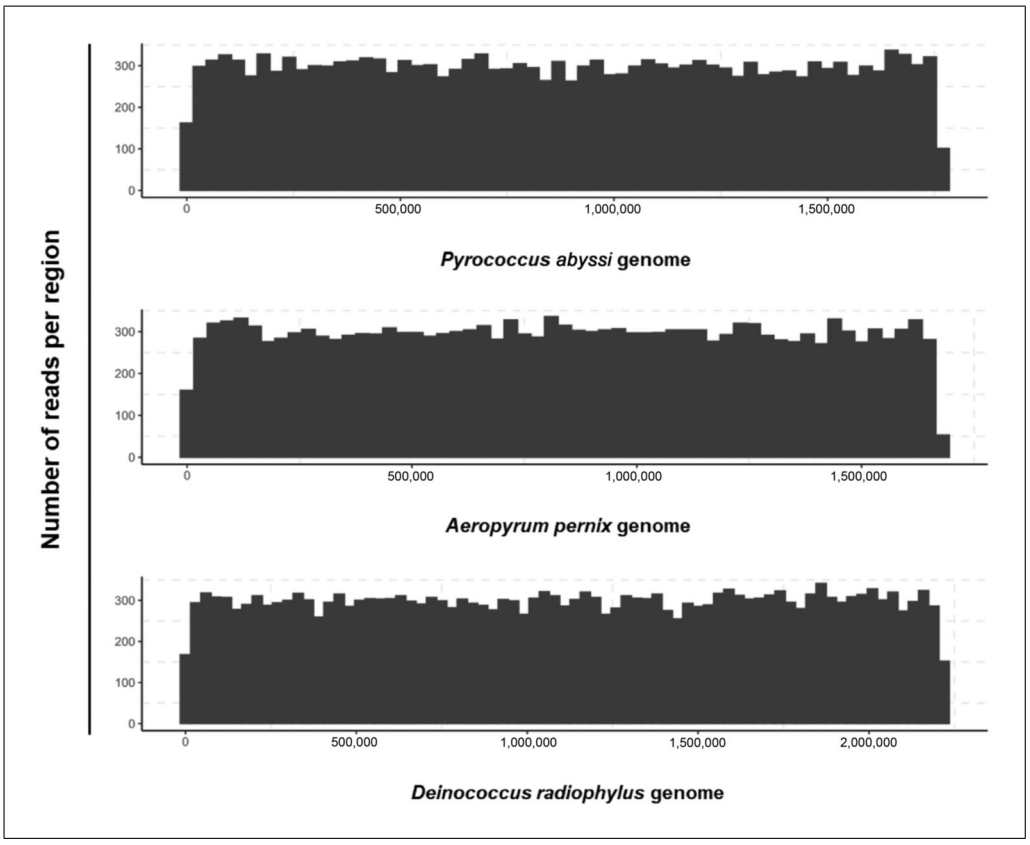

Figure 4 Genome coverage of plasmid gDNA libraries. Reads from the plasmid gDNA library were mapped against the corresponding genomes, referenced in GenBank as *P. abyssi* GCF_000195935.2, *A. pernix* BA000002.3, and *D. radiophilus* CO086380.1. Data are condensed to show genome regions corresponding to 30,000 bp to be grouped. The number of reads per region represents the sequencing depth.

final library. The NGS protocol is described below.

Briefly, 1.5 µg of the plasmid gDNA library was digested with enzyme NmeAIII (NEB, cat. no. R0711S) to linearize the plasmids. Using this linearized plasmid, a HiFi SMRTbell library was prepared using the SMRTbell prep kit 3.0 (PacBio). This library was sequenced using the Sequel II system. The resulting HiFi reads were mapped to the source organism's chromosome using the SMRTlink HiFi mapping tool. All three libraries had near-complete coverage of the genomes (Fig. 4). Next, the sequencing coverage was calculated for each library. First, the sequencing depth for each nucleotide in the genome was determined. A minimum threshold of 10× read depth was used to calculate coverage over each genome. The number of nucleotide positions represented in ten or more reads was divided by the total number of nucleotides in the reference genome, resulting in the percent coverage. Based on these parameters, the gDNA libraries had the following percent coverage: *P. abyssi*, 97%; *D. radiophilus*, 99%; *A. pernix*, 100%.

Some regions of the chromosome were under-represented. This could be due to sequencing bias, toxicity of the gene product, or the presence of repetitive sequences that serve as poor substrates for the native *E. coli* polymerases when the gene products are expressed at high numbers from multi-copy plasmids. To investigate these possibilities, we analyzed a region of *P. abyssi* genome (residues 762590-762882) that had only 1× coverage. This region is an intragenic space containing adenine and thymine repetitive sequences. To improve coverage of these repetitive regions, one can choose an expression vector with low plasmid copy number and/or a weaker promoter.

Acknowledgments

The authors would like to thank to James B. Eaglesham for his help in sequence analysis of the genome; the Agencia Nacional de Investigación e Innovación, Uruguay, grant FCE_1_2021_1_166635 to BM and fellowship POS_FCE_2021_1_1010803 to VF; the Fondo para la Convergencia Estructural del Mercosur COF 03/11-Institut Pasteur of Montevideo; and the late Don Comb for the creative, collaborative environment of New England Biolabs, which permitted this study.

Author Contributions

Valeria Florez: Data curation; investigation; methodology; validation; writing—original draft; writing—review and editing. **Jessica Khani:** Data curation; investigation; methodology; validation. **Emily McNutt:** Methodology; project administration; supervision; writing—review and editing. **Bruno Manta:** Conceptualization; funding acquisition; supervision; writing—review and editing. **Mehmet Berkmen:** Conceptualization; supervision; writing—review and editing.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data available on request from the authors.

Literature Cited

Asakawa, S., Abe, I., Kudoh, Y., Kishi, N., Wang, Y., Kubota, R., Kudoh, J., Kawasaki Minoshima, S., & Shimizu, N. (1997). Human BAC library: Construction and rapid screening. *Gene*, *191*, 69–79.

Bar-Joseph, Z., Gerber, G. K., Lee, T. I., Rinaldi, N. J., Yoo, J. Y., Robert, F., Gordon, D. B., Fraenkel, E., Jaakkola, T. S., Young, R. A., & Gifford, D. K. (2003). Computational discovery of gene modules and regulatory networks. *Nature Biotechnology*, *21*(11), 1337–1342.

Behera, A. K., Schlund, K. J., Mason, A. J., Alila, K. O., Han, M., Grout, R. L., & Baum, D. A. (2013). Enhanced deoxyribozyme-catalyzed RNA ligation in the presence of organic cosolvents. *Biopolymers*, *99*(6), 382–391.

Benoit, R. M., Ostermeier, C., Geiser, M., Li, J. S., Widmer, H., & Auer, M. (2016). Seamless insert-plasmid assembly at high efficiency and low cost. *PLoS One*, *11*(4), e0153158.

Bircher, L., Geirnaert, A., Hammes, F., Lacroix, C., & Schwab, C. (2018). Effect of cryopreservation and lyophilization on viability and growth of strict anaerobic human gut microbes. *Microbial Biotechnology*, *11*(4), 721–733.

Biver, S., Steel, S., Portetelle, D., & Vandelbol, M. (2013). *Bacillus subtilis* as a tool for screening soil metagenomic libraries for antimicrobial activities. *Journal of Microbiology and Biotechnology*, *23*(6), 850–855.

Chow, P. (2005). Cloning of λ DNA fragments into pUC19 vector to study the ligation efficiency of NdeI-digested and HindII-digested pUC19 by T4 DNA ligase. *Journal of Entrepreneurship, Management, and Innovation*, *8*, 8–13.

Chuzel, L., Ganatra, M. B., Rapp, E., Henrissat, B., & Taron, C. H. (2018). Functional metagenomics identifies an exosialidase with an inverting catalytic mechanism that defines a new glycoside hydrolase family (GH156). *Journal of Biological Chemistry*, *293*(47), 18138–18150.

Craig, J. W., Chang, F., Kim, J. H., Obiajulu, S. C., & Brady, S. F. (2013). Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse

Proteobacteria. *Applied and Environmental Microbiology*, *76*(5), 1633–1641.

Covaris. (2020). g-TUBE for DNA shearing. M020006_Rev_Jan2020.

Díaz-Rullo, J., Rodríguez-Valdecantos, G., Torres-Rojas, F., Cid, L., Vargas, I. T., González, B., & González-Pastor, J. E. (2021). Mining for perchlorate resistance genes in microorganisms from sediments of a hypersaline pond in Atacama desert, Chile. *Frontiers in Microbiology*, *12*, 723874.

Dobrijevic, D., di Liberto, G., Tanaka, K., de Wouters, T., Dervyn, R., & Boudebbouze, S. (2013). High-throughput system for the presentation of secreted and surface-exposed proteins from gram-positive bacteria in functional metagenomics studies. *PLoS ONE*, *8*(6), e65956.

Durin, G., Boles, C., & Ventura, P. (2012). Complementary DNA shearing and size selection tools for mate-pairs library construction. *Journal of Biomolecular Techniques*, *23*, (suppl), S36–S37.

Escuder-Rodríguez, J. J., González-Suarez, M., deCastro, M. E., Saavedra-Bouza, A., Becerra, M., & González-Siso, M.-I. (2022). Characterization of a novel thermophilic metagenomic GH5 endoglucanase heterologously expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. *Biotechnol Biofuels*, *15*(1), 76–83.

Farrar, K., & Donnison, I. (2007). Construction and screenings of BAC libraries made from Brachypodium genomic DNA. *Nature Protocols*, *2*, 1661–1674.

Felczykowska, A., Dydecka, A., Bohdanowicz, M., Gasior, T., Sobon, M., Kobos, J., Bloch, S., Nejman-Falenczyk, B., & Wegrzyn, G. (2014) The use of fosmid metagenomic libraries in preliminary screening for various biological activities. *Microbial Cell Factories*, *13*, 105.

Gallagher, S. R., & Desjardins, P. R. (2006). Quantitation of DNA and RNA with absorption and fluorescence spectroscopy. *Current Protocols in Molecular Biology*, *76*(1), A.3D.1–A.3D.21.

Gansauge, M. T., Gerber, T., Glocke, I., Korlević, P., Lippik, L., Nagel, S., Riehl, M., Schmidt, A., & Meyer, M. (2017). Single-stranded DNA library preparation from highly degraded DNA using *T4* DNA ligase. *Nucleic Acids Research*, *45*(10), e79.

González-Pastor, J. E., & Mirete, S. (2010). Novel metal resistance genes from microorganisms: A functional metagenomic approach. *Methods in Molecular Biology*, *668*, 273–285.

Green, J., & Rao, V. B. (1998). Recombinant DNA principles and methodologies. In *Formulation of cloning strategies*. Marcel Dekker, Inc.

Green, M. R., & Sambrook, J. (2020a). Alkaline phosphatase. *Cold Spring Harbor Protocols*, *2020*(8), 100768.

Green, M. R., & Sambrook, J. (2020b). Dephosphorylation of DNA fragments with alkaline phosphatase. *Cold Spring Harbor Protocols*, *2020*(8), 100669.

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., & Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chemistry & Biology*, *5*(10), 245–249.

Herskovits, T. T., Singer, S. J., & Geiduschek, E. P. (1961). Nonaqueous solutions of DNA. Denaturation in methanol and ethanol. *Archives of Biochemistry and Biophysics*, *94*, 99–114.

Horspool, D. R., Coope, R. J., & Holt, R. A. (2010). Efficient assembly of very short oligonucleotides using T4 DNA ligase. *BMC Research Notes*, *3*, 291.

Janner, C. R., Brito, A. L., Moraes, L. M., Reis, V. C., & Torres, F. A. (2013). pPCV, a versatile vector for cloning PCR products. *Springerplus*, *2*, 441.

Liu, Z. G., & Schwartz, L. M. (1992). An efficient method for blunt-end ligation of OCR products. *Biotechniques*, *12*(1), 28–30.

Loock, M., Antunes, L. B., Heslop, R. T., de Lauri, A., Brito-Lira, A., & Cestari, I. (2023). High-efficiency transformation and expression of genomic libraries in yeast. *Methods and Protocols*, *6*(5), 89–94.

McNutt, E., Ke, N., Thurman, A., Eaglesham, J. B., & Berkmen, M. (2024). SAS: Split antibiotic selection for identifying chaperones that improve protein solubility. *Heliyon*, *10*(5), e26996. https://doi.org/10.1016/j.heliyon.2024.e26996

Motohashi, K. (2019). A novel series of high-efficiency vectors for TA cloning and blunt-end cloning of PCR products. *Scientific Reports*, *9*, 6417.

Nakayama, Y., Yamaguchi, H., Eigana, N., & Esumi, M. (2016). Pitfalls of DNA quantification using DNA-binding fluorescent dyes and suggested solutions. *PLOS ONE*, *11*(3), e0150528.

Nilewski, S., Varatnitskaya, M., Masuch, T., Kusnezowa, A., Gellert, M., Baumann, A. F., Lupilov, N., Kusnezow, W., Koch, M. H., Eisenacher, M., Berkmen, M., Lillig, C. H., & Leichert, L. I. (2022). Functional metagenomics of the thioredoxin superfamily. *Journal of Biological Chemistry*, *296*, e100247.

Osburn, E. D., McBride, S. G., & Strickland, M. S. (2024). Microbial dark matter could add uncertainties to metagenomic trait estimations. *Nature Microbiology*, (9), 1427–1430.

Panaro, J., Po, K., Yuen, I., Sakazume, T., Fortina, P., Kricka, L., & Wilding, P. (2000). Evaluation of DNA fragment sizing and quantification by the Agilent 2100 Bioanalyzer, *Clinical Chemistry*, *46*(11), 1851–1853.

Pascal, A. V., Augustijn, H. E., van den Berg, K., van der Hooft, J., Fischbach, M. A., & Medema, M. H. (2021). BiG-MAP: An automated pipeline to profile metabolic gene cluster abundance and expression in microbiomes. *mSystems*, *6*(5), e0093721.

Pheiffer, B. H., & Zimmerman, S. B. (1983). Polymer-stimulated ligation: Enhanced blunt- or cohesive-end ligation of DNA or

deoxyribooligonucleotides by T4 DNA ligase in polymer solutions. *Nucleic Acids Research*, *11*(22), 7853–7871.

Potapov, V., Ong, J., Kucera, R., Langhorst, B., Bilotti, K., Pryor, M., Cantor, E., Canton, B., Knight, T., Evans, T., & Lohman, G. (2018). Comprehensive profiling of four base overhang ligation fidelity by T4 DNA ligase and application to DNA assembly. *ACS Synthetic Biology*, *7*(11), 2665–2674.

Rich, M. H., Sharrock, A. V., Mulligan, T. S., Matthews, F., Brown, A. S., Lee-Harwood, H. R., Williams, E. M., Copp, J. N., Little, R. F., Francis, J. J., Horvat, C. N., Stevenson, L. J., Owen, J. G., Saxena, M. T., Mumm, J. S., & Ackerley, D. F. (2023). A metagenomic library cloning strategy that promotes high-level expression of captured genes to enable efficient functional screening. *Cell Chemical Biology*, *30*(12), 1–12.

Rossi, R., Montecuco, A., Ciarocchi, G., & Biamonti, G. (1997). Functional characterization of the T4 DNA ligase: A new insight into the mechanism of action, *Nucleic Acids Research*, *25*(11), 2106–2113.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular cloning: A laboratory manual. In *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press.

Sévin, D., Fuhrer, T., Zamboni, N., & Sauer, U. (2017). Nontargeted *in vitro* metabolomics for high throughput identification of novel enzymes in *Escherichia coli*. *Nature Methods*, *14*, 187–194.

Simon, C., & Daniel, R. (2017). Construction of small-insert and large-insert metagenomic libraries. *Methods in Molecular Biology*, *1539*, 1–12.

Villamizar, G. A., Nacke, H., & Daniel, R. (2017). Function-based metagenomic library screening and heterologous expression strategy for genes encoding phosphatase activity. *Methods in Molecular Biology*, *1539*, 249–260.

Wagner, A. O., Praeg, N., Reitschuler, C., & Illmer, P. (2015). Effect of DNA extraction procedure, repeated extraction and ethidium monoazide (EMA)/propidium monoazide (PMA) treatment on overall DNA yield and impact on microbial fingerprints for bacteria, fungi and archaea in a reference soil. *Applied Soil Ecology*, *93*, 56–64.

Walker, M., Simpson, J., & Redinbo, M. (2022). A structural metagenomics pipeline for examining the gut microbiome. *Current Opinion in Structural Biology*, *75*, e102416.

Wan, K. H., Yu, C., George, R. A., Carlson, J. W., Hoskins, R. A., Svirskas, R., Stapleton, M., & Celniker, S. E. (2006). High-throughput plasmid cDNA library screening. *Nature Protocols*, *1*, 624–632.

Wei-Po, L., & Wen-Shyong, T. (2009). Computational methods for discovering gene networks from expression data. *Briefings in Bioinformatics*, *10*(4), 408–423.

Wilson, G. G., & Murray, N. E. (1979). Molecular cloning of the DNA ligase gene from bacteriophage T4, *Journal of Molecular Biology*, *132*(3), 471–491.

Wilson, R., Morton, S., Deiderick, H., Gerth, M., Paul, H., Gerber, I., Patel, A., Ellington, A., Hunicke-Smith, S., & Patrick, W. (2013). Engineered DNA ligases with improved activities *in vitro*. *Protein Engineering, Design and Selection*, *26*(7), 471–478.

Yaung, S., Deng, S., Li, N., Braff, J. L., Church, G., Bry, L., Wang, H., & Gerber, G. (2015). Improving microbial fitness in the mammalian gut *in vivo* temporal functional metagenomics. *Molecular System Biology*, *11*(3), 788.

Zhou, M., Abid, M., Yin, H., Wu, H., Teklue, T., Qiu, H. J., & Sun, Y. (2018). Establishment of an efficient and flexible genetic manipulation platform based on a fosmid library for rapid generation of recombinant pseudorabies virus. *Frontiers in Microbiology*, *9*, e21 32.