nature protocols

PROTOCOL UPDATE

https://doi.org/10.1038/s41596-020-0338-0

Check for updates

Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA

Marie-Theres Gansauge ✉, Ayinuer Aximu-Petri, Sarah Nagel and Matthias Meyer ✉

It has been shown that highly fragmented DNA is most efficiently converted into DNA libraries for sequencing if both strands of the DNA fragments are processed independently. We present an updated protocol for library preparation from single-stranded DNA, which is based on the splinted ligation of an adapter oligonucleotide to the 3′ ends of single DNA strands, the synthesis of a complementary strand using a DNA polymerase and the addition of a 5′ adapter via blunt-end ligation. The efficiency of library preparation is determined individually for each sample using a spike-in oligonucleotide. The whole workflow, including library preparation, quantification and amplification, requires two work days for up to 16 libraries. Alternatively, we provide documentation and electronic protocols enabling automated library preparation of 96 samples in parallel on a Bravo NGS Workstation (Agilent Technologies). After library preparation, molecules with uninformative short inserts (shorter than ~30–35 base pairs) can be removed by polyacrylamide gel electrophoresis if desired.

This protocol is an update to *Nat. Protoc.* 8, 737–748 (2013): https://doi.org/10.1038/nprot.2013.038

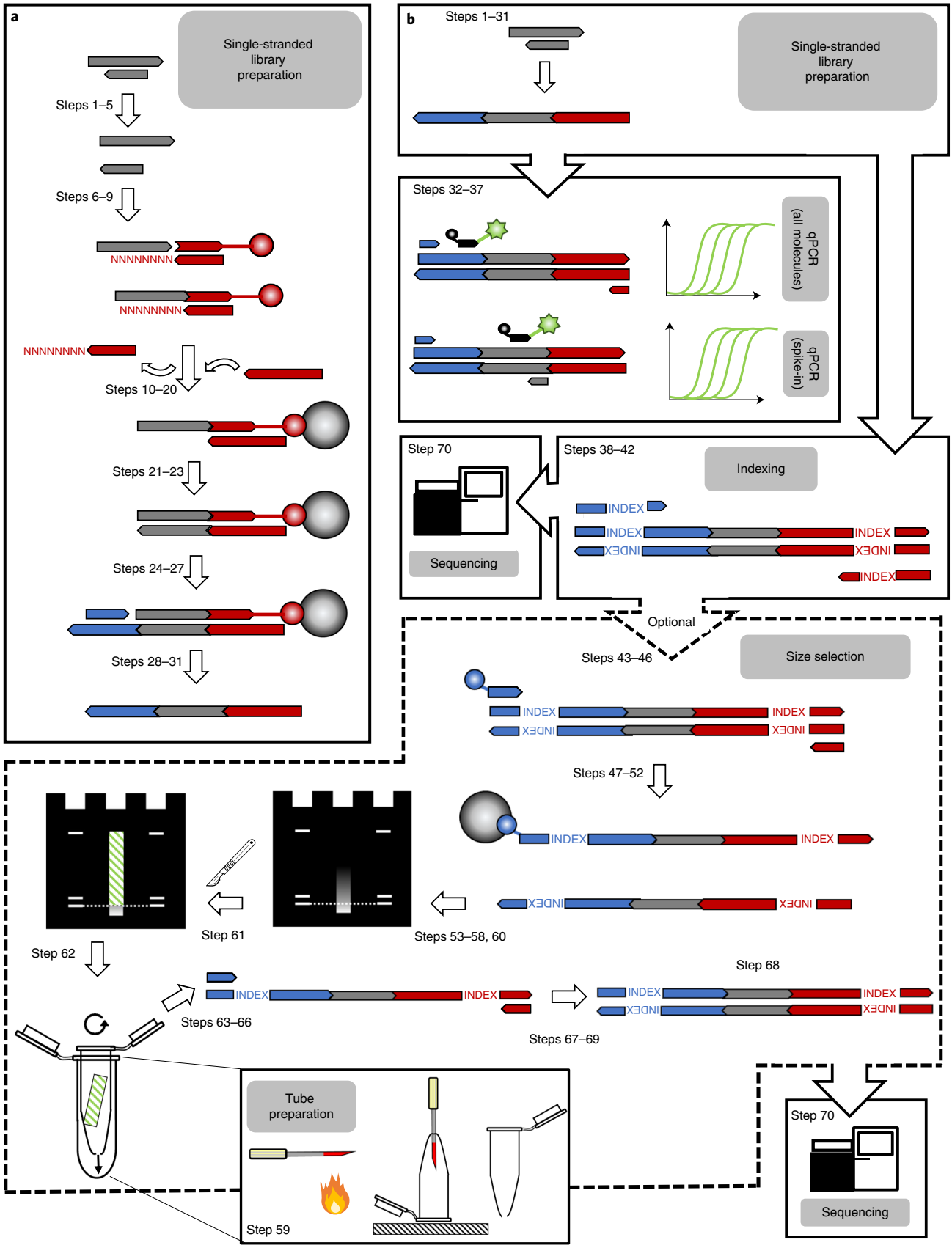
Introduction

Development of the protocol

The preparation of DNA libraries is a critical step in the preparation of samples for high-throughput DNA sequencing, particularly if only small quantities of highly degraded DNA are present, such as in ancient biological material. In 2012, we introduced a library preparation method that differed fundamentally from previously existing methods in that it relies on the heat denaturation of double-stranded DNA fragments and the subsequent conversion of single DNA strands into library molecules[1,2]. Single-stranded library preparation minimizes the loss of short DNA fragments and those with single-stranded breaks, thereby increasing the number of library molecules that can be retrieved from highly degraded DNA. The power of the method was first demonstrated by the generation of a high-coverage genome sequence from just a few milligrams of bone powder obtained from a finger bone of a Denisovan individual, an extinct Pleistocene hominin[1]. Since then, it has made the recovery of additional high-quality genomes from ancient DNA possible, including those of three Neandertals[3–5] and an early modern human[6]. Single-stranded DNA library preparation was also instrumental for the retrieval of DNA sequences from the ~430,000-year-old bear and hominin remains from Sima de los Huesos in Spain[7–9], which are by far the oldest remains outside of permafrost that have been sequenced to date. More recently, it has allowed the recovery of Neandertal and Denisovan DNA from Pleistocene cave sediments[10].

A schematic overview of the single-stranded library preparation protocol provided here is shown in Fig. 1a. Briefly, after heat denaturation of the double-stranded DNA fragments, a biotinylated adapter is ligated to the 3′ end of each single-stranded molecule. After immobilizing the ligation products on beads, the template strand is copied by extending a primer hybridized to the adapter using a DNA polymerase, which creates a double-stranded DNA fragment to which the second adapter is joined by blunt-end ligation. The library is then released from the beads at a high

Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. ✉e-mail: marie_gansauge@eva.mpg.de; mmeyer@eva.mpg.de


a
Single-stranded library preparation
Steps 1–5
Steps 6–9
NNNNNNNN
NNNNNNNN
NNNNNNNN
Steps 10–20
Steps 21–23
Steps 24–27
Steps 28–31
b
Steps 1–31
Single-stranded library preparation
Steps 32–37
qPCR (all molecules)
qPCR (spike-in)
Step 70
Sequencing
Steps 38–42
Indexing
INDEX
Optional
Steps 43–46
Size selection
Steps 47–52
Steps 53–58, 60
Step 61
Step 62
Steps 63–66
Step 68
Steps 67–69
Tube preparation
Step 59
Step 70
Sequencing

◀ **Fig. 1 | Schematic overview of the workflow described in this protocol. a**, Single-stranded library preparation is initiated by heat denaturation of the sample DNA and the removal of 5′ and 3′ phosphate groups (if present). Subsequently, an adapter oligonucleotide (red) carrying a 5′ phosphate and a 3′ biotin (indicated by a circle) is ligated to the 3′ end of each DNA fragments using *T4* DNA ligase. This reaction is enabled by a splinter oligonucleotide carrying a stretch of eight random nucleotides at its 3′ end. The ligation products as well as excess adapters are immobilized on streptavidin-covered magnetic beads (gray circles), and splinter oligonucleotides are removed by a wash step at an elevated temperature. A primer is hybridized to the adapter oligonucleotide and a copy of the template strand synthesized using the Klenow fragment of *E. coli* DNA polymerase I. The primer contains phosphorothioate backbone modifications to prevent its exonucleolytic degradation (not depicted). Unincorporated primers are removed through a bead wash at an elevated temperature, preventing the formation of adapter dimers during the subsequent blunt-end ligation of a second, double-stranded adapter (blue), a reaction again catalyzed by *T4* DNA ligase. One of the adapter strands carries a 3′ dideoxy modification to prevent self-ligation of adapters. The library strand is released from the beads by heat denaturation. **b**, After library preparation, the yield of library molecules as well as the number of successfully converted control oligonucleotides, which were spiked-into the sample DNA before library preparation, is determined using two probe-based qPCR assays. Libraries are amplified and indexed via PCR with 5′ tailed primers using optimal cycle numbers inferred from qPCR. Libraries are sequenced directly or undergo an optional size-selection step. For size selection, one of the library strands is biotinylated using a four-cycle PCR, the library is immobilized on streptavidin-coated beads and the non-biotinylated strand is isolated by incubation with an alkaline solution. Size separation is performed on a denaturing polyacrylamide gel. Markers loaded to the left and right of the library guide the excision of gel slice containing library molecules of desired lengths. The gel slice is crushed by centrifugation through a perforated tube and the DNA is extracted from the gel, amplified by PCR and sequenced.

temperature, and its concentration is determined using quantitative PCR (qPCR). A final amplification by PCR introduces the full-length adapter sequences required for sequencing using Illumina technology as well as pairs of sample-specific indices.

Since the first detailed protocol for single-stranded library preparation was published in *Nature Protocols* in 2013[2], the method has been continuously refined. Some of the previously reported improvements are as follows. (1) *Bst* DNA polymerase was replaced by the Klenow fragment of *Escherichia coli* DNA polymerase I in the primer extension step, which eliminated the blunt-end repair step that was previously required before the blunt-end ligation of the second adapter[11]. (2) The ligation of the first adapter to the single-stranded sample molecules is now performed using splinter-mediated ligation with *T4* DNA ligase at 37 °C instead of using CircLigase at 60 °C[12,13]. The change in ligase greatly reduced reagent costs, improved the robustness of the protocol and made it compatible with pipetting on automated liquid handling systems[10] where tubes cannot be sealed to avoid evaporation. (3) An oligonucleotide spike-in was introduced that allows monitoring the efficiency of library preparation for each sample and detecting inhibition of enzymes that might result from impurities in the sample DNA[14].

In addition to the previously published modifications above, the protocol provided here encompasses modified oligonucleotide sequences, which make the method robust to fluctuations in the synthesis quality of the oligonucleotides used. This has been achieved by replacing spacer modifications in the adapter and splinter oligonucleotides by 2′-*O*-methyl-ribonucleotides, which are cheaper and easier to synthesize and effectively prevent incorporation of fragmented oligonucleotides into the library. Beyond other small optimizations that improve ease of use, we also describe an optional size selection procedure that allows for efficient removal of library molecules with extremely short inserts (<35 base pairs (bp)) after library amplification. Such molecules, which can result from the extraction of extremely short DNA fragments from highly degraded samples, are usually uninformative and consume unnecessary sequencing capacity if they make up a substantial proportion of the library[15]. Finally and importantly, we provide electronic protocol files (https://zenodo.org/record/3631147) and a Supplementary Manual that allow automated library preparation in a 96-well format on a Bravo NGS Workstation (Agilent Technologies). Automation of sample preparation is becoming increasingly important for studies where the analysis of hundreds or thousands of samples is desirable, such as the characterization of ancient environmental DNA.

Applications and comparison with other methods

Several studies comparing library preparation methods have confirmed that yields of library molecules from ancient biological material are substantially higher with the single-stranded method than with double-stranded library preparation[12,16,17]. The average improvement over double-stranded methods was estimated to ~ten-fold, but gains can be expected to be even higher with particularly heavily degraded material. We therefore recommend the use of single-stranded library preparation for all cases where destructive sampling should be kept to a minimum or where DNA preservation is

expected to be very poor. This might include, for example, material older than ~50,000 years that was not preserved under permafrost conditions or material from warm environments. For relatively well-preserved material, simpler double-stranded methods remain a viable alternative[18,19]. As lab automation eliminates most of the manual handling steps, there is no benefit, in our view, to performing double-stranded library preparation if liquid handling systems are used.

An additional benefit of single-stranded library preparation is that it retains the full-length sequences and original strand orientation of the DNA molecules. This becomes apparent when comparing the effect of the presence of uracils on sequences generated with single- and double-stranded library preparation. Uracils result from cytosine deamination in ancient DNA and accumulate preferentially at the ends of molecules. Whereas uracils manifest as cytosine-to-thymine (C-to-T) substitutions in single-stranded libraries[1], libraries prepared with double-stranded methods carry both C-to-T and guanine-to-adenine (G-to-A) substitutions[20]. The latter are caused by blunt-end repair, in which a DNA polymerase with 3′−5′ exonuclease activity is used to add or remove nucleotides from the 3′ ends of molecules. Although deamination-induced sequence differences complicate the differentiation between base damage and real sequence differences, they can also be exploited to distinguish authentic ancient DNA from recent contamination in silico during sequence analysis[8,21] or even physically at the stage of library preparation using uracil selection methods[22,23]. The latter methods are also based on single-stranded library preparation and have proven useful for massive sequence generation from samples that are particularly heavily contaminated with recent human DNA[23]. However, these methods are not a suitable alternative to regular single-stranded library preparation for routine applications.

The utility of single-stranded library preparation is not confined to ancient DNA research. Studies by us and other groups have shown that single-stranded library preparation increases the number of DNA molecules from formalin-fixed tissue that can be made accessible for sequencing by several hundred-fold or even thousand-fold[12,24]. In addition, the method has been shown to improve the recovery of sequences from cell-free DNA and to more accurately retain the size distribution and base composition of the sample DNA[12,25,26]. Single-stranded library preparation is thus gaining substantial popularity also in biomedical research. Over the years, variations of the original single-stranded library preparation method, as well as new approaches to single-stranded library preparation, have been reported for various applications, such as cell-free DNA sequencing[27,28], methylation mapping[29], chromatin and DNA replication analyses[30–32] or Rad-Seq of museum samples[33]. In addition, commercial kits for single-stranded library preparation are available from Swift Biosciences (e.g. the Accel-NGS 1S Plus DNA Library Kit) and Claret Bioscience (e.g. the SRSLY NGS Library Prep Kit). Whereas the Swift Biosciences kit relies on a workflow similar to the one described here (addition of a first adapter and primer extension and ligation of a second adapter), the SRSLY method provides a simpler workflow that is based on the simultaneous splinted ligation of two adapters to both ends of molecules[34]. In contrast to the protocol presented here, both commercial systems include size-selective purification steps after adapter ligation, which are likely to reduce the yield of library molecules, especially those with short inserts. This might be particularly detrimental for very heavily degraded material such as the Sima de los Huesos fossils, where almost all surviving molecules are shorter than 45 bp[8]. However, no direct comparisons have been made between those kits, or any other of the aforementioned methods, and single-stranded library preparation as described here, and none of them has been applied to highly degraded ancient DNA in published research. As it is unclear whether there are large differences in performance, we do not advise against using a kit, especially if only a small number of reactions need to be performed. In this case, the initial investment in reagents for the protocol described here might be higher than the price of a kit. For large numbers of samples (hundreds or thousands), our protocol allows single-stranded library preparation, indexing, library quantification and purification at a total cost of only ~12.50 € per sample for reagents (based on German list prices for reagents).

Experimental design

Laboratory environment

It is common practice that sample preparation for genetic analysis of ancient biological material is performed in a dedicated ancient DNA clean room to minimize contamination with exogenous DNA from humans and other sources[35]. Such facilities do not usually exist in laboratories working with other types of material. As single-stranded library preparation was developed to enable sequencing of trace amounts of highly degraded DNA, it is critical, irrespective of the material under study, to create

a workspace for DNA extraction and library preparation that reduces the risk of DNA contamination as much as possible. This could, for example, be a fume hood that was thoroughly cleaned with a sodium hypochlorite (bleach) solution[36] before starting the experiments and that is located in a laboratory where no large quantities of DNA, such as PCR products, are handled. In addition, clean pipettes, filter tips and fresh reagents should be used and the negative controls closely monitored.

Preparing sample DNA for library preparation

The protocol described here was optimized for the conversion of up to 50 ng of short single- or double-stranded DNA molecules (smaller than ~250 bp) into DNA libraries. The efficiency of library preparation can decrease if larger quantities or longer DNA fragments are used as input for library preparation. However, this is hardly ever a concern when working with ancient bones, teeth or sediments. For this type of material, we recommend DNA extraction from 50 mg of sample powder, ideally using a silica-based method, such as specified in the protocol by Rohland et al.[37], and the subsequent conversion of 20% of the DNA extract into a single-stranded library (e.g. using 10 µl of input if the extract volume is 50 µl). We have processed thousands of samples with this strategy over the last years and rarely observed inefficiencies in library preparation that appeared to be due to a saturation with input DNA. Quantification of the DNA extract is therefore unnecessary, in our experience. In fact, it is a strength of the method presented here that it allows successful library preparation and sequencing also in cases where DNA concentrations are so low that they cannot be accurately determined using spectrophotometry or fluorescence measurements with intercalating dyes, which is frequently the case with ancient biological material. Instead, the protocol presented here provides an indirect measure of the DNA content of the sample material through counting the number of library molecules that are obtained by qPCR.

Controls

To verify successful library preparation, we recommend the inclusion of several controls with each experiment. These are, first, a positive control consisting of 0.1 pmol (~6×10^{10} molecules) of a 40-nucleotide oligonucleotide (CL304, see Table 1 for oligonucleotide sequences). If library preparation was successful, this control should yield at least 6×10^9 library molecules, indicating a conversion rate of molecules into the library of greater than 10%. Second, one or more library negative controls without sample DNA should be included, as well as an extraction negative control consisting of a mock reaction that was carried through the DNA extraction process. Depending on the synthesis quality of the adapter and splinter oligonucleotides, the negative controls typically yield between 10^7 and 5×10^8 library molecules and are dominated by artifacts generated during library preparation, for example through the incorporation of adapter or splinter oligonucleotides into the library. All negative controls should be included in sequencing to estimate the level of contamination with human or other undesired DNA (see 'Anticipated results' section below). Third, in addition to dedicated positive control reactions, ~6 million copies of oligonucleotide CL304 are spiked into each library preparation reaction. The number of spike-in molecules is so small that their contribution to the final library is negligible. However, through a qPCR assay that measures the number of control library molecules generated in each reaction, the spike-in makes it possible to determine the efficiency of library preparation separately for each sample relative to the negative controls (see Box 1). This allows for the detection of sporadic inefficiencies in library preparation, which can result, for example, from pipetting errors or the presence of inhibitory substances that were co-purified during DNA extraction.

Uracil-DNA-glycosylase treatment

Although deamination-aware algorithms are available that allow highly accurate genotype calling from high-coverage sequence data[38], the analysis of low-coverage sequence data from ancient DNA is complicated by the occurrence of deamination-induced C-to-T substitutions. In the previous library preparation protocol[2], we suggested treatment of the sample DNA with uracil-DNA-glycosylase (UDG) before library preparation to reduce the effect of deamination on sequence analysis. Two enzymes are available for this purpose. The first, *E. coli* UDG, excises uracils from the interior of molecules but shows reduced activity for uracils located at the 5′ terminus or within the two terminal 3′ nucleotides of molecules[1,39]. The second, *Archaeoglobus fulgidus* (*Afu*) UDG, is more efficient in the removal of terminal uracils[2].

Even though *E. coli* UDG treatment is presented as an option in the current protocol, in recent years we have omitted UDG treatment of sample DNA. This decision was based on the notion that

Table 1 | Oligonucleotide sequences used in the protocol

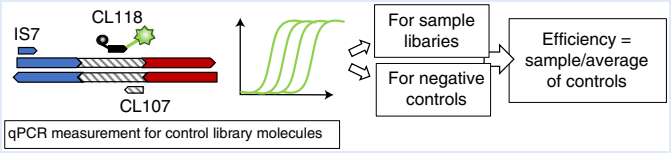
Name	Description	Sequence (5′–3′)[a]	Stock/working dilution [μM], (purification)[b]
Library preparation oligonucleotides (dissolve and dilute in TE buffer)			
TL181	1st adapter	Phosphate-AGATCGGAAGAAA[A][A][A][A][A][A][A]-TEG-Biotin	100/- (Des.)
TL159	Splinter	SpacerC12-[A][A][A]CTTCCGATCTNNNNNNNN[A]-AminoC6	100/- (Des.)
CL128	Extension primer	GTGACTGGAGTTCAGACGTGTGCTCTTCC*G*A*T*C*T	100/- (Des.)
CL53	2nd adapter, strand 1	CGACGCTCTTC-ddC	500/- (HPLC)
TL178	2nd adapter, strand 2	Phosphate-GGAAGAGCGTCGTGTAGGGAAAGAGTGTA	500/- (Des.)
Positive control template (dissolve and dilute in TET buffer)			
CL304	Control DNA	Phosphate-ATTCAGCTCCGGTTCCCAACGATCAAGGCGAGT TACATGA-Phosphate	100/0.1 (HPLC)
Library amplification primers (dissolve and dilute in water)			
P5 index primers	Forward primers	See Supplementary Table 2 in ref. [2]	100/10 (RPC)
P7 index primers	Reverse primers	See Supplementary Table 2 in ref. [2]	100/10 (RPC)
IS5	Forward primer	AATGATACGGCGACCACCGA	100/10 (Des.)
IS5 biotinylated	Forward primer	Biotin-AATGATACGGCGACCACCGA	100/10 (Des.)
IS6	Reverse primer	CAAGCAGAAGACGGCATACGA	100/10 (Des.)
Primers used for preparing the qPCR standard (dissolve and dilute in water)			
CL105	Forward primer	ACACTCTTTCCCTACACGACGCTCTTCCTCGTCGTTTGGTATGG CTTC	100/10 (Des.)
CL106	Reverse primer	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCATGTAACT CGCCTTGATCGT	100/10 (Des.)
Oligonucleotides used for qPCR (dissolve and dilute in TE buffer)			
IS7	Forward primer A&B	ACACTCTTTCCCTACACGAC	100/10 (HPLC)
IS8	Reverse primer A	GTGACTGGAGTTCAGACGTGT	100/10 (HPLC)
IS10	Probe A	FAM-A{G}A{T}C{G}GAAGAGC{A}CAC-BHQ1	100/10 (RP-HPLC)
CL107	Reverse primer B	TCATGTAACTCGCCTTGATCGT	100/10 (HPLC)
CL118	Probe B	FAM-TTCAGCTCCGGTTCCCAACGAT-BHQ1	100/10 (HPLC)
Primers used for preparing the markers for size selection (dissolve and dilute in water)			
i30F	Forward primer	Biotin-TACAATCTGCTCTGATGCCG	100/10 (Des.)
i30R	Reverse primer	GGTGATGACGGTGAAAACCT	100/10 (Des.)
i35F	Forward primer	Biotin -GTGGTTTGTTTGCCGGATCA	100/10 (Des.)
i35R	Reverse primer	AACAGGATTAGCAGAGCGAG	100/10 (Des.)
i40F	Forward primer	Biotin -TACCTCGCTCTGCTAATCCT	100/10 (Des.)
i40R	Reverse primer	TCAGTTCGGTGTAGGTCGTT	100/10 (Des.)
i150F	Forward primer	Biotin -ACATTTCCGTGTCGCCCTTA	100/10 (Des.)
i150R	Reverse primer	TGAGAATAGTGTATGCGGCG	100/10 (Des.)
Sequencing primer (dissolve and dilute in water)			
CL72	Sequencing primer	ACACTCTTTCCCTACACGACGCTCTTCC	100/- (IE-HPLC)

[a]ddC, dideoxycytidine; TEG, triethylene glycol spacer; *, phosphothioate linkage; [N], 2′-*O*-methyl-RNA; {N}, locked nucleic acid (LNA); SpacerC12, 12-carbon spacer; FAM, fluorescein amidite; BHQ1, Black Hole Quencher-1. [b]HPLC, high-performance liquid chromatography; Des., Desalted; RP(C), reverse phase (cartridge); IE, ion exchange.

uracil excision fragments the DNA further, thereby reducing library yields. In addition, it is often preferable to retain the full deamination signal to maximize the power for enriching sequences from deaminated molecules in downstream analyses to deplete contamination. In many cases, population genetic analyses focus on the analysis of variants that were pre-ascertained in high-quality genomes. If these variants are transversions, the effect of deamination in the analysis is usually negligible. Furthermore, as single-stranded library preparation fully retains the information about which strand was sequenced, the orientation of sequence reads can be taken into account to distinguish between C-to-T substitutions that are possibly caused by deamination and G-to-A substitutions that are not, allowing transition variants to be confidently identified in ancient DNA sequences. On the other hand, UDG treatment can help to reduce mapping bias by minimizing the number of differences between the sequenced fragments and the reference genome, which might be particularly useful when working with organisms for which no good reference genome is available.

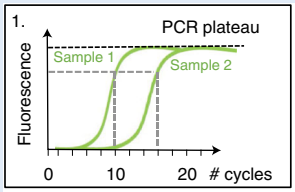
Box 1 | Detecting sporadic inefficiencies in library preparation

The probe-based qPCR assay depicted below (Steps 32–37 of the protocol) allows for independently determining the number of library molecules obtained from the control oligonucleotide in each reaction, including the negative controls containing no sample DNA. In theory, the number of control library molecules should be uniform across all reactions if library preparation was correctly performed for all samples. However, impurity of the sample DNA might occasionally lead to an inhibition of enzymatic reactions during library preparation, reducing library preparation efficiency. Library preparation efficiency can be estimated by comparing the number of control library molecules generated in the sample libraries to those in the negative controls, as indicated below.



Box 2 | Determining the optimal cycle number for library amplification using qPCR amplification plots

1 Analyze the qPCR amplification plot independently for each sample to determine the number of cycles until which amplification remains in the exponential phase.
2 Determine the relative quantity of sample DNA that will be used for indexing PCR, taking differences in reaction volume, input volume and the library dilution into account.
3 Assuming full PCR efficiency (i.e. a doubling of molecules in each cycle), determine the number of cycles that have to be subtracted to compensate for the above differences in the amplification reaction.



2.	Reaction volume (µl)	Input of DNA (µl)	Dilution
qPCR	25 (4)	1 (20)	1:50 (50)
Indexing PCR	100	20	-

3. e.g. Sample 2 (16 cycles in qPCR)

4× the reaction volume	$2^n = 4$, $n = 2$	(2 cycles more)
20× more sample DNA	$2^n = 20$, $n = 4.3$	(~4 cycles less)
50× less diluted	$2^n = 50$, $n = 5.6$	(~6 cycles less)

Optimal cycle number of library amplification:
16 – 8 = 8

Library amplification

The protocol described here involves the amplification of libraries by PCR. This step is necessary to prevent the loss of molecules during sequencing and to 'immortalize' the content of a library so that it remains available for future use—for example, to target parts of the genome by hybridization capture[40–42]. When amplifying libraries for the first time, pairs of sample-specific indices are introduced so that libraries from multiple samples and controls can be sequenced simultaneously[43]. Owing to the low concentration of template DNA in these indexing PCR reactions, a hot-start enzyme, AccuPrime *Pfx* DNA polymerase, is used to suppress the formation of primer dimers. The proofreading activity of the enzyme does not interfere with the amplification of uracil-containing library molecules, as the template strands created during library preparation are copies of the original sample molecules. Subsequent re-amplification of libraries, which is necessary, for example, as part of the size-selection protocol or hybridization capture, is performed with Herculase II Fusion DNA polymerase, a less costly enzyme with proofreading but no hot-start activity.

Although the two DNA polymerases included in this protocol have been chosen to minimize length and base composition biases that are associated with PCR amplification of heterogeneous template molecules[44], it is advisable, if possible, to perform only as many PCR cycles as necessary to obtain large enough quantities of library for downstream experiments. For manual library preparation, we suggest to determine the optimal PCR cycle number separately for each sample based on the amplification plots obtained from the qPCR measurement (see Box 2 for instructions). For automated library preparation, where all libraries are amplified simultaneously in plate format, it is inevitable to set a fixed cycle number for all libraries. To be compatible with high-throughput screening, we suggest that this number is chosen so that all reactions reach PCR plateau (see Supplementary

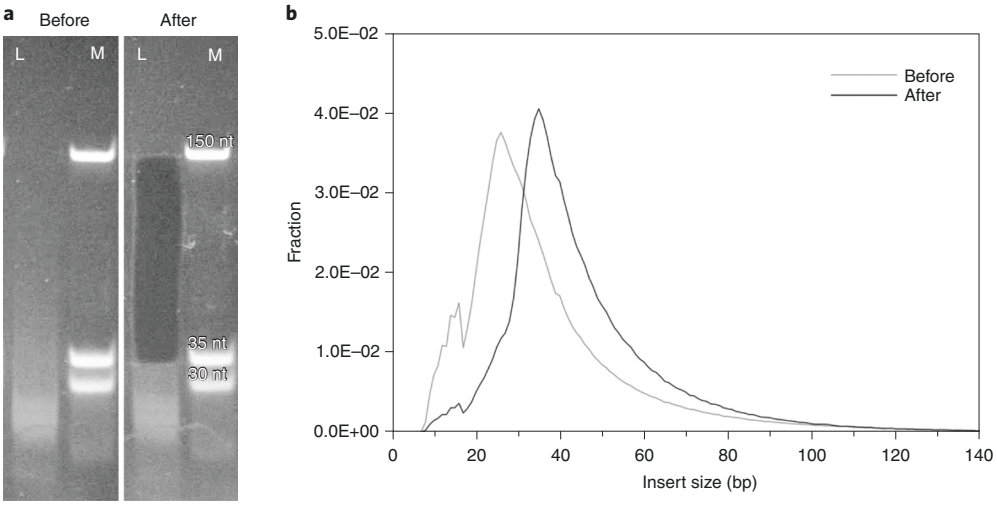


Fig. 2 | Size selection of amplified libraries. a, Image of a denaturing 5% polyacrylamide gel before and after the isolation of the desired library fraction. The size of the markers (M) corresponds to library molecules (loaded on lane L) carrying inserts of size 30, 35 and 150 bp, respectively. Gel excision was performed targeting an insert size range between 35 and 150 bp. **b**, Insert size distribution obtained from sequences generated from a library before and after gel excision.

Manual). This strategy accepts a small increase in PCR bias but enables pooling of all libraries in equal volumes before sequencing, as all libraries reach similar DNA concentrations in PCR plateau.

If libraries are amplified into PCR plateau, heteroduplexes are formed[45]—that is, molecules with single- and double-stranded regions that result from re-hybridization of the two library strands via their adapter sequences. These structures do not interfere with hybridization capture or sequencing, but they prevent the use of capillary gel electrophoresis for quantifying libraries before sequencing and might also distort concentration measurements based on fluorescent double-stranded DNA binding dyes or spectrophotometry. If necessary, heteroduplexes can be removed by subjecting 500 ng of a purified library, or a pool of purified libraries, to a PCR with primers IS5 and IS6 as described in Steps 66 and 68 but using only a single PCR cycle (see Extended Data Fig. 1 for an overview of library amplification schemes used in this protocol).

Size selection of molecules

One of the strengths of single-stranded library preparation lies in the efficient recovery of short DNA fragments. However, whereas DNA strands as short as 17 nucleotides are retained in library preparation[14], sequences shorter than 30−35 bp do not usually allow secure identification of molecules originating from the organism under study[15]. The extent to which library molecules with such short inserts are present depends on the extraction method used to prepare the sample DNA[14]. If short uninformative molecules make up a substantial proportion of the library, it might be preferable to remove them to reduce sequencing costs. Although size separation on agarose gels is, in principle, suitable for this purpose[1,2,23], we have found this strategy to be ineffective if the fraction of short molecules is greater than ~50%. The incomplete separation of short and long molecules on agarose gels is explained both by the limited resolution of agarose gels and by the formation of heteroduplexes that cannot be entirely avoided, even if library amplification is halted before PCR plateau. We have therefore developed a size-selection protocol for amplified libraries that relies on the gel excision of only one of the two library strands from denaturing polyacrylamide gels. To make this possible, amplified libraries are made single stranded by incorporating a biotin into one of the two library strands by PCR with one biotinylated primer (Fig. 1b). The library is then bound to streptavidin-coated magnetic beads, which allows isolation of the non-biotinylated strand for size separation. It should be noted that, whereas this procedure produces very precise size cutoffs (see Fig. 2 for an example), it cannot easily be scaled for many samples.

Automation

Despite improvements over the previous version of the protocol in terms of ease of use, single-stranded library preparation remains relatively time-consuming if performed manually. We have

therefore developed a version of the protocol that is compatible with liquid handling on the Bravo NGS Workstation (Agilent Technologies). The electronic protocol files and instructions on how to use them are provided in a Supplementary Manual and on the Zenodo website (https://zenodo.org/record/3631147). However, we caution that, apart from the investment required for acquisition of this liquid handling system, successful implementation and subsequent maintenance of the automated protocol requires substantial expertise in protocol development and troubleshooting on the system. The reason for this is that some parameters, such as pipetting distances and the force of the grippers, have to be adjusted slightly differently on each machine to ensure consistent and reliable performance (see the Supplementary Manual for advice). In addition, open liquid handling on 96-well plates greatly increases the risk of cross-contamination among samples. We strongly advise to test for this possibility after the implementation of the protocol and whenever further adjustments are made. This can be achieved, for example, by generating libraries from the positive control oligonucleotide CL304 as well as water controls that are alternatingly distributed across the entire plate in a chess-board-like setup. A possible carryover of the oligonucleotide to neighboring wells can then be detected using the qPCR assay B described in Steps 32–37.

Limitations of the protocol and expertise required

Although the implementation of automated single-stranded library preparation is challenging and should be attempted only as part of a sustained effort to develop a large-scale sample preparation pipeline, a laboratory technician with experience in performing qPCR and library preparation should not face problems implementing the manual version of single-stranded library preparation described here. One general constraint of the protocol is that the adapter sequences were specifically designed for Illumina sequencers and that libraries cannot easily be sequenced using other technologies. In addition, one of the adapter sequences differs from standard Illumina adapters by a 5-bp deletion, which requires the use of a custom sequencing primer for the first insert read (see Extended Data Fig. 1 for adapter sequences and binding sites of primers used for library amplification and sequencing). We have not encountered difficulties with replacing sequencing primers on our in-house HiSeq and MiSeq machines as well as when outsourcing projects to large sequencing centers. Nonetheless, it is advisable to clarify upfront with the sequencing unit whether the necessary modifications can be made.

Materials

Reagents

- Ancient or damaged DNA or oligonucleotide sample to be sequenced (see 'Reagent setup' section below)
- Oligonucleotide solutions for library preparation and amplification (see Table 1 and 'Reagent setup' section below)
- Homemade buffers (see 'Reagent setup' section below)
- Water, HPLC-grade (Merck, cat. no. 270733)
- 5 M NaCl solution (Merck, cat. no. S5150-1L)
- 1 M Tris-HCl solution, pH 8.0 (AppliChem, cat. no. A4577,0500)
- 0.5 M EDTA solution, pH 8.0 (AppliChem, cat. no. A4892,1000)
- Tween-20 (Merck, cat. no. P5927-100ML); also prepare a 2% (vol/vol) solution in water
- 20× SSC buffer (Thermo Fisher Scientific, cat. no. AM9763)
- 20% (wt/vol) SDS solution (Thermo Fisher Scientific, cat. no. AM9820) **! CAUTION** Can cause skin and eye irritation. Wear gloves and eye protection.
- 3 M Sodium acetate, pH 5.2 (Merck, cat. no. S7899-500ML) **! CAUTION** Can cause skin and eye irritation. Wear gloves and eye protection.
- *T4* RNA ligase reaction buffer (NEB, cat. no. B0216L), including 50% (wt/vol) PEG 8000
- ATP solution (100 mM; Thermo Fisher Scientific, cat. no. R0441)
- Klenow fragment, including 10× reaction buffer (Thermo Fisher Scientific, cat. no. EP0052)
- *T4* polynucleotide kinase 10 U μl^{-1} (Thermo Fisher Scientific, cat. no. EK0031)
- FastAP thermosensitive alkaline phosphatase (Thermo Fisher Scientific, cat. no. EF0651)
- *T4* DNA ligase, high concentrated 30 U μl^{-1} (Thermo Fisher Scientific, cat. no. EL0013)
- *T4* DNA ligase, 5 U μl^{-1} (Thermo Fisher Scientific, cat. no. EL0012), including 10× reaction buffer and 50% (wt/vol) PEG-4000
- Herculase II fusion DNA polymerase, including 5× Herculase II reaction buffer (Agilent, cat. no. 600679)

- Optional: USER enzyme (NEB, cat. no. M5505L)
- Dynabeads MyOne Streptavidin C1 (Life Technologies, cat. no. 65001)
- 25 mM each dNTP (Thermo Fisher Scientific, cat. no. R1121)
- Maxima Probe qPCR Master Mix (2×; Thermo Fisher Scientific, cat. no. K0261)
- AccuPrime *Pfx* DNA polymerase (Thermo Fisher Scientific, cat. no. 12344-024), including 10× AccuPrime reaction mix
- MinElute PCR Purification Kit (Qiagen, cat. no. 28006)
- QIAquick Nucleotide Removal Kit (Qiagen, cat. no. 28306)
- pUC 19 vector (NEB, cat. no. N3041S)
- SYBR Gold Nucleic Acid Gel Stain (Thermo Fisher Scientific, cat. no. S11494)
- 20/100 Ladder (IDT, cat. no. 51-05-15-02)
- 5× Novex TBE Running Buffer (Thermo Fisher Scientific, cat. no. LC6675)
- 2× TBE-Urea Sample Buffer (Bio-Rad, cat. no. 1610768)
- Agilent DNA 1000 Kit (Agilent, cat. no. 5067-1504)

Equipment

- Latex gloves (Roth, cat. no. L949.1)
- Nitrile gloves (Ansell Health Care, cat. no. 112-2688)
- 2-ml Safe-lock tubes (Eppendorf, cat. no. 0030120094)
- 1.5-ml Safe-lock LoBind tubes (Eppendorf, cat. no. 0030120086)
- 0.5-ml Safe-lock LoBind tubes (Eppendorf, cat. no. 0030121023)
- 0.2-ml PCR eight-tube strips (Eppendorf, cat. no. 0030124359)
- 96-well semi-skirted PCR plate (Thermo Fisher Scientific, cat. no. AB2400)
- MicroAmp optical eight-cap strips (Thermo Fisher Scientific, cat. no. 4323032)
- 50-ml tubes (Merck, cat. no. CLS430921)
- Scalpel (Braun, cat. no. 10567364)
- Sterican needle, 0.9 × 40 mm (Braun, cat. no. 2050798)
- Racks for 0.2-ml, 0.5-ml and 1.5-ml tubes (many suppliers)
- Magnetic-ring stand (96-well; Thermo Fisher Scientific, cat. no. AM10050)
- Magnetic rack for 1.5-ml tubes (MagJET Separation Rack; Thermo Fisher Scientific, cat. no. MR02)
- Rotator for 1.5-ml tubes (VWR, cat. no. 13916-822)
- 5% Criterion TBE-urea polyacrylamide gel (Bio-Rad, cat. no. 3450086)
- Criterion cell and PowerPac basic power supply (Bio-Rad, cat. no. 165-6019)
- Rocking shaker (Grant-Bio PMR-30, Kisker, cat. no. 144201)
- Dark reader transilluminator (Clare Chemical Research, cat. no DR196)
- Microcentrifuge with adapters for 0.2-ml, 0.5-ml, 1.5-ml and 2-ml tubes (NeoLab, cat. no. D-8550)
- Vortex mixer (Scientific Industries, cat. no. SI-0256) and, optionally, Centrifuge/Vortex Multispin MSC-6000 (BioSan, cat. no. BS-010211-AAL)
- Benchtop centrifuge for 1.5-ml and 2-ml tubes (Eppendorf, cat. no. 5420000318)
- Plate centrifuge for 96-well PCR plates (Eppendorf, cat. no. 5948000913)
- Thermal cyclers with lid heating—one for 0.2-ml PCR tubes and one for 0.5-ml tubes (DNA Engine Thermal CyclerPTC-200 Thermo Cycler, MJ Research, cat. no. 8252-30-0001)
- Benchtop thermo mixer (ThermoMixer C; Eppendorf, cat. no. 5382000015)
- 96-well qPCR system (CFX96 Touch Real-Time PCR Detection System; Bio-Rad, cat. no. 1855195)
- 2100 Bioanalyzer (Agilent, cat. no. G2939BA)
- Spectrophotosmeter (NanoDrop 2000/2000c, cat. no. ND-2000)
- UV cross-linker (Bio-Link BLX 254, Vilber, cat. no. 611110811)
- Illumina sequencing instrument (MiSeq, HiSeq, NovaSeq platforms) and related sequencing chemistry

Reagent setup

▲ **CRITICAL** All buffers and reagents below will suffice for at least 48 reactions. ▲ **CRITICAL** Steps 1 and 2 of the protocol are sensitive to contamination with exogenous DNA. We recommend UV-decontamination of the TET buffer and water used in these steps as well as when preparing the control oligonucleotide dilutions. For this purpose, prepare at least three 1-ml aliquots of TET buffer and one of water in 1.5-ml tubes and UV-C (254-nm wavelength) irradiate them with a dose of 7 J/cm^2

in a cross-linker. Note that decontamination in 50-ml Falcon tubes with closed caps is ineffective owing to the thickness of the plastic walls.

Bind and wash (B&W) buffer I (50 ml): Combine 47.125 ml of water, 1 ml of 5 M NaCl, 500 µl of 1 M Tris-HCl (pH 8.0), 100 µl of 0.5 M EDTA (pH 8.0), 25 µl of Tween 20 and 1.25 ml of 20% (wt/vol) SDS. Store at room temperature (20−25 °C). Shelf life is 2 months.

Wash buffer II (50 ml): Combine 48.375 ml of water, 1 ml of 5 M NaCl, 500 µl of 1 M Tris-HCl (pH 8.0), 100 µl of 0.5 M EDTA (pH 8.0) and 25 µl of Tween 20. Store at room temperature for up to 1 year.

Stringency wash buffer (50 ml): Combine 49.5 ml of water, 250 µl of 20% (wt/vol) SDS and 250 µl of 20× SSC buffer. Store at room temperature. Shelf life is 2 months.

TE buffer (50 ml): Combine 49.4 ml of water, 500 µl of 1 M Tris-HCl (pH 8.0) and 100 µl of 0.5 M EDTA (pH 8.0). Store at room temperature for up to 1 year.

TET buffer (50 ml): Combine 49.375 ml of water, 500 µl of 1 M Tris-HCl (pH 8.0), 100 µl of 0.5 M EDTA (pH 8.0) and 25 µl of Tween 20. Store at room temperature for up to 1 year.

Gel elution buffer (50 ml): Combine 44.375 ml of water, 5 ml 5 M NaCl, 500 µl of 1 M Tris-HCl (pH 8.0), 100 µl of 0.5 M EDTA (pH 8.0) and 25 µl of Tween 20. Store at room temperature for up to 1 year.

Melt buffer (1 ml): Combine 850 µl water, 125 µl 1M NaOH and 25 µl 2% (vol/vol) Tween 20. Always prepare freshly.

Decontamination and hybridization of adapter/splinter mix for single-stranded ligation (60 µl)

In the first tube of a 0.2-ml PCR eight-tube strip, combine 18 µl of water, 3 µl of 10× *T4* RNA ligase reaction buffer, 6 µl of 100 µM adapter oligonucleotide TL181, 1.5 µl of 10 U μl^{-1} Klenow fragment and 1.5 µl of 10 U μl^{-1} *T4* polynucleotide kinase. In another tube of the strip, combine 12 µl of water, 3 µl of 10× *T4* RNA ligase reaction buffer, 12 µl of 100 µM splinter oligonucleotide TL159, 1.5 µl of 10 U μl^{-1} Klenow fragment and 1.5 µl of 10 U μl^{-1} *T4* polynucleotide kinase. Vortex the tubes and briefly spin down in a microcentrifuge. Incubate the tube for 20 min at 37 °C in a thermal cycler, followed by 1 min at 95 °C to inactivate the enzyme. The oligonucleotides are now purified from undesired synthesis artifacts and contaminant DNA. Combine the contents of both tubes in a tube of a fresh eight-tube strip and incubate the mix at 95 °C for 10 s, followed by a ramp to 10 °C at 0.5 °C s^{-1} in a thermal cycler. Transfer 60 µl of adapter/splinter mix with a final concentration of 10/20 µM into a 1.5-ml tube. Store at −20 °C and thaw at room temperature before use.

Hybridization of double-stranded adapter for second ligation (100 µl)

In one tube of an 0.2-ml PCR eight-tube strip, combine 9.5 µl of TE buffer, 0.5 µl of 5 M NaCl, 20 µl of 500 µM oligonucleotide CL53 and 20 µl of 500 µM oligonucleotide TL178. Vortex the tubes and briefly spin down in a microcentrifuge. Hybridize the adapter oligonucleotides by incubating the reaction mix in a thermal cycler at 95 °C for 10 s, followed by a ramp to 14 °C at 0.1 °C s^{-1}. Add 50 µl of TE for a final double-stranded adapter concentration of 100 µM. Store at −20 °C for up to 1 year and thaw at room temperature when used.

Preparation of positive control oligonucleotide aliquots

Dilute 100 µM stock solution of oligonucleotide CL304 to 0.1 µM by combining 999 µl of TET (UV-decontaminated) and 1 µl of stock solution. Mix well and prepare 10-µl aliquots in eight-tube strips. Store at −20 °C for up to 1 year. Before starting a library preparation experiment, thaw one tube at room temperature, vortex the tube, spin the liquid down and prepare a 0.5 nM dilution by combining 5 µl of the diluted oligonucleotide with 995 µl of TET. Vortex the tube, spin the liquid down and prepare a final 10 pM oligonucleotide dilution by combining 2 µl of the previous dilution with 98 µl of TET. Vortex the tube and spin the liquid down.

Preparation of qPCR standard

Perform a 50-µl PCR assay with AccuPrime *Pfx* DNA polymerase using 10 ng of pUC19 plasmid DNA as the template and primers CL105 and CL106 (Table 1). Follow the manufacturer's instructions and choose an annealing temperature of 60 °C for 30 cycles. Purify the PCR product using the MinElute PCR Purification Kit. Elute the DNA in 20 µl of TET and determine its concentration using a spectrophotometer. The expected concentration range is 50–100 ng μl^{-1}. Use 10 ng of the purified PCR product as template for a second amplification reaction with an arbitrary combination of P5 and P7 indexing primers. Purify the PCR product using the MinElute PCR Purification Kit and measure the DNA concentration on a DNA 1000 chip using Bioanalyzer 2100. Prepare a ten-fold dilution series in TET ranging from 10^9 to 10^2 copies per µl. Store at −20 °C for up to 1 year.

Preparation of single-stranded markers for gel excision

For each desired size marker, perform a 100-µl PCR assay with AccuPrime *Pfx* DNA polymerase using the respective primer pair provided in Table 1 (one biotinylated and one non-biotinylated primer) and 10 ng of pUC19 plasmid as the template. Follow the manufacturer's instructions and choose an annealing temperature of 60 °C for 30 cycles. Purify the PCR product with the MinElute PCR Purification Kit. Elute the DNA in 20 µl of TET and determine its concentration using a spectrophotometer. The expected concentration range is 100−200 ng μl^{-1}. Transfer 2 µg of DNA to a fresh tube and fill up to 20 µl with water. Resuspend the stock solution of Dynabeads MyOne Streptavidin C1 by vortexing. Transfer 100 µl of bead suspension into a 1.5-ml tube, pellet the beads using a magnetic rack and discard the supernatant. Perform two bead washes by resuspending the beads in 500 µl of B&W buffer I, pelleting the beads in a magnetic rack and discarding the supernatant. Resuspend the beads in 250 µl of B&W buffer I. Add the prepared PCR product (20 µl) to the bead suspension and repeatedly invert the tubes upside down on a rotator for 15 min at room temperature. Pellet the beads in a magnetic rack, discard the supernatant and wash the beads with 500 µl of wash buffer II as described above. Pellet the beads again, discard the supernatant and resuspend the beads in 50 µl of melt buffer. Let the tubes stand at room temperature for 5 min, pellet the beads and collect the supernatants. Add 10 µl of sodium acetate (3 M, pH 5.2) and purify the DNA using the Nucleotide Removal Kit following the manufacturer's instructions but using MinElute columns (from the MinElute PCR Purification Kit) and eluting in 20 µl of TET. Measure the DNA concentration using a spectrophotometer. The expected concentration range is 20–40 ng μl^{-1}. Dilute the marker with TET to 20 ng μl^{-1}. Store at −20 °C for up to 1 year.

Procedure

Single-stranded DNA library preparation

Uracil removal, dephosphorylation and heat denaturation ● Timing 30 min

1 Prepare the DNA samples in 0.5-ml tubes by pipetting the desired volume of DNA extract and filling with TET to 30 µl. We recommend preparation of no more than 16 libraries in parallel, including at least one positive control (CL304) and one library preparation negative control with water instead of sample DNA.

Optional: If removal of internal uracils is desired, fill tubes to only 29 µl and add 1 µl of USER enzyme mix (1 U μl^{-1}). Mix gently by flicking the tubes with a finger, spin tubes briefly in a microcentrifuge and incubate them in a thermal cycler for 1 h at 37 °C. Then continue with the next step.

2 In a 1.5-ml tube, prepare a master mix containing the reagents below. Multiply the volumes by the number of reactions that are to be performed plus one in excess to ensure that the master mix suffices for all samples. Mix the reagents by vortexing before adding the enzyme and mix gently by flicking the tube with a finger thereafter.

▲ **CRITICAL STEP** Always prepare a fresh 10 pM dilution of CL304 (see 'Reagent setup').

Reagent	Volume (µl) per reaction	Final concentration in reaction after Step 3/Step7
Water	3.6	—
T4 RNA ligation buffer (10×)	8	1.75×/1×
Tween 20 (2% (vol/vol))	2	0.09%/0.05%
Spike-in positive control CL304 (10 pM)	1	0.22 pM/0.125 pM
Fast AP (1 U μl^{-1})	1	0.02 U μl^{-1}/inactive

3 Add 15.6 µl of the reaction master mix to the prepared DNA to obtain a total volume of 45.6 µl.

4 Mix the tube contents by firmly flicking the side of each tube. Spin the tubes briefly in a microcentrifuge and transfer to a PCR cycler.

5 Incubate reactions at 37 °C for 10 min and at 95 °C for 2 min. Hold reactions at 4 °C afterwards. Continue with the next step during the incubation phase.

Splinted ligation of first adapter ● Timing 1.5 h

6 In a 2-ml tube, prepare a master mix containing the reagents below. Multiply the volumes by the number of reactions that are to be performed plus one in excess. Pre-mix the reagents for 5 min by

repeatedly inverting the tube upside down on a rotator before adding the enzyme. Repeatedly invert the tubes upside down on a rotator at least 10 min thereafter to ensure that the reagents are properly mixed.

▲ **CRITICAL STEP** PEG-8000 is highly viscous; pipette slowly. Vortexing instead of rotation is not recommended, as it does not achieve homogenous mixing of the reagents.

Reagent	Volume (µl) per reaction	Final concentration in reaction
PEG-8000 (50% (wt/vol))	32	20%
ATP (100 mM)	0.4	0.5 mM
TL181/TL159 (10/20 µM)	1	0.125/0.25 µM
T4 DNA ligase high conc. (30 U $µl^{-1}$)	1	0.375 U $µl^{-1}$
Total	34.4	—

7 Add 34.4 µl of the reaction master mix to each sample from Step 5 to obtain a final reaction volume of 80 µl.

8 Mix the tube contents repeatedly by firmly flicking the side of each tube. Spin tubes briefly in a microcentrifuge and transfer to a PCR cycler.

▲ **CRITICAL STEP** Proper mixing is critical. Visually inspect each tube and ensure that there are no streaks before moving on to the next step.

9 Incubate at 37 °C for 1 h and at 95 °C for 2 min. Hold at 10 °C afterwards.

■ **PAUSE POINT** Samples can be stored at −20 °C for several days.

Immobilization of ligation products on beads ● Timing 45 min

10 Resuspend the stock solution of Dynabeads MyOne Streptavidin C1 by vortexing. For each reaction, transfer 20 µl of bead suspension into a 2-ml tube (e.g. 320 µl for 16 reactions). Pellet the beads using a magnetic rack and discard the supernatant.

11 Add 500 µl of B&W buffer I and resuspend the beads by vortexing. Spin the tubes briefly in a microcentrifuge, place on a magnetic rack and discard the supernatant. Repeat this step for a total of two washes.

12 Resuspend the beads in a volume of B&W buffer I corresponding to 100 µl per reaction plus 100 µl in excess (e.g. 1.7 ml for 16 reactions). Split the suspension into 100-µl aliquots in 1.50-ml tubes

13 Add 100 µl of B&W buffer I to the ligation reactions from Step 9, mix by pipetting up and down and transfer the diluted ligation reaction to the bead suspension prepared in Step 12. The final volume of the bead suspension is 280 µl.

14 Repeatedly invert the tubes upside down on a rotator for 20 min at room temperature.

First bead wash ● Timing 20 min

15 Spin the beads briefly in a microcentrifuge. Pellet the beads in a magnetic rack and discard the supernatant.

16 Add 200 µl of B&W buffer I and resuspend the beads by vortexing. Briefly spin the tubes in a microcentrifuge, pellet the beads in a magnetic rack and discard the supernatant.

17 Add 100 µl of Stringency wash buffer and resuspend the beads by vortexing. Incubate the tubes for 3 min at 45 °C in a thermo mixer without shaking. Pellet the beads in a magnetic rack and discard the supernatant.

▲ **CRITICAL STEP** In this step, vortex slowly to avoid spilling liquid on the tube walls and cap. If spilling occurs, spin the tubes in a microcentrifuge and resuspend the beads more gently. We recommend using a spin-mix-spin device (see Equipment) for spinning and vortexing to perform both actions in the most time-effective manner.

18 Add 200 µl of wash buffer II and resuspend the beads by vortexing. Briefly spin the tubes in a microcentrifuge and leave them in a rack on the bench.

Second-strand synthesis ● Timing 45 min

19 In a 2-ml tube, prepare a master mix containing the reagents below. Multiply the volumes by the number of reactions that are to be performed plus one in excess. Mix the reagents by vortexing before adding the enzyme and mix gently by flicking the tube with a finger thereafter.

Reagent	Volume (µl) per reaction	Final concentration in reaction
Water	40.35	—
Klenow reaction buffer (10×)	5	1×
dNTP (25 mM each)	0.4	200 µM each
Tween 20 (2% (vol/vol))	1.25	0.05%
CL128 (100 µM)	1	2 µM
Klenow fragment (10 U μl^{-1})	2	0.4 U μl^{-1}
Total	50	—

20 Place the tubes containing the beads in a magnetic rack and discard the wash buffer.
? TROUBLESHOOTING

21 Add 50 µl of the reaction master mix from Step 19 to the beads and resuspend the beads by pipetting up and down.
? TROUBLESHOOTING

22 Incubate the tubes at 35 °C for 20 min in a thermo mixer with shaking at 800 r.p.m.

Second bead wash ● Timing 20 min

23 Perform bead washes with B&W buffer I, Stringency wash buffer and wash buffer II exactly as described in Steps 15–18.

Ligation of second adapter ● Timing 1.5 h

24 In a 2-ml tube, prepare a master mix containing the reagents below. Multiply the volumes by the number of reactions that are to be performed plus one in excess. Mix the reagents by vortexing before adding the enzyme and mix firmly by flicking the tube with a finger thereafter.

Reagent	Volume (µl) per reaction	Final concentration in reaction
Water	73.5	—
T4 DNA ligase buffer (10×)	10	1×
PEG-4000 (50% (wt/vol))	10	5%
Tween-20 (2% (vol/vol))	2.5	0.05%
CL53/TL178 (100 µM each)	2	2 µM each
T4 DNA ligase (5 U μl^{-1})	2	0.1 U μl^{-1}
Total	100	—

25 Place the tubes containing the beads in a magnetic rack and discard the wash buffer. Add 100 µl of the reaction master mix to the beads and resuspend by pipetting up and down.
? TROUBLESHOOTING

26 Incubate tubes at room temperature for 1 h in a thermo mixer at 22 °C with shaking at 800 r.p.m.

Third bead wash ● Timing 20 min

27 Perform bead washes with B&W buffer I, Stringency wash buffer and wash buffer II exactly as described in Steps 15–18.

Elution ● Timing 15 min

28 Place the tubes in a magnetic rack and discard the wash buffer.
? TROUBLESHOOTING

29 Resuspend the beads in 50 µl of TET buffer by vortexing and transfer bead suspensions to the tubes of 0.2-ml PCR eight-tube strip. Spin the strip briefly in a microcentrifuge.
? TROUBLESHOOTING

30 Incubate the bead suspensions for 1 min at 95 °C in a thermal cycler with heated lid, followed by cooling to 25 °C.

31 Place the strip in a ring magnet plate and transfer the supernatant, which is the final library, to fresh 1.5-ml tubes.
■ PAUSE POINT Samples can be stored at −20 °C for at least half a year.

Quantification of libraries and controls ● Timing 2.5 h

32 Prepare a 50-fold dilution of each library in a tube of a 0.2-ml PCR eight-tube strip by combining 1 µl of library with 49 µl of TET buffer. Mix by vortexing and spin the tubes in a microcentrifuge.

33 Prepare two master mixes, each in a 2-ml tube, containing the reagents for two qPCR assays: the first (assay A) measures the total yield of library molecules; the second (assay B) measures the number of library molecules obtained from the spike-in control oligonucleotide. Multiply the reagent volumes below by the number of measurements that are to be performed and include an excess of 10%. Mix the reagents by vortexing. When determining the number of reactions required, note that seven dilutions of the qPCR standard and at least one qPCR negative control have to be included, and that all measurements should be performed in replicates (e.g. 48 reactions are needed to measure 16 sample libraries together with seven standard dilutions and one qPCR negative control in duplicate).

	Assay A (all molecules)		Assay B (control molecules)	
Reagent	**Volume (µl) per reaction**	**Final conc.**	**Volume (µl) per reaction**	**Final conc.**
Water	10	—	10	—
Maxima probe qPCR master mix (2×)	12.5	1×	12.5	1×
IS7 (10 µM)	0.5	200 nM	0.5	200 nM
IS8 (10 µM)	0.5	200 nM	—	—
IS10 (10 µM)	0.5	200 nM	—	—
CL107 (10 µM)	—	—	0.5	200 nM
CL118 (10 µM)	—	—	0.5	200 nM
Total	24		24	—

34 For each measurement, dispense 24 µl of the master mix to a well of an optical 96-well PCR plate. Both assays can be performed on the same plate if the number of measurements does not exceed 96 or successively on separate plates.

35 Add 1 µl of the diluted library, qPCR standard or water to each well. Seal the plate with optical caps or adhesive foil and mix by vortexing and briefly spin the plate in a plate centrifuge.
▲ **CRITICAL STEP** Do not discard the library dilutions. Store them at −20 °C to allow additional measurements in the future if needed.

36 Place the plate in a qPCR machine and perform a cycling protocol consisting of an initial activation step at 95 °C for 10 min, followed by 45 cycles at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s. Fluorescence is measured at the end of the annealing step.

37 Determine the number of library and control molecules in each sample using the software provided with the qPCR system.
? TROUBLESHOOTING

Indexing and amplification ● Timing 3.5 h

38 In a 2-ml tube, prepare a master mix containing the reagents below. Multiply the volumes by the number of reactions that are to be performed and include an excess of 10%. Mix the reagents by vortexing before adding the enzyme and mix firmly by flicking the tube with a finger thereafter.

Reagent	Volume (µl) per reaction	Final conc.
Water	20	—
AccuPrime *Pfx* buffer (10×)	10	1×
AccuPrime *Pfx* DNA polymerase (2.5 U μl^{-1})	1	0.025 U μl^{-1}
Total	31	—

39 For each reaction, dispense 31 µl of the master mix to a tube of a 0.2-ml PCR eight-tube strip.

40 Add 10 µl of P7 indexing primer (10 µM) and 10 µl of P5 indexing primer (10 µM) to each well so that each sample receives a unique pair of indices. Add 49 µl of the library from Step 31, mix by vortexing and spin the tubes briefly in a microcentrifuge.

41 Incubate the reactions in a thermal cycler at 95 °C for 2 min, followed by a selected number of cycles at 95 °C for 20 s, 60 °C for 30 s and 68 °C for 1 min, followed by an additional 5 min at 68 °C. ▲ **CRITICAL STEP** Determine the optimal cycle number for each library by following the instructions provided in Box 2. If only a single thermal cycler is available, separate the tubes of the strip before cycling and remove them from the cycler a few seconds before the end of the 68-°C elongation step when the respective cycle number has been reached. Alternatively, all reactions can be cycled into PCR plateau using a fixed cycle number—for example, if large quantities of amplified libraries are needed for hybridization capture. Note that heteroduplexes are formed in PCR plateau that interfere with library quantification using electrophoresis-based systems.

42 Purify the amplified libraries using the MinElute PCR Purification Kit. The pH of the binding buffer has to be adjusted by adding 2.4 µl of sodium acetate (3 M, pH 5.2) to each reaction, as indicated by the color change of the indicator included in the PB binding buffer. Elute in 20 µl of TET.
■ **PAUSE POINT** Samples can be stored at −20 °C for at least 1 year.

Size selection on denaturing polyacrylamide gels (optional)

Biotinylation of indexed libraries ● Timing 1 h

43 Determine the concentration of the library that requires gel excision using a spectrophotometer.

44 Prepare a PCR mix according to the scheme below.

Reagent	Volume (µl) per reaction	Final conc.
Water	x	—
Herculase II reaction buffer (5×)	20	1×
dNTP mix (25 mM each)	1	250 µM each
IS5 biotinylated (10 µM)	10	1 µM
IS6 (10 µM)	10	1 µM
100 ng indexed library	x	1 ng μl^{-1}
Herculase II fusion DNA polymerase (1 U μl^{-1})	2	0.01 U μl^{-1}
Total	100	—

45 Incubate the reactions in thermal cycler at 95 °C for 2 min, followed by 4 cycles at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s.

46 Purify the amplified libraries using the MinElute PCR Purification Kit. The pH of the binding buffer has to be adjusted by adding 2.4 µl of 3 M sodium acetate (pH 5.2) to each reaction. Elute in 20 µl of TET.

Denaturation of non-biotinylated library strand ● Timing 1 h

47 Resuspend the stock solution of Dynabeads MyOne Streptavidin C1 by vortexing. Transfer 100 µl bead suspension into a 1.5-ml tube.

48 Add 500 µl of B&W buffer I and resuspend the beads by vortexing. Spin the tube briefly in a microcentrifuge, place on a magnetic rack and discard the supernatant. Repeat this step for a total of two washes

49 Add 20 µl of the biotinylated library from Step 46 and repeatedly invert the tubes upside down on a rotator at room temperature for 15 min.

50 Spin the tube briefly using a microcentrifuge. Pellet the beads on a magnetic rack and discard the supernatant. Add 200 µl of wash buffer II and resuspend the beads by vortexing. Briefly spin the tube in a microcentrifuge, place on a magnetic rack and discard the supernatant.

51 Resuspend the beads in 50 µl of melt buffer by pipetting up and down. Incubate at room temperature for 5 min. Briefly spin the tube in a microcentrifuge, pellet the beads on a magnetic rack and transfer the supernatant containing the single-stranded amplified library to a fresh 1.5-ml tube.

52 Add 10 µl of sodium acetate (3 M, pH 5.2) to each melted product for neutralization. Purify the DNA with the Nucleotide Removal Kit. Follow the manufacturer's instructions, but replace the QiaQuick spin columns supplied in the kit with MinElute spin columns (from the MinElute PCR Purification Kit). Elute in 20 µl of TE.
■ **PAUSE POINT** Samples can be stored at −20 °C for at least half a year.

Size selection on a denaturing polyacrylamide gel ● Timing 2 d

53 Take a precast 5% polyacrylamide urea gel from the fridge and let it equilibrate to room temperature.

54 Prepare the sample for loading on a denaturing gel by combining the reagents below in a tube of a 0.2-ml PCR eight-tube strip.

Reagent	Volume (µl) per reaction	Final conc.
Water	7	—
Single-stranded amplified library	3	—
TBE-Urea sample buffer (2×)	10	1×
Total	20	—

55 Prepare a marker DNA mix by combining the reagents below in another tube of the eight-tube strip.

Reagent	Volume (µl) per reaction	Final conc.
Water (to 40 µl)	18.8	—
30-nucleotide marker (20 ng μl^{-1})	0.4	0.4 ng μl^{-1}
35-nucleotide marker (20 ng μl^{-1})	0.4	0.4 ng μl^{-1}
150-nucleotide marker (20 ng μl^{-1})	0.4	0.4 ng μl^{-1}
TBE-Urea sample buffer (2×)	20	1×
Total	40	—

▲ **CRITICAL STEP** The markers suggested here are optimal for isolating library molecules with insert sizes between 35 and 150 bp. The 30-bp marker is included to allow a visual control of size separation. Adjust the composition of markers as needed if other insert sizes should be selected.

56 Incubate the DNA and marker mixes at 95 °C for 2 min in a thermal cycler.

57 Mount the gel in the electrophoresis chamber and add TBE buffer to the top and bottom reservoirs. Keep the packaging of the gel for later use. Remove the plastic comb and flush the gel pockets with TBE buffer by pipetting up and down. Load 20 µl of sample DNA into one of the pockets and 20 µl of marker DNA into each of the neighboring pockets (see Fig. 2a).

▲ **CRITICAL STEP** Load the sample and markers to the center of the gel, as the DNA does not always run straight at the sides.

58 Apply a current of 200 V for 55 min (~12 V cm^{-1}).

59 In the meantime, heat up a needle or syringe with a diameter of approximately 0.9 mm using a Bunsen burner or lighter and pierce a hole into the bottom of a 0.5-ml tube (see Fig. 1b, Tube preparation, Step 59)

60 Use the packaging of the gel as reservoir for staining by adding 100 ml of TBE buffer and 10 µl of SYBR Gold dye (10,000×). Gently shake the reservoir to mix dye and buffer on a rocking shaker. Break the gel cartridge open and remove the front cover containing the top reservoir. Submerge the gel in the reservoir while still attached to the plastic back of the cartridge. Stain for 5 min by gently shaking the reservoir on a rocking shaker.

61 Place the gel on the Dark Reader Transilluminator and make two horizontal cuts through the sample and marker lanes, cutting along the marker bands. Then make two vertical cuts flanking the sample lane. Halve the gel slice lengthways and transfer both pieces into the 0.5-ml tube prepared in Step 59. Possible carryover of marker DNA is not critical, as the marker cannot be amplified and sequenced.

? TROUBLESHOOTING

62 Place the pierced 0.5-ml tube containing the gel slice inside a 2-ml tube (see Fig. 1b for an illustration) and spin for 1 min at 10,000*g* in a benchtop centrifuge to break the gel into small pieces and collect them in the 2-ml tube.

63 Discard the empty 0.5-ml tube. Add 300 µl of gel elution buffer to the mashed gel and incubate overnight at room temperature in a thermo mixer with constant shaking at 800 r.p.m.

■ **PAUSE POINT** Overnight incubation

64 Spin the tube for 1 min at 10,000*g* in a benchtop centrifuge to pellet the gel. Transfer the supernatant to a 15-ml Falcon tube and add 2.4 ml of PN buffer included with the Nucleotide Removal Kit. Add 10 µl of 3 M sodium acetate (pH 5.2). Vortex the mixture and centrifuge the tube for 1 min at 1,500*g* at room temperature. Residual gel will pellet at the bottom of the tube. Carefully transfer 700 µl of supernatant to a MinElute column and centrifuge for 30 s at 10,000*g*. Discard the flow-through and

repeat the transfer of supernatant to the column and centrifugation three times. Continue DNA purification from the PE buffer wash steps as detailed in the manual of the Nucleotide Removal Kit. Elute in 20 µl of TE buffer.

▲ **CRITICAL STEP** Avoid carryover of gel to the silica spin column. Leave some buffer in the Falcon tube rather than disturbing the pellet.

65 Determine the concentration of library molecules by qPCR as described in Steps 32–37.

66 To amplify the gel-excised library, prepare a PCR mix in one tube of a 0.2-ml PCR eight-tube strip by combining the reagents below.

Reagent	Volume (µl) per reaction	Final conc.
Water	59	—
Herculase II reaction buffer (5×)	20	1×
dNTP (25 mM each)	1	250 µM
IS5 (10 µM)	4	0.4 µM
IS6 (10 µM)	4	0.4 µM
Gel-excised library	10	—
Herculase II fusion DNA Polymerase (1 U μl^{-1})	2	0.01 U μl^{-1}
Total	100	—

67 Incubate the reactions in a thermal cycler at 95 °C for 2 min, followed by a selected number of cycles at 95 °C for 10 s, 60 °C for 30 s and 72 °C for 30 s.

▲ **CRITICAL STEP** Determine the optimal cycle number that avoids PCR plateau by following the instructions provided in Box 2.

68 Purify the amplified libraries using the MinElute PCR Purification Kit. The pH of the binding buffer has to be adjusted by adding 2.4 µl of 3 M sodium acetate (pH 5.2) to each reaction. Elute in 20 µl of TET.

Library quantification and sequencing ● Timing 1–3 d

69 Determine the concentration of each library using a DNA 1000 chip on the Bioanalyzer 2100 or similar capillary electrophoresis systems. Alternatively, pool the libraries if multiplex sequencing is desired and determine the concentration of the pool.

▲ **CRITICAL STEP** If indexing PCR has reached plateau, it is necessary to remove heteroduplexes before performing capillary gel electrophoresis (see 'Experimental design' section in the Introduction for details).

? TROUBLESHOOTING

70 For sequencing, follow the instructions provided by Illumina for multiplexed sequencing, but replace the first read sequencing primer by CL72. A fresh dilution of this primer should be prepared before each sequencing run by combining 10 µl of 100 µM CL72 with 1.99 ml of hybridization buffer provided in the kit containing the sequencing reagents. We recommend paired-end sequencing (e.g. 2 × 75 bp) for ancient DNA. Approximately 1 million sequence reads per sample library are usually sufficient to determine the proportion of mapped sequences, their length distribution and DNA damage profiles, and to estimate the content of informative molecules in the library (see 'Anticipated results' section). Fewer sequences (~200,000) are needed from the negative controls that were included during DNA extraction and/or library preparation.

Troubleshooting

Troubleshooting advice can be found in Table 2.

Table 2 | Troubleshooting table

Step	Problem	Possible reason	Solution
21, 25, 29	Beads appear clumpy	Beads dried out	Usually not critical. Continue library preparation
20, 25, 28	Foam remains on the beads after removal of the wash buffer and before adding reaction mix or elution buffer	Detergent in the wash buffer causes foaming if beads are resuspended too harshly	Does not affect library yield; try to resuspend beads more gently (centrifuge/vortex multispin device recommended)

Table continued

Table 2 (continued)

Step	Problem	Possible reason	Solution
37	Control oligonucleotide yield in sample libraries much lower than in library negative controls	Impurities in sample DNA extracts or saturation of library preparation (indicated by yields greater than ~2×10^{11} molecules)	Repeat library preparation with smaller volume of DNA extract
	Yield of molecules in the sample libraries does not exceed library negative controls	Too little DNA in sample; library is dominated by artifacts	Repeat library preparation with more DNA extract
	Low yield of library molecules, including in the library positive controls	No proper mixing of reagents in the reaction steps	Visually inspect the success of mixing for each tube, especially in the ligation reactions, to ensure proper mixing
61	Poor separation of markers	Gel not equilibrated to room temperature or run time too short	Remove artifacts by gel excision before sequencing or perform target enrichment via hybridization capture instead of sequencing the library directly
	Poor separation of markers	Gel not equilibrated to room temperature or run time too short	Allow gel to come room temperature before starting electrophoresis; increase run time by up to 70 min
	Sample library not visible on gel	Library concentration is low or the room not dark enough	Not critical; continue with gel excision guided by the markers
	Fragment size distribution does not show desired distribution after gel excision	Cutting was not precise enough	Perform several cuts with the same library on one gel and then choose the best library for sequencing
69	Fragment size distribution shows a secondary peak of unexpectedly long library molecules	PCR plateau was reached and heteroduplexes have formed	Perform a single cycle of PCR with primers IS5 and IS6 to remove heteroduplexes

Timing

Steps 1–5, uracil removal, dephosphorylation and heat denaturation: 30 min
Steps 6–9, splinted ligation of first adapter: 1.5 h
Steps 10–14, immobilization of ligation products on beads: 45 min
Steps 15–18, first bead wash: 20 min
Steps 19–22, second strand synthesis: 45 min
Step 23, second bead wash: 20 min
Steps 24–26, ligation of second adapter: 1.5 h
Step 27, third bead wash: 20 min
Steps 28–31, elution: 15 min
Steps 32–37, quantification of libraries and controls: 2.5 h
Steps 38–42, indexing and amplification: 3.5 h
Steps 43–46, biotinylation of indexed libraries: 1 h
Steps 47–52, denaturation of non-biotinylated library strand: 1 h
Steps 53–68, size selection on a denaturing polyacrylamide gel: 2 d
Steps 69 and 70, library quantification and sequencing: 1–3 d

Anticipated results

An example for how the protocol provided here can be used to determine DNA preservation in ancient biological material is provided in Supplementary Table 1. In this experiment, DNA was extracted from ancient hominin bones and sediments using an automated version of a silica-based protocol described earlier[37]. Single-stranded libraries were prepared in 96-well format as detailed in the Supplementary Manual and pooled and sequenced on one lane of a HiSeq 2500 (Illumina) in paired-end mode. Overlapping forward and reverse reads were merged into full-length molecule sequences using leeHom[46] and mapped against the human reference genome using BWA[47].

According to the qPCR measurements that were performed on the libraries, 1.5×10^8 library molecules were obtained on average in the library negative controls, and 1.1×10^{10} library molecules were obtained on average in the library positive controls (Fig. 3). Both values are within the range of numbers expected for successful library preparation (see Introduction, 'Controls' section). Likewise,

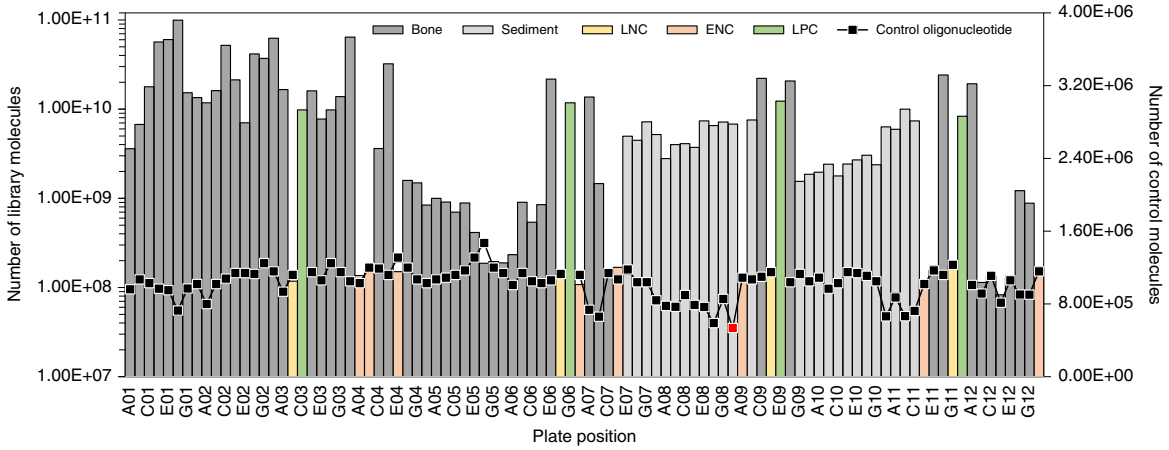


Fig. 3 | Quantitative PCR results obtained after the automated preparation of 96 single-stranded DNA libraries. The experiment shown here included sample DNA from ancient bones and sediments as well as extraction negative controls (ENCs), library preparation positive controls (LPCs) and library preparation negative controls (LNCs). A sample with poor control oligonucleotide yield (<50% compared to the LNC and ENC reactions) is highlighted in red.

the number of control library molecules from the spike-in oligonucleotide is similar between the sample libraries and the negative controls with the exception of one library prepared from a sediment sample, where the number of control molecules dropped to less than 50% of those in the negative controls. This observation indicates that the respective sample might contain inhibitory substances. Less DNA input should be used for library preparation from this sample in the future should additional libraries be needed.

Supplementary Table 1 provides further information about the quality of library preparation and the characteristics of the libraries that can be gathered from the sequence data. One of the most important measures is the total number of nucleotides with similarity to the reference genome in each library (also referred to as 'informative sequence content'[12]). This number is determined by multiplying the fraction of sequenced molecules that are at least 35 bp long and produced an alignment to the human reference genome by the total number of unique molecules in each library (as estimated by qPCR) and the average length of the mapped sequences. For the library and extraction negative controls, the informative sequence content corresponds to the amount of human contamination that was introduced during library preparation and DNA extraction, which is generally less than 5 million bp in our hands. For the sample libraries, it represents an estimate of the genomic coverage in the library, which, together with other library characteristics, such as the fraction of mapped sequences and the frequency of deamination-induced C-to-T substitutions, allows one to assess the state of DNA preservation in the sample and to determine the best strategy for further data acquisition.

Reporting Summary

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

Data availability

All raw data points underlying Fig. 3 are listed in Supplementary Table 1.

Code availability

Electronic protocol files for automated library preparation and auxiliary files are available at the Zenodo website (https://zenodo.org/record/3631147).

References

1. Meyer, M. et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).
2. Gansauge, M. T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737–748 (2013).

3. Prüfer, K. et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).
4. Prüfer, K. et al. A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* **358**, 655–658 (2017).
5. Mafessoni, F. et al. A high-coverage Neandertal genome from Chagyrskaya Cave. Preprint at https://www.biorxiv.org/content/10.1101/2020.03.12.988956v1 (2020).
6. Fu, Q. et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).
7. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
8. Meyer, M. et al. A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* **505**, 403–406 (2014).
9. Meyer, M. et al. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature* **531**, 504–507 (2016).
10. Slon, V. et al. Neandertal and Denisovan DNA from Pleistocene sediments. *Science* **356**, 605–608 (2017).
11. Korlević, P. et al. Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *Biotechniques* **59**, 87–93 (2015).
12. Gansauge, M. T. et al. Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Res.* **45**, e79 (2017).
13. Gansauge, M. T. & Meyer, M. A method for single-stranded ancient DNA library preparation. *Methods Mol. Biol*. 1963, 75–83 (2019).
14. Glocke, I. & Meyer, M. Extending the spectrum of DNA sequences retrieved from ancient bones and teeth. *Genome Res.* **27**, 1230–1237 (2017).
15. de Filippo, C., Meyer, M. & Prüfer, K. Quantifying and reducing spurious alignments for the analysis of ultra-short ancient DNA sequences. *BMC Biol.* **16**, 121 (2018).
16. Bennett, E. A. et al. Library construction for ancient genomics: single strand or double strand? *Biotechniques* **56**, 289–290 (2014).
17. Wales, N. et al. New insights on single-stranded versus double-stranded DNA library preparation for ancient DNA. *Biotechniques* **59**, 368–371 (2015).
18. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc*. **2010**, pdb prot5448 (2010).
19. Caroe, C. et al. Single-tube library preparation for degraded DNA. *Methods Ecol. Evol.* **9**, 410–419 (2018).
20. Briggs, A. W. et al. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl Acad. Sci. USA* **104**, 14616–14621 (2007).
21. Skoglund, P. et al. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl Acad. Sci. USA* **111**, 2229–2234 (2014).
22. Gansauge, M. T. & Meyer, M. Selective enrichment of damaged DNA molecules for ancient genome sequencing. *Genome Res.* **24**, 1543–1549 (2014).
23. Bokelmann, L. et al. A genetic analysis of the Gibraltar Neanderthals. *Proc. Natl Acad. Sci. USA* **116**, 15610–15615 (2019).
24. Stiller, M. et al. Single-strand DNA library preparation improves sequencing of formalin-fixed and paraffin-embedded (FFPE) cancer DNA. *Oncotarget* **7**, 59115–59128 (2016).
25. Snyder, M. W., Kircher, M., Hill, A. J., Daza, R. M. & Shendure, J. Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. *Cell* **164**, 57–68 (2016).
26. Burnham, P. et al. Single-stranded DNA library preparation uncovers the origin and diversity of ultrashort cell-free DNA in plasma. *Sci. Rep*. **6**, 27859 (2016).
27. Turchinovich, A. et al. Capture and Amplification by Tailing and Switching (CATS). An ultrasensitive ligation-independent method for generation of DNA libraries for deep sequencing from picogram amounts of DNA and RNA. *RNA Biol.* **11**, 817–828 (2014).
28. Karlsson, K. et al. Amplification-free sequencing of cell-free DNA for prenatal non-invasive diagnosis of chromosomal aberrations. *Genomics* **105**, 150–158 (2015).
29. Raine, A., Manlig, E., Wahlberg, P., Syvanen, A. C. & Nordlund, J. SPlinted Ligation Adapter Tagging (SPLAT), a novel library preparation method for whole genome bisulphite sequencing. *Nucleic Acids Res.* **45**, e36 (2017).
30. Wu, J., Dai, W., Wu, L. & Wang, J. SALP, a new single-stranded DNA library preparation method especially useful for the high-throughput characterization of chromatin openness states. *BMC Genomics* **19**, 143 (2018).
31. Ding, J., Taylor, M. S., Jackson, A. P. & Reijns, M. A. Genome-wide mapping of embedded ribonucleotides and other noncanonical nucleotides using emRiboSeq and EndoSeq. *Nat. Protoc.* **10**, 1433–1444 (2015).
32. Petryk, N. et al. Replication landscape of the human genome. *Nat. Commun.* **7**, 10208 (2016).
33. Tin, M. M., Economo, E. P. & Mikheyev, A. S. Sequencing degraded DNA from non-destructively sampled museum specimens for RAD-tagging and low-coverage shotgun phylogenetics. *PLoS ONE* **9**, e96793 (2014).
34. Troll, C. J. et al. A ligation-based single-stranded library preparation method to analyze cell-free DNA and synthetic oligos. *BMC Genomics* **20**, 1023 (2019).
35. Knapp, M., Clarke, A. C., Horsburgh, K. A. & Matisoo-Smith, E. A. Setting the stage - building and working in an ancient DNA laboratory. *Ann. Anat.* **194**, 3–6 (2012).

36. Kampmann, M. L., Borsting, C. & Morling, N. Decrease DNA contamination in the laboratories. *Forensic Sci. Int. Genet.* **6**, E577–E578 (2017).
37. Rohland, N., Glocke, I., Aximu-Petri, A. & Meyer, M. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat. Protoc.* **13**, 2447–2461 (2018).
38. Prüfer, K. snpAD: an ancient DNA genotype caller. *Bioinformatics* **34**, 4165–4171 (2018).
39. Varshney, U. & van de Sande, J. H. Specificities and kinetics of uracil excision from uracil-containing DNA oligomers by *Escherichia coli* uracil DNA glycosylase. *Biochemistry* **30**, 4055–4061 (1991).
40. Burbano, H. A. et al. Targeted investigation of the Neandertal genome by array-based sequence capture. *Science* **328**, 723–725 (2010).
41. Fu, Q. et al. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl Acad. Sci. USA* **110**, 2223–2227 (2013).
42. Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
43. Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**, e3 (2012).
44. Dabney, J. & Meyer, M. Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *Biotechniques* **52**, 87–94 (2012).
45. Thompson, J. R., Marcelino, L. A. & Polz, M. F. Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nucleic Acids Res.* **30**, 2083–2088 (2002).
46. Renaud, G., Stenzel, U. & Kelso, J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res.* **42**, e141 (2014).
47. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

Acknowledgements

We thank L. Bokelmann, E. Essel, L. Lippik, J. Richter, B. Schellbach and A. Weihmann for help in the lab, D. Massilani, S. Pääbo, B. Vernot and E. Zavala for helpful discussions, I. Bünger for help with installing software, J. Kelso and J. Visagie for help with data processing and L. Jauregui for comments on the manuscript. This work was funded by the Max Planck Society.

Author contributions

M.-T.G. and M.M. developed the manual protocol. A.A.-P. and M.M. developed the automated protocol version, which was tested and optimized by S.N. M.M. and M.-T.G. wrote the paper with help from A.A-P. and S.N.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41596-020-0338-0.

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-020-0338-0.

Correspondence and requests for materials should be addressed to M-T.G. or M.M.

Peer review information *Nature Protocols* thanks Eva-Maria Geigl and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Received: 19 November 2019; Accepted: 17 April 2020;
Published online: 1 July 2020

Related links

Key references using this protocol

Slon, V. et al. *Science* **356**, 605–608, (2017): https://doi.org/10.1126/science.aam9695
Hajdinjak, M. et al. *Nature* **555**, 652–656 (2018): https://doi.org/10.1038/nature26151
Slon, V. et al. *Nature* **561**, 113–116 (2018): https://doi.org/10.1038/s41586-018-0455-x

Protocol update to:

Gansauge, M. & Meyer, M. *Nat. Protoc.* **8**, 737–748 (2013): https://doi.org/10.1038/nprot.2013.038

Extended Data Fig. 1 | Overview of the adapter sequences and library amplification primers used in this protocol. All sequences are shown in 5′–3′ orientation.

nature research

Corresponding author(s): Matthias Meyer

Last updated by author(s): Feb 3, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data processing was performed using the software supplied with the qPCR and DNA sequencing instruments specified in the manuscript. Summary statistics were generated using bwa, samtools and other software specified in the manuscript. Electronic protocol files for automated library preparation are published as part of this study.
Data analysis	No special code or software was required for data analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All molecule and sequence counts presented in Figure 3 are available in Supplementary Table 1.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The paper details a method. The performance of the method was demonstrated by providing results from an experiment with 96 samples, which is the number of samples that fit on a microtiter plate.
Data exclusions	No data were excluded, results from all samples are shown.
Replication	Positive and negative controls were performed in several replicates to demonstrate the reproducibility of the method.
Randomization	Not relevant, no comparisons were made between groups of samples.
Blinding	Not relevant, all qPCR and sequence data were obtained using the instruments' software regardless of sample type or origin.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☐	☒ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Palaeontology

Specimen provenance	DNA extracts from ancient bone and sediment samples were used solely to exemplify the technical feasibility of data generation using the method presented. Only technical parameters are shown, with no reference to the age or provenance of the material, which are irrelevant in this context and are not reported in the manuscript.
Specimen deposition	Material, or leftovers thereof, are stored at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany.
Dating methods	No dates were obtained for this study.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.