

Preparation of DNA Sequencing Libraries for Illumina Systems—6 Key Steps in the Workflow



Preparation of libraries for DNA sequencing for Illumina systems involves multiple steps. In a general workflow, purified DNA is fragmented, end-repaired, and A-tailed; adapters are ligated to the DNA fragments; libraries are amplified if necessary; and the prepared libraries are cleaned, quantitated, and normalized before loading onto a flow cell (**Figure 1**). Since library preparation plays a critical role in obtaining high-quality data [1], researchers should understand the underlying principles and considerations for the key steps in the workflow.

On this page

1. DNA sequencing methods

a. Whole-genome sequencing

b. *De novo* sequencing

c. Targeted sequencing

d. Exome sequencing
2. DNA fragmentation strategies

a. Mechanical shearing

b. Enzymatic digestion

c. Transposon-based fragmentation
3. End repair and adapter ligation

a. End repair

b. Adapter ligation

c. Index hopping and unique dual indexes
4. Library amplification considerations

a. PCR-free libraries

b. PCR-based libraries
5. Size selection and cleanup

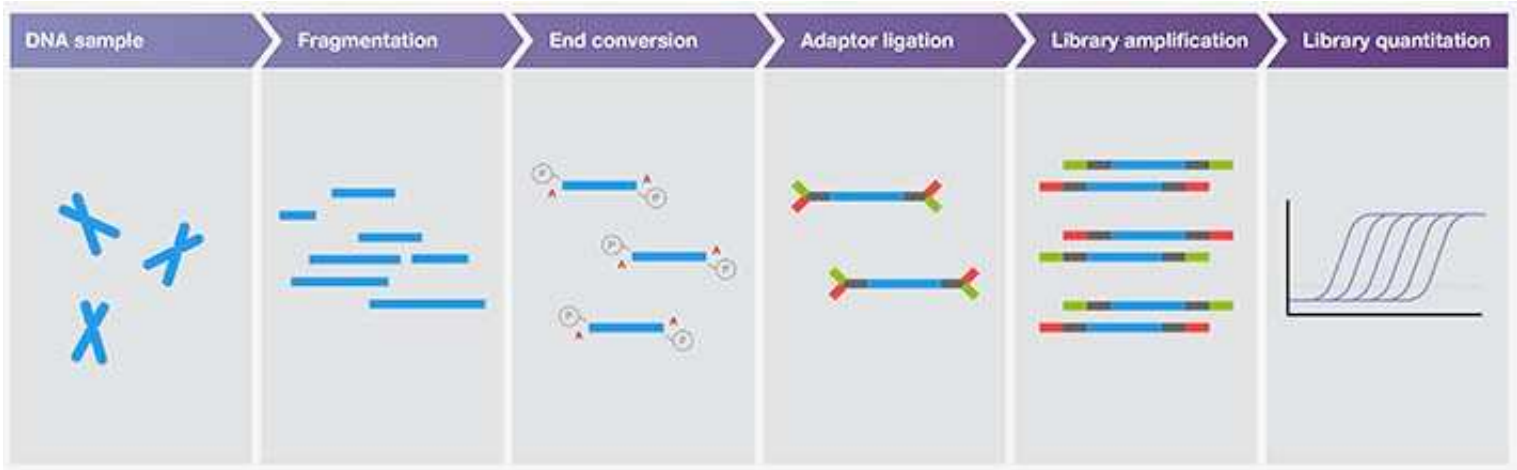
a. Importance

b. Methods
6. Library quantitation approaches

a. Microfluidics

b. Fluorometry

c. qPCR



Click to enlarge image

Figure 1. DNA sequencing library preparation.

1. DNA sequencing methods

Limited-time promotions on consumables and equipment [See offer details >](#)

Common DNA sequencing methods include whole-genome sequencing, *de novo* sequencing, targeted sequencing, and exome sequencing (discussed below) (Figure 2). Exome sequencing can also be sequenced for epigenetic studies—e.g., methylation analysis (also known as **bisulfite sequencing** or **Bis-Seq**) and DNA–protein interaction sequencing (commonly known as ChIP-Seq), which are not covered in this section. The method of choice depends on the research goals and biological questions to address [2-4].

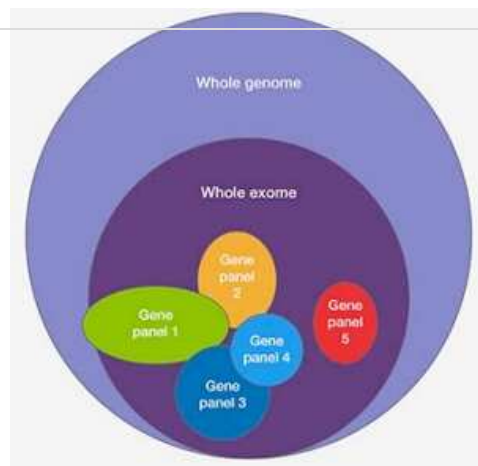


Figure 2. Common DNA sequencing methods. Exome and gene panel sequencing are considered targeted methods, since they only include subsets of the whole genome. Some gene panels may include promoter sequences.

a. Whole-genome sequencing

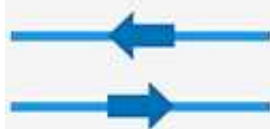
Whole-genome sequencing, or **WGS**, is performed to sequence the entire genome of an organism using the total genomic DNA. WGS data of a sample is then compared to a reference sample or control—for instance, comparison between cancer cells and normal cells—for small and large genetic variations. Examples of these genetic variations include single nucleotide polymorphisms (SNPs); single nucleotide variations (SNVs); nucleotide insertions and deletions (indels); structural rearrangements such as inversions, duplications, and translocations; and copy number variations (CNVs) (Figure 3).

ThermoFisher
SCIENTIFIC

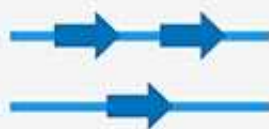
Single nucleotide polymorphism (SNP)

Insertion and deletion (indel)

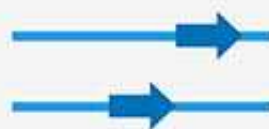
Structural rearrangements



Inversion



Duplication



Translocation



Copy number variation (CNV)

Figure 3. Common genetic variations.

WGS is useful for uncovering genetic mutations in an unbiased and detailed manner. However, it requires a large amount of sample input and involves extensive data processing, especially when analyzing the human genome, which is large and complex.

[Top](#)b. *De novo* sequencing

When genomic data for a particular organism are either unavailable or of insufficient quality, ***de novo* sequencing** (meaning “from the beginning”) is a method of building or updating the **reference genome**. Although a whole-genome sample may be used in sequencing, the lack of a reference sequence necessitates assembling overlapping short sequencing reads into longer contiguous sequences (**contigs**) (**Figure 4A**) using computational tools. The main goal is to generate an overall physical map that represents the whole genome without (large) gaps.

De novo sequencing usually relies on a hybrid approach for assembling the genome: reads from long-insert paired-end sequencing, referred to as **mate-pair sequencing** (with higher error rate), are used to build a scaffold, and reads from short-insert paired-end sequencing (with lower error rate) are used to fill in and improve the quality of a new genome map [5] (**Figure 4B**).

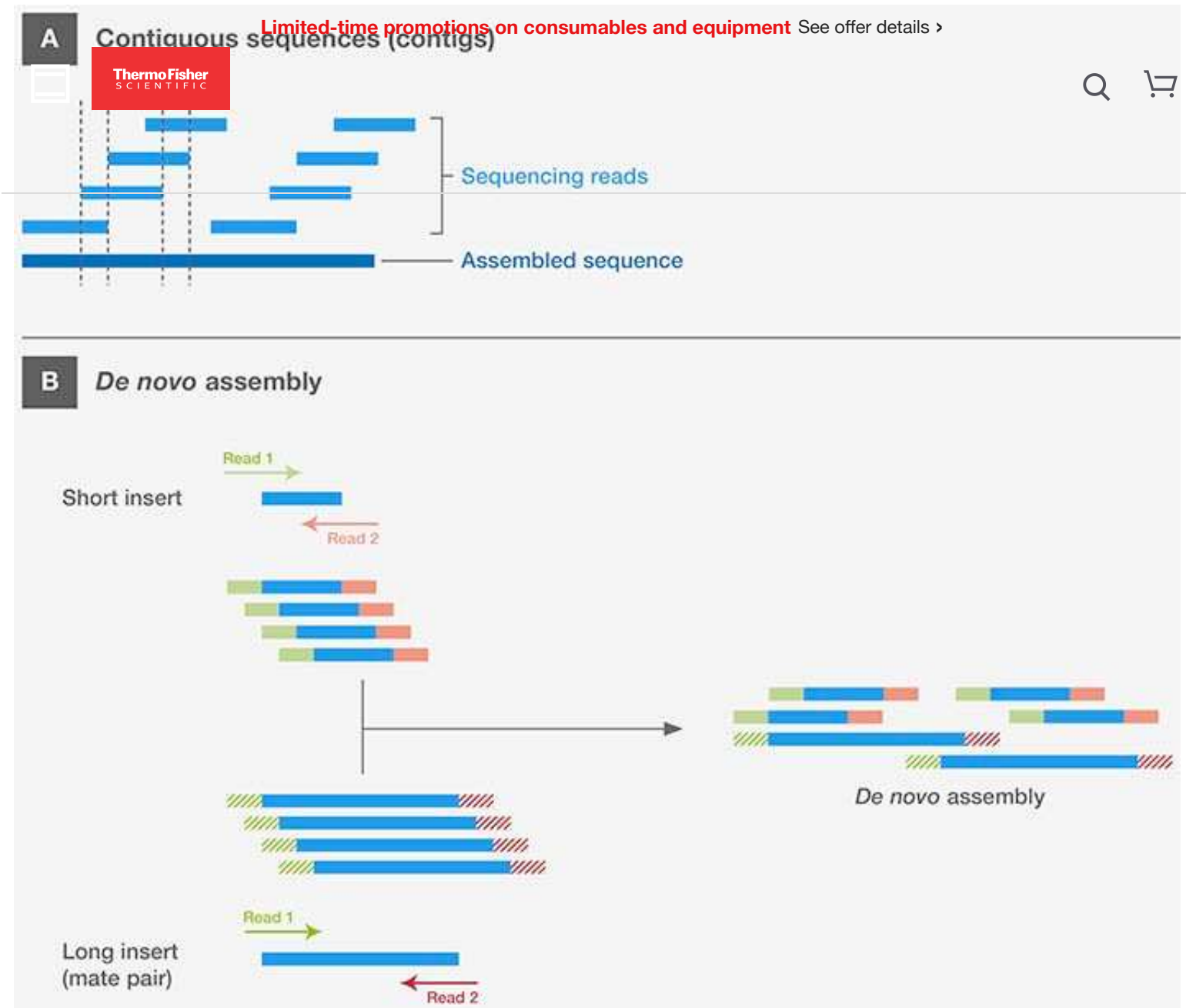


Figure 4. *De novo* sequencing and assembly. (A) Alignment of contiguous sequences. (B) Assembly of short-insert and long-insert paired-end reads into a reference genome.

Top

c. Targeted sequencing

Targeted sequencing (instead of WGS) is used when the goal of the experiment is to sequence specific genes, sets of related genes, or targeted regions of a genome. An example of targeted sequencing is screening for known cancer genes in different types of cancer cells. Therefore, targeted sequencing is **hypothesis-driven** and requires knowledge of the sequence of the reference genes or genomic regions. Since targeted sequencing does not require analysis of the whole genome (e.g., 3.2×10^9 base pairs for human), it allows more reads, better coverage, and higher depth, and therefore improved detection of rare variants at a lower cost than WGS.

To perform targeted sequencing, samples are enriched for the sequences of interest. Among methods available for enrichment of target sequences, the two most common approaches are hybrid capture and PCR amplification [6].

- The **hybrid capture strategy** utilizes a set of oligonucleotide probes that are complementary to the target sequences. Probes are usually coupled to magnetic or biotinylated beads so that target sequences hybridized to the probes can be selected from the mixture. After removing the unbound sequences, target sequences are released from the probes and prepared as a sequencing library (**Figure 5**). Target enrichment by hybrid capture usually requires higher sample input and a longer workflow, but it may yield more uniform coverage and higher data quality over the PCR enrichment method.



Click to enlarge image

Figure 5. Target enrichment by hybrid capture. Blue = desired sequences, red = magnetic bead-bound probes.

- **Target enrichment by PCR**, also known as **amplicon sequencing**, relies on highly **multiplexed PCR** to amplify DNA sequences corresponding to target regions. As many as 24,000 primer pairs, each pair designed to amplify a specific region, may be used to capture hundreds of sequences in one PCR run (**Figure 6**). PCR enables limited sample input and a faster workflow. However, the quality and coverage of the data obtained may be impacted by primer design, PCR efficiency, amplification bias, etc.

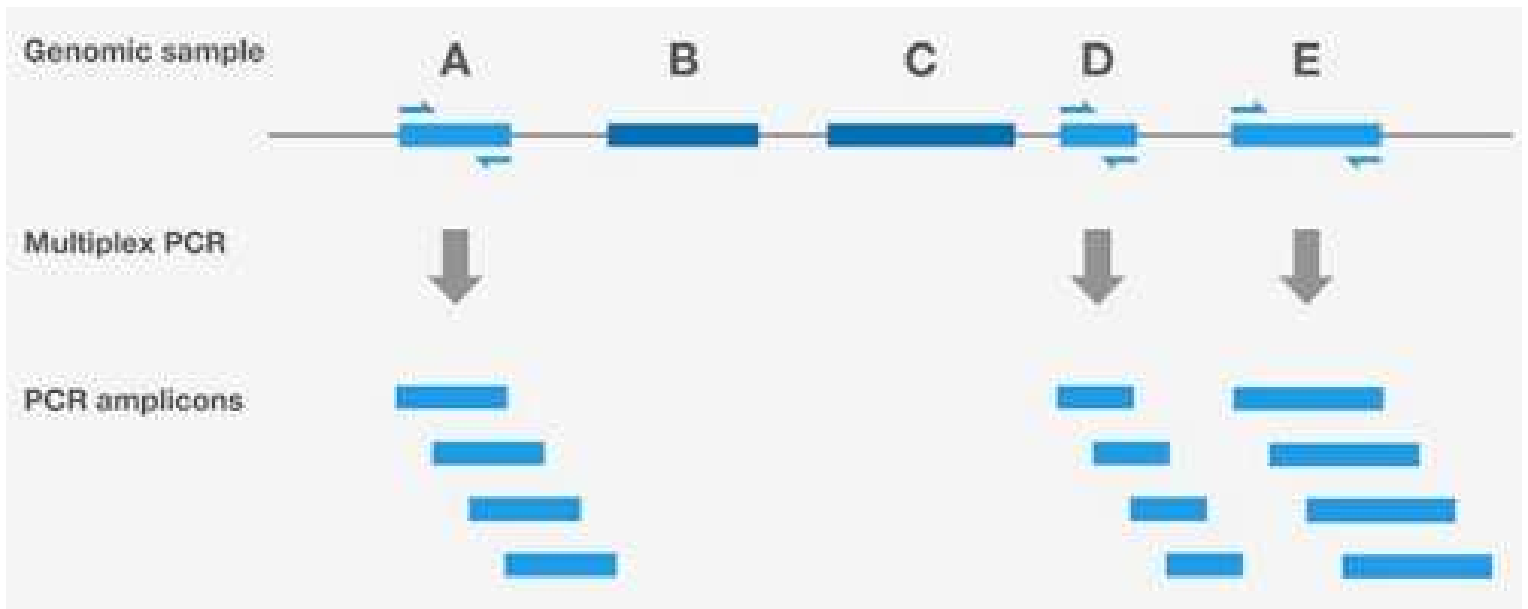


Figure 6. Target enrichment by PCR amplification.

d. Exome sequencing

Exome sequencing is a special type of targeted method to sequence protein-coding regions of the genome, called the exome [7]. While making up only about 1–2% of the human genome, the exome harbors approximately 85% of known disease-causing mutations. Therefore, whole-exome sequencing (**WES**) enables researchers to focus on identifying genetic mutations and variations that are significantly implicated in diseases.

Top

2. DNA fragmentation strategies

The first step in NGS library preparation for Illumina systems is fragmentation of DNA into the desired size range, typically 300–600 bp depending on the application. Traditionally, two methods have been employed for DNA fragmentation: mechanical shearing and enzymatic digestion. Typically, 1–5 mg of input DNA is required for fragmentation, but often less is needed for enzymatic fragmentation approaches.

Between the two methods, **mechanical shearing** is more widely used because of its unbiased fragmentation and ability to obtain more consistent fragment sizes (**Figure 7**). On the other hand, **enzymatic digestion** requires lower DNA input and offers a more streamlined library preparation workflow.

Limited-time promotions on consumables and equipment. See offer details >

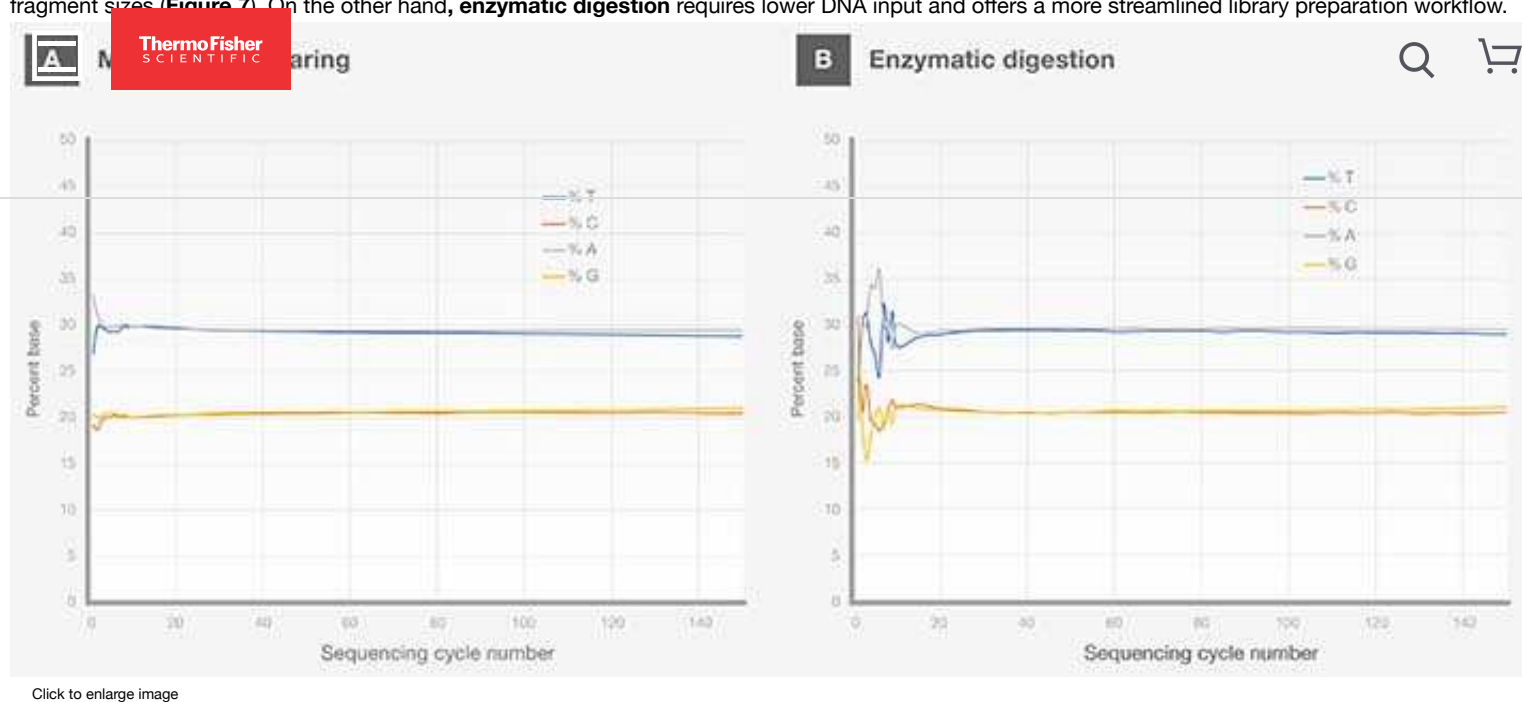


Figure 7. Comparison of percentage of each base at each position in sequencing of samples prepared by mechanical shearing vs. enzymatic digestion. Mechanical shearing shows very little bias in base representation at the beginning of reads, but enzymatic digestion shows some base imbalance at this stage.

Top

a. Mechanical shearing

Mechanical shearing involves breakage of phosphodiester linkages of DNA molecules by applying shear force. Widely used methods include high-power unfocused sonication, nebulization, and focused high-frequency acoustic shearing.

- Sonication** is the simplest method among the three and uses a sonicator (probe- or waterbath-based) to emit low-frequency acoustic waves for shearing. Although **probe-based** sonication delivers more focused energy towards the sample, the samples are in an open container, directly in contact with the probe, and thus are at a high risk of contamination. **Waterbath-based** sonication, on the other hand, keeps the samples within a closed system but usually requires higher energy due to energy dissipation/dispersion and low output. In either approach, optimization is needed to obtain the desired fragment lengths (**Figure 8**). Resting/cooling periods between sonication cycles should be incorporated to keep the samples from overheating, which necessitates a longer workflow and wait time.

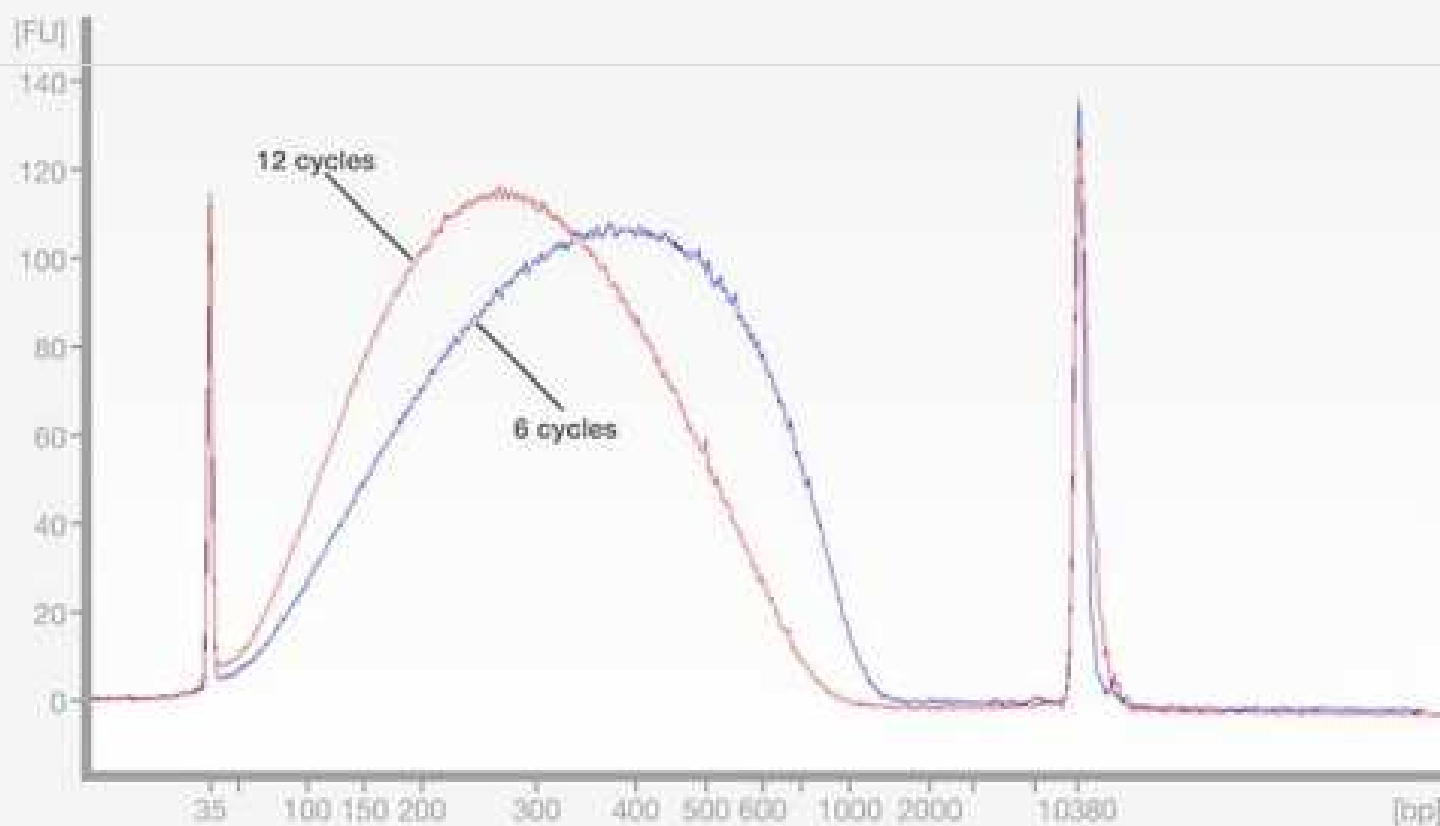


Figure 8. Dependence of average fragment length distribution on number of sonication cycles (1 sonication cycle = 30 sec).

- **Nebulization** creates shear force with compressed gas, forcing a nucleic acid solution through a small hole in a nebulizer. The aerosolized sample with fragmented DNA is then collected. The level of fragmentation can be controlled by the compressed gas pressure and can also be affected by the solution's viscosity and temperature. This method requires a large sample input and often results in high sample loss (low recovery).
- The **focused acoustic method** (developed by Covaris) uses high-frequency ultrasonic waves to shear DNA. High-frequency waves concentrate high energy on the sample within a small enclosed tube while minimizing heat generation. It has become a preferred method of mechanical shearing among NGS users because of its advantages over traditional sonication and nebulization, such as minimal sample loss, low risk of contamination, and better control over uniform fragmentation. However, the special equipment needed and the associated cost often limit its usage.

Top

b. Enzymatic digestion

Enzymatic digestion is an effective alternative to the mechanical shearing methods. **Endonucleases** and nicking enzymes are usually employed to cleave both strands of DNA or nick individual strands to generate double-stranded breakage. To avoid sequence bias, enzymes with less cleavage specificity and/or cocktails of enzymes are used for fragmentation. The enzymatic digestion approach typically requires lower DNA input than mechanical shearing and thus is a method of choice when you have limited samples. In addition, enzymatic digestion and then downstream library preparation steps can be done in the same tube, thus enabling automation, streamlining the workflow, minimizing sample loss, reducing contamination risks, and decreasing hands-on time.

Top

c. Transposon-based fragmentation

Some users may follow transposon-based library preparation as an alternative to mechanical shearing and enzymatic digestion (**Figure 9**) [8]. Using transposons, **this approach** fragments DNA templates and simultaneously tags them with transposon sequences, generating blunt DNA fragments with transposed sequences at both ends. Adapters (and indexes) are added via adapter-addition PCR. Therefore, some steps of the conventional workflow, such as traditional DNA fragmentation, end conversion, and adapter ligation, are circumvented when following this approach.



Figure 9. Fragmentation and tagging by transposons.

Top

3. End repair and adapter ligation

a. End repair

Following the fragmentation step, DNA samples are subjected to end repair (also called **end conversion**). DNA fragments produced by mechanical shearing or enzymatic digestion have a mix of 5' and 3' protruding ends that need repair or conversion for ligation with the adapters. The following are key steps in the process to blunt, phosphorylate, and adenylate the termini (**Figure 10**) [2].

- 5' overhangs are filled in by 5'→3' polymerase activity of an enzyme such as **T4 DNA polymerase** or **Klenow fragment**
- 3' overhangs are removed by 3'→5' exonuclease activity of an enzyme such as **T4 DNA polymerase**
- 5' ends of the blunted DNA fragments are phosphorylated (for efficient subsequent ligation) by an enzyme such as **T4 polynucleotide kinase**
- 3' ends of the blunted DNA fragments are adenylated (A tailing), which is required for T–A ligation with **Illumina adapters**, by an enzyme such as **Klenow fragment (exo⁻)** or **Taq DNA polymerase**

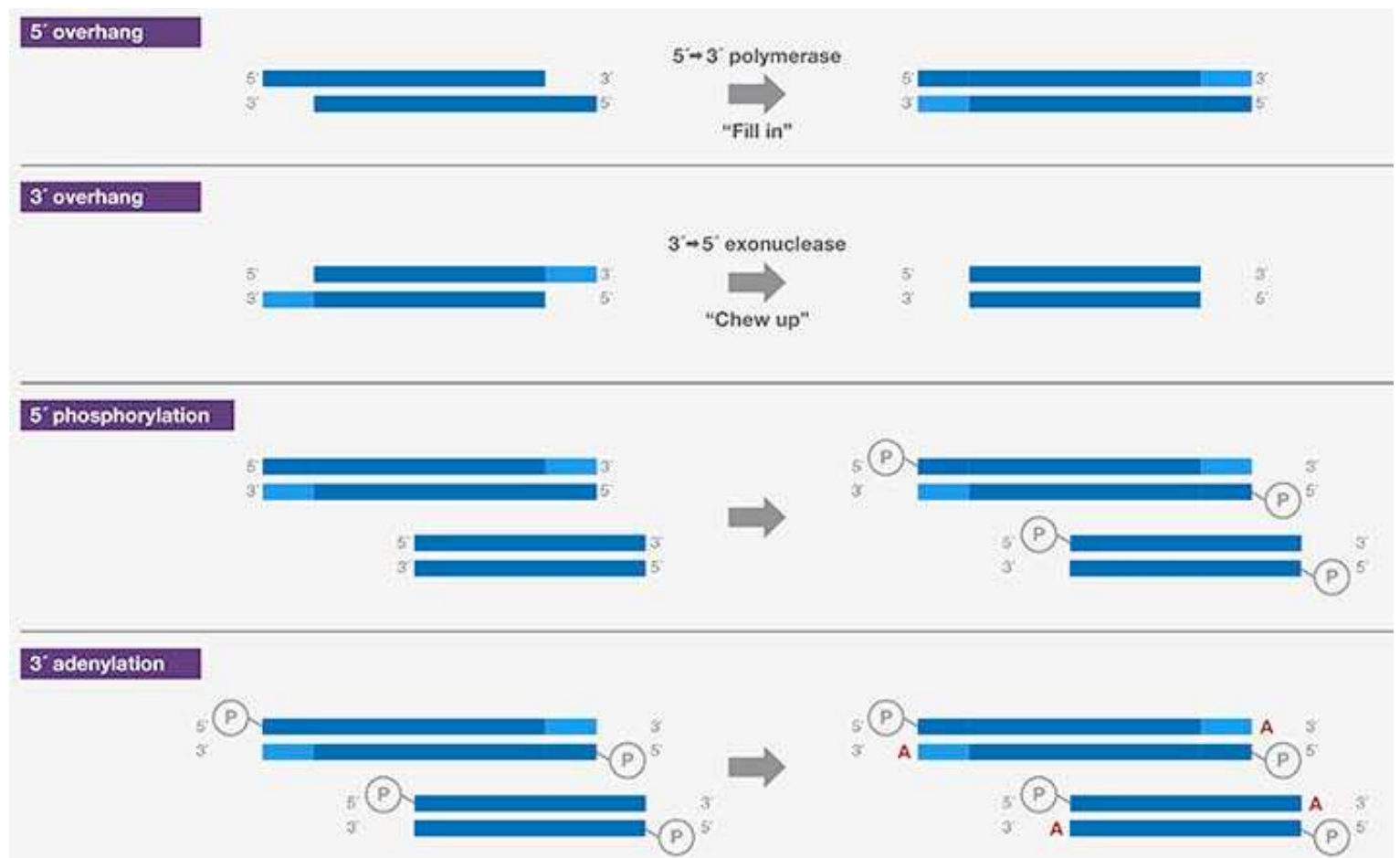


Figure 10. End conversion process.

The end conversion process involves a number of enzymatic steps, but some commercially available kits are designed to run all these reactions in a single tube, saving time and sample loss.

Top

b. Adapter ligation

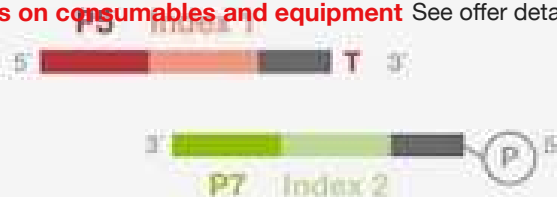
Limited-time promotions on consumables and equipment [See offer details >](#)

Adapters are short, single-stranded oligonucleotides that facilitate **clonal amplification and sequencing reactions**. Identical duplex adapters are ligated to both ends of DNA fragments so that oligos on the flow cell can recognize them for sequencing. In library preparation, a stoichiometric excess of adapters relative to sample DNA is used to help drive the **ligation reaction** to completion. Ligation efficiency is critical for conversion of DNA fragments into sequenceable molecules and thus impacts **conversion rate and yield** of the libraries. Because library fragments are flanked by adapters, they are sometimes called **inserts**.

During formation of the adapter duplexes, two strands of oligos called P5 and P7 are annealed. The **P5 and P7 adapters** are named after their sites of binding to the flow cell oligos. The adapters are noncomplementary at their ends to prevent their self-ligation and thus form a Y shape after annealing. This Y shape is no longer maintained if **library amplification** is subsequently performed (**Figure 11**).



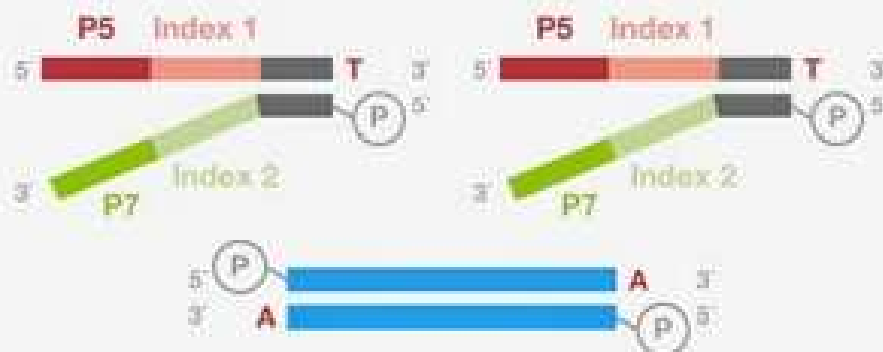
ThermoFisher
SCIENTIFIC



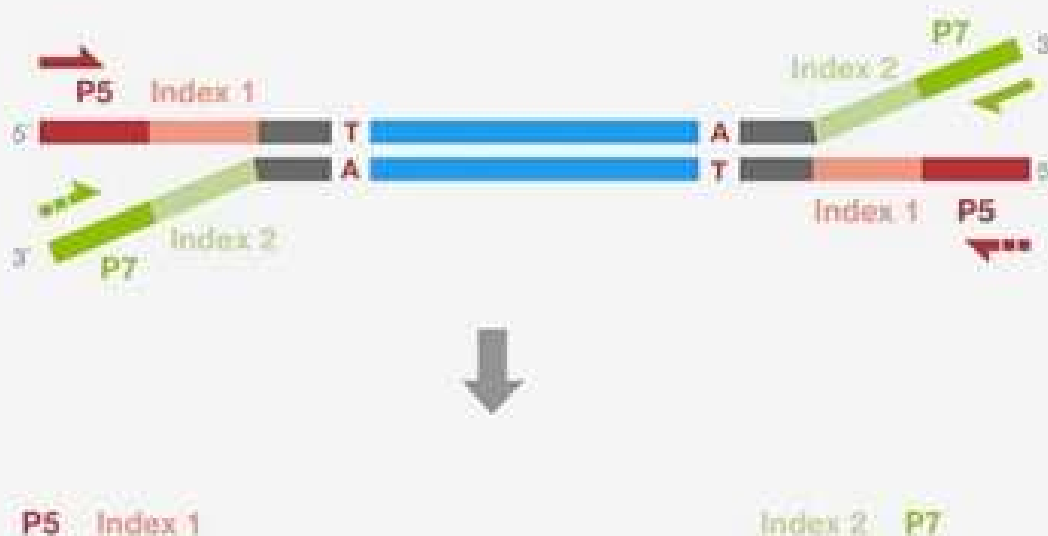
Annealing



Ligation



PCR



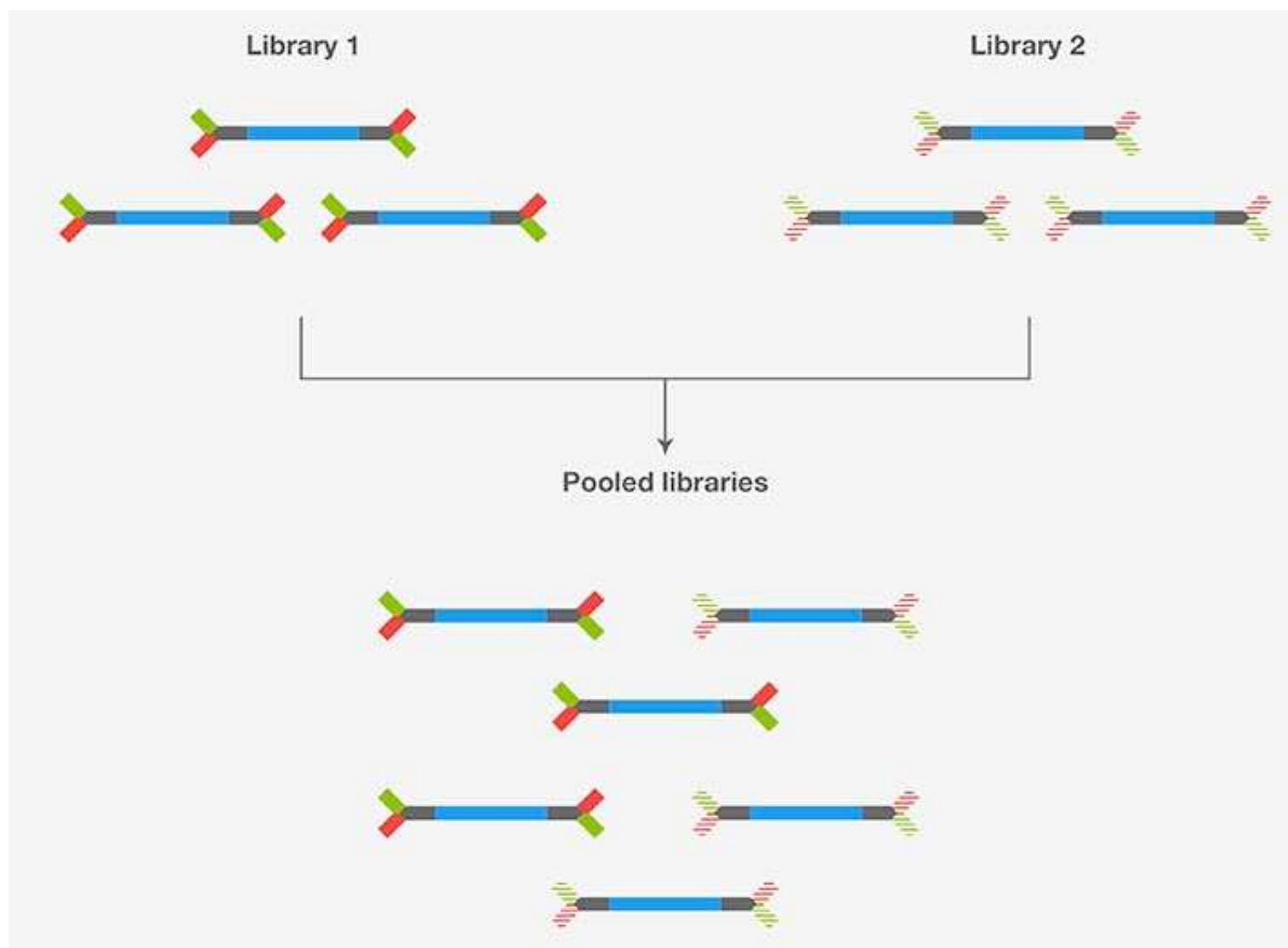
Looking for more information? Library adapters are usually 50–60 nucleotides long and often consist of the features described below (**Figure 12**) [9-10].



Figure 12. Sequencing adapters. (* = phosphorothioate linkage)

- Sites of binding to **P5** or **P7** oligos on the flow cells and to the sequencing primers
- **Index sequences** composed of specific 6–8 nucleotides to distinguish one sample from another. Index sequences enable multiplexing, a process of sequencing multiple libraries in one flow cell, and **dual-indexed** libraries are commonly employed for multiplex sequencing (**Figure 13**).
- Additional **T on the 3' end of the P5** adapter to prevent formation of adapter dimers and facilitate ligation with the 3' A of library fragments (similar to **TA cloning**). Since a missing 3' T would lead to adapter dimer formation, the more stable phosphorothioate linkage (instead of phosphodiester) is usually used to attach the 3' T to the adapters.
- **Phosphate on the 5' end of the P7** adapter for **ligation** with the 3' end of library fragments

For PCR-amplified libraries and RNA-Seq libraries, **unique molecular identifiers** (UMI) may be included to enable tracking of every library fragment and monitoring of deviations during **library amplification** [11].



c. Index hopping and unique dual indexes

Index hopping is a phenomenon associated with multiplexing or pooling of library samples. When two or more libraries are sequenced together in the same flow cell, one of the indexes assigned to one library may become swapped with that of another library (**Figure 14**). Index hopping has always affected multiplex libraries (e.g., from cross-contamination of indexes) but has become more prominent when sequencing is performed on **patterned flow cells with exclusion amplification chemistry** [12]. Index hopping has serious implications in subsequent **data analysis**, such as incorrect assignment of sequencing data from one sample (library) to another.

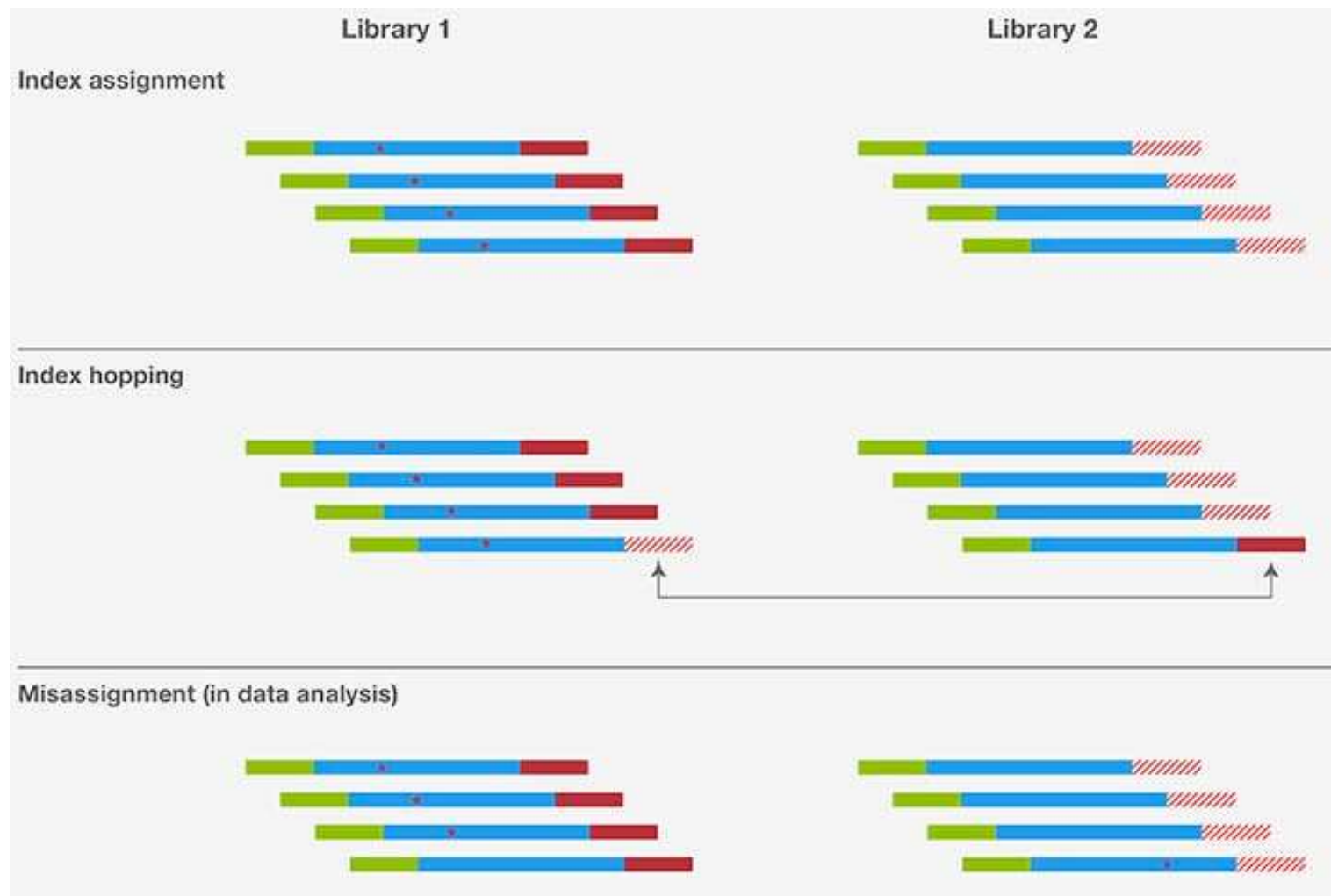


Figure 14. Index hopping. (* = mutation of interest from Library 1)

Two main strategies have been employed to minimize the effect of index hopping during sequencing.

- Using **unique dual indexes** (UDIs) instead of **combinatorial dual indexes** (CDIs) (**Figure 15**) [13-14]. Assigning a set of UDIs to each library in the sequencing pool helps ensure that index 1 and index 2 sequences be designated only once during sample pooling prior to loading of the sequencer.
- **Minimizing the amount of free, unligated adapter** in the samples. Removal of unligated adapters from the libraries helps minimize index hopping. Possibly for that reason, **PCR-free libraries** are reported to be more susceptible to index hopping than PCR-amplified libraries [12], because fewer cleanup steps are usually performed to remove unligated adapters. The amount of unligated adapters can be measured by **microfluidics-based electrophoresis**.

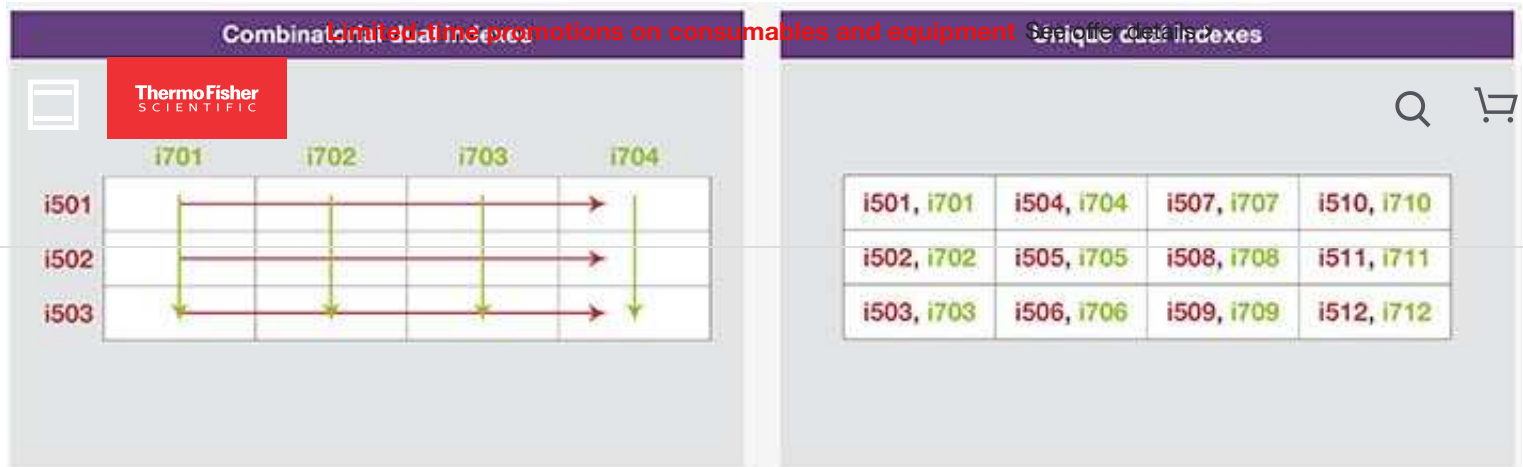


Figure 15. Combinatorial dual indexes (CDI) vs. unique dual indexes (UDI).

Top

4. Library amplification considerations

Depending on the need for amplification, DNA library preparation methods can be categorized as PCR-free or PCR-based. In either method, care should be taken to follow protocols that yield highly diverse and representative libraries of input samples from different amounts to help generate high-quality data.

a. PCR-free libraries

Since PCR amplification can contribute to GC bias, PCR-free library preparation is usually the preferred method to create libraries covering **high-GC or high-AT** sequences, to help ensure library diversity [1, 15]. Note that even with PCR-free library preparation methods, bias can be introduced during **cluster generation** and from the chemistry of the sequencing step itself.

Compared to PCR-based methods, PCR-free libraries require higher input amounts of starting material (although improvements have been made in lowering the input requirements). This can be challenging in scenarios such as using limited or precious samples and highly degraded nucleic acids. With PCR-free libraries, accurate assessment of library quality and quantity may be difficult, compared to PCR-amplified libraries [16].

Nevertheless, better representation and balanced coverage offered by PCR-free libraries make them attractive for the following applications:

- Studies of population-scale genomics and molecular basis of a disease
- Investigation of promoters and regulatory regions in the genome, which often are high in GC or AT content
- Whole-genome sequencing analysis and variant calling for single-nucleotide polymorphisms (SNPs) and small insertions or deletions (indels)

Top

b. PCR-based libraries

The PCR-based method is a popular strategy for constructing NGS libraries, since it allows lower sample input and selective amplification of inserts with adapters at both ends. However, PCR can introduce **GC bias**, leading to challenges in **data analysis**. For example, GC bias may hinder **de novo genome assembly** and single-nucleotide polymorphism (SNP) discovery.

A number of factors can impact GC bias, and the following factors should be considered to achieve balanced library coverage [17]:

- PCR enzyme and **master mix** used (**Figure 16**)
- Number of PCR cycles run, and cycling conditions
- PCR additives or enhancers in the reaction



ThermoFisher
SCIENTIFIC



Figure 16. Varying levels of GC bias in libraries amplified with different PCR enzyme master mixes.

With a given **PCR enzyme or master mix**, an increase in the number of PCR cycles usually increases GC bias. Therefore, a general recommendation is to run the **minimum number of cycles** (e.g., 4–8) that generates sufficient library yields for sequencing.

Decreasing the number of PCR cycles also reduces **PCR duplicates** and improves library complexity. PCR duplicates are defined as sequencing reads resulting from two or more PCR amplicons of the same DNA molecule. Although bioinformatic tools are available to identify and remove PCR duplicates during **data analysis** [18], minimizing PCR duplicates is important for efficient use of the flow cell in sequencing.

Other PCR artifacts can also result in reduced library quality and complexity. These artifacts include **amplification bias** (due to PCR stochasticity), **nucleotide errors** (from **enzyme fidelity**), and **PCR chimeras** (due to enzyme's template switching) (**Figure 17**) [19].

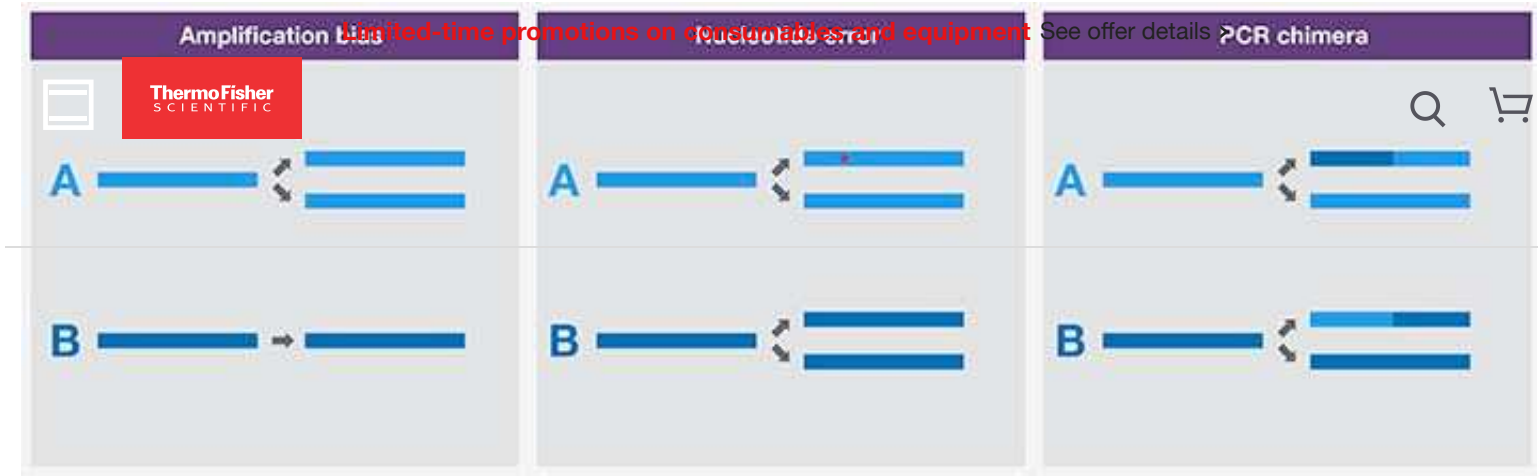


Figure 17. Common PCR artifacts.

Top

5. Size selection and cleanup

An important step in NGS library preparation is **size selection** and/or **cleanup**. Depending on the library preparation protocol, it may be performed following **fragmentation**, **adapter ligation**, or **PCR amplification**. As its name implies, the process selects the desired fragment size range, while removing unwanted components such as excess adapters, adapter dimers, and primers.

a. Importance of size selection and cleanup

In NGS libraries, **uniformity of fragment sizes** is critical to enable maximum data output and reliable **data analysis** because there are limitations to sequencing read length as dictated by NGS applications. If DNA inserts are much longer than recommended, some portions of the inserts remain unsequenced. On the other hand, inserts shorter than recommended result in suboptimal use of sequencing reagents and resources. A mix of short and long inserts could lower sequencing efficiency and pose challenges in **data analysis**.

Removal of unligated adapters and adapter dimers (two adapters ligated to each other) is crucial to improve data output and quality. Excess adapters often compete with library fragments in binding to the flow cell, lowering data output. Even worse, adapter dimers can also clonally amplify and generate sequencing “noise”, which must be filtered out during the **data analysis**. With the introduction of patterned flow cells, excess unligated adapters make the libraries more prone to **index hopping** during sequencing [12].

Top

b. Methods for size selection and cleanup

Among methods used for size selection, agarose gel-based and magnetic bead-based are two of the most popular. Sample amounts, sample throughput, protocol time, and size range of the libraries may determine the suitability of either method [20].

Size selection from agarose gels is essentially a **gel purification** process in which DNA fragments separated through the gel according to size are collected (**Figure 18**). In addition to being simple and effective, the method allows flexibility in gel percentages for separation and collection of fragments in a narrow range. However, it requires large amounts of sample and a long processing time, although specialized gels are **available** to simplify the process [21-22].

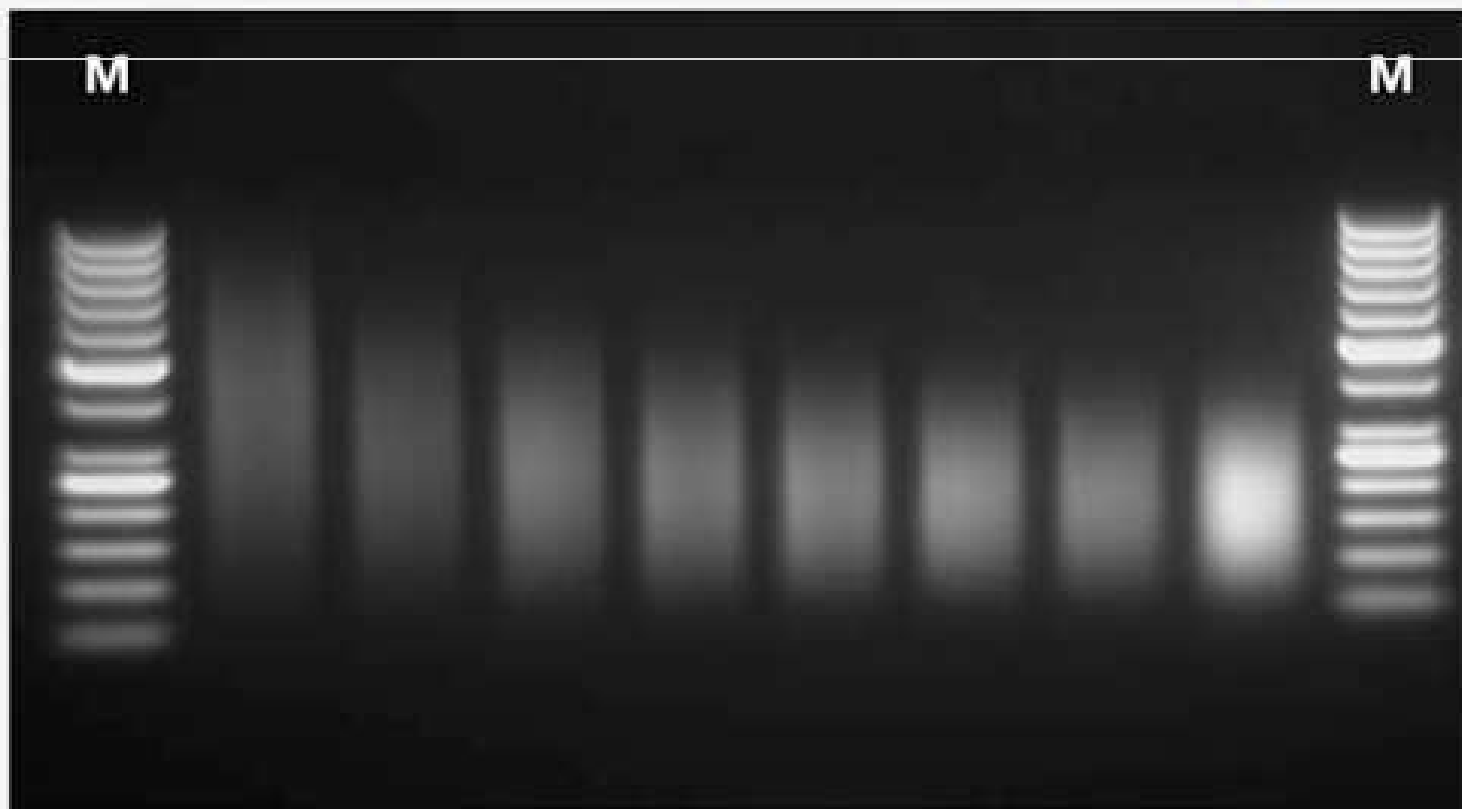


Figure 18. Size selection by agarose gel.

Size selection by magnetic beads is widely used in NGS library preparation. This method relies on binding and unbinding of DNA fragments of different lengths to the magnetic beads, which is controlled by the ratio of beads to DNA and by buffer composition (**Figure 19**) [20,23]. Suitability with low sample amounts, high recovery of DNA, ability to automate, and flexibility to select the desired fragment size range make this method attractive to NGS users. Nevertheless, the method may not be suitable to separate fragments that are very close in molecular weights.



Top

Before NGS libraries are loaded onto the sequencer, they should be **quantified** and normalized so that each library is sequenced to the desired depth with the required number of reads. Concentrations of prepared NGS libraries can vary widely because of differences in the amount and quality of nucleic acid input, as well as the **target enrichment** method that may be used. While **underclustering** due to overestimated library concentrations can result in diminished data output, **overclustering** can result in low quality scores and problematic downstream analysis (**Figure 20**).

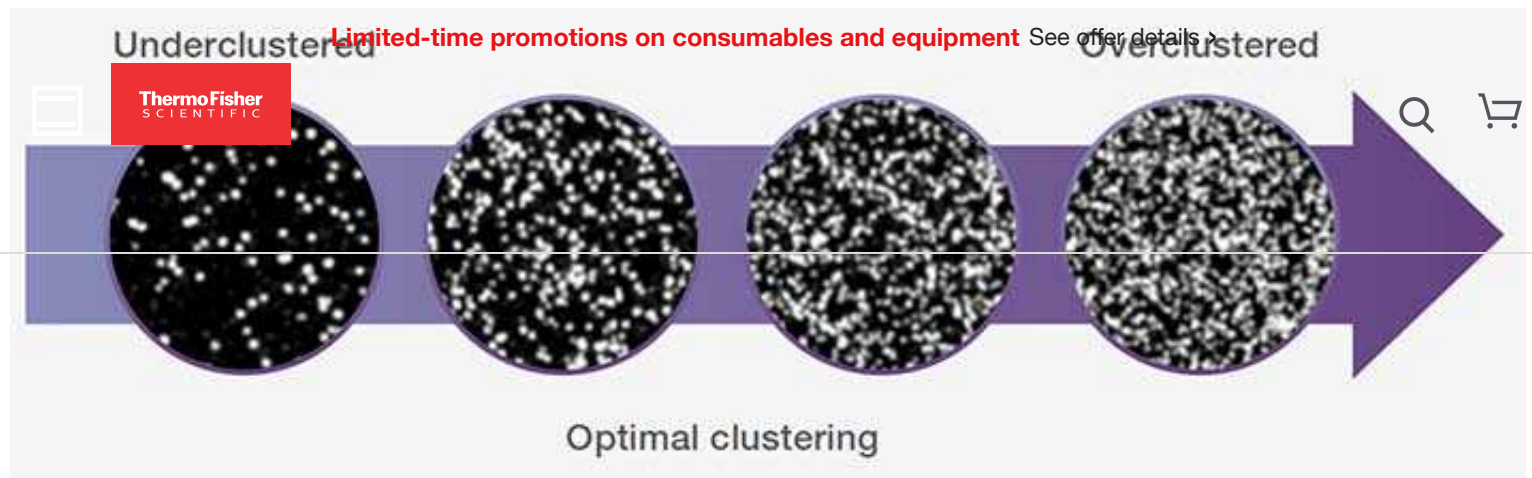


Figure 20. Library clustering on a flow cell.

Top

a. Microfluidics-based quantitation

Microfluidic electrophoresis separates fragments in NGS libraries based on size and can estimate the quantity of different size ranges using a reference standard (**Figure 21**). More commonly, however, the results of fragment analysis obtained by this method are used in conjunction with the two other methods listed below for more accurate quantitation of NGS libraries.

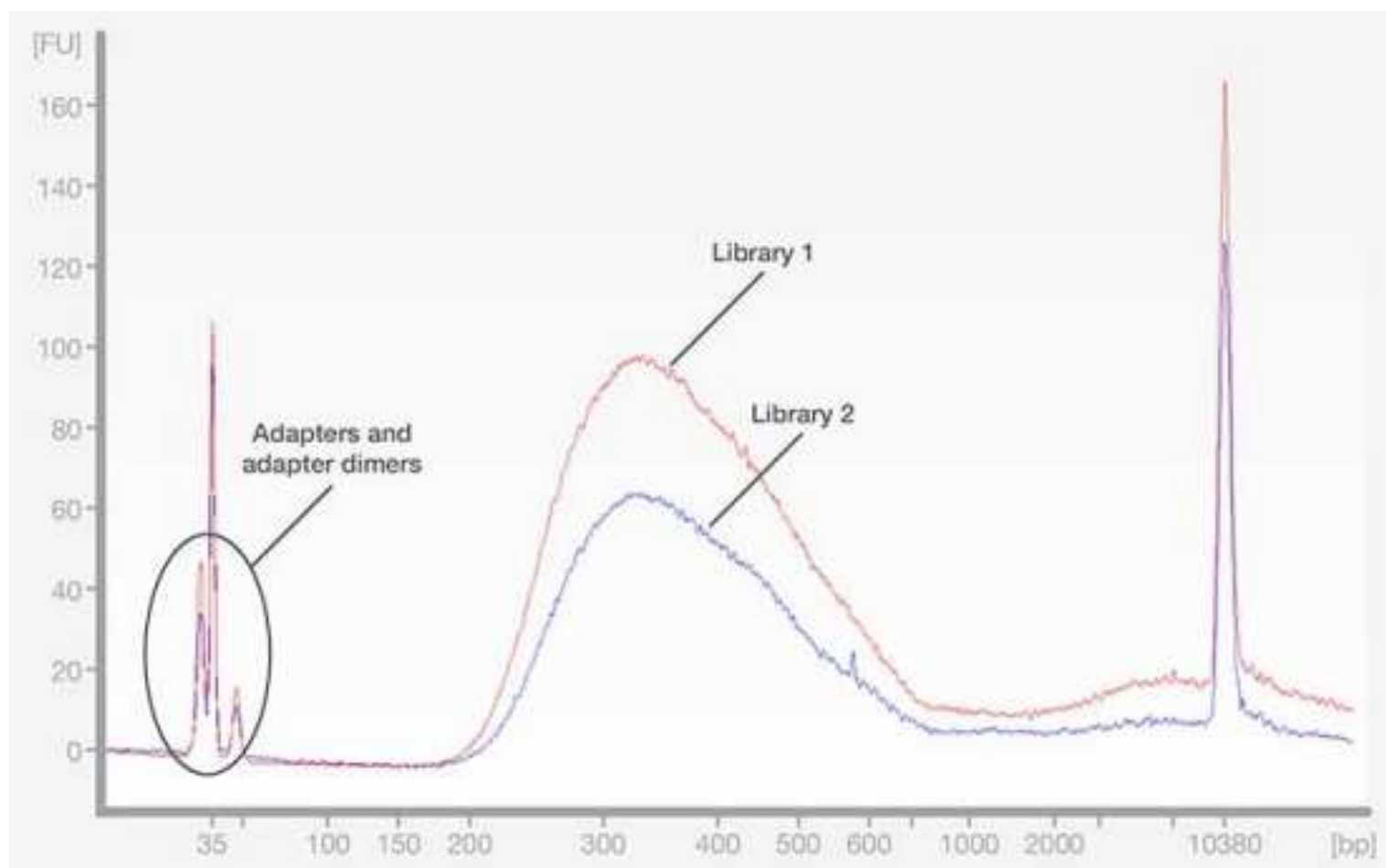


Figure 21. Fragment analysis of libraries by microfluidics-based electrophoresis.

Top

b. Fluorometry-based quantitation

Limited-time promotions on consumables and equipment See offer details >

The **fluorometric** assay uses fluorescent dyes that bind specifically to double-stranded DNA (dsDNA) to determine library concentration [24]. After a short incubation with the dye, the samples are read in a **fluorometer**, and library concentrations are calculated by (built-in) analysis software. Although the workflow is simple and takes only a few minutes per sample, this method may not scale well above 20–30 samples because samples are often read one at a time. Nevertheless, flexible input volumes and short incubation times allow for quick and easy testing of prepared libraries for concentrations. Since the measured concentration is for total dsDNA, the average size distribution of the libraries should be taken into account for accurate quantitation.

Video: Quantitation of DNA, RNA, and protein using a fluorometer



Top

c. qPCR-based quantitation

The **qPCR**-based assay quantifies NGS libraries by amplifying DNA fragments with the P5 and P7 adapters (**Figure 22**) [25]. A qPCR standard curve is used to determine a **broad range of library concentrations**, even as low as femtomolar. Since the PCR primers are designed specifically to bind to the adapter sequences, the **qPCR assays** detect only properly adapted, amplifiable libraries that can form **clusters** during sequencing. Note, though, that qPCR can also amplify adapter dimers; therefore, melting curve analysis and/or fragment size analysis should be performed to assess specificity and accuracy of quantitation by qPCR. The final library concentration is calculated based on the following formula.

$$\frac{\text{Length of qPCR DNA standard fragments (bp)}}{\text{Average lengths of library fragments (bp)}} \times \text{Library concentration (determined from qPCR)} = \text{Final library concentration}$$



ThermoFisher
SCIENTIFIC



PCR-amplified libraries

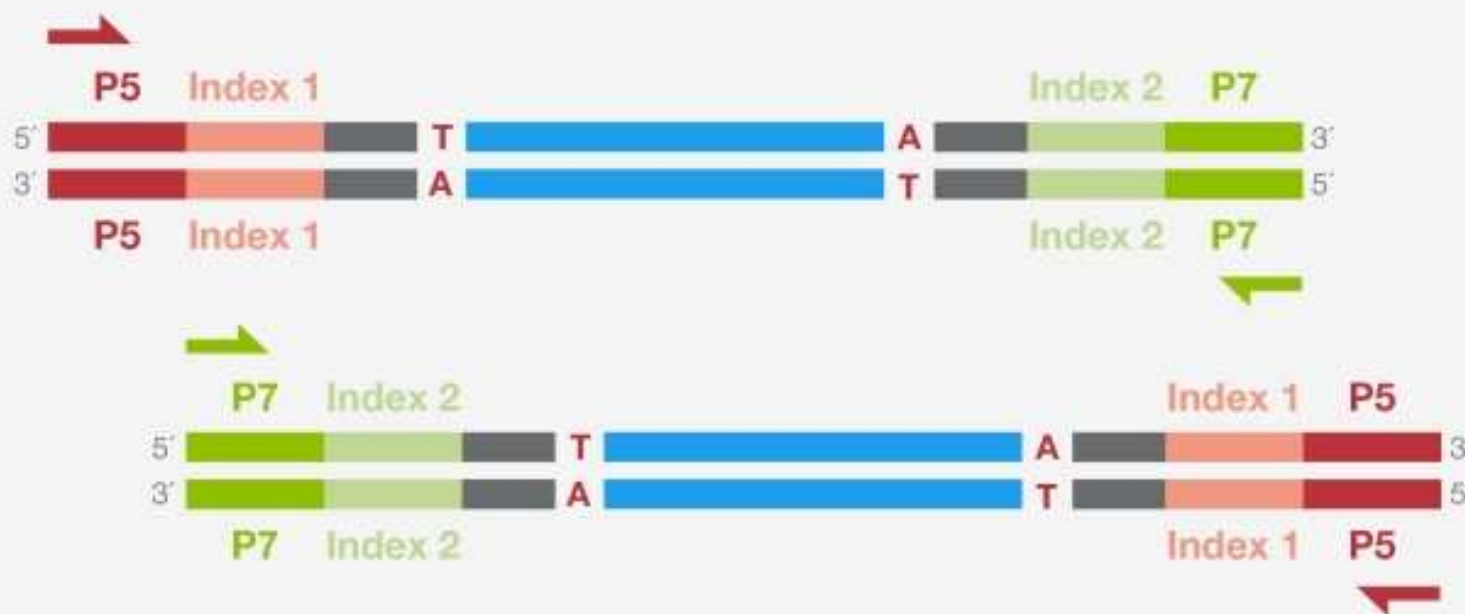


Figure 22. Schematic of primer binding in library quantitation by qPCR.

After preparation and quantitation, libraries of desired quantity and quality are ready to load on a flow cell for subsequent **clonal amplification and sequencing**.

[Top](#)

Resources

1. Jones MB, Highlander SK, Anderson EL et al. (2015) Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proc Natl Acad Sci U S A* 112(45):14024-14029.
2. Head SR, Komori HK, LaMere SA et al. (2014) Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* 56(2):61-64, 66, 68.
3. Sun Y, Ruivenkamp CA, Hoffer MJ et al. (2015) Next-generation diagnostics: gene panel, exome, or whole genome? *Hum Mutat* 36(6):648-655.
4. Belkadi A, Bolze A, Itan Y et al. (2015) Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci U S A* 112(17):5473-5478.
5. Sohn JI, Nam JW (2018) The present and future of de novo whole-genome assembly. *Brief Bioinform* 19(1):23-40.
6. Kozarewa I, Armisen J, Gardner AF et al. (2015) Overview of Target Enrichment Strategies. *CurrProtoc Mol Biol* 112:7.21.1-23.

Limited-time promotions on consumables and equipment See offer details >

7. Isakov O, Perrone M, Shomron N (2013) Exome sequencing analysis: a guide to disease variant detection. *Methods Mol Biol* 1038:137-158.
8. Adey A, Morrison HG, Asan et al. (2010) Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol* 11(12):R119.
9. Illumina. (2018) Sequencing Library Preparation Sequences Document.
10. Illumina. (2018) Illumina TruSeq DNA Adapters De-Mystified.
11. Kivioja T, Vähärautio A, Karlsson K et al. (2011) Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods* 9(1):72-74.
12. Illumina, Inc. (2018) Effects of index misassignment on multiplexing and downstream analysis.
13. MacConaill LE, Burns RT, Nag A et al. (2018) Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics* 19(1):30.
14. Costello M, Fleharty M, Abreu J et al. (2018) Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. *BMC Genomics* 19(1):332.
15. van Dijk E, Jaszczyszyn Y, Thermes C (2014) Library preparation methods for next-generation sequencing: tone down the bias. *Exp Cell Res* 322(1):12-20.
16. UC Davis Genome Center. NGS Library Construction.
17. Aird D, Ross MG, Chen W-S et al. (2011) Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol* 12:R18.
18. Ebbert MT, Wadsworth ME, Staley LA et al. (2016) Evaluating the necessity of PCR duplicate removal from next-generation sequencing data and a comparison of approaches. *BMC Bioinformatics* 17 Suppl 7:239.
19. Kebschull JM, Zador AM (2015) Sources of PCR-induced distortions in high-throughput sequencing data sets. *Nucleic Acids Res* 43:e143.
20. Bronner IF, Quail MA, Turner DJ et al. (2014) Improved Protocols for Illumina Sequencing. *Curr Protoc Hum Genet* 80:18.2.1-42.
21. Gibson JF, Kelso S, Skevington JH (2010) Band-cutting no more: A method for the isolation and purification of target PCR bands from multiplex PCR products using new technology. *Mol Phylogenet Evol* 56(3):1126-1128.
22. Quail MA, Gu Y, Swerdlow H et al. (2012) Evaluation and optimisation of preparative semi-automated electrophoresis systems for Illumina library preparation. *Electrophoresis* 33(23):3521-3528.
23. Hawkins TL, O'Connor-Morin T, Roy A et al. (1994) DNA purification and isolation using a solid-phase. *Nucleic Acids Res* 22(21):4543-4544.
24. Thermo Fisher Scientific, Inc. (2018) Qubit dsDNA assay specificity in the presence of single-stranded DNA. (Application note)
25. Buehler B, Hogrefe HH, Scott G et al. (2010) Rapid quantification of DNA libraries for next-generation sequencing. *Methods* 50(4):S15-8.

Learn more

Molecular cloning

PCR

qPCR

Electrophoresis

Related products

NGS library preparation

NGS library amplification

NGS library quantitation

Share

For Research Use Only. Not for use in diagnostic procedures.