Paired-End Sequencing
Sample Preparation Guide

FOR RESEARCH USE ONLY

Topics

3 Introduction
5 Best Practices
6 DNA Input Recommendations
7 Paired-End Sample Preparation Kit Contents
9 User-Supplied Consumables and Equipment
11 Fragment DNA
16 Perform End Repair
18 Adenylate 3' Ends
19 Ligate Adapters
21 Purify Ligation Products
24 Enrich DNA Fragments
26 Purify Final Product
28 Validate Library

Catalog # PE-930-1001
Part # 1005063 Rev. B
September 2009

illumina®

Introduction

This protocol explains how to prepare libraries of genomic DNA for paired-end sequencing analysis. The goal of this protocol is to add adapter sequences onto the ends of DNA fragments to generate the following sequencing library format:

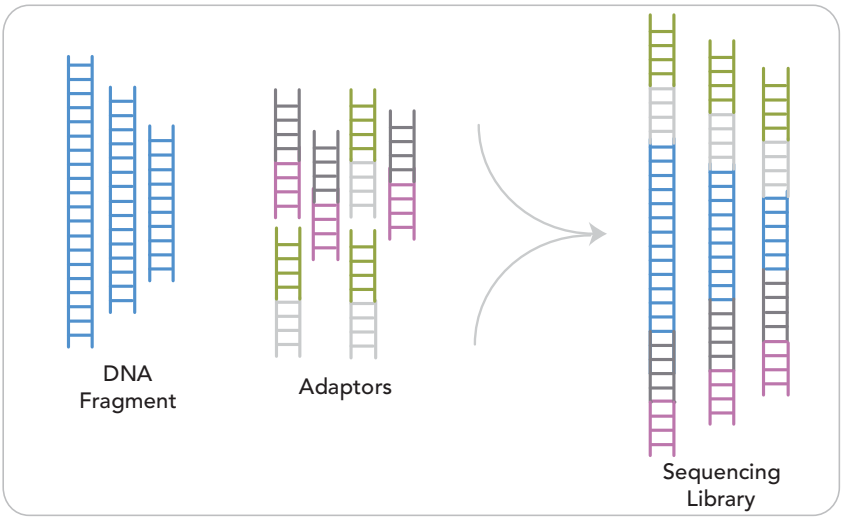

Figure 1 Sequencing Library after Sample Preparation

The two adapters contain sequences that are complimentary to the two surface-bound amplification primers on the flow cells.

Depending on the number of samples being prepared, you need one or more of the Illumina® kits in Table 1.

Table 1 Paired-End DNA Sample Preparation Kits

Sample Preparation Kit	Number of Samples	Catalog #
Paired-End Sample Prep Kit	10	PE-102-1001
Paired-End Sample Prep Kit	40	PE-102-1002

Sample Prep Workflow

The following figure illustrates the steps in the paired-end sample preparation protocol. The total time required is approximately 6 hours.

Input genomic DNA (1–5 µg) is fragmented by hydrodynamic shearing to generate <800 bp fragments. The fragments are blunt ended and phosphorylated, and a single 'A' nucleotide is added to the 3' ends of the fragments in preparation for ligation to an adapter that has a single-base 'T' overhang. Adapter ligation at both ends of the genomic DNA fragment confers different sequences at the 5' and 3' ends of each strand in the genomic fragment. The products of this ligation reaction are purified and size-selected by agarose gel electrophoresis. Size-selected DNA is PCR amplified to enrich for fragments that have adapters on both ends. The resulting sample library is again purified and size-selected by agarose gel electrophoresis. The final purified product is then quantitated prior to seeding clusters on a flow cell.

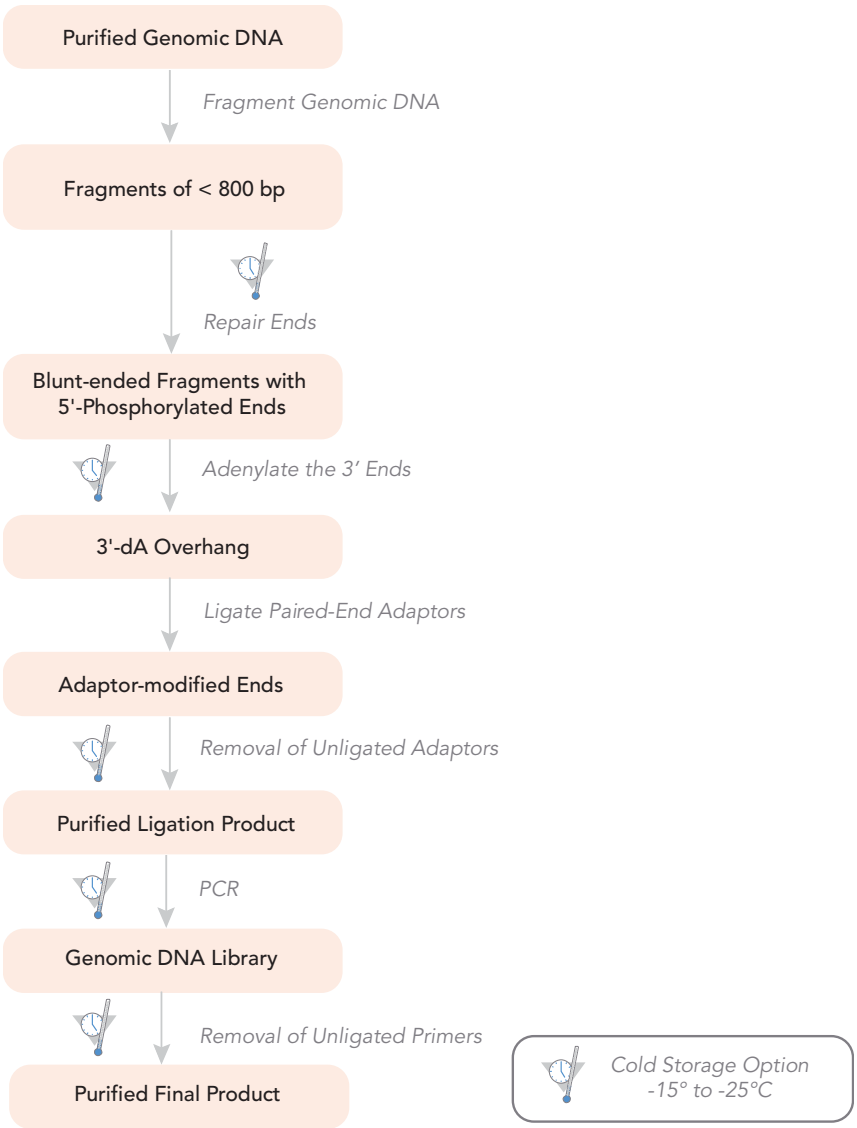

Figure 2 Sample Preparation Workflow

Best Practices

When preparing genomic DNA libraries for sequencing, you should always adhere to good molecular biology practices.

Liquid Handling

Good liquid handling measures are essential, particularly when quantifying libraries or diluting concentrated libraries for making clusters. Small differences in volumes (±0.5 µl) can sometimes give rise to very large differences in cluster numbers (~100,000). Small volume pipetting can also be a source of potential error in protocols that require generation of standard curves, such as PicoGreen assays or qPCR, or those that require small but precise volumes, such as the Agilent BioAnalyzer. If small volumes are unavoidable, then due diligence should be taken to ensure that pipettes are correctly calibrated. Also, care should be taken, because solutions of high molecular weight dsDNA can be viscous and not evenly dispersed, resulting in aliquot measurements that are not representative of the true concentration of the solution.

Potential DNA Contaminants

Undesirable library artifacts may result from incorrect DNA quantitation, including interference from superfluous nucleic acids in a sample (e.g., RNA, small nucleic acid fragments, nucleotides, single-stranded DNA), excess proteins, or other contaminating materials. DNA quality may also affect the quantity of usable DNA in a sample. For example, if the DNA is damaged (e.g., heavily nicked or containing extensive apurinic/apyrimidinic sites), then many of these fragments may fail during library preparation. High molecular weight dsDNA derived from host genomes can also confound accurate quantitation. For example, bacteria artificial chromosomes (BACs) and other bacterially-derived plasmids usually contain a few percent of the chromosomal DNA from the host cells, despite the best purification efforts. These sequences may ultimately give rise to unwanted clusters on a flow cell lane. However, this contamination can be accurately quantified by analyzing aligned reads generated during sequencing against known bacterial sequences and subtracting these out. High molecular weight contamination may also be estimated prior to library preparation using qPCR assays designed to target unique chromosomal markers.

Temperature Considerations

Temperature is another important consideration for making genomic DNA libraries. Elevated temperatures should be particularly avoided in the steps preceding the adapter ligation. DNA fragments that have a high AT content are more likely to denature into single strands than GC-rich fragments, which can result in an increased probability of creating a bias in the sequencing coverage. As a general rule, libraries should be kept at temperatures below 20°C. Temperature is less of an issue after the adapters have been ligated onto the ends of the DNA, although care should be taken not to denature the library prior to the agarose gel electrophoresis process, because single-stranded DNA runs at a different position.

DNA Input Recommendations

Input DNA Quantitation

The ultimate success or failure of a library preparation strongly depends on using the correct amount of input DNA, particularly when starting quantities are <1 µg. Thus, correct quantitation of genomic DNA is essential.

Assessing DNA Quantity and Quality

Absorbance measurements at 260 nm are commonly used to quantify DNA. The ratio of absorbance at 260 nm to absorbance at 280 nm is used as an indication of sample purity, and values of 1.8–2.0 are considered indicative of relatively pure DNA. However, both measurements can be compromised by the presence of RNA or small nucleic acid fragments such as nucleotides. Thus, genomic DNA samples should be carefully collected to ensure that they are free of contaminants, and the most accurate spectrophotometric method available should be used to quantify the input genomic DNA.

DNA quantitation methods that rely on intercalating fluorescent dyes measure only double-stranded DNA and are less subject to excess nucleic acids. However, these methods require the preparation of calibration curves and are highly sensitive to pipetting error. Ensure that pipettes are correctly calibrated and are not used at the volume extremes of their performance specifications.

Gel electrophoresis is a powerful means for revealing the composition of low molecular weight fragmented DNA samples smaller than 10,000 bp. Protein contamination can cause the DNA to smear, RNA is clearly visible at the bottom of the gel, and DNA can be visualized at its expected size. However, intact genomic DNA is generally of high molecular weight and is therefore difficult or impossible to analyze by electrophoresis.

Consistency of Results

Gel electrophoresis is a powerful means for revealing the condition (even the presence or absence) of DNA in a sample. Impurities, such as detergents or proteins, can be revealed by smearing of DNA bands. RNA, which interferes with 260 nm readings, is often visible at the bottom of a gel. A ladder or smear below a band of interest may indicate nicking or other damage to DNA. Where possible, or necessary, a gel should be run to assess the condition of the DNA sample.

Paired-End Sample Preparation Kit Contents

Check that you have all of the reagents identified in this section before proceeding to sample preparation.

Kit Contents, Box 1

Store at -15° to -25°C

This box is shipped on dry ice. As soon as you receive it, store the components at -15° to -25°C.

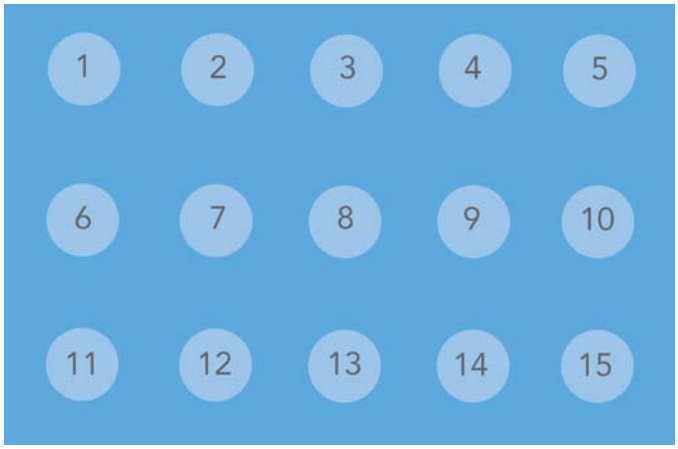

Figure 3 *Paired-End Sample Prep Kit, Box 1*, part # 1001809

1. T4 DNA Ligase Buffer with 10 mM ATP (100 µl), part # 1000534
2. Klenow Enzyme (10 µl), part # 1000515
3. Klenow Buffer (100 µl), part # 1000535
4. DNA Ligase Buffer 2X (250 µl), part # 1000523
5. Phusion™ DNA Polymerase (Finnzymes Oy) (250 µl), part # 1000524
6. 10 mM dNTP Mix (20 µl), part # 1001932
7. T4 PNK (50 µl), part # 1000519
8. 1 mM dATP (100 µl), part # 1000520
9. PE Adapter Oligo Mix (100 µl), part # 1001782
10. PCR Primer PE 1.0 (10 µl), part # 1001783
11. T4 DNA Polymerase (50 µl), part # 1000514
12. Empty
13. Klenow Exo - (30 µl), part # 1000536
14. DNA Ligase (50 µl), part # 1000522
15. PCR Primer PE 2.0 (10 µl), part # 1001784

Kit Contents, Box 2 Store at Room Temperature

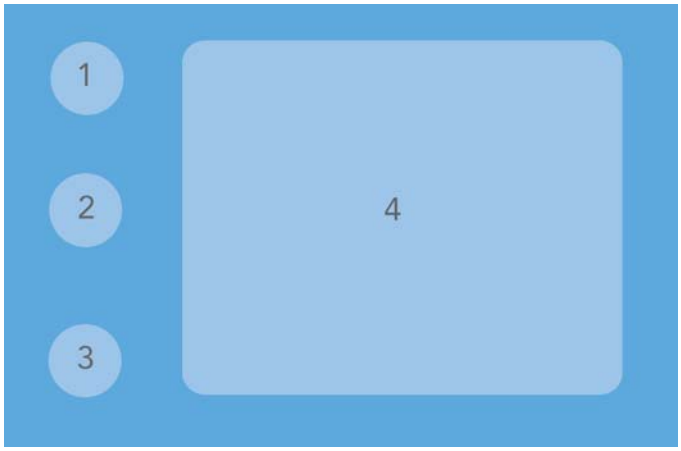

Figure 4 Paired-End Sample Prep Kit, Box 2, part # 1000182

1. Nebulization Buffer (7 ml), part # 1000466
2. TE Buffer (10 ml), part # 1000465
3. Ultra Pure Water (10 ml), part # 1000467
4. Nebulizer Kit (10 each), part # 1000541

User-Supplied Consumables and Equipment

Check to ensure that you have all of the necessary user-supplied consumables and equipment before proceeding to sample preparation.

Table 2 *User-Supplied Consumables*

Consumable	Supplier
50X TAE Buffer	BIO-RAD, part # 161-0743
100 bp Redi-Load DNA ladder	Invitrogen, part # 10488-058
Certified low range ultra agarose	BIO-RAD, part # 161-3106
Compressed air source of at least 32 psi (Do not use CO_2 which could alter the pH of the nebulizer buffer)	General lab supplier
Disposable scalpels	General lab supplier
Distilled water	General lab supplier
Loading buffer (50 mM Tris pH 8.0, 40 mM EDTA, 40% (w/v) sucrose)	General lab supplier
MinElute PCR Purification Kit	QIAGEN, part # 28004
Purified DNA (1–5 µg, 5 µg recommended) DNA should be as intact as possible, with an OD260/280 ratio of 1.8–2.0	General lab supplier
PVC tubing	Intersurgical, part # 1174-003
QIAquick Gel Extraction Kit (for 50 samples) or QIAquick Gel Extraction Kit (for 250 samples)	QIAGEN, part # 28704 (50 samples) QIAGEN, part # 28706 (250 samples)
QIAquick PCR Purification Kit	QIAGEN, part # 28104
SYBR Green	Invitrogen, part # S7563

Table 3 *Equipment Checklist*

Consumable	Supplier
Benchtop microcentrifuge	General lab supplier
Benchtop centrifuge with swing-out rotor (e.g., Sorvall Legend RT)	General lab supplier
Dark reader transilluminator or a UV transilluminator	Clare Chemical Research, part # D195M
Electrophoresis unit	General lab supplier
Gel trays and tank	General lab supplier
Thermal cycler	General lab supplier

Fragment DNA

DNA is fragmented using a nebulization technique, which breaks up DNA into pieces less than 800 bp in minutes using a disposable device. This process generates double-stranded DNA fragments containing 3' or 5' overhangs.

A compressed air source is used, to force a solution of DNA through a narrow orifice. The solution follows a cyclical route from a collection chamber via a siphon tube back to the sample reservoir and through the orifice again. After 5–6 minutes of repeated passing through the orifice, the sample is fragmented to its final size. Nebulization is a very reproducible process and produces random fragments that approximate a Poisson distribution when sequenced. Alternative fragmentation methods may be more appropriate based on sample availability (see *Alternative Fragmentation Methods*).

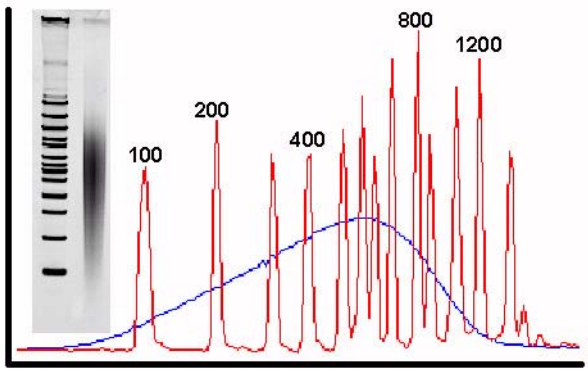

Figure 1 *Fragment Genomic DNA*

Consumables

Illumina-Supplied

- Nebulizers (box of 10 nebulizers and vinyl accessory tubes)
- Nebulization Buffer (7 ml)
- TE Buffer

User-Supplied

- QIAquick PCR Purification Kit (QIAGEN, part # 28104)
- Purified DNA (1–5 µg, 5 µg recommended)
  DNA should be as intact as possible, with an OD260/280 ratio of 1.8–2.0
- Compressed Air of at least 32 psi
  Do not use CO_2 which could alter the pH of the nebulizer buffer
- PVC tubing
  Dimensions: 1/4 inch ID, 3/8 inch OD, 1/16 inch wall, 1 meter length

CAUTION

If you intend to nebulize DNA that could possibly contain any pathogenic sequences such as pathogenic viral DNA, perform the nebulization process under containment conditions (e.g., a biosafety cabinet) to prevent exposure to aerosols.

Procedure

The DNA sample to be processed should be highly pure, having an OD260/280 ratio of between 1.8 and 2.0, and should be as intact as possible.

NOTE

If you are not familiar with this shearing method, Illumina recommends that you test this procedure on test samples and practice assembling the nebulizer before proceeding with your sample DNA.

1. Remove a nebulizer from the plastic packaging and unscrew the blue lid.

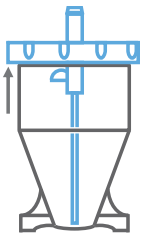

Figure 2 Remove the Nebulizer Lid

2. Using gloves, remove a piece of vinyl tubing from the packaging and slip it over the central atomizer tube. Push it all the way to the inner surface of the blue lid.

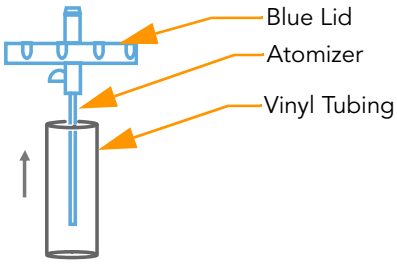

Figure 3 Assemble the Nebulizer

3. Add 1–5 µg of purified DNA in a total volume of 50 µl of TE buffer to the nebulizer.
4. Add 700 µl nebulization buffer to the DNA and mix well.
5. Screw the lid back on (finger-tight).

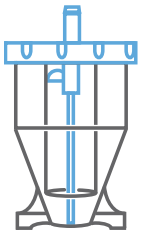

Figure 4 Replace the Nebulizer Lid

6. Chill the nebulizer containing the DNA solution on ice while performing the next step.

7. Connect the compressed air source to the inlet port on the top of the nebulizer with the PVC tubing, ensuring a tight fit. To do so, apply even pressure onto the circumference of the neck of the nebulizer to ensure that a small gap is not created which could enable air to leak.

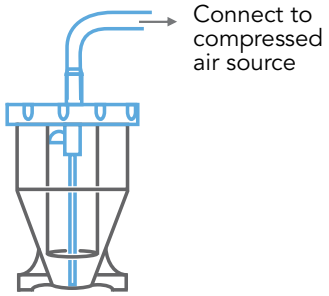

Figure 5 Connect Compressed Air

8. Bury the nebulizer in an ice bucket and place it in a fume hood.
9. Use the regulator on the compressed air source to ensure the air is delivered at 32–35 psi.
10. Nebulize for 6 minutes. You may notice vapor rising from the nebulizer; this is normal. Also, the nebulization buffer may turn white or appear frozen.
11. Centrifuge the nebulizer at 450 xg for 2 minutes to collect the droplets from the side of the nebulizer. If necessary, use an old nebulizer as a counter-balance.
12. If a centrifuge is not available, then use 2 ml of the binding buffer (PB or PBI buffer) from the QIAquick PCR Purification Kit to rinse the sides of the nebulizer and collect the DNA solution at the base of the nebulizer.
13. Measure the recovered volume. Typically, you should recover 400–600 µl.
14. Follow the instructions in the QIAquick PCR Purification Kit to purify the sample solution and concentrate it on one QIAquick column, eluting in 30 µl of QIAGEN EB.
15. Perform a quality control step on the recovered DNA to ensure the presence of sufficient DNA (i.e., quantify the DNA by a 260 nm reading, or by a PicoGreen assay, bioanalyzer, or equivalent). Ensure that there is at least 0.5 µg of DNA. If not, then fragment more DNA.

If you do not plan to proceed to *Perform End Repair* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Alternative Fragmentation Methods

Genomic DNA sample preparation must generate dsDNA fragments of a narrow distribution around a desired median size. The fragmentation method should not introduce bias in the fragment sequence, should be able to recover ~100% of the DNA, and should be automatable, inexpensive, and simple to use. Illumina recommends nebulization for most cases because it meets many of these criteria. Alternative DNA fragmentation methods are outlined below:

Illumina does not recommend these methods, but if you choose to use them, use them with caution.

Covaris

Covaris has developed an instrument that adopts ultrasound technology to fragment nucleic acids in a controlled and precise manner. DNA can be sheared in a small volume (50 µl) without sample loss due to evaporation and without significant bias in the resulting fragments.

The Covaris instrument can achieve fragments with a smaller end-point size and a narrower range than with a nebulizer (Figure 6). However, this is only a benefit if fragment sizes of ~100 bp are desired. With the read lengths currently available using Illumina's genetic analyzers, ~100 bp fragments are not typically desired for paired-end reads. For example, a median fragment size of 300 bp (having a standard deviation of 30 bp or 10% variance) is required in the library generation protocol for paired 100 base reads. This would result in 99% of the fragments being >210 bp. Likewise, a median fragment size of 450 bp is needed for paired 150 bp reads.

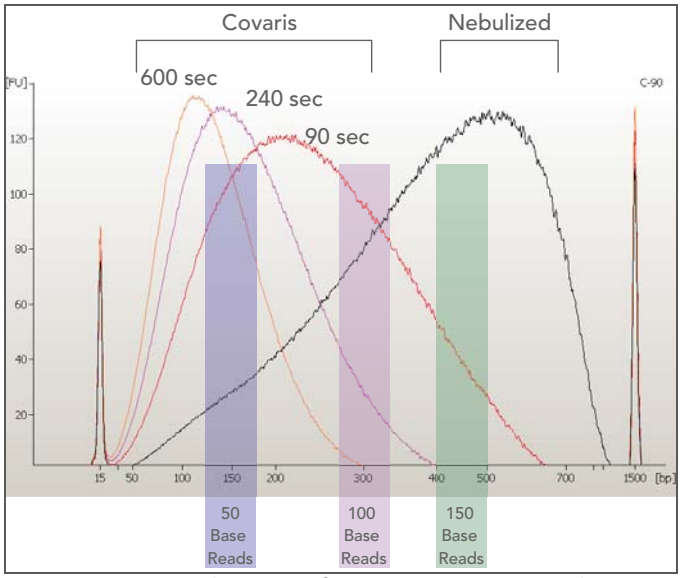

Figure 6 Distribution Of Fragment Sizes Resulting From Nebulization Or Covaris

When the Covaris instrument is used to shear DNA to a median fragment size of ≥200 bp and greater, the distribution of fragment sizes broadens and is comparable to that seen when using a nebulizer.

Sonication

Sonication is sometimes used as an alternative to nebulization, since sonicators are relatively common in molecular biology labs. Care must be taken to avoid raising the temperature of the DNA solution above ~20°C. Heating small fragments with high AT content may result in denaturation. If this occurs, single stranded fragments will not ligate to the double stranded adaptors and these will be under-represented in the final library. During sonication, DNA fragments are typically generated in a decreasing but broad range of fragment sizes. Therefore, sonication settings must be chosen carefully to reproducibly generate a majority of fragments in the desired size range.

A probe-based sonicator wherein the probe tip is directly immersed into the DNA sample may be used, although problems with sample heating, vaporization, and contamination may arise. Alternatively, a cup horn sonicator filled with ice and water will allow the tube containing the DNA solution to remain sealed during fragmentation.

Hydroshear

Hydroshearing fragments DNA by a similar mechanism to nebulization. A syringe pump is used to pass the DNA solution back and forth through a narrow orifice. As the liquid streams through the orifice, it becomes laminar (i.e., the center of the stream moves faster than the stream closer to the walls of the orifice) and DNA molecules are stretched and snapped at the midpoint of their length. Hydroshearing is best for generating fragments with a two-fold size range of 1,000-40,000 bp. However, it is not effective for generating fragments <1,000 bp in length, and so is not suitable for generating standard Illumina single read and short-insert paired-end sequencing libraries.

Perform End Repair

DNA fragmentation by physical methods produces heterogeneous ends, comprising a mixture of 3’ overhangs, 5’ overhangs, and blunt ends. The overhangs will be of varying lengths and ends may or may not be phosphorylated. This step converts the overhangs resulting from fragmentation into blunt ends using T4 DNA polymerase and Klenow enzyme. The 3' to 5' exonuclease activity of these enzymes removes 3' overhangs and the polymerase activity fills in the 5' overhangs. In addition, the T4 PNK in this reaction phosphorylates the 5' ends of the DNA fragments.

To minimize pipetting errors, especially with the small volume enzyme additions, prepare the reagents for multiple samples simultaneously. As a result, you pipette once from the reagent tubes with a larger volume, rather than many times with 1 µl volumes. Prepare a master mix of enzymes, water, buffer, etc. and aliquot this in a single pipetting movement to individual samples to standardize across multiple samples.

Consumables

Illumina-Supplied

- T4 DNA Ligase Buffer with 10 mM ATP
- 10 mM dNTP Mix
- T4 DNA Polymerase
- Klenow Enzyme
- T4 PNK
- Water

User-Supplied

- QIAquick PCR Purification Kit (QIAGEN, part # 28104)

Procedure

1. Prepare the following reaction mix on ice, in the following order:
   - Water (45 µl)
   - DNA sample (30 µl)
   - T4 DNA ligase buffer with 10 mM ATP (10 µl)
   - 10 mM dNTP mix (4 µl)
   - T4 DNA polymerase (5 µl)
   - Klenow enzyme (1 µl)
   - T4 PNK (5 µl)

   The total volume should be 100 µl.
2. Mix gently, but thoroughly, and centrifuge briefly.
3. Incubate in a thermal cycler for 30 minutes at 20°C.

4. Follow the instructions in the QIAquick PCR Purification Kit to purify on one QIAquick column, eluting in 32 µl of QIAGEN EB.

If you do not plan to proceed to *Adenylate 3' Ends* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Adenylate 3' Ends

A single 'A' nucleotide is added to the 3' ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. A corresponding single 'T' nucleotide on the 3' end of the adapter provides a complementary overhang for ligating the adapter to the fragment. This strategy ensures a low rate of chimera (concatenated template) formation.

To minimize pipetting errors, prepare the reagents for multiple samples simultaneously. As a result, you pipette once from the reagent tubes with a larger volume, rather than many times with small volumes. Prepare a master mix of enzymes, water, buffer, etc. and aliquot this in a single pipetting movement to individual samples to standardize across multiple samples.

Consumables

Illumina-Supplied

- Klenow Buffer
- 1 mM dATP
- Klenow Exo -

User-Supplied

- MinElute PCR Purification Kit

This process requires a QIAquick MinElute column rather than a normal QIAquick column.

Procedure

1. Prepare the following reaction mix on ice, in the following order:
   - DNA sample (32 µl)
   - Klenow buffer (5 µl)
   - 1 mM dATP (10 µl)
   - Klenow exo - (3 µl)

   The total volume should be 50 µl.
2. Incubate in a thermal cycler for 30 minutes at 37°C.
3. Follow the instructions in the MinElute PCR Purification Kit to purify on one QIAquick MinElute column, eluting in 10 µl of QIAGEN EB.

Ligate Adapters

This process ligates adapters to the ends of the DNA fragments, preparing to generate sequencing libraries complimentary to the amplification primers on the flow cell.

For adapter ligation to blunt ended DNA, this adapter adds distinct sequences to the 5′ and 3′ ends of each strand in the genomic fragment. Additional sequences are added by tailed primers during PCR, which are necessary for amplification on the flow cell during cluster formation.

To minimize pipetting errors, prepare the reagents for multiple samples simultaneously. Pipette a large volume once from the reagent tubes, rather pipetting small volumes multiple times. Prepare a master mix of enzymes, water, buffer, etc. and aliquot this in a single pipetting movement to individual sample in order to standardize across multiple samples.

Consumables

Illumina-Supplied

- DNA Ligase Buffer, 2X
- PE Adapter Oligo Mix
- DNA Ligase

User-Supplied

- QIAquick PCR Purification Kit (QIAGEN, part # 28104)

Procedure

This procedure uses a 10:1 molar ratio of adapter to genomic DNA insert, based on a starting quantity of 5 µg of DNA before fragmentation. If you started with less than 5 µg, reduce the volume of adapter accordingly to maintain the 10:1 ratio of DNA.

High adapter ratios can lead to the generation of adapter-dimers, whereas low adapter ratios can result in incomplete ligation.

1. Prepare the following reaction mix on ice, in the following order:
   - Water (6 µl)
   - DNA sample (10 µl)
   - DNA ligase buffer, 2X (25 µl)
   - PE adapter oligo mix (4 µl)
   - DNA ligase (5 µl)

   The total volume should be 50 µl.
2. Incubate for 15 minutes at 20°C.

3. Follow the instructions in the QIAquick PCR Purification Kit to purify on one QIAquick MinElute column, eluting in 10 µl of QIAGEN EB.

If you do not plan to proceed to *Purify Ligation Products* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Purify Ligation Products

This process purifies the products of the ligation reaction on a gel and removes unligated adapters, removes any adapters that may have ligated to one another, and selects a size-range of sequencing library appropriate for cluster generation.

Size Selection

Illumina recommends gel electrophoresis and band excision after adapter ligation to remove excess adapter and adapter dimers and to tighten the range of fragment sizes. Ligation reaction products are separated on an agarose gel and a ~2 mm wide gel slice containing DNA of the desired size is excised.

Excision Range

Illumina recommends a 200 bp insert size target (+/- 1 standard deviation of 20 bp, i.e, a 10% variance) for read lengths of 2×75 bp or shorter. In practice, this translates to a 2 mm gel slice at ~300 bp to account for the length of the adapter sequences flanking the inserts. For read lengths of 2×100 bp or longer, an insert size range of 300 bp or greater should be targeted for excision from the gel, unless users intentionally want to sequence overlapping read pairs.

Consumables

User-Supplied

- Certified Low Range Ultra Agarose
- 50x TAE Buffer
- Distilled Water
- SYBR Green
- Loading Buffer
- 100 bp Redi-Load DNA Ladder
- QIAquick Gel Extraction Kit (QIAGEN, part # 28704 or 28706)

Procedure

Illumina does not recommend purifying multiple samples on a single gel due to the risk of cross-contamination between libraries.

It is important to perform this procedure exactly as described, to ensure reproducibility.

It is important to excise as narrow a band as possible from the gel during gel purification. Paired-end libraries should consist of templates of the same size or nearly the same size, and as narrow a size range as possible. Illumina recommends that a Dark Reader is used to visualize DNA on agarose gels.

1. Prepare a 40 ml, 2% agarose gel with distilled water and TAE according to the manufacturer's instructions. The final concentration of TAE should be 1X.
2. Cast the gel in a tray to accommodate 56 µl.
   Recommended well size: 1 mm (length) x 8 mm (width) x 7 mm (height).
3. Add 3 µl of loading buffer to 8 µl of the 100 bp Redi-Load ladder.
4. Add 3 µl of loading buffer to 10 µl of the DNA from the purified ligation reaction.
5. Dilute 1 µl of SYBR Green 1:100 with water, and add 1 µl of this dilution to each sample and to the ladder that was prepared in step 3.
6. Load 8 µl of the ladder onto one lane of the gel.
7. Load the entire sample onto another lane of the gel, leaving a gap of at least one empty lane between ladder and sample.

Flanking the library on both sides with ladders may make the library excision easier.

8. Run gel at 120 V for 60 minutes.

The voltage and run time can vary with different electrophoresis equipment.

9. View the gel on a Dark Reader transilluminator or a UV transilluminator. To minimize DNA damage, limit the exposure time to UV radiation and use a long wavelength UV lamp, if possible.
10. Place a clean scalpel vertically above the sample in the gel at the desired size of the template.
11. Excise a 2 mm slice of the sample lane at approximately 400 bp using the markers as a guide.

Cutting a band of 400 bp will result in an insert size of about 300 bp. Adapters add about 80 bp to each fragment. The sequencing read length should be considered when cutting fragment sizes. Sequencing reads in to the adapter will cause chimeric reads, unalignable to the reference sequence.

Using a separate scalpel, you can also slice a 600 bp fragment out of the gel as a back up library and process this separately alongside the 400 bp library well. You then get two libraries from one sample preparation.

12. Follow the instructions in the QIAquick Gel Extraction Kit to purify on one QIAquick column, eluting in 30 µl of QIAGEN EB.

13. Discard the scalpel to avoid sample cross-contamination.

If you do not plan to proceed to *Enrich DNA Fragments* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Enrich DNA Fragments

This process uses PCR to selectively enrich those DNA fragments that have adapter molecules on both ends, and to amplify the amount of DNA in the library for accurate quantification. The PCR is performed with two primers that anneal to the ends of the adapters. The number of PCR cycles is minimized to avoid skewing the representation of the library.

PCR Amplification

The PCR amplification step of the protocol performs four key functions:

- **Add Sequences** - Additional sequences are added to the ends of the adapters so that the final amplified templates contain sequences to enable hybridization with primers bound to the flow cell surface for cluster generation.
- **Enrich Fragments** - PCR enriches for fragments that have adapters ligated on both ends. Fragments with only one or no adapters on their ends are by-products of inefficiencies in the ligation reaction. Neither species can be used to make clusters, as fragments without any adapters cannot hybridize to surface-bound primers in the flow cell, and fragments with an adapter on only one end can hybridize to surface bound primers but cannot form clusters. However, the presence of these incomplete ligation products can lead to an underestimation of the final quantity of library DNA.
- **Enrich Templates** - PCR enriches for templates that include the non-template 'A' nucleotide added during the A-tailing step of the protocol and therefore eliminates adapter dimers. This is accomplished by using proprietary, modified primers that are completely resistant to the 3'–5' exonuclease activity of the Phusion polymerase used for PCR. These primers reach all the way to the non-templated 'A' before the start of the genomic sequence. If the 'A' is not present (as in the case for an adapter dimer), then the terminal 'T' on the primer will mismatch and not extend.
- **Provides Material** - PCR provides enough material to enable reliable quantitation of the final library.

Consumables

Illumina-Supplied

- Phusion DNA Polymerase
- PCR Primer PE 1.0
- PCR Primer PE 2.0
- Ultra Pure Water

User-Supplied

- QIAquick PCR Purification Kit (QIAGEN, part # 28104)

Procedure

This process assumes 5 µg of DNA input into library preparation. If you are using 0.5 µg, adjust the process as described in the following table.

Input of DNA to Library Prep	Volume of Purified Library into PCR	Volume of Water	Number of PCR Cycles
5 µg	1 µl	22 µl	10
0.5 µg	10 µl	13 µl	12

CAUTION

To avoid sample cross-contamination, set up PCR reactions (all components except the template DNA) in a designated clean area, preferably a PCR hood with UV sterilization and positive air flow.

1. Prepare the following reaction mix on ice, in the following order. The volume of each mix should be 50 µl:
   - PCR primer PE 2.0 (1 µl)
   - Ultra pure water (22 µl)
   - DNA (1 µl)
   - Phusion DNA polymerase (25 µl)
   - PCR primer PE 1.0 (1 µl)

   The total volume should be 50 µl.
2. Mix gently, but thoroughly, and centrifuge briefly.
3. Aliquot the master mix out to the desired number of PCR tubes.
4. Transport the PCR tubes, on ice, to the bench.
5. Add the template DNA.
6. Amplify using the following PCR process:
   a. 30 seconds at 98°C
   b. 10 or 12 cycles of:
      40 seconds at 98°C
      30 seconds at 65°C
      30 seconds at 72°C
   c. 5 minutes at 72°C
   d. Hold at 4°C
7. Follow the instructions in the QIAquick PCR Purification Kit to purify on one QIAquick MinElute column, eluting in 30 µl of QIAGEN EB.

If you do not plan to proceed to *Purify Final Product* immediately, the protocol can be safely stopped here. If you are stopping, store the samples at -15° to -25°C overnight or longer. When proceeding, thaw the samples on ice.

Purify Final Product

This process further purifies the products of the PCR ligation reaction on a gel to select a size-range of templates appropriate for subsequent cluster generation.

Consumables

User-Supplied

- Certified Low Range Ultra Agarose
- 50x TAE Buffer
- Distilled Water
- SYBR Green
- Loading Buffer (50 mM Tris pH 8.0, 40 mM EDTA, 40% (w/v) sucrose)
- 100 bp Redi-Load DNA ladder
- QIAquick Gel Extraction Kit (QIAGEN, part # 28704 or 28706)

Procedure

Illumina does not recommend purifying multiple samples on a single gel due to the risk of cross-contamination between libraries.

It is important to perform this procedure exactly as described, to ensure reproducibility.

It is important to excise as narrow a band as possible from the gel during gel purification. Paired-end libraries should consist of templates of the same size or nearly the same size, and as narrow a size range as possible.Illumina recommends that a Dark Reader is used to visualize DNA on agarose gels.

1. Prepare a 40 ml, 2% agarose gel with distilled water and TAE. The final concentration of TAE should be 1X.
2. Cast the gel in a tray that is approximately 8 cm in length.
3. Add 3 µl of loading buffer to 8 µl of the ladder.
4. Add 3 µl of loading buffer to 10 µl of the DNA from the purified PCR reaction.
5. Dilute 1 µl of SYBR Green 1:100 with water, and add 1 µl of this dilution to each sample and the ladder prepared in step 3.
6. Load 8 µl of the 100 bp Redi-Load ladder onto one lane of the gel.

7. Load the entire sample onto another lane of the gel, leaving a gap of at least one empty lane between ladder and sample.

Flanking the library on both sides with ladders may make the library excision easier.

8. Run gel at 80 V for 2 hours at 4°C.

The voltage and run time can vary with different electrophoresis equipment.

9. View the gel on a Dark Reader transilluminator or a UV transilluminator.
10. Place a clean scalpel vertically above the sample in the gel at the desired size of the template.
11. Excise a 2 mm slice of the sample lane at approximately 400 bp using the markers as a guide.

Cutting a band of 400 bp will result in an insert size of about 300 bp. Adapters add about 80 bp to each fragment. The sequencing read length should be considered when cutting fragment sizes. Sequencing reads in to the adapter will cause chimeric reads, unalignable to the reference sequence.

Using a separate scalpel, you can also slice a 600 bp fragment out of the gel as a back up library and process this separately alongside the 400 bp library well. You then get two libraries from one sample preparation.

12. Follow the instructions in the QIAquick Gel Extraction Kit to purify on one QIAquick column, eluting in 15 µl of QIAGEN EB.
13. Discard the scalpel.

Validate Library

Illumina recommends performing the following quality control steps on your DNA library.

1. Check the template size distribution by running an aliquot of the library on a gel or an Agilent Bioanalyzer.
2. If validating by gel, load 10% of the volume of the library on a gel and check that the size range is as expected: a narrow smear similar in size to the DNA excised from the gel after the ligation.

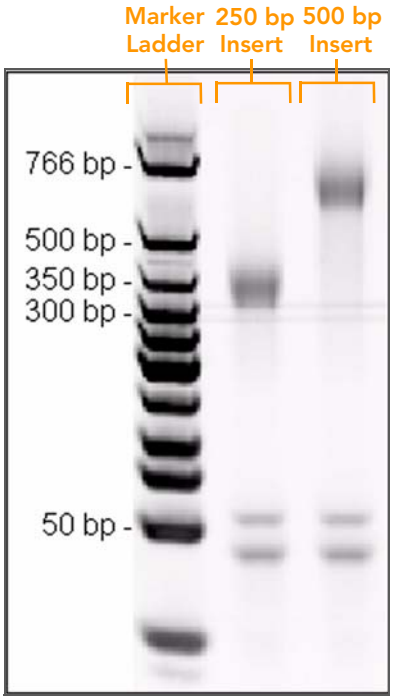

Figure 7 Library Validation Gel

This example shows a library run on a 4–20% TBE polyacrylamide gel, stained with Vistra Green (GE Healthcare # RPN5786) and visualized on a fluorescence scanner. Lane 1 shows a marker ladder. Lane 2 and Lane 3 show paired-end libraries with insert sizes of approximately 250 bp and 550 bp, respectively. The two bands around 50 bp in size are primers from the enrichment PCR step and have no effect on the subsequent formation of clusters.

3. If using Agilent Bioanalyzer, load 10% of the volume of the library on Agilent chip.

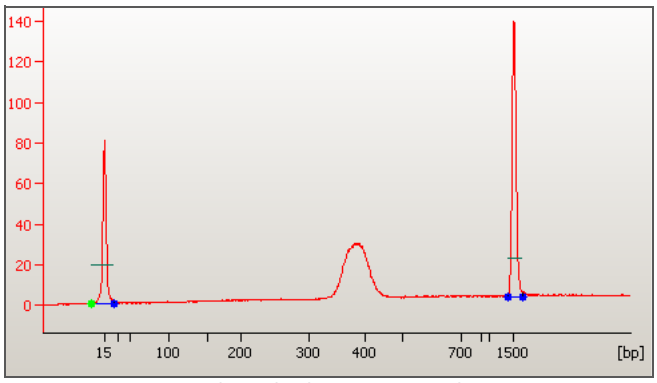

Figure 8 Paired-End Library Example

An example of a library prepared using this protocol is shown in Figure 8. The library has insert sizes ranging from 250–350 bp (i.e., a template size range of 350–450 bp).

NOTE

Cutting a band of 350–450 bp will result in an insert size of about 250–350 bp. Adapters add about 80 bp to each fragment. Sequencing read length should be considered when cutting fragment sizes. Sequencing reads in to the adapter will cause chimeric reads, unalignable to the reference sequence.

NOTE

The sequenced Paired-End library reports the average insert size and distribution metrics in the Summary.htm output in the GERALD sub-directory of the run folder.

Library Quantitation

In order to achieve the highest quality of data on Illumina sequencing platforms, it is important to create optimum cluster densities across every lane of every flow cell. This requires accurate quantitation of DNA library templates. As described for sample input quantitation, any method of measuring DNA concentration has certain advantages and potential drawbacks.

By Optical Density or Intercalating Dye

Absorbance-based quantitation can erroneously measure the presence of residual PCR primers or may not be sensitive enough to make an accurate measurement in circumstances where the final library yield is low. Measurements based on intercalating dyes such as PicoGreen may also be used, but care should be taken to minimize pipetting errors and to ensure that the library is not denatured to avoid inaccuracies.

Bioanalyzer

Automated analysis using an instrument like the Agilent Bioanalzyer can be used to quantify libraries and is an excellent alternative to traditional gel electrophoresis to determine the size of the library fragments. As with intercalating dye assays, single-stranded templates will not be detected nor contribute to the library quantitation even though these templates will form clusters. Therefore, care must be taken to ensure that the library is not denatured and that pipetting errors are minimized.

qPCR

qPCR is an alternative method of quantifying DNA that measures the relationship between the initial concentration of a template and how its concentration changes during progressing cycles of thermal amplification. As in conventional PCR, a polymerase, dNTPs, and two primers are designed to match sequences within a template. In SYBR Green-based qPCR, the amplicon quantity is measured by binding of an intercalating dye. In contrast, TaqMan-based qPCR employs a fluorescent probe that hybridizes to amplicons downstream of one of the amplification primers. During extension, the polymerase encounters the hybridized probe and removes the fluorescent moiety that then fluoresces in solution. For qPCR library quantitation, the amplification primers and TaqMan probes are designed to match sequences within the Illumina adapters. Thus, qPCR offers advantages for measuring library concentration because it will only measure templates that have adapter sequences on both ends. Moreover, qPCR sensitivity allows accurate quantitation of dilute libraries with concentrations below the threshold of detection of conventional spectrophotometric methods.

qPCR quantitates DNA of unknown concentration by comparison to a control sample of known concentration. The control sample should be a library that gives a known number of clusters when titrated on a flow cell. Although any library prepared for sequencing on the Illumina platform can be used as a control for qPCR, the control library should be as similar as possible in terms of template size, GC content, and library type (e.g., genomic DNA or ChIP-Seq) to the experimental libraries for quantitation. Control libraries can be easily prepared using the Illumina Paired-End Library Generation kit.

To correlate library concentration with cluster number, it is necessary to generate a titration flow cell for the control template. This can be achieved by preparing serial dilutions of 'template hyb' from the control library and counting the number of clusters following sequencing. A minimum of two sequencing cycles is required to achieve an accurate cluster count because Illumina's Pipeline software uses the first two cycles to identify individual clusters in a full length sequencing run (Figure 9). A cluster titration curve for the control template should be linear to the point at which the clusters become too dense to count accurately. Once the control library has been correlated with cluster number, this library can be used as a qPCR control.

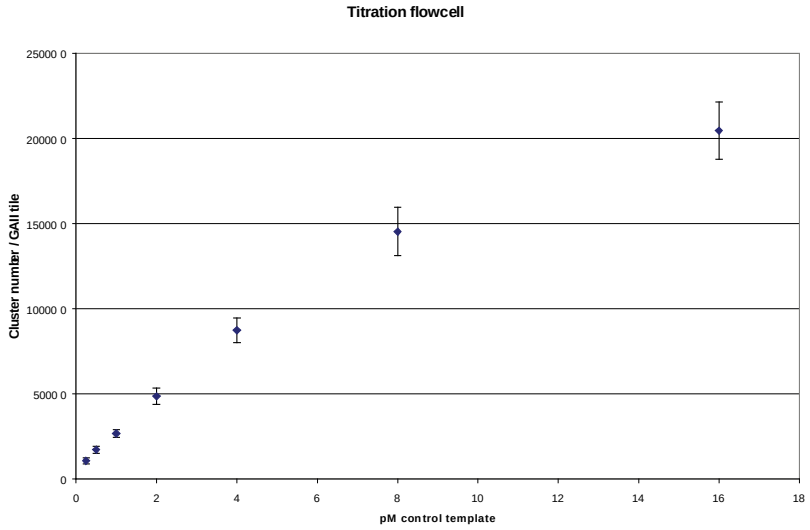

Figure 9 Titration Of A Control 300 bp Insert E. Coli Library

Figure 10 shows an example of the cluster numbers achieved with 21 independent *E. coli* libraries that were quantified by qPCR against a control library. All libraries, including the control library, were then hybridized to lanes of a flow cell at 10 pM concentration. All libraries gave similar cluster numbers.

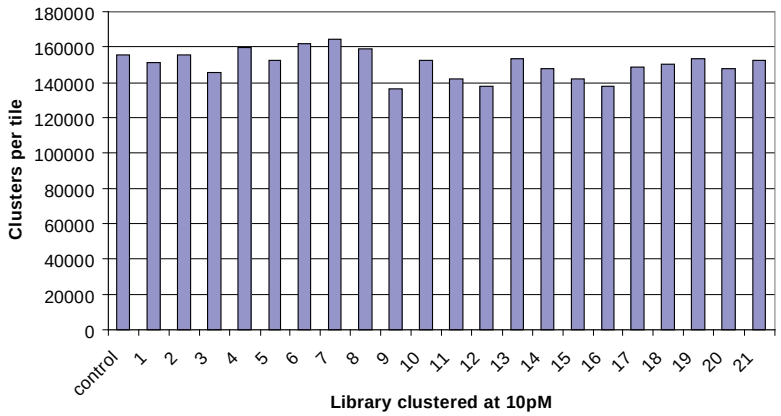

Figure 10 Example of Cluster Numbers Achieved

For more information on using qPCR to quantify libraries, reference the Illumina *qPCR Quantification Protocol Guide*.

Illumina, Inc.
9885 Towne Centre Drive
San Diego, CA 92121-1975
+1.800.809.ILMN (4566)
+1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com

illumina®