Protocols for Metagenomic Library Generation: Truseq

NGS Core Facility at Argonne National Laboratory
Stephanie M. Greenwald & Sarah M. Owens

Materials

Kits, Reagents, and Consumables

1. Qubit HS Assay, for Invitrogen's Qubit Fluorometer. Catalog number: Q32851 or Q32854.
2. Agencourt Ampure XP Beads. Catalog Number A63880..
3. microTube AFA Fiber Pre-Slit with Snap-Cap 6X16mm, for Covaris S-series. Catalog Number: 5520045
4. Kapa Biosystems Library Amplification kits catalog number: KK2701
5. Wafergen's PrepX Complete ILMN DNA library Kits catalog number: 400075
6. Absolute Ethanol
7. Blue Pippin Prep DNA Gel Cassettes (3%, 2%, 1.5%, or 0.75%) catalog numbers: BDF3010, BDF3010, BDF2010, BDF1510
8. Bio-O Scientific's Nextflex DNA Barcodes and PCR mixture catalog numbers: 5140-08 and 520999
9. Agilent DNA HS kit catalog number: 5067-4626

Equipment

1. Invitrogen's Qubit Fluorometer
2. Covaris S-series system
3. Wafergen's Apollo 324 system
4. Magnetic Stand or Rack (holds 1.5ml or 96 well plates)
5. Thermocycler
6. Sage Science's Blue Pippin Prep
7. Agilent 2100 Bioanalyer

Methods

Quantification for Library Preparation

We start with 500ng of high-quality DNA for TruSeq metagenomic library prep, although lower quality and concentrations may be used. However, this may lead to lower quality sequencing data and more observed sequencing bias.

We use the following protocol:

1. Make a Qubit working solution by diluting the Qubit DNA reagent 1:200 in Qubit DNA buffer using a sterile plastic tube.
2. Load 190ul of Qubit working solution into tubes labeled standard 1 and 2.
3. Add 10ul of standard 1 solution and standard 2 solution to the appropriate tube and mix by vortexing for 2-3 seconds.
    - Note: These are positive and negative controls used to calibrate the instrument.
4. Load 198ul of Qubit working solution into each individual assay tube.
5. Add 2ul of DNA to each assay tube and mix by vortexing 2-3 seconds. The final volume of this solution should equal 200ul.
    - **Note**: The amount of sample and working solution added to each assay tube can vary depending on concentration of the sample. The sample can vary between 1-20ul, and the working solution can vary between 199-180ul with the final volume equaling 200ul. It is recommended to use 2ul of sample to produce the most accurate results.
6. Allow all the tubes to incubate at room temperature for 2 minutes.
7. Select DNA assay on the Qubit Fluorometer. Select run a new calibration.
8. Insert the tube containing Standard 1, close lid and press read.
9. Remove standard 1 and repeat step 8 for standard 2.
10. Insert sample, close lid and press read.
11. Calculate concentration using dilution calculation on Qubit Fluorometer by selecting original volume of sample added to the assay tube.
12. Repeat step 10 and 11 until all samples have been quantified.

TruSeq Metagenomic Library Prep

Shearing of Libraries

For TruSeq libraries we use the Covaris S-series system for mechanical shearing. We use a minimum of 500ng sample in 50ul-100ul. If there is not at least 500ng in 100ul, we use Agencourt Ampure XP Beads to concentrate the sample down to a smaller volume using a

1.8X bead ratio. The conditions set on the Covaris are directly related to the preferred insert size of the final library, we use the following table to determine what conditions to use for shearing.

Target Base Pair (Peak)	150	200	300	400	500	800	1000	1500
Duty Cycle	10%	10%	10%	10%	5%	5%	5%	2%
Intensity	5	5	4	4	3	3	3	4
Cycles per Burst	200	200	200	200	200	200	200	200
Time (seconds)	430	180	80	55	80	50	40	15

We use the following protocol:

1. After determining an appropriate insert size for the libraries use the above table to program the Covaris with the size specific operating conditions.
2. Bring the water bath to 6-8°C while letting the Covaris de-gas for at least 30 minutes
3. Aliquot 500ng of each sample in 50ul-100ul into separate Covaris microTubes.
4. Insert each Tube into the Covaris adaptor on the head of the instrument, submerge the fully enclosed samples into the head of the instrument into the water-bath.
5. Hit the start button on the control software in order to initiate the given operating conditions, and let the program run its duration.
6. Take out the microTube and set it aside on ice
7. Repeat steps 4-5 for all samples

End repair, A-tailing, and Adaptor Ligation

1. Take the product from the shearing protocol and bring all samples to 15ul with Agencourt Ampure XP Beads at a 1.8X ratio.
    a. Transfer the samples into separate microcentrifuge tubes or plate wells from the covaris microTubes

b. Vortex the AMPure XP Beads until they are resuspended at room temperature.
c. Add 1.8Xul of the mixed AMPure XP Beads to each sample. Mix thoroughly by gently pipetting the entire volume 10 times.
d. Incubate at room temperature for 15 minutes. During incubation, prepare an 80% ethanol solution.
e. Place the tubes or plate on the magnetic stand at room temperature for at least 5 minutes, until the liquid appears clear, and the beads are magnetically separated.
f. Remove and discard the supernatant from each tube. Do not disturb the beads.
g. With the samples still on the magnetic stand, add 200ul of freshly prepared 80% ethanol to each sample without disturbing the beads.
h. Incubate at room temperature for at least 30 seconds while still on the magnetic stand, then remove and discard all of the supernatant from each tube. Again, do not disturb the beads.
i. Repeat steps 6 and 7 one more time for a total of two 80% ethanol washes.
j. Allow the samples to air dry on the magnetic stand at room temperature for 15 minutes or until the beads no longer appear wet.
k. Add 15ul of nuclease-free water to each tube.
l. Thoroughly resuspend the beads by gently pipetting 10 times.
m. Incubate the tubes at room temperature for 2 minutes.
n. Place the tubes back onto the magnetic stand at room temperature for at least 5 minutes, until the liquid appears clear.
o. Transfer the clear supernatant from each tube to an appropriate collection tube. Leave at least 1ul of the supernatant behind to avoid carryover of magnetic beads.

2. Prepare the indexes that have been assigned to each sample.
    a. Dilute 2ul of index from a 25um stock with 13ul of nuclease-free water in a new PCR-strip tube.
3. Prepare the ligation mix add the following to a new PCR-strip tube:
    a. 12ul of T4 Ligase Buffer
    b. 1ul of T4 Ligase Enzyme

    c. 2ul of nuclease-free water
4. Prepare 70% Ethanol
5. Chose the ILM program that is closest to the chosen insert size for the given samples. The user can choose from 220bp, 320bp, 520bp, and 870bp insert sizes.
6. Load consumables, reagents, and samples onto the Apollo 324 following the deck layout given on the Apollo's screen
7. Let the Apollo's program run its course. The Apollo will run samples through an end-repair reaction, a-tail ligation reaction, adaptor-ligation reaction, and size-selection reaction.
8. The samples are libraries at this point – aliquot them from the Apollo labeled product tubes into a new plate or micocentrifuge tubes.

3.4.5 PCR and Size Selection

PCR and further size selection is not always necessary. For some applications the wide size distribution generated during library prep is sufficient. If the libraries are at least 2nM concentration then PCR is unnecessary. If size selection is unnecessary start this protocol at step 9. If PCR is necessary we use using KAPPA's PCR mixture. We use 10-15 cycles of PCR to achieve at least 2nM concentration. If a tighter size distribution is necessary, we further size selection with the Blue Pippin Prep (Sage Science, Inc., Beverly, MA). We use the following protocol:

1. Choose the appropriate cassette to the given insert size and library size
    a. 3% cassette ranges from 90-200bp
    b. 2% cassette ranges from 100-600bp
    c. 1.5% cassette ranges from 250bp-1.5kb
    d. 0.75% cassette ranges from 1-50kb
2. Program the Pippin
    a. In the Blue Pippin software go to the Protocol Editor tab
    b. Click on the Cassette folder that matches the appropriate cassette for the given library size
    c. Select either "range" or "tight" and enter in the given base pair range or peak
    d. Click the "use internal standards" button
3. Calibrate the Optics

a. Place the calibration fixture in the optical nest, close the lid, and hit "calibrate."
b. Continue only if it passes, if it does not pass, try again.

4. Load the Cassette
    a. Inspect the cassette from bubbles, breakage of agarose column, and equal buffer levels.
    b. Dislodge any bubbles from the elution chamber
    c. Place the cassette into the optics nest
    d. Fill the sample well to the top with buffer
    e. Remove any buffer from the elution well and fill it with 40ul of fresh buffer.
    f. Place a seal over the elution wells to keep them from overflowing during the run
    g. Run a continuity test and continue only if it passes. Try again if it fails.
5. Mix the library and dye
    a. Mix at least 30ul of library with 10ul of dye. If there is less then 30ul use nuclease-free water to dilute the libraries to 30ul.
    b. Vortex the libraries and dye well and spin the mixture down
6. Load the samples
    a. Remove 40ul of buffer from the sample well and replace it with the 40ul mixture of sample and dye.
    b. Repeat for each sample
    c. Close the lid and hit the "start" button
7. The Blue Pippin will run for 30-56 minutes depending on the given program
8. Open the lid, remove the samples from the elution wells and place into a collection tube
9. Check the concentration of the samples with a DNA HS assay on the Qubit Fluorometer
    a. Refer to the Qubit Protocol mentioned into the Quantification section
10. Use the Qubit concentration and estimated library size (100bp + insert size) to calculate the molarity of the sample with the following equation (with X= ng/ul concentration and Y= estimated size of fragment in bp): Molarity in nM = $[X/1*10^{-6}]/[Y*660]$

11. If the estimated molarity is less the 2nM then proceed to PCR in step 12. If it is 2nM or higher proceed to *final library quantification*
12. PCR using Bio-O Scientific's Nextflex DNA Barcodes and PCR mixture
    a. Mix 7.5ul of the library, 29.5ul of nuclease-free water, 12ul of NEXTflexTM PCR master mix, and 2ul NEXTflex Primer Mix in a well of a PCR strip tube or plate.
    b. Set a pipette to 50ul and mix by pipetting up and down 10 times
    c. PCR on a thermocycler under the following settings
        i. 2 minutes at 98°C
        ii. 10-15 cycles of: 30 seconds at 98°C, 30 seconds at 65°C, 60 seconds at 72°C
        iii. 4 minutes at 72°C
    d. Add 44ul of AMPure XP Beads
    e. Incubate at room temperature for 15 minutes. During incubation, prepare an 80% ethanol solution.
    f. Place the tubes or plate on the magnetic stand at room temperature for at least 5 minutes, until the liquid appears clear.
    g. Remove and discard the supernatant from each tube. Do not disturb the beads.
    h. With the samples still on the magnetic stand, add 200ul of freshly prepared 80% ethanol to each sample, without disturbing the beads.
    i. Incubate at room temperature for at least 30 seconds while still on the magnetic stand, then remove and discard all of the supernatant from each tube. Again, do not disturb the beads.
    j. Repeat steps 6 and 7 one more time for a total of two 80% ethanol washes.
    k. Allow the tubes to air dry on the magnetic stand at room temperature for 15 minutes or until the beads no longer appear wet.
    l. Add 15ul of nuclease-free water to each tube.
    m. Thoroughly resuspend the beads by gently pipetting 10 times.
    n. Incubate the tubes at room temperature for 2 minutes.
    o. Place the tubes back onto the magnetic stand at room temperature for at least 5 minutes, until the liquid appears clear.

p. Transfer the clear supernatant from each tube to an appropriate collection tube. Leave at least 1ul of the supernatant behind to avoid carryover of magnetic beads.

13. Proceed to *Final Library Quantification*

Final Library Quantification

It is very important to assess the quality of a library before it is sequenced. The molarity and library size are critical for down-stream steps. Figure 2 shows examples of completed library types. Illumina recommends that completed libraries achieve a molarity of at least 2nM or greater in order to be sequenced with quality results on an Illumina Sequencer. It is important to remove any primer-dimers that may be present. Primer-dimers will be visible on a Bioanalyzer electropherogram between bases 0-100 depending on the length of PCR primers you are using. If primer-dimers are present use a 1X ratio of AMPure XP Beads to remove them. To assess the quality of the completed library, we recommend the following protocol:

1. Use the Qubit Fluorometer to determine the concentration of libraries in ng/ul
    a. Reference the Qubit Protocol in the Quantify section
2. Use the Agilent 2100 Bioanalyzer to determine the library insert size and length
    a. Check to make sure that the Chip Priming Station has its base adjusted to position C and the syringe clip is adjusted to the bottom-most position.
    b. The Electrodes on the head of the Bioanalyzer should be cleaned with nuclease-free water.
    c. To prepare the gel for a DNA HS Chip, all reagents should be left at room temperature for at least 30 minutes to equilibrate.
    d. 15ul of Agilent DNA HS dye concentrate should be added to a vial of High Sensitivity gel matrix and vortexed for at least 10 seconds.

e. The entire dye-concentrate and gel-matrix should be transferred into a spin filter and spun for 10 minutes at 2240 g +/-20%. **Note:** this mixture should be used within 30 days.
f. One DNA HS chip should be placed on the priming station.
g. 9.0ul of gel-dye mix should be pipetted into the well-marked "G." The plunger should be positioned at 1 ml. The chip priming station should be closed and the plunger should be depressed until the clip holds it. The plunger should remain pressurized for 60 seconds then released. If the plunger does not move back up to the 1 ml position after it is released the plunger needs to be manually moved to the 1 ml position.
h. 9ul of gel-dye mix should be pipetted into the remaining three gel-dye wells.
i. 5ul of DNA HS marker should be pipetted into all of the sample wells and the ladder well.
j. 1ul of each DNA sample should be added into the 11 sample wells.
k. 1ul of DNA HS ladder should be added to the ladder well
l. The chip should be vortexed with an IKA vortex mixer for 1 minute at 2400 rpm.
m. The chip should be immediately loaded into the Bioanalyzer, the corresponding chip type selected in the software, and "Start run" selected.
n. To check the quality of the run, 13 markers should be seen on the electropherogram of the ladder at 35, 50, 100, 150, 200, 300, 400, 500, 600, 700, 1000, 2000, and 3000 base pairs. **Note**: If these peaks are not seen the run has failed, and the data is not accurate.
o. The Bioanalyzer will produce DNA concentrations in pg/ul, the base pair peak, a molarity estimate in pmol/l, and the base pair range that the library covers

3. Use the concentration from the Qubit and peak base pair size generated by the Bioanalyzer to calculate the molarity of the sample with the equation provided above.

4. The libraries are considered complete and ready for Illumina sequencing if the molarity is 2nM or greater.

Sequencing and Data Analysis

The output from an Illumina Next Generation sequencing run is ultimately a fastq.qz file. Sequences with low quality scores are filtered out before analysis. Metagenomes will be analyzed using the available online resources (e.g. IMG/M, MGRAST, CAMERA, EBIs Metagenomics portal, etc.) providing annotation by comparing transcripts to different functional gene databases (e.g. using BLAST to assign functions against M5NR, SFams and SEED). For more detailed descriptions of potential functional pipelines and analyses of these data see Thomas, 2012 **(8)** and Myer et al. 2008 ***(9).***

References

1.Lorenz P and Eck J. (2005). Metagenomics Industrial Applications. *Nature Reviews Microbiology* **3**, 510-516. doi:10.1038/nrmicro1161

2. Segata, N., Boernigen D., Tickle TL., Morgan XC., Garrett WS, Huttenhower C. (2013) Computational Meta'omics for Microbial Community Studies. Molecular Systems Biology, doi: 10.1038/msb.2013.22

3. Hampton-Marcell JT., Moormann SM., Owens SM., Gilbert JA. (2013) Preparation and Metatranscriptomic Analyses of Host-Microbe Systems in Microbial Metatgenomics, Metatranscrtomics, and Metaproteomics volume 531 issue 169. Academic Press. ISBN 978-0-12-407863-5. ISSN 0076-6879.

4. Mason, O., Hazen T., Borglin, S., Chain, P., Dubinsky, E., Fortney, J., et al. (2012). Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. The ISME journal, 6(9), 1715-1727

5. Rubin BER., Gibbons SM., Kennedy S., Hampton-Marcell J., Owens S., Gilbert J. (2013) Investigating the Impact of Storage Conditions on Microbial Community Composition in Soil Samples**. PLOSone.** DOI: 10.1371/journal.pone.0070460

6. Wang LY., Ke WJ., Sun XB., Liu JF., Gu JD., Mu BZ. (2013). Comparison of bacterial community in aqueous and oil phases of water-flooded petroleum reservoirs using pyrosequencing and clone library approaches. Appl Microbiol Biotechnol. DOI 10.1007/s00253-013-5472-y

7. Korenblum E., Souza DB., Penna M, Seldin L. (2013). Molecular Analysis of the Bacterial Communities in Crude Oil Samples from Two Brazilian Offshore Petroleum Platforms. International Journal of Microbiology Volume 2012. doi:10.1155/2012/156537

8. Thomas, T., Gilbert, J., Meyer, F. (2012). Metagenomics – a guide from sampling to data analysis. Microbial Informatics and Experimentation, 2, doi:10.1186/2042-5783-2-3

9. Myer F., Paarmann D., D'Souza., Olson R., Glass EM., Kubal M., Paczian T., Rodriguez A., Stevens R., Wilke A., Wilkening J., Edwards RA. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 2008, **9**:386 doi:10.1186/1471-2105-9-386