**Sequencing Workflow Overview**

**Self-Service Platform**

* Check availability on the Outlook Calendar: ‘CVMBS-MIP NGS Access’
* Contact Marylee at [marylee.layton@colostate.edu](mailto:marylee.layton@colostate.edu) to reserve a time slot
  + You must provide an account # and the PI name in order to reserve a time slot
* Make sure you have building access. Contact Marylee if you do not.
* Load runs during normal work hours: M-F 8-5.
  + Contact Marylee via email or phone at 970-657-1507 during normal work hours
* If you run into an issue with the sequencer, contact Marylee, she will communicate with Illumina for troubleshooting.
* Instrument Operation:
  + You are responsible for loading the sequencer
  + Marylee will perform all post-run and maintenance washes
* Reagents Supplied:
  + 1N NaOH
  + 20pM Denatured and Diluted (D+D) PhiX
* Reagents you need to supply:
  + Qubit reagents (Thermo)
  + Tapestation reagents (Agilent)
  + Kapa Library Quantification Kit (Roche: KK4824)
  + Sequencing Kit (see CSU PriceList from Illumina for discounted prices)
  + Consumables:
    - 1.5ml DNA lo-bind tubes
    - 1.5ml tubes
    - Pipet tips
    - Gloves
    - Thermocycler plates and film
* Register for a BaseSpace account here: login.illumina.com
  + You will need a basespace account in order to run the instrument.

**Pooling**

The goal is to make a multiplexed library with equal amounts of each library into one final pool. You can do this 2 ways:

1. Mass based: pool an equal *mass* of each sample for even sequencing coverage
   1. Determine concentration (ng/µl) of each sample using Qubit
   2. Target the same number of ng for each sample.
   3. Determine volume of each library to add.

Example data:

|  |  |  |  |
| --- | --- | --- | --- |
| Sample name | Concentration (ng/µl) | Target ng | Volume to pool |
| 1 | 5.2 | 10 | 1.92 |
| 2 | 10 | 10 | 1 |
| 3 | 20 | 10 | 0.5 |
| 4 | 2 | 10 | 5 |
| 5 | 7 | 10 | 1.42 |

Note: Pipetting volumes smaller than 2µl is not accurate. Determine a mass that will allow you to pipet a minimum of 2µl for each sample.

1. Molarity based: pool an equal *molarity* of each sample for even sequencing coverage.
   1. Determine concentration (ng/µl) of each sample using Qubit
   2. Dilute samples to be 1ng/µl or less.
   3. Determine average size and molarity using Agilent High Senstivity D1000 tape.
   4. Set a region to determine estimated molarity.
   5. Use the following equation to determine what volume to add of each sample:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sample Name |  |  |  |  | Volume to pool |
| Lung 1 | 10 nM | 2 | 4 nM | 20 µl | 4 µl |
| Lung 2 | 15 nM | 2 | 4 nM | 20 µl | 2.67 µl |

* Add water to 20µl so the final pool is 4nM in 20µl.

**Notes on Pooling:**

* Benefits of mass based pooling:
  + Quick and easy to do
  + Cheaper, no tapestation required
* Drawbacks of mass based pooling:
  + Pooling may not be even because you don’t know the average size of products and whether there are primer dimers, etc. that will skew the pooling
* Benefits of Molarity based pooling:
  + Accurate pooling
* Drawbacks of molarity based pooling:
  + More expensive because you have to run a tape
  + More time/math involved.

**QC**

The goal of QC is to determine the exact molarity of your final library before loading onto the sequencer. If your estimation is off, you will overload or underload the run. This can result in run failure or suboptimal run metrics. ***This is the trickiest part of the entire process, so take your time and make sure you are confident in the estimated final molarity before loading!***

1. Determine concentration (ng/µl) using Qubit.
2. Dilute samples to ~1ng/µl. By doing this, the library is within the quantification range of the tapestation and should be ~4nM. For example: A typical amplicon library that is ~500bp is 4nM when diluted to 1ng/µl.
3. Determine average size of the library using the Agilent D1000 High Sensitivity tape.Chart

   Description automatically generated with medium confidence

**Figure 1: This is what a good amplicon library looks like. There aren't any adapter dimers and there is one clean peak.**

Diagram

Description automatically generated

**Figure 2:How many adapter dimers are okay? The rule of thumb is <10% of the total library.**

A picture containing diagram

Description automatically generated

Figure : This is a good shotgun library. There is a nice broad size range.

A picture containing diagram

Description automatically generated

Figure : This is not ideal! The library is small and there are a lot of adapter dimers.

Diagram

Description automatically generated

Figure : This is an mRNA prep. The broad size distribution may make it difficult to get an accurate cluster density on the flow cell.

Diagram, engineering drawing

Description automatically generated

Figure : This is a good mRNA prep with a nice size distribution and no adapter dimers.

1. Perform Library Quantification using Kapa Library Quantification Kit.

* Important Documents:
  + kapa\_library\_quantification\_technical\_guide.pdf
  + ngs\_libqc\_template.doc
  + ngs\_libqc\_calculator.xls
  + library\_concentration\_calculator.xltx
* Kit: Kapa Library Quantification Kit
* You can use the ngs QuantStudio3
* It’s a good idea to use a run-to-run control to give confidence that the standards worked well. We use a 1:1000 dilution of 4nM PhiX. You can use whatever works for you, maybe a successful library that you are confident in the loading concentration.

**SIGNS THERE IS AN ISSUE WITH YOUR LIBRARY:**

1. The average size is not what you expected
2. The calculated molarity by qPCR is not what you expected: mainly, it is extremely low, i.e. less than 0.5nM
3. If a library is 0.5nM or less, I would not load it.
4. It is NOT unusual for the qPCR to be different than the estimated molarity as determined by the tapestation.

**Loading the MiSeq**

* Follow Illumina’s Denature and Dilute protocol: miseq-denature-dilute-libraries-guide-15039740-10.pdf
* For a quick protocol use: load\_sequencer\_template
* Prepare a sample sheet. This is how the instrument knows how many cycles for R1, R2 and indexes to do.
  + If you have single indexes use this sample sheet template: single\_index\_samplesheet\_template.csv
  + If you have dual indexes use this sample sheet template: dual\_index\_samplesheet\_template.csv
  + Do not use these symbols: (.,-) or any spaces!
  + Only open the sample sheet in excel!
  + Make sure you have the correct # of reads and # of bases in indexes.
* Make sure you have your BaseSpace username and password!

NOTES:

**Data Transfer**

* BaseSpace:
  + Use your account to monitor run performance
  + You can download fastq files from basespace
  + You have 1TB of storage, so this is limited storage.
* ngsdata01 server
  + Send your sample sheet to Marylee: [marylee.layton@colostate.edu](mailto:marylee.layton@colostate.edu)
  + Marylee will email you with directory information when the data is demultiplexed.
* Data Security:
  + Runs automatically go to your basespace, the ngsdata01 server, and a drobo.