Deep Sequencing Core Labs
Guidelines for the Preparation and Submission of Samples

Your sample (i.e. your "library") must be sized, quantified, and proofed. These are described below. You will need a Sample Ticket and Sample Information Sheet for each group of samples that will be run together. Each sample or sample mix is treated as an individual project for cluster formation, sequencing, and data retrieval. Please assist us by providing a ticket for each sample mix when it is submitted.

Uniform Size:

- The clusters are formed by seeding a lawn of library fragments onto the flow cell, then attaching and amplifying each sequence by bridge amplification. For optimal cluster formation, library fragments should be close to the same size (within 50bp).
- DNA libraries should have a median fragment size under 1Kb (500bp ± 50bp is recommended); Chromatin IP libraries are often in the 300bp range, and small RNA libraries are usually under 100bp. The Core cannot guarantee performance for libraries with fragments over 800bp.
- For cluster formation, we need to know the median size (not the mean). Please put the size information on the Service Ticket.

Quantification:

- Optimal seeding of the clusters also depends upon delivering the correct amount of your library to the flow cell. In order to do this, we need an accurate quantification of the material in your library. Standard spectrophotometric measurement of material after it is gel-purified (sized) is recommended.
- If you need the Fragment Analyzer results to get an accurate concentration of your library, please provide us beforehand with your best estimate of concentration, so that we can choose the correct assay or sample dilution needed to get accurate results.
- The final material should be delivered to the core at 10-20nM in 20-25 µl of sterile water or EB. If you aren't able to do this, please let us know, and provide us with the concentration in ng/µl or nM and the median fragment size.

Proofing:

- If you are not using a commercial kit or an already-established library prep method, this step is critical for two reasons. (1) demonstrates that your library has fragments which can be used to create clusters and (2) indicates whether your library contains fragments with inserts from your target genome. By cloning an aliquot of your library and sequencing 20 clones, you will know if the linkers for attachment are present, you will confirm which sequencing primer should be used for cluster formation, and you will know if your library has any inserts from the genome of interest or if the bulk of the clones contain primer-dimers or other unwanted material.
- Please remember that only a few unwanted clones here will represent potentially millions of sequences after the analysis. Please note on your Service Ticket what number of clones you sequenced and how many had attachment linkers (if only half had linkers, then we need to adjust the amount of material used to form clusters to compensate for this).

  Attachment sequences/linkers:
  P5: AATGATACGGCGACCACCGA
  P7: CAAGCAGAAGACGGCATACGA

- For adapters/sequencing primers, PLEASE indicate which one to use for your sample (**ex: Illumina TruSeq, TruSeq small RNA, Nextera, NEB, Takara/Clontech, Other Library Kit, or Custom adapters**). All adapters must contain the attachment sequences and priming site(s). The Core must know which sequencing primer to use, as not all primers are in the standard mixes for Illumina instruments. In the case of a custom sequencing primer, please discuss it with the Core before submitting, as varying Tm and sequence may make the primer incompatible with some Illumina systems.

Data:

- Each run generates a large number of image files. These are transferred to the High Performance Computing Cluster where the data analysis pipeline reduces the images to sequences. Demultiplexing (sorting by index) is also done in the analysis pipeline. At the time of sample submission, please note the following items on the Service Ticket:
    1) Globus Account: Name of person with account, and email assigned to the account. This person should have a folder in their account entitled **DeepSEQ** into which the data from the pipeline will be deposited.
    2) Whether you want md5sum values (to confirm all data has been transferred intact)
    3) Email contacts to arrange the data transfer and to notify when data is ready. If you do not have access to a Globus account, the Data Transfer Contact must arrange for an alternative. Portable drives of any sort are NOT feasible due to size, speed, and security limitations.
- DeepSEQ users must transfer data to their own storage within 5 days of being notified that the data is ready. The Core Lab cannot archive data for customers.

Sample Drop-Off:

- There are locations for sample drop-off in the ASC, the NERB, and the LRB (see our website for current specifics). Bags and labels are available near the drop-off fridges. Tickets should be turned in with your samples. Please do NOT leave samples in the “freezer” portion of the dropoff fridges as that section is not temperature-stable. If your sample must remain frozen, please package it appropriately in a container with ice.
- Additionally, the Core is located in the Reed-Rose-Gordon building on the Maple Ave campus in Shrewsbury. Samples can be dropped off there if you are not on the Worcester campus. Please let us know when you are coming over, and contact Daniella Wilmot or Ellie Kittler for sample login.

Please check
http://www.umassmed.edu/nemo/sample-submission
for the most up-to-date information on sample submission procedures.