UF | ICBR
Gene Expression & Genotyping

DNA Sequencing Frequently Asked Questions

Q: Which ICBR core conducts DNAseq projects?

A: Both the NextGen DNA Sequencing core (projects with <48 samples) and Gene Expression & Genotyping core (projects with ≥48 samples) can process DNAseq projects. The larger sample sizes (≥48) allow for the Gene Expression & Genotyping core to utilize a miniaturized approach that significantly reduces costs for customers. This FAQ will outline the process for projects submitted through the Gene Expression & Genotyping core. DNA-seq projects should be submitted using the ICBR iLab portal, and once libraries have been constructed, the pooled library will be transferred to the NextGen DNA Sequencing core for sequencing. No separate request in iLab is necessary if sequencing is requested along with library construction. To start the DNAseq project, please contact ICBR-Geneexpression@ad.ufl.edu or yanp@ufl.edu for more details and for planning your experiment.

If you need data analysis, please fill out the data analysis form. You can also directly the ICBR Bioinformatics core (ICBR-Bioinformatics@ad.ufl.edu).

Q: What is the Illumina DNA-seq workflow using NEB reagents?

1. **Start with genomic DNA.** The "NEBNext Ultra II FS DNA Library Prep Kit for Illumina" (New England Biolabs) doesn't give a DNA quality (DIN) recommendation. For DNA input, the kit recommends 100pg – 500ng of purified, genomic DNA.
2. **DNA fragmentation.** High quality, genomic DNA needs to be fragmented to provide usable data from short-read Illumina platforms. The "NEBNext Ultra II FS DNA Library Prep Kit for Illumina" (New England Biolabs) uses an enzymatic fragmentation step that also incorporates end repair to prepare for ligation of the Illumina adaptors. We have found that a minimal fragmentation time allows us to accommodate a range in genomic DNA quality as submitted by customers.
3. **Ligation of Illumina adaptors.** The NEBNext adaptors have a hairpin loop that is then cleaved by USER enzyme, following ligation. When you get to the final amplification step with barcoded primers, you'll need to use primers from the "NEBNext Multiplex Oligos for Illumina" (96 Unique Dual Index Primer Pairs, Set 2, Set 3, Set 4, and Set 5).
4. **Library QC.** TapeStation to estimate average library size and Qubit or qPCR for library quantification. With mass and size information, calculate molarity. In order to provide cost savings for customers, the only project-wide DNAseq library QC done by the Gene Expression & Genotyping core is measuring the DNA concentration with Qubit. A couple of the lowest and highest concentration libraries are run on Tapestation to verify the libraries look as expected. Libraries are pooled equimolar using the same average size (averaged from the check of lowest and highest concentration libraries) for all libraries. The pool is submitted for an initial, shallow sequencing run. Based on the sequencing output from the shallow run, libraries are repooled for a second run to provide the remaining, customer requested reads per sample.
5. Depending on your familiarity with the library construction protocol, qPCR quantification may be skipped. However, we strongly suggest quantifying the final library pool using qPCR (the Gene Expression core can do it for $18.59 / sample or pool). qPCR measures the "functional" molecules in the library (i.e., molecules that contain full p5 and p7 adaptors).

6. Normalize concentrations and pool equimolarly.
7. The Gene Expression & Genotyping core will pass the pooled library to the NextGen DNA Sequencing core for sequencing. The number of reads targeted per sample will vary depending on genome size and desired coverage. One lane of the 10B flowcell on the NovaSeq X produces ~1250 million reads, so the number of samples (barcoded!!) that can be run on 1 lane = 1250 / # million reads per sample. ICBR also offers the 25B flowcell for the NovaSeq X for larger projects, generating ~3125 million reads per lane.
8. For pre-constructed DNAseq libraries (i.e., libraries that have NOT been made in the Gene Expression & Genotyping core), please contact the NextGen DNA Sequencing core for requested volumes and concentrations of your pool and submit directly to the NextGen DNA Sequencing core.

Q: How much DNA is needed for DNA-Seq library prep?

Library complexity is very important in a DNAseq experiment. If the starting DNA is limited, the coverage of the genome is likely to suffer. The standard protocol for library construction requires between 100pg and 500ng of genomic DNA. If possible, 250 to 500ng of genomic DNA is preferred for sample QC and library prep. The miniaturized approach used by the Gene Expression & Genotyping core only allows for DNA input volumes of 5.2µl per sample. It is critical to have **high DNA concentrations** to maximize DNA input for the DNAseq library construction process.

Q: What are the DNA quality Requirements?

A: The kit makes no recommendations for genomic DNA quality. But high molecular weight DNA (with little/no visible degradation) will provide the best chance for success.

Q: Will the GE Core send QC results before they start my library prep?

A: Yes, the GE core will email QC results to customers. QC results include genomic DNA quality (DIN values as estimated by Tapestation Genomic DNA assay) and quantity (DNA concentration as estimated by Qubit). We will highlight potential issues with quality and quantity of customers' samples as they relate to the library prep and give customers some possible options to make a decision on how to move forward with the submitted samples.

Q: How do customers submit replacement samples?

A: If after initial QC, customers want to submit additional samples to the core, please upload the sample info of replaced sample(s) to iLab, and clearly indicate which samples replaced which samples. Because the charge for library construction includes the initial sample QC, the QC cost for additional samples will be added to the project at a cost of $11.51 per sample.

Q: Can customers get back the samples?

A: Yes. Please make a note in the comment box in iLab submission form that you want the samples back upon the completion of the project. We keep customer's samples up to 6 months. It is customers' responsibility to retrieve the samples on time.

Q: How long does the core keep customer's leftover samples for?

A: Due to limited space in the GE core freezers, we only keep customer's samples for 6 months. If a customer would like samples returned to them, they should contact the core to schedule a day and time to come and pick them up.

Q: What is the turnaround time for an DNA-seq project?

A: Turnaround time depends largely on the volume of activity at the time of sample submission and the complexity of the project. Given typical workloads, library construction requires approximately 2-3 weeks, and sequencing often requires approximately 2-3 weeks after that. If data analysis service is requested, this would add an additional 1-2 weeks.

Q: How should I design my DNA-seq experiment?

A: You can schedule a consultation meeting by contacting .ICBR-GeneExpression@ad.ufl.edu or yanp@ufl.edu. We will schedule a meeting with the participation of members from the Gene Expression & Genotyping, NextGen DNA Sequencing, and Bioinformatics cores.