03/2019

Batch-Tag-Seq - Gene Expression Profiling
UC Davis Genome Center - DNA Technologies and Bioinformatics Cores
UC Rates

3'-Tag-Seq is a protocol to generate exceptionally low-noise and low-cost gene expression profiling data. For more details on the technology please see our FAQs on **3'-Tag-RNA-Seq**.

We now offer *Batch-Tag-Seq* packages that include 3'-Tag-RNA-Seq library preparations & sequencing & optional data analysis at low per-sample rates. The differential gene expression (DGE) data analysis is performed by the **Bioinformatics Core** (please contact them at bioinformatics.core@ucdavis.edu).

For *Batch-Tag-Seq* we will collect samples and process them in larger batches and sequence the barcoded libraries together on sequencing runs, allowing us to offer affordable recharge rates on a per-sample basis. This also simplifies the budgeting and planning of experiments since scientists do not have to adjust their experiments to the sequencing capacity of the sequencers.

Each Batch-Tag-Seq project requires a minimum of 8 samples.

The UC recharge rates for Batch-Tag-Seq package is:

➢ 3'-Tag-Seq Library Prep & Sequencing	$104 (per sample)
➢ DGE Data Analysis through the Bioinformatics Core	$80 (per sample)

Please note that samples should be submitted with Bioanalyzer traces (or equivalent). Alternatively we can also run the RNA sample QC. There are three options:

➢ Bioanalyzer RNA PICO assay (2-5 ng/ul total RNA conc.)	$12 (per sample)
➢ Bioanalyzer RNA NANO assay (30-300ng/ul total RNA conc.)	$12 (per sample)
➢ Batch LabChip RNA QC (15-150 ng/ul total RNA conc,)	$5 (per sample)

In order to enable low cost DGE data analyses and short turnaround times, we will process the Batch-Tag-Seq samples in **high throughput fashion**. We will not spend time customizing the protocol for individual samples (for example we will not run sample cleanups, sample concentrations, or repeat the library preps or PCR amplification with varying cycle numbers). The 3'-Tag-RNA-Seq library prep protocol is very robust, therefore no problems should be expected as long as the RNA samples fulfill the sample requirements. **Please note that the customer is responsible for the sample quality**.

Further considerations:

- We require 20 ul of each total RNA sample at a concentration of 25 to 100 ng/ul for batch-processing as quantified by Qubit (**dnatech.genomecenter.ucdavis.edu/sample-requirements/**).
- We suggest the RNA sample concentrations to be **normalized** for a project. A Qubit instrument is available in the Core for accurate RNA concentration measurements.
- The RNA samples need to be dissolved in molecular biology grade H2O or EB buffer. As always RNA-seq samples need to **be DNA-free**.
- The sample concentration must be determined by fluorometry (e.g. Qubit; plate-reader with Ribo-Green), as spectrometry quantifications (e.g. Nanodrop) are very unreliable.
- To assure the chemical purity of the samples the absorbance ratios, as measured by Nanodrop, should be between 1.8 and 2.1 (260/280 nm ratio) and above 1.5 (260/230 nm ratio).
- 3'-Tag-RNA-Seq is only suitable for **eukaryotic** total RNA-samples.
- For each condition at least 3 biological replicate samples need to be sequenced to allow for a meaningful DGE analysis.
- If RNA-isolation protocols involving TriZol are used, the RNA should then be purified via a spin column kits (e.g. Zymo RNA clean & concentrate) to remove any solvent traces.
- 3'-Tag-Seq has a relatively high tolerance for RNA integrity variation. Nevertheless, we do not recommend using RNA samples with a RIN-score lower than 6 for batch processing. We cannot guarantee for the outcome of the library prep and the data for lower quality samples or samples without Bioanalyzer QC.
- The libraries will be sequenced on Illumina HiSeq 4000 or NextSeq 500 sequencers with single-end 84 or 90 bp reads (SE84 or SE90). Please note that for some analysis pipelines it is recommended to trim off the first 11 bases from the reads. We will provide the full length data. Trimming is not necessary if you are using a local aligner (like **STAR** or **BBmap**). The sequences can be trimmed easily, for example with the "`reformat`" command from **BBTools**.
- You will receive 3 to 6 million reads per sample. For typical experiments about 2 million reads per sample are required for the DGE analysis of highly- and medium-expressed genes.

Large-Scale Tag-Seq projects:

For large scale projects you might consider customizing the scale of experiments. The recharge rates broken down into individual services are:

➢ 3'-Tag-Seq library prep and pooling only	$80 (per sample)
➢ Custom (low input) 3'-Tag-Seq library prep and pooling only	$99 (per sample)
➢ SR90 HiSeq sequencing lane	$1,120
➢ NextSeq 75 cycles HIGH output run	$1,620

We will require the same sample QC as outlined before.

***Custom* 3'-Tag RNA-Seq:**

In contrast to the batch processing described here we can adjust the library prep parameters for 3'-Tag-Seq when running custom library preps. The *custom protocols* can generate usable data for inputs as low as 10 ng total and can also work with degraded RNA samples. *Custom protocols* do require additional labor and are thus priced slightly higher. Please note that we can't vouch for the results for marginal samples. Please inquire with us with a description of your samples.

We will use a UMI protocol for low input samples enabling the removal of some low input artifacts sample induced biases.

- Custom (low input) 3'-Tag-Seq library prep and pooling only $99 (per sample)

Bioinformatics:

A prerequisite for bioinformatic data analysis is the availability of a well annotated reference genome (including UTRs; to be provided by the customer). The deliverables will include data QC, read-counts-per-gene tables, and DGE analysis for simple comparisons. At least 3 biological replicate samples need to be sequenced for each condition. Depending on the nature of the samples and the phenotypes, meaningful experiments might require higher replicate numbers. Please inquire with the Bioinformatics Core staff (**bioinformatics.core@ucdavis.edu**) for details and order the analysis with them. The optional DGE analysis service will include:

- QC and pre-processing of data
- Mapping to genome reference
- Analysis in Bioconductor/R: normalized read counts
- Include QA/QC metrics from Preprocessing and Mapping
- Single-factor DGE analysis
- MDS plot and any other QA/QC plots from DE analysis
- Differential Expression Table
- Full description of processing method

- DGE Data Analysis through the Bioinformatics Core $80 (per sample; minimum 12 samples)

Please let us know if we can help with any questions.

Best regards,
Lutz Froenicke
dnatech.genomecenter.ucdavis.edu
email: dnatech@ucdavis.edu; lfroenicke@ucdavis.edu