Frequently Asked Questions

ChIP-Seq

Q1. What is meant by ChIP-Seq ?

A1. From Illumina [1]: "*Chromatin immunoprecipitation (ChIP) is a powerful method to selectively enrich for DNA sequences bound by a particular protein in living cells. ChIP-Seq on Illumina sequencing systems supports virtually unconstrained selection of any ChIP-able protein and/or modification to be studied. These include transcription factors, polymerases and transcriptional machinery, structural proteins, protein modifications, and DNA modifications. … The ChIP process enriches specific crosslinked DNA protein complexes using an antibody against a protein of interest. Unique oligonucleotide adapters are then added to the small stretches of DNA that are bound to the protein of interest to enable massively parallel sequencing.*" Some applications of this technology include [2]:

- Discovery of transcription factor binding sites
- Identification of genes regulated by known transcription factors and co-regulators
- Analysis of epigenetic events
- Direct comparison of regulatory events in different cell states (i.e. normal vs. disease)
- Investigation of drug effects and other stimuli on regulatory pathways

Q2. What material should I send to be analyzed by ChIP-Seq ?

A2. Generally, we start with 10-50 ng of ChIP-enriched DNA. Samples should be submitted in 1.5-1.7 ml microfuge tubes (example: VWR cat. no.89000-028) or 2 ml screw cap tubes (example: Sarstedt cat. no. 72.694.007). Please DO NOT send samples in 0.5 or 0.2 ml tubes. If possible, the sample should be evaluated by the investigator by testing for the relative enrichment of a relevant gene. The best control material is an unprocessed aliquot of the input DNA that went into the ChIP enrichment step. A light sequencing of this sample in parallel can reveal potential false-positives.

Q3. What data are returned by NISC ?

A3. Typically, NISC returns to the investigator a file containing basecalls and quality scores (BAM files) for each sample. The investigator is expected to provide data analyses; this is not offered by NISC. Data files are each about 1 – 3 GB in size. NISC typically aims for 15 million single-end reads.

Q4. How does ChIP-Seq compare to ChIP-Chip ?

A4. NISC has not performed this comparison, but Illumina provides the following list of features [2]:

- **High-Quality Data:** Positional resolution of mapped binding sites ± 50 bp.
- **Sensitivity:** Robust quantification for determining binding specificities of varying strengths; increase amount of sequencing to increase sensitivity.
- **Specificity:** Higher quality data even in complex genomes; lower background than ChIP-Chip, no cross hybridization.
- **Genome-Wide Analysis:** Identifies any binding sites, not limited to ChIP-Chip array features or candidate sequences.

Q5. How long do the reads need to be for ChIP-Seq analysis ?

A5. Typically, read lengths are 50 bases. This length should be sufficient for mapping of most reads to the reference genome. Some investigators are exploring the utility of longer reads and paired-end reads for advanced analyses.

Q6. How many reads are used for a mammalian ChIP-Seq analysis ?

A6. NISC usually targets 15 million reads per library. ChIP-Seq libraries are constructed with indexed adapters, which allow many barcoded libraries to be sequenced as a pool. Greater efficiency is achieved when a pool of 8 libraries are run per lane in the Rapid Run mode on the HiSeq 2500 instrument [3].

Reference:

1. Illumina, Inc. (2014): “ChIP-Seq DNA Sample Prep Kit” http://www.illumina.com/products/chip-seq_dna_sample_prep_kit.ilmn
2. Illumina, Inc. (2014) “Whole-Genome Chromatin IP Sequencing (ChIP-Seq) http:// http://support.illumina.com/sequencing/literature.ilmn
3. Sims, D., *et al.* (2014) “Sequencing depth and coverage: key considerations in genomic analyses.” *Nature Rev. Genetics* **15**: 121-132.