Custom Capture Sequencing

Q1. What is custom capture sequencing ?

A1. Custom Capture Sequencing (CCS) is an efficient strategy to selectively sequence specific (targeted) regions of interest in the genome to discover rare or common variants associated with a disorder or phenotype [1, 2]. By focusing sequence production on targeted regions of the genome, many more individuals can be examined at significantly reduced cost and time compared to sequencing their entire genomes. Custom capture sequencing gives the option of targeting exons in selected genes or tiling across an entire region of the genome. The cumulative size of the targeted regions of interest typically ranges from ~500 kb – 50 Mb.

Q2. How is CCS performed at NISC ?

A2. The standard starting point for a custom capture project is the identification of a region or regions of interest. The NISC staff then uses this information to develop a custom capture design which typically consists of oligonucleotide probes for hybridization-based capture or multiplex PCR primers that can be used to enrich for the targeted region(s) of the genome. NISC provides the investigator a summary of the design coverage for review.. Based on feedback, the design is modified as necessary. Once the investigator has given final approval of the design, then the custom capture kit is ordered. Note that the time from beginning the design process to receipt of a custom capture kit at NISC is on the order of 10 weeks. Only then can NISC start sequencing samples. Custom capture projects typically include 24-576 samples.

Q3. What material should I send for CCS ?

A3. We need a minimum of 150 ng of highly-purified genomic DNA (0.5 μg preferred) in a volume of 50 μl or less for CCS. 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. To ensure that each sample is uniformly pure and free of infectious agents, we strongly recommend that all DNAs be phenol:chloroform extracted before submission. A simple protocol is available from NISC. Ref:
www.nisc.nih.gov/docs/gDNA_submission_exome_cc.pdf

Frequently Asked Questions

Q4. How should the DNA be qualified ?

A4. The investigator must submit an image of an analytical agarose gel or a trace as evidence the DNA is of good integrity and the appropriate molecular weight for the sequencing approach. We highly recommend Qubit for quantitation of the DNA sample, since it uses a double-strand DNA-specific method. UV absorption methods, e.g., using a NanoDrop spectrophotometer, can drastically overestimate the concentration of DNA due to RNA and small molecule contamination.

Q5. How long are the reads for CCS analyses ?

A5. Typically, NISC generates read lengths of 150 bases on a NovaSeq X Plus. Paired-end reads generate a total of 300 bases of sequence (150b from each end) from each fragment in the library.

Q6. How many reads are required for CCS analyses ?

A6. We target the number of reads that are calculated to produce an average read-depth typically greater than 100X coverage [3].

Q7. How are variants called ?

A7. For human samples, we use NVIDIA Parabricks software for variant calling. Parabricks is a GPU-accelerated alignment and variant calling pipeline that uses BWA and GATK algorithms under the hood. The pipeline follows the GATK Best Practices for Variant calling and Filtration.

Q8. What data are returned by NISC ?

A8. We return either raw FASTQ files or aligned BAM files and variant calls (in GVCF format).

References:

1. NimbleGen (2013) "SeqCap EZ Library."
www.nimblegen.com/products/lit/05227887001_SeqCapBroch_Jan2013.pdf
2. Illumina (2013) "An Introduction to Next-Generation Sequencing Technology."
www.illumina.com/documents/products/Illumina_Sequencing_Introduction.pdf
3. Sims, D., et al. (2014) "Sequencing depth and coverage: key considerations in genomic analyses." Nature Rev. Genetics 15: 121-132.