The Children's Hospital of Philadelphia®
RESEARCH INSTITUTE

CENTER FOR APPLIED GENOMICS

Whole Exome Sequencing Submission/guideline

Whole Exome Sequencing (WES) utilizes the latest in library prep capture methods providing crucial information regarding the expression of protein-coding regions within a genome. WES technology allows for variant identification and help a vast array of applications from cancer genomics to population genetics. Investigating the coding regions of the genome can be a more economic approach to scaling studies or discovery research as well as reducing the data analysis load as opposed to whole genome sequencing. With reliable and sensitive detection of variants, this method is ideal for getting a targeted view of protein-coding regions. With 99.9% sensitivity covered by 100X coverage, WES provides a much more affordable and size efficient (8Gb at 100X) method to variant identification[1]. The process typically starts with several genomes pooled into 8-plex pools hybridized with a panel and blockers to prevent non-specific capture. This panel can be customized or an off the shelf pre-formulated panel for general species like mouse, human, dog, and etc. The capture product is amplified and then sequenced with up to 384 uniquely indexed samples depending on desired coverage.

Requirements:

1. Isolation of DNA

The first step in Whole Exome Sequencing is the isolation of DNA. DNA quality is typically measured with Nanodop spectrophotometer for concentration and purity check (260/280 ratio = ~1.8). The Gold Standard method for concentration is qubit to identify the amount of volume needed to reach the total DNA ng input threshold to generate libraries. Note that some factors concerning sample storage, type of source material, and extraction method can lead to abnormal Nanodrop ratios as well as reduced concentration and quality. In the case of extracting DNA using non-standard methods or kits please discuss with the core beforehand.

1: *Find the Method That's Right for Your Research*. Illumina, 2016.

CAG standard quantity and purity criteria for DNA is:

a. The nanodrop 260/280 ratio should be around ~1.8 and 260/230 ratio should be above ~2.0
    - Lower than 1.8 there can be protein or phenol contaminants
b. Some DNA extraction protocols may have some residues that absorb at 230nM like EDTA and Guanidine HCL.
c. Concentrations and volume for our automated library preparation:

Automated Library Preparations			
Library Preparation	Concentration	Volume	Comments
Twist Whole Exome	1.5ng/uL – 5.0ng/uL	40uL	DNA should be eluted in Water, 10mM Tris-HCL pH 8.0, or EB

2. Preparing your samples for CAG

- Isolated DNA should be transferred to a clean 96-well plate or low-bind 1.5mL tubes.
- To avoid seeing batch differences in the final data, it is important to put the samples on the plate in a random order, starting with position A1, then B1, etc. Do not leave any blank wells between the samples. Ensure the correct seal is used when utilizing plates.
- Record sample IDs and do not use patient identifiers or names in an xlsx file according to the criteria established by CAG.
- Annotate plate if possible, with marker.
- Plates should be stored appropriately before delivering to CAG
- Aliquot around 5uL of sample for QC in case QC has not been performed by your lab.
- Communicate with CAG team with any questions or concerns.

1: *Find the Method That's Right for Your Research*. Illumina, 2016.

CAG address: (Hours M-F 9am-5pm)
Center for Applied Genomics (CAG) – NGS Lab
The Children's Hospital of Philadelphia (CHOP)
3615 Civic Center boulevard, Room 1014
Philadelphia, PA 19104-4318
(267) 426-0695
billingsj@chop.edu
CAG NGS iLab page

3. Data Analysis

Data analysis plays an important role in NGS Whole Exome pipeline. CAG has a very skilled team of Bioinformaticians who provide comprehensive data analysis appropriate for specific experiments and project goals. The bioinformatics team is an integrated part of our NGS platform and our scientists have a strong background in both the experimental and analytical aspects of NGS. With that in mind, we consider each projet independently to determine the most appropriate analysis tools and pipeline to answer the relevant biological questions. Our services are offered as a fee for service model or collaboration efforts.

Our data analysis includes:

- Comprehensive QC of sequencing data
- Mapping: Alignment of reads to the specified reference genome (Human, Mouse, Rat, drosophilia, etc.)
- Variant calling (SNVs, indels, CNVs)
- Variant annotation (dbSNP, ClinVar, GENCODE, etc.)
- Somatic variant analysis (Mutect2)
- Additional analysis types can be discussed on a per-project basis

1: *Find the Method That's Right for Your Research*. Illumina, 2016.