4553 Winters Chapel Rd. #100
Atlanta, GA 30360 USA
+1-855-otogene
sales@otogenetics.com
www.otogenetics.com

Otogenetics/CE Advanced Human Exome data analysis specifications

Upon completion, please send this form via email to support@otogenetics.com or by fax to +1-206-339-8150.

Quote #: _____ Institution Name: _____

First Name: _____ Last Name: _____

Data Delivery
Email Address: _____ Genome Selection: _____
(gmail preferred) (hg19 default)

Inclusion of all analyses included in the advanced human exome data analysis specifications? (see Appendix I): YES NO If No, please specify:

Optional Data Analysis:

❶ Somatic mutation analysis (tumor versus normal) and Trio analysis: YES NO
IF YES, please specify the analysis and the sample groups:

❷ Identification of common and unique SNPs/InDels within a group of samples: YES NO
If YES, please specify the analysis and the sample groups:

❸ Somatic CNV detection*: YES NO
If YES, please specify the analysis and the sample groups.

*The analysis will be carried out with normal/tumore pair and tumor-only. The current algorithms are not robust and request for such analysis will be considered on a case-by-case basis.

4553 Winters Chapel Rd. #100
Atlanta, GA 30360 USA
+1-855-otogene
sales@otogenetics.com
www.otogenetics.com

Otogenetics/CE Advanced Human Exome data analysis specifications

Appendix I

Otogenetics Advanced Exome Data Analysis Specifications:

The raw reads are mapped to human genome sequences (default hg19, hg18 available upon request) and further refined by PCR duplication marking, local realignment around InDels, and quality score recalibration. Called SNPs and small InDels are subject to comprehensive functional annotation, including dbSNP, affected gene, affected amino acid, and PolyPhen/SIFT/PhyloP/LRT/MutationTaster/GERP++ scores.

Deliverable
1. Alignment to reference sequences in BAM, which can be used to visualize the alignment to hg19 (or hg18 per special request) using genomic browsers, such as IGV.

2. Mapping statistics at genome, chromosome, and targeted regions in EXCEL

3. Coverage, mapping quality, and duplication* histograms in PNG
*duplication rate: two reads are considered duplicate if their mapping positions are identical. In the Duplication Rate Histogram, the x-axis is the level of duplication and the y-axis is the count of loci – sets of start/end genomic positions. The duplication rate is (# loci with >1 duplication level)/ (# total loci).

4. Called SNPs and small InDels** in the VCF format in EXCEL ; initial all inclusive***;
report tab: SNP_VCF;
**small InDels: <10-15 bp
*** No filter is applied. The preference is not to remove low confidence calls from the report, but to flag them, because they might be significant in certain experimental/biological context.

5. The above all-inclusive report with annotated SNPS and small InDels in EXCEL***;
Report tab: SNP_Annotation_Comprehensive

6. Deleterious SNPs and small InDels listed by gene and by locus in EXCEL**** ;
Report tabs: Significant_SNPs_by_Gene and Significant_SNPs_By_Locus
****A series of filters are applied, on the basis of the comprehensive functional annotation, to identify the most disease relevant SNPs.

7. Confidence-filtered SNPs and small Indels in VCF format in Excel*****;
Report tab: SNP_Confidence_VCF
*****Standard confidence filters applied to the initial all inclusive SNP/InDel calls in the VCF report. The filters include: Hard_To_Validate (MQ0 >= 4 && ((MQ0 / (1.0 * DP)) > 0.1)), LowCoverage (DP < 5), VeryLowQual (QUAL < 30.0), LowQual (QUAL > 30.0 && QUAL < 50.0), LowQD (QD < 1.5), and StrandBias (SB > -10.0).

Additional Analysis:
1. Somatic mutation analysis (normal vs. tumor; or different disease stages) and Trio analysis.
2. Identification of common and unique SNPs/InDels within a group of samples.
3. Somatic CNV detection will be carried out with normal/tumor pair and tumor-only. The current algorithms are not robust and request for such analysis will be considered on a case-by-case basis.

Rev. 110615