**Annotation and filtering workshop: question sheet**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Q1** | From the ANNOVAR log file, how many refGene transcripts were used for annotation? | | | |
|  | | | |
| **Q2** | Is there anything else that would be useful for annotation? | | | |
|  | | | |
| **Q3** | Which variant is the most likely to cause the patients disease? | | | |
|  | | | |
| **Q4** | For the causal variant, use Unified Genotyper VCF and information from the variant calling lecture to populate the list quality control parameters and categorise the value of each as either good, intermediate or poor | | | |
| **Quality control parameter** | **Value** | **Category** | |
| Depth |  |  | |
| Allelic depth (AD) |  |  | |
| Strand bias (FS) |  |  | |
| Variant confidence (QUAL) |  |  | |
| Quality by depth (QD) |  |  | |
| Mapping quality (MQ) |  |  | |
| Mapping quality bias (MQRankSum) |  |  | |
| Base quality bias (BaseQRankSum) |  |  | |
| **Q5** | Is your variant present in [ClinVar](http://www.ncbi.nlm.nih.gov/clinvar/) and [OMIM](http://www.omim.org/) and if it is does the expected phenotype match your patient? | | | |
|  | | | |
| **Q6** | Has your variant been published and if so what is the citation? | | | |
|  | | | |
| **Q7** | Look up the function of your gene. Eg. [GeneCards](http://www.genecards.org/). What is known about it? | | | |
|  | | | |
| **Q8** | If the variant is real, what disease do you think the patient may have? | | | |
|  | | | |
| **Q9** | Repeat the filtering process to determine the number of variants that remain after each of the filters in the table is applied. | | | |
| **Filtering criteria** | | | **No. variants** |
| None | | |  |
| Exonic | | |  |
| Exonic and absent from public databases (dbSNP non flagged, 1000 genomes, ESP) | | |  |
| Exonic, absent from public databases and located in a candidate gene | | |  |
| **Q10** | How manyvariants are flagged by dbSNP as clinically associated? | | | |
|  | | | |
| **Q11** | How many variants are present in OMIM? | | | |
|  | | | |

Turn over for last question

|  |  |  |  |
| --- | --- | --- | --- |
| **Q12** | For the causal variant you have chosen, use the result from wANNOVAR to populate the list of pathogenicity predictors and select a category for each value. | | |
| **Pathogenicity predictor** | **Value** | **Category** |
| [SIFT](http://sift.jcvi.org/) |  | Deleterious/Tolerated/Missing |
| [Polyphen-2 HDIV](http://genetics.bwh.harvard.edu/pph2/) |  | Probably damaging/Possible damaging/Benign/Missing |
| Polyphen-2 HVAR |  | Probably damaging/Possible damaging/Benign/Missing |
| [LRT score](http://www.genetics.wustl.edu/jflab/lrt_query.html) |  | Deleterious/Neutral/Unknown/Missing |
| [Mutation Taster](http://www.mutationtaster.org/info/documentation.html) |  | Disease\_causing\_automatic/Disease\_causing/Polymorphism/ Polymorphism\_automatic/Missing |
| [Mutation Assessor](http://mutationassessor.org/) |  | Predicted functional (high, medium)/Predicted non-functional (low, neutral)/Missing |
| [FATHMM score](http://fathmm.biocompute.org.uk/) |  | Damaging/Tolerated/Missing |
| [RadialSVM](http://genomics.usc.edu/members/15-member-detail/36-coco-dong) |  | Deleterious/Tolerated/Missing |
| [LR score](http://genomics.usc.edu/members/15-member-detail/36-coco-dong) |  | Deleterious/Tolerated/Missing |
| [VEST3 score](http://karchinlab.org/apps/appVest.html) |  | None, likelihood of functional effect increase with score |
| [CADD raw](http://cadd.gs.washington.edu/home) |  | None, likelihood of damaging effect increase with score |
| [CADD phred](http://cadd.gs.washington.edu/home) |  | None, ranked and phred scaled CADD score |
| [GERP++RS](http://mendel.stanford.edu/SidowLab/downloads/gerp/) |  | None, conservation increases with score (range -12.3 to 6.17). Scores >2, high sensitivity for truly constrained sites |
| [phyloP46way placental](http://compgen.cshl.edu/phast/help-pages/phyloP.txt) |  | None, the larger the score, the more conserved the site |
| [phyloP100way vertebrate](http://compgen.cshl.edu/phast/help-pages/phyloP.txt) |  | None, the larger the score, the more conserved the site |
| [SiPhy 29way logOdds](https://www.broadinstitute.org/genome_bio/siphy/index.html) |  | The larger the score the more conserved the site |