

Tempus Bioinformatics Technical Challenge

For this challenge, you are asked to prototype a variant annotation tool. We will provide you with a VCF file, and you will create a small software program to annotate each variant in the file. You may use whichever language you like.

Each variant must be annotated with the following pieces of information:

1. Depth of sequence coverage at the site of variation.
2. Number of reads supporting the variant.
3. Percentage of reads supporting the variant versus those supporting reference reads.
4. Using the Ensembl VEP API, get the gene of the variant, type of variation (substitution, insertion, CNV, etc.) and their effect (missense, silent, intergenic, etc.). The API documentation is available here: [Ensembl GRCh37 Rest API](#)
5. The minor allele frequency of the variant if available.
6. Any additional annotations that you feel might be relevant.

Please upload all relevant code to a public repo on GitHub. Be sure to include the annotated variants in a csv/tsv file. Note that work will be assessed based on quality of code, documentation, and problem solving more-so than the annotations themselves.

Feel free to reach out to bioinformatics-code-challenge@tempus.com with questions.