Dear applicant,

To assess your general programming level, we would like you to provide a solution to the problem stated below. You have 48h to accomplish this task. Take special care for readability and flexibility of your solution.

Generate a program that takes in a list of genomic variants in a tabular format (see attachment, Variants.txt). The program should execute the following actions:

**1. Annotate variants for effect**

- use myvariant.info to get snpEff annotations variant by variant. An example call is :   
 *http://myvariant.info/v1/variant/chr16:g.28883241A>G?fields=snpeff*

- Parse the results, keep variants with a “HIGH” putative impact, add annotations as extra columns (also adjust the headerline in the output). Remark, if variant is annotated in multiple transcripts, keep this variant when it has a ‘HIGH’ impact in one or more transcripts.

**2. Annotate variants for frequency**

- For the HIGH impact variants, also fetch minor allele frequencies over all populations from exac:  
 [*http://myvariant.info/v1/variant/chr10:g.61802416T%3EC?fields=exac.alleles,exac.af*](http://myvariant.info/v1/variant/chr10:g.61802416T%3eC?fields=exac.alleles,exac.af)  
- Parse results, keep variants with MAF < 0.001. Add MAF as extra column to output.

Remark: if multiple alleles exist on the position of the variant, be sure you extract the correct MAF for that specific variant in the list!

Remark2: variants unknown to Exac are novel variants, and those should definitely be kept in the generated output-file.

**3. Print out the resulting variants to a user-provided output file**

The resulting table should consist of rare/novel high-impact variants only.

**Some structural requirements:**

- Programming language doesn’t matter.

- Use functions/routines for repeated tasks.

- Provide a short how-to when ‘-h’ is provided to the program.

- Provide error-handling :

- myvaraint.info failure: try again 1 time, else skip variant with error message  
 - input file is not found : print help and exit.