Please create a Python or R script to do transcript ID matching between RefSeq ID (NM\_\*) and ENSEMBL ID (ENST\*). Specifically, this tool should retrieve ENSEMBL transcript IDs and gene symbols when a list of RefSeq isoform IDs is provided and vice versa.

Hint 1:

This tool will do the similar thing with the function “getBM” in the R package called biomRt. The R package biomaRT provides a powerful way to access BioMart datadaset and enables a wide range of powerful online queries from gene annotation to database mining.

Hint 2:

Alternatively, download the ENSEMBL to RefSeq mapping from UCSC Genome Browser at <https://hgdownload.soe.ucsc.edu/goldenPath/hg38/database/kgXref.txt.gz>, to be used in your script.

Ans: I have generated an Rshiny app based on the inputs. When quering the website , You will get outputs based on three queries, Genes, RefSeq ID, Transcript ID

Website can be found here: <https://apampana.shinyapps.io/map_ids/>

I am also including Code snippet I used to generate this app.