The desired editing efficiencies of HOPE (substitution/insertion/deletion) are calculated by CRISPResso2 (<https://github.com/pinellolab/CRISPResso2>), a software designed to enable rapid and intuitive interpretation of genome editing experiments.

Herein, we use CRISPResso2 to calculate the efficiencies of desired substitution and insertion/deletion with “PE (Prime-edited) mode” and “HDR mode”, respectively.

For desired substitution events analysis, following parameters are required: “—prime editing pegRNA spacer seq”, “—prime editing pegRNA extension seq”, “—prime editing pegRNA scaffold seq”, “—prime editing nicking guide seq”. And for desired insertion and deletion events, the HDR amplicon sequence was used as the “-e” parameter, “—prime editing pegRNA spacer seq”, “—prime editing nicking guide seq” are also required. Specifically, “—discard indel reads” parameter is used for desired editing events calculation, while “—ignore substitutions” is used for undesired indels calculation. The precise editing frequencies are calculated based on the “CRISPResso quantification of editing frequency” files obtained from CRISPResso2. The corresponding formulas are shown as below:

To evaluate the detailed undesired indel types of HOPE, we carry out the indel analysis with “HOPE\_indel\_analysis.py” (Python3).

The script needs two inputs. The "-ifa" parameter is the input of the corresponding amplicon sequence, and the "-iat" is the input filtered from the "Alleles\_frequency\_table.txt" file analyzed by CRISPResso2. For example, filter the column” Reference\_Name“ as ”Prime-edited“ and column ”Read\_Status“ as ”MODIFIED“.

The details of INDEL can be obtained through the pipeline which is mentioned above, including the positions of insertion, length of insertion sequence, reads statistics and insertion sequence.

Furthermore, you can also obtain the start and end positions of the deletion sequence, deletion sequence length, and the reads statistics.

The results of cumulative statistical reads of insertion and deletion events at each location of amplicon reference can be obtained.

In addition, you can also obtain statistics on reads that begin to occur at certain locations of the amplicon reference for insertion and deletion.

Cas-OFFinder (https://github.com/snugel/cas-offinder) is used to predict the off-target sites. The substitution or indel ratio at the predicted off-target editing sites, or a ± 25 bp region surrounding the off-target sites could be calculated as descripted above. Average substitution/indel ratios observed out of potential off-target editing regions (± 25 bp surrounding the predicted off-target sites) are defined as “BackGround” (excluding the ± 15 bp from the two side of the amplicon sequence). Note that SNPs are not taken into the calculation (serve as mutation ratio less than 10%).

The R script shows an example for calculating the BG ratio. The input “.bmat” file contain the information of substitution, insertion, deletion and ambiguous statical reads for each location index of amplicon reference, which could be found in Detect-seq pipeline