**1.To digest genome of a species using single enzyme cutter in silico:**

Firstly, copying the genome-sequence files in FASTA format of a species into the folder “INPUT\_CHR” under the directory of “single\_cutting\_scripts”. And changing names of the files (chr1.fa, chr2.fa, chr3.fa…).

Secondly, opening the script of “pro\_complete.pl” and changing the values of the hash varable “%hash\_1” in line5 according to the enzymes that need to be tested in silico. Note that the keys of the hash are recognition sequences and the values of the hash are cutting positions.

Thirdly, type “perl pro\_complete.pl” in command line of Linux system to run the script. During the script running, you need to input the number of chromosomes in the tested species and press “enter” key when the sentence of “input the number of chromsomes” shows in command line. The result files will be outputted into the folder “OUTPUT\_STEP1” and named as “STEP3\*.txt”, such as “STEP3CTGCAG5.txt”. In each “STEP3\*.txt” file, two columns are included. The first column is the lengths of fragments from sigle-enzyme-cutting. The other is the number of the fragments with the same length.

**2. To digest genome of a species using double enzyme cutters in silico:**

Firstly, copying the genome-sequence files in FASTA format of a species into the top directory of “double\_cutting\_scripts” folder. And changing names of the files (chr1.fa, chr2.fa, chr3.fa…).

Secondly, opening “dd\_list.txt” file and filling it with enzymes information. There are three columns in the file. These columns are delimited with Tab. The first column is the names of double enzymes joined with “\_” symbol. The other columns are composed with recognition sequence and cutting position of each enzyme cutter, respectively. The recognition sequence and cutting position are joined with “/” symbol. An example of “dd\_list.txt” file can be find in “example” folder.

Thirdly, opening “dd\_autorun.pl” and changing the value of varable in line32 (my $x\_files =5) into the number of chromosomes in tested species. And then typing “perl dd\_autorun.pl” in command line of Linux system to run the program. As a result, two types of output files can be found. In one type , every file names include “dd”, such as “AflII\_BfaI\_RESULT\_dd.txt”. In the other type, every file names include “total”, such as “AflII\_BfaI\_RESULT\_total.txt”.

Fourthly, the files with “dd” are moved into the folder “dd\_results”, while the files with “total” are moved into the folder “total\_results”. Then running the scripts of “dd\_results\_further\_analysis.pl” and “total\_results\_further\_analysis.pl”, respectively. As a result, two result files named as “dd\_further\_results\_100-600bp.txt” and “total\_further\_results.txt” will be outputted. In the file of “dd\_further\_results\_100-600bp.txt”, five columns will be outputted. First column includes the names of every file with “dd”. Second column is the number of fragments with two different cutting ends. Third column is the total length of these fragments. Fourth column is the number of fragments with two different cutting ends and 100-600bp length. Fifth column is the total length of the fragments in fourth column. Similarly, three columns will be found in the file of “total\_further\_results.txt”. First column is the name of each file with “total”. Next column is the number of all fragments from double-enzyme-cutting. Last column is the total length of all fragments.

**3. To identify genome locations of selected double-enzyme-cutting fragments (300-600bp) in silico:**

Firstly, copying the genome-sequence files in FASTA format of a species into the top directory of “dd\_enzyme\_cutting\_sites” folder. And changing names of the files (chr1.fa, chr2.fa, chr3.fa…). Then, opening “dd\_list.txt” file and filling it with enzymes information. There are three columns in the file. These columns are delimited with Tab. The first column is the names of double enzymes who are joined with “\_” symbol. The other columns are composed with recognition sequence and cutting position of each enzyme cutter, respectively. The recognition sequence and cutting position are joined with “/” symbol. An example of “dd\_list.txt” file can be find in “example” folder. Lastly, typing “perl dd\_cuttingsites\_of\_frag300-600.pl” in command line of Linux system to run the program. The results can be outputted into the folder “results”.

**Note:** Output results of the above programs can be further processed in EXCEL according to your special purpose