SALESIAN
JOURNAL OF HUMANITIES & SOCIAL SCIENCES

TECHNOLOGY
AND
SOCIETY

VOLUME IV, NO. 1 MAY 2013

RDAT: A TOOL FOR DNA ANALYSIS

Amit Ghosh Roy
Ardhendu Mandal

A Restriction Digest is a procedure used in molecular biology to prepare DNA for analysis or other processing. This enzymatic technique can be used for cleaving DNA molecules at specific sites, ensuring that all DNA fragments that contain a particular sequence have the same size; furthermore, each fragment that contains the desired sequence has the sequence located at exactly the same position within the fragment. These enzymes are called restriction endonucleases or restriction enzymes,[1] and they are able to cleave DNA molecules at the positions at which particular short sequences of bases are present.

RDAT (Restriction Digestion Analysis Tool), is a program that will accept an input DNA sequence and produce a comprehensive report of the restriction enzymes that will cleave the sequence. It produces a variety of outputs including restriction enzyme maps, theoretical digests and links into the restriction enzyme database, REBASE.

Key Words: Restriction digest, Restriction enzymes, DNA sequence, Enzyme database

Background

A restriction enzyme (or restriction endonuclease) is an enzyme that cuts double-stranded or single stranded DNA at specific recognition nucleotide sequences known as restriction sites. Such enzymes, found in bacteria and archaea, are thought to have evolved to provide a defense mechanism against invading viruses. Inside a

[1] A restriction enzyme (or restriction endonuclease) is an enzyme that cuts DNA at or near specific recognition nucleotide sequences known as restriction sites. URL:http://en.wikipedia.org/wiki/Restriction_enzyme (accessed on 5/Apr/2013).

bacterial host, the restriction enzymes selectively cut up foreign DNA in a process called restriction; host DNA is methylated by a modification enzyme (a methylase) to protect it from the restriction enzyme's activity. Collectively, these two processes form the restriction modification system. To cut the DNA, a restriction enzyme makes two incisions, once through each sugar-phosphate backbone (i.e. each strand) of the DNA double helix.

Recognition site

Restriction enzymes recognize a specific sequence of nucleotides and produce a double-stranded cut in the DNA. While recognition sequences vary between 4 and 8 nucleotides, many of them are palindromic, whichcorrespond to nitrogenous base sequences that read the same backwards and forwards. In theory, there are twotypes of palindromic sequences that can be possible in DNA.The mirror-like palindrome is similar to those found in ordinary text, in which a sequence reads the same forward and backwards on the same DNA strand (i.e., single stranded) as in GTAATG. The inverted repeat palindrome is also a sequence that reads the same forward and backwards, but the forward and backward sequences are found in complementary DNA strands (i.e., double stranded) as in GTATAC (Notice that GTATAC is complementary to CATATG). The inverted repeat is more common and has greater biological importance than the mirror-like.

EcoRI[2] digestion produces "sticky" ends,

G^AATT

CTTAA^G

Whereas, SmaI restriction enzyme[3] cleavage produces "blunt" ends

CCC^GGG

GGG^CCC

Recognition sequences in DNA differ for each restriction enzyme,

[2] EcoRI (pronounced, "eco R one") is an endonuclease enzyme isolated from strains of E. Coli, and is part of the restriction modification system. URL: http://en.wikipedia.org/wiki/EcoRI (accessed on 5/Apr/2013).

[3] SmaI is a type II restriction enzyme; it is an isoschizomer of XmaI. URL: https//en.wikipedia.org/wiki/SmaI (accessed on 5/Apr/2013).

producing differences in the length, sequence and strand orientation (5' end or the 3' end) of a sticky-end "overhang" of an enzyme restriction.

Electrophoreses[4] mobility factor

When a potential difference (voltage) is applied across the electrod; it generates a potential gradient E, which is applied voltage V divided by the distance between the electrodes.

Potential gradient (E) =V/d.

When potential gradient is applied, the force on a molecule is F=Eq where E=potential gradient & q=total charge. However, here is also a frictional resistance that retards the movement of the charged module.

Frictional force depends on the following four criteria:-

- Hydrodynamic size of the molecule.
- The shape of the molecule
- Pore size of the medium
- Viscosity of the buffer.

The velocity v of a charged molecule in an electric field is:

V=Eq/f, where f is the frictional coefficient.

More commonly, the term electrophotic mobility (mu) of an ion is used, which is the ratio of the velocity of the ion to field strength i.e. mu=V/E, depending on the value (supplied by the client) of E, q, f we have to compute.

The movement of DNA fragments through agarose gel, frictional force increases as the size of the fragment increases. Therefore, the shorter fragments will move faster rather than bigger fragments.

Importance of DNA Analysis

Isolated restriction enzymes are used to manipulate DNA for different scientific applications. Restriction enzymes can also be used

[4] Electrophoresis is the motion of dispersed particles relative to a fluid under the influence of a spatially uniform electric field. URL: https://en.wikipedia.org/wiki/Electrophoresis (accessed on 5/Apr/2013).

to distinguish gene alleles[5] by specifically recognizing single base changes in DNA known as single nucleotide polymorphisms (SNPs). In a similar manner, restriction enzymes are used to digest genomic DNA for gene analysis by Southern blot. This technique allows researchers to identify how many copies (or paralogues) of a gene are present in the genome of one individual, or how many gene mutations (polymorphisms) have occurred within a population.

The resulting digested DNA is very often selectively amplified using PCR,[6] making it more suitable for analytical techniques such as agarose gel electrophoresis, and chromatography. It is used in genetic fingerprinting, and RFLP analysis.[7]

Problem Analysis

The Type II restriction enzymes are among the most valuable tools available to researchers in molecular-biology. These enzymes recognize short DNA sequences and cleave at, or close to, their recognition sites. A comprehensive database (REBASE) contains information about these enzymes including their recognition specificities and their sensitivity to DNA methylation.

Some software which cuts DNA with restriction enzymes, is also available, on the internet Such as NEBcutter,[8] Webcutter[9] which cuts DNA with restriction enzymes, but there are some problems:

* Availability of internet for 24 x 7.

[5] An allele is one of a number of alternative forms of the same gene or same genetic locus (generally a group of genes). URL: http://en.wikipedia.org/wiki/Allele (accessed on 7/Apr/2013).

[6] The polymerase chain reaction (PCR) is a biochemical technology in molecular biology to amplify a single or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence.
URL: http://en.wikipedia.org/wiki/Polymerase_chain_reaction (accessed on 7/Apr/2013).

[7] In RFLP analysis, the DNA sample is broken into pieces (digested) by restriction enzymes and the resulting restriction fragments are separated according to their lengths by gel electrophoresis.

[8] URL: http://en.wikipedia.org/wiki/Restriction_fragment_length_polymorphism (accessed on 7/Apr/2013).
URL: http://tools.neb.com/NEBcutter2/ (accessed on 7/Apr/2013).

[9] URL: http://rna.lundberg.gu.se/cutter2/ (accessed on 7/Apr/2013).

* They are not freely available to use in a PC.
* Those softwares are very costly.
* Most of the software deals with single DNA sequence. For analysis multiple sequences must be cut simultaneously so that they can be compared.
* Need for development a virtual gel viewer.
* Their maintenance status is unclear.

Restriction Digestion Analysis Tool is a PC version software, which requires no internet connection and which analyzes DNA sequences for the presence of restriction enzyme sites in a convenient and easy to use manner.

Pseudo code

RDAT is consist of a number of modules, here one of the module's pseudo code has been discussed

Module for Non blunt end process:

1. Start
2. INPUT DNA sequence in Fasta format
3. SCAN DNA Sequence and CHECK SPECIFICITY till the END of Sequence
4. If NOT found then GOTO step 10 otherwise
5. SELECT ENZYME FROM THE REBASE
6. SHOW MULTIPLE ENZYME & SINGLE ENZYME
7. If SINGLE ENZYME Selected then Do Step A, B & C

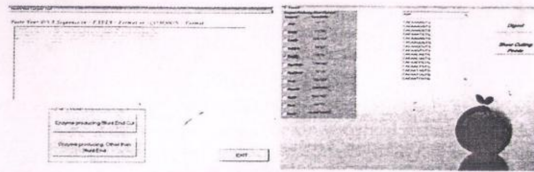
A. ITERATION UPTO TO THE LENGTH OF DNA SEQUENCE
B. CHECKING SPECIFICITY IN THE SEQUENCE
C. Check which all ENZYME(or ENZYMES) has CUTTING point

8. IF MULTIPLE ENZYME Selected then Repeat Step 7 for each ENZYME
9. List all CUTTING points
10. STOP

Results

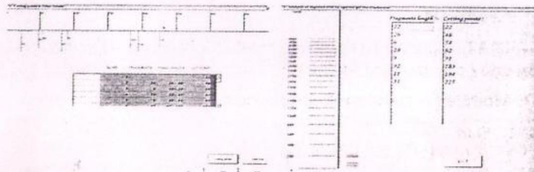
In this section some snapshots of the program are shown. These snapshots are of results, which are got after Sequencing and Analysing various DNA sequence.

Pasting DNA sequence in Fasta format[10] on the editor Click on the procedure "Enzyme Producing Blunt End Cut."



Show the cut points on a linear Scale Electrophoresis

DNA digestion using Agarose gel



Conclusion

* The "Restriction Digestion Analysis Tool" is the GUI based and easy to handing specially of Student of the college and University, the researchers. Especially researchers will be very much benefited by this tool as they have analyzed and extract formation from the report of restriction enzymes that will cleave the DNA sequence.
* This software will help in grouping of different DNA sequences according to their restriction.

[10] FASTA format is a text-based format for representing either nucleotide sequences or peptide sequences, in which nucleotides or amino acids are represented using single-letter codes.
URL: http://en.wikipedia.org/wiki/FASTA_format http://en.wikipedia.org/wiki/FASTA_format

* We are unable to provide the analysis report for multiple DNA sequence due to the lack of time and limitations of our ordinary computer system (PC), because analysis for multiple DNA sequence will require very complex computations, which can be done by a computer with high computational power. The graphical representation of virtual gel view of multiple sequences is compared to extract useful information about gene and generating dendrogram.[11]
* Restriction Digestion Analysis Tool will work on only valid DNA sequence. If there is any error in DNA sequence then system will detect it and generate a dialog massage. But Restriction Digestion Analysis Tool can't rectify that incorrect DNA sequence.[12]

[11] Dendrograms are often used in computational biology to illustrate the clustering of genes or samples. URL: https://en.wikipedia.org/wiki/Dendrogram

[12] Reference
- David W. Mount (CBs), *Bio Informatics Sequence & Genomes* - for DNA cleavage and REABASE analysis.
- Alexander N. Glazer, *Microbial Biotechnology*, Freeman & company - for Restriction analysis.