Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 7, Issue, 01, pp.11260-11263, January, 2015

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

GENOME SEQUENCING IN NEMATODES AND ITS APPLICATION

***Kshetrimayum Sumita and Debanand Das**

Department of Nematology, Assam Agricultural University, Jorhat-785013, Assam, India

ARTICLE INFO

Article History:
Received 21st October, 2014
Received in revised form
14th November, 2014
Accepted 25th December, 2014
Published online 23rd January, 2015

Key words:

Caenorhabditis elegans,
Maxam-Gilber,
PCR.

ABSTRACT

Every organism, may be an animal or a plant, has different gene within them. Genes are the locatable regions of genomic sequences, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions, and or other functional sequence regions (Pearson, 2006 and Pennisi, 2007). The sum total of the entire gene in an organism is known as Genome. It is encoded either in DNA or, for many types of viruses in RNA. The process of determining the precise order of nucleotides within a DNA molecule is called as genome sequencing. The first DNA sequencing method was Maxam-Gilbert sequencing method (Maxam and Gilbert, 1977) and later on many methods have been developed by many scientists for genome sequencing. Among them, Shotgun sequencing (Staden, 1979) is one of the important and widely used methods in the field of Nematology. *Caenorhabditis elegans* is the first multicellular organism whose genome is completely sequenced having 97Mb (now 100Mb) (The *C. elegans* Sequencing Consortium, 1998). With the help of genome sequencing, detection of particular nematode and study on genetic variations can be done with accurate result and, it also helps in generating a new PCR primer specific for a nematode. The functional gene identified in genome sequencing helps to know the mechanism of parasitism in the host which will play an important role in nematode disease management; ultimately it will give a boon in the science of agriculture.

INTRODUCTION

Every organism is controlled by different genes, where it holds the information to build and maintained an organism's cells and passes genetic traits to offspring. The word 'gene' is derived from the Greek word *genesis* meaning "birth", or *genos* meaning "origin". A gene is a molecular unit of heredity of a living organism which is made up of nucleotide sequences where a combination of 3 nucleotide (codon) that codes for one amino acid. According to Pearson (2006) and Pennisi (2007), gene is defined as "locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions, and/or other functional sequence regions. Since genes are nothing but segments of nucleic acid that codes for a particular specific polypeptide, thus these nucleic acids are nucleotide polymers which are linear and unbranched. The functions of nucleic acid are to store and transfer information needed for the synthesis of specific proteins that determines the structure and functions of cell. Carbon, Hydrogen, Oxygen, Nitrogen and Phosphorus forms a nucleotide. A nucleotide molecule consists of three smaller molecules of different types *viz*., phosphoric acid, a pentose sugar and a molecule of nitrogenous organic base.

Corresponding author: Kshetrimayum Sumita
Department of Nematology, Assam Agricultural University, Jorhat-785013, Assam, India.

a. **Phosphoric acid** (H_3PO_4): It is biologically called as phosphate group as it dissociates under physiological conditions freeing hydrogen ions. This phosphate group gives acidic character to the nucleic acids.
b. **Pentose sugar**: It may be ribose ($C_5H_{10}O_5$) or deoxyribose ($C_5H_{10}O_4$). Ribose molecule differs from deoxyribose molecule in having a hydroxyl group (-OH) instead of a hydrogen a Carbon2. Thus, the deoxyribose sugar has one oxygen atom less than ribose sugar, hence its name. To reflect this unusual sugar component, chromosomal nucleic acid is called deoxyribonucleic acid abbreviated as DNA and analogous nucleic acid in which the sugar component is ribose are termed as ribonucleic acids, abbreviated as RNA.
c. **Nitrogenous bases**: The organic base of a nucleotide is either a pyrimidine or purine. They are flat heterocyclic compound having nitrogen (N) and carbon (C) in the ring structure. Nitrogen in the ring tends to take up H^+ from the solution and give the molecules their basic nature, hence the term nitrogenous bases.

Genome: It is the sum total of all the genes in an organism. In modern molecular biology and genetics, the genome is the genetic material of an organism. It is encoded either in DNA or for many types of virus, in RNA. The genome includes both genes and non-coding sequences of DNA/RNA (Riddle, M., 2006).

The term "genome" was coined by Hans Karl Albert Winkler, who is a professor of Botany at University of Hamburg in 1920. The dictionary meaning of 'genome' is blend of gene and chromosome.

GENOME SEQUENCING/ DNA sequencing

It is the process of determining the precise order of nucleotide within a DNA molecule. It includes any method or technology that is used to determine the sequence of individual gene, large genetic region (operons), full chromosomes or entire genome and order of four bases in a strand of DNA.

History of genome sequencing

Year	Development of genome sequencing
1869	Meischer discovered DNA
1940	Avery propose DNA as "genetic material"
1950-1960	Discovery of DNA structure by Watson and Crick (1953)
1970-1980	Development of Maxam-Gilbert sequencing by Maxam and Gilbert (1977) Development of Sanger sequencing by Federick Sanger (1977) Development of Shotgun sequencing by Staden (1979)
1981-1990	Polymerase chain reaction (PCR) by Kary Mullis (1983)
1991-2000	454 Pyrosequencing developed by Mostafa Ronaghi and Pal Nyren (1996)
2001-2010	Solexa or Illumina sequencing (2006) by Balasubramanium and Klenerman

There are different methods of genome sequencing *viz*., Maxam-Gilbert sequencing (Chemical sequencing), Chain-termination method (Sanger sequencing), Shotgun sequencing, Bridge PCR, Massively parallel signature sequencing (MPSS), Polony sequencing, 454 pyrosequencing, Illumina (Solexa) sequencing, Solid sequencing, Ion Torrent semiconductor sequencing, DNA nanoball sequencing, Heliscope single molecule sequencing, Single molecule real time (SMRT) sequencing, Nanopore DNA sequencing, Tunnelling currents DNA sequencing, Sequencing by hybridization, Sequencing with mass spectrometry, Microfluidie Sanger sequencing, Microscopy- based techniques, RNAP sequencing, *in vitro* high-throughput sequencing etc. Out of these sequencing methods, the most widely used methods for genome sequencing are Sanger sequencing, Shotgun sequencing, 454sequencing (pyrosequencing), Illumina (Solexa) sequencing and Bridge PCR sequencing. Among these four sequencing method, Shotgun sequencing is commonly used in the field of Nematology.

Shotgun sequencing

It is a sequencing method designed for analysis of DNA sequences longer than 1000base pairs, up to and including entire chromosomes. This method requires the target DNA to be broken into random fragments. The broken fragments are then; clone so that sequencing can be done. After sequencing individual fragments, the sequences reassembled on the basis of their overlapping regions (Staden, 1979).

Steps of cloning

1. Isolation of desired Genomic DNA
2. Amplification: to produce many exact copies of desired DNA
3. Gel electrophoresis - to detect the presence / absence of DNA, size and quality of DNA.
4. Selection of good vector for cloning
5. Digestion/ restriction of vector- breaking of certain part of vector are done by digestion enzyme (ECoR1, Hind 1, Hind III etc.) so that desired DNA can be inserted
6. Insert of desired DNA in the vector.
7. Ligation- it is the process to join the cut ends of vector after DNA is inserted using ligase enzymes
8. Transformation- it is the change of vector host. After the entry of desired DNA, the vector cell detoriates so, the inserted DNA is transfer to another host (specially used vector = *Escherichia coli*) known as competent cell, and are grown in suitable media. Lastly, screening is done using Blue-White method so as to check whether the clone is transformed or not. White colony = cell transformed; Blue colony = cell not transformed.

After Sequencing is completed, it may be stored in Gene bank

Gene bank is an open access, sequence database annotated collection of all publicly available nucleotide sequences and their protein translations.

Genome sequencing completed

Organism type	Organism	Genome size	Note
Virus	Bacteriophage MS2	3.5kb	First sequence RNA genome (Fiers *et al*.,1976)
Virus	Phage Φ-X174	5.4kb	First sequence DNA genome (Sanger *et al*.,1977)
Bacterium	*Hamophilus influenzae*	1.8Mb	First genome of living organism sequence (Fleischmann, *et al*., 1995)
Plant	*Arabidopsis thaliana*	157Mb	First plant genome sequence (Greihuber, *et al*., 2000)
Moss	*Physcomitrella patens*	480Mb	First genome of a bryophytes sequence (Lang, *et al*., 2008)
Yeast	*Saccharomyces cereyisiae*	12.1Mb	First eukaryotic genome sequence, 1996
Nematode	*Caenorhabditis elegans*	100Mb	First multicellular organism genome sequence, The *C. elegans* Sequencing Consortium, 1998
Nematode	*Meloidogyne hapla*	54Mb	Smallest nematode genome completed (Opperman *et al*., 2008)

Nematode

Nematodes are lower invertebrates, highly diversified, perhaps the most numerous multicellular animals on the earth. Generally, they are found in marine and fresh waters (free-living nematodes) or in soil. They occur at the bottom of the lakes, rivers, at enormous depths in the ocean and all types of soil. Some species can survive temperature constantly below freezing point while, others live in the waters of hot springs and, still others can withstand complete dryness on the surface of rocks during hot summer, reviving again with the onset of

the rainy season. A very large number of species are parasites of different kinds of plants and animals. The parasitic species are of considerable agricultural, clinical and veterinary importance as pests of plants and parasites of man and livestock. Some species of nematodes are microphagous, saprophagous, and many species are herbivorous or phytophagous obtaining nourishment directly from plants. The word "nematode" is derived from the Greek words, *nema* - means "threads", and *oides* meaning "form". Nematode is defined as triploblastic, bilaterally symmetrical, unsegmented, pseudocoelomate animal lacking respiratory and circulatory system.

Some basic information on genomic status of *C. elegans*, *M. hapla* and *Brugia malayi*

(Smant *et al*., 1998)

(*C. elegans* =free-living; *M. hapla* =phytonematodes; *Brugia malayi* =human parasitic nematode)

	C. elegans	*M. hapla*	*Brugia malayi*
Genome size, Mb	100	54	90-95
Sequence coverage, %	100	99.2	76.5
Gene density	235	270	162
Exon per gene	6	6	7
G+C, %	35.4	27.4	30.5
Chromosomes	6	16	6
Predicted proteins	23,662	16,676	11,500

Application of genome sequencing in nematodes

a. Identification of particular nematode
b. Study of genetic variation
c. Identification of functional gene
d. Role in disease management

Identification of particular nematode

Identification of nematode can be done by both conventional as well as non-conventional method. To identify a particular nematode, disease sample should be collected which contain soil and plant roots. After collection (for non-conventional) we have to go for certain steps, they are as follows:

1. Nematode sample
2. Isolation of DNA
3. Gel electrophoresis for genomic DNA
4. PCR
5. Gel electrophoresis
6. Partial Sequencing
7. Data analysis
8. Phylogenetic tree

Study of genetic variation

It can also be done by conventional method based on host differential, but accurate result can be obtained when we go for conventional followed by non-conventional method. The steps required for this are:

1. Nematode sample
2. Isolation of DNA
3. Gel electrophoresis for genomic DNA
4. PCR
5. Gel electrophoresis
6. Data analysis
7. Phylogenetic tree

Identification of functional gene (Plant-parasitic nemtode)

Most of the plant parasitic nematode under the class Secernentea has three glands present in the basal bulb of esophagus, one dorsal and two sub-ventral glands. Earlier during 1990s, it was thought that the location of the dorsal gland valve near the stylet knobs readily allows dorsal gland cell secretions to be released through the stylet. Labeling of stylet secretions produced by adult females of *M. incognita* with a monoclonal antibody specific for dorsal gland granules (Hussey *et al*., 1990) and video-enhanced observation of nematode secretory activity *in vivo* (Wyss and Zunke, 1986) provide evidence that dorsal gland cell secretions can be released through stylet of plant parasitic nematodes. In contrast, the release of secretory proteins from the sub-ventral gland cells into the esophageal lumen directly behind the triaradiate pump chamber in the metacarpus and the rigid and circular lumen of the esophagus anterior to the pump chamber was considered to restrict anterior flow of sub-ventral secretion (Doncaster, 1971). Therefore, secretions of the sub-ventral gland cell were originally thought to pass only posterior in the esophageal lumen to the intestine to function in intracorporeal digestion.

With recent discoveries of genes present in the sub-ventral glands, it has been confirmed that not only dorsal gland plays important role, sub-ventral gland is also responsible in early parasitism.

Following are some of the genes responsible in early parasitism present in both sub-ventral and dorsal glands of root knot and cyst nematodes (Hussey *et al*., 2002).

1. ***Hg-eng-1*** and ***Gr-eng-1*** expressed in sub-ventral gland (*Heterodera glysine* and *Globodera rostochiensis* respt.*)* is responsible in plant/host cell wall disintegration.
2. ***Mj-cm-1*** expressed in sub-ventral glands of parasitic *Meloidogyne javanica* is responsible in alteration of cell wall, hormone as well as defense compound in plant
3. ***Mi-cbp-1*** expressed in sub-ventral gland of *M. incognita* is responsible in modification of plant cell**.**
4. ***Mi-msp-1*** expressed in dorsal gland of *M. incognita* plays important role in early parasitism

Identification of functional gene (Free-living nematode)

Free-living nematodes are those types of nematodes that is non-parasitic to plants. They feed on other micro-organisms that dwells in the soil ecosystem. With the help of genome sequencing, many genes have been identified from *Caenorhabditis elegans* which is also known as biological model organism. Following are some of the functional genes that have been identified.

1. *nuc*-1= nuclease deficient, degradation of cell corpes
2. *ced*-1 and *ced*-2 = removal of cell corpse

3. *ced*-5, *ced*-6, *ced*-7, *ced*-10, *ces*-1 and *ces*-2 = cell death specification (control the death fate of specific set of cells)
4. *ced*-9 = protects cells from programmed cell death
5. *daf*-2 and *age*-1 = controls the entry of developing nematodes into the alternate dauer juvenile stage

Role in disease management

Genome sequencing in nematode can be used as useful tool for reducing diseases caused by nematodes. Proving transgenic plant will serve better in managing nematode diseases.

Conclusion

Few nematode genomes have been completely sequenced such as *Caenorhabditis. elegans* with genome size 100Mb and *Meloidogyne hapla* having 54Mb. With the help of genome sequencing, similarities between nematodes as well as with other different organisms can be identified. It has been reported that nematode (*C. elegans*) is similar to human (*Homo sapiens*) upto 36% while from human it is 74%, likewise nematode is similar to yeast (*Saccharomyces cerevisiae*) and bacteria (*Escherichia coli*) upto 26% and 9.1%, respectively and from yeast and bacteria it was found to be 51% and 26%, respectively.

Identification of both functional and non-functional genes provides knowledge based on mechanism of parasitism for plant parasitic nematode, in case of free-living nematode (*C. elegans*) it helps to know which type of gene is responsible for different physiological function of the nematode. In future, when transgenic plants will be commercially available, then it can reduce the disease cause by nematode alone and can also reduce the disease intensity caused by association with other organisms. If this progress is made in the future, it can help atleast 60% to the field of agriculture ultimately, proving boon to our country's economy.

REFERENCES

Annonymous 2009. "Saccharomyces Genome Database". Yeastgenome.org.

Doncaster, C.C. 1971. Feeding in plant parasitic nematodes: Mechanisms and behavior. Pp.137-157.In B. M. Zuckerman, W. F. Mai, and R. A. Rohde, eds. Plant Parasitic nematodes Vol II. New York: Academic Press

Fiers, W., Contreras, R., Duerinck, F., Haegeman, G., Iserentant, D., Merregaert, J., Min Jou, W., Molemans, F., Raeymaekers, A., Van Den Berghe, A., Volckaert, G. and Ysebaert, M. 1976. "Complete nucleotide-sequence of bacteriophage MS2-RNA - primary and secondary structure of replicase gene". *Nature,* 260 (5551): 500–507.

Fleischmann, R., Adams, M., White, O., Clayton, R., Kirkness, E., Kerlavage, A., Bult, C., Tomb, J., Dougherty, B. and Merrick, J. 1995. "Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd". *Science,* 269 (5223): 496–512.

Greilhuber, J., Borsch, T., Müller, K., Worberg, A., Porembski, S. and Barthlott, W. 2006. "Smallest angiosperm genomes found in Lentibulariaceae, with chromosomes of bacterial size". *Plant Biology,* 8 (6): 770–777.

Hussey, R.S., Davis, E.L. and Baum, T.J. 2002. Secrets in secretions: genes that control nematode parasitism of plants. *Braz. J. Plant Physiol*., 14(3): 183-194.

Hussey, R.S., Paguio, O.R. and Seabury, F. 1990. Localization and purification of a secretory protein from the esophageal glands of *Meloidogyne incognita* with monoclonal antibodies. *Phytopatho*., 80: 709-714.

Lang, D., Zimmer, A.D., Rensing, S.A. and Reski, R. 2008. "Exploring plant biodiversity: the *Physcomitrella* genome and beyond". *Trends Plant Sci.,* 13 (10): 542 549.

Maxam, A.M. and Gilbert, W. 1977. "A new method for sequencing DNA". *Proc. Natl. Acad. Sci. U.S.A.* 74 (2): 560–564.

Pearson, H. 2006. "Genetics: what is a gene?". *Nature,* 441 (7092): 398– 401.

Pennisi, E. 2007. "DNA Study Forces Rethink of What It Means to Be a Gene". *Science,* 316 (5831): 1556–1557.

Ridley, M. 2006. *Genome*. New York, NY: Harper Perennial. ISBN 0-06-019497-9.

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchison, C.A., Slocombe, P. M. and Smith, M. 1977. "Nucleotide sequence of bacteriophage phi X174 DNA". *Nature,* 265 (5596): 687–695.

Smant, G. *et al*. 1998. Endogenous cellulose in animals: Isolation of beta-1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proc Natl Acad Sci*., USA 95: 4906-4911

Staden, R. 1979. "A strategy of DNA sequencing employing computer programs." *Nucleic Acids Research*, 6 (7): 2601–2610.

The *C. elegans* Sequencing Consortium 1998. Genome Sequence of the nematode *C. elegans*: A Platform for investigating biology. *Science,* 282: 2012-2017.

Wyss, U. and Zunke, U. 1986. Observation on the behavior of second stage juveniles of *Heterodera schachatii* inside host roots. *Review de Nematologie,* 9: 153-165.

****