Gene 499 (2012) 202–208

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short Communication

Genome wide computational analysis of *Brugia malayi* helicases: A comparison with human host

Renu Tuteja [a,*], Abulaish Ansari [a], Anita [b,1], Manish Kumar Suthar [b,1], Jitendra Kumar Saxena [b,1]

[a] *Malaria Group, International Centre for Genetic Engineering and Biotechnology, P.O. Box 10504, Aruna Asaf Ali Marg, New Delhi-110067, India*
[b] *Division of Biochemistry, CSIR-Central Drug Research Institute, Lucknow, India*

ARTICLE INFO

Article history:
Accepted 20 February 2012
Available online 28 February 2012

Keywords:
DEAD box
DEAH box
Helicase
Parasite
Brugia malayi
Unwinding

ABSTRACT

The availability of *Brugia malayi* genome sequence has paved ways for the search of homologues for a variety of genes. Helicases are ubiquitous enzymes involved in all the nucleic acid metabolic pathways and are essential for the development and growth. The genome wide analysis of *B. malayi* for different helicases showed the presence of a number of DEAD box helicases, 7 DEAH box helicases, RecQ helicases, repair helicases, super killer helicases, MCM2-7 complex, Rad54 and two subunits of Ku helicase. The comparison of protein sequence of each helicase with its human counterpart indicated characteristic differences in filarial helicases. There are noticeable differences in some of the filarial helicases such as DHX35, RecQL1 and Ku. Further characterization of these helicases will help in understanding physiological significance of these helicases in filarial parasites, which in future can be utilized for chemotherapy of parasitic infection.

1. Introduction

Human lymphatic filariasis (LF), an infectious disease caused by lymph-dwelling nematode, is transmitted by mosquitoes. About 1.2 billion people living in tropical and subtropical areas of more than 83 countries worldwide are affected by this disease (Blaxter et al., 2002; WHO, 2010). The Global Programme to Eliminate Lymphatic Filariasis (GPELF) was established in 1999 with the objective of interrupting transmission of the parasites in all endemic countries by 2020 (Addiss, 2010; Freedman, 1998; Michael, 1999).

Many species of filarial parasites are known, each relatively specific for its host. *Wuchereria bancrofti*, *Brugia malayi*, *Onchocerca volvulus*, *Dipetalonema perstans*, *Dipetalonema streptocerca*, *Loa loa* and *Mansonella ozzardi* are the species responsible for producing infestations in man (Manson-Bahr and Apted, 1982). The life cycle of filarial parasites includes three distinct phases viz. microfilariae (mf), infective larva (La) and adults. *Mansonia* species of mosquito propagates *B. malayi*, while *Culex quinquefasciatus* is responsible for transmission of *W. bancrofti*. The adult parasites reside in connective tissues, muscles, circulatory or lymphatic system of the host. The microfilariae (sheathed/unsheathed) released into peripheral blood of the host, have a life span of 14–70 days and exhibit nocturnal/diurnal periodicity.

The filarial genome project has used *B. malayi* as model filarial nematode. Many genes discovered as EST are now being utilized for immunodiagnostic, vaccine and drug target potential. The analysis of filarial ESTs is aided by the complete genome sequence of *Caenorhabditis elegans* which is a good model for the basic biochemistry of a nematode (Blaxter, 1998). Filarial parasites, including *B. malayi*, carry three genomes viz; nuclear, mitochondrial (available at GenBank accession no. AF538716) and that of an alpha proteobacterial endosymbiont, *Wolbachia*. The *B. malayi* nuclear genome is organized as five chromosomes and estimated to be 90 to 95 Mb (Blaxter et al., 2002; Ghedin et al., 2007).

Numerous molecular targets are known for filarial parasites and few of them have been utilized as potent chemotherapeutic target (Mishra et al., 2005; Muller et al., 1988; Nagarajan et al., 1988; Singh et al., 2007; Singh et al., 2008). DNA topoisomerases are involved in managing the topological problems associated with DNA replication, transcription, recombination and chromatin remodeling by introducing temporary single or double stranded breaks in the DNA. The presence of DNA topoisomerases in various filarial parasites and their different life stages viz. *B. malayi*, *Setaria cervi* and *Acanthocheilonema viteae* has been demonstrated and used as target for development of new antihelminthic compounds (Chandra et al., 2004; Kumar et al., 2008; Pandya et al., 1999). Recently hexokinase has been shown as potential chemotherapeutic target (Singh et al., 2008).

RNA/RNA and RNA/DNA helicases are essential for many cellular processes such as transcription, translation, replication, repair and

Abbreviations: ATP, adenosine tri-phosphate; HRDC, helicase and RNaseD C-terminal; La, infective larva; LF, lymphatic filariasis; MDA, Mass Drug Administration; mf, microfilariae; MCM, minichromosome maintenance; NER, nucleotide excision repair; RTS, Rothmund–Thomson syndrome; ss, single-stranded; SF, superfamilies; SKIV2L, superkiller viralicidic activity 2-like; BS, Bloom's syndrome; WS, Werner syndrome; GPELF, The Global Programme to Eliminate Lymphatic Filariasis; TFIIH, transcription factor II H; XPB, xeroderma pigmentosum group B; XPD, xeroderma pigmentosum group D.

* Corresponding author. Tel.: +91 11 26741358; fax: +91 11 26742316.
E-mail addresses: renu@icgeb.res.in, renututeja@gmail.com (R. Tuteja).
[1] Tel.: +91 0522 2625932; fax: +91 0522 2623405, 2623938.

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2012.02.036

RNA splicing processes requiring at least a transient unwinding of the nucleic acids to provide an appropriate substrate (Matson and Kaiser-Rogers, 1990; Matson et al., 1994; Sancar and Sancar, 1988; Tuteja and Tuteja, 2004a, 2004b; Umate et al., 2011). RNA helicases, which unwind helical secondary structures of RNA molecules, are classified into families viz. DEAD-box proteins, DEAH-box proteins and DExH-box proteins and all of them contain at least eight characteristic sequence motifs including the ATP-hydrolysis motif II (Jankowsky and Jankowsky, 2000). The increasing number of proteins in the DEAD/H family may reflect different functions, RNA substrates, or tissue specificities of RNA helicases. Glh1 protein of *C. elegans* is a type of DEAD-box helicase which shares 399 amino acid conserved region (Roussell and Bennett, 1993). The presence of a zinc finger domain at amino terminal distinguishes glh1 from other known RNA helicases. Dna2 helicase was first identified in *Saccharomyces cerevisiae* which plays a key role in DNA replication as it possesses single-stranded (ss) DNA-dependent ATPase, DNA helicase and ssDNA-specific endonuclease activities. Dna2 proteins are well conserved throughout eukaryotes. A null mutation of *C. elegans dna2* causes a delayed lethality, allowing survival of some mutant *C. elegans* adults to F2 generation (Kim et al., 2005). It is well known that helicases are also present in genomes of a number of cellular pathogens, viruses and malaria parasite (Frick and Lam, 2006; Tuteja, 2010). These enzymes are utilized for control of viral and malarial infections and can serve as feasible drug targets for the filaria (Frick and Lam, 2006; Tuteja, 2007).

For the last 20 years the drug treatments for filariasis have not changed considerably but the drug resistance is rising with time (Campbell, 1982). Currently in Mass Drug Administration (MDA) program a single dose of 6 mg/kg DEC plus 400 mg albendazole is recommend by WHO (Bhumiratana et al., 2010). Moxidectin, a drug to be used for controlling onchocerciasis might be effective against lymphatic filariasis (Bockarie and Deb, 2010).

Given the importance of helicases in various systems and in order to obtain an insight of helicases in *B. malayi* genome, we have carried out genome wide computational analysis using the bioinformatics approaches. In the following sections we report the detailed analysis of almost all the helicases from *B. malayi* and their comparison with the human host.

2. Materials and methods

All the sequence data used in the analysis were downloaded from *B. malayi* genome database available at www.genome.jp/dbget-bin/www_bfind?B.malayi. The downloaded sequences were used as query to match with the human homologue using BLAST search (www.ncbi.nlm.nih.gov). The corresponding human sequence was retrieved and various domains were manually assigned. Similarly the domains were also assigned manually in *B. malayi* sequence and the data are presented in figures. The domain analysis was done using Scan Prosite at (http://expasy.org). The domain structure was taken from the results and used in the figures. Similarly the human sequence was also analyzed and the results are presented in figures.

3. Results and discussion

3.1. All helicases

The genome of filarial nematode parasite *B. malayi* available at www.genome.jp/dbget-bin/www_bfind?B.malayi was investigated using the word 'helicase' as query. The results of this search presented in Supplementary Table 1 showed a total of 63 hits. Further evaluation of this list indicated that it included some non helicase genes also and only 44 genes of this list are real helicases. All the 44 helicase members from this list and some additional helicases identified using BLAST search and their further analysis are presented under different categories in the following sections. Gorbalenya and Koonin (1993) classified RNA and DNA helicases into superfamilies (SF) 1–5, based on the presence of conserved motifs. Two essential motifs known as Walker A or motif I and Walker B or motif II are characteristic of all of these proteins. Members of SF 1 and 2 contain 6 additional motifs and are active as monomers or dimers as opposed to SF 3, 4 and 5 members which are active as hexamers (Tanner and Linder, 2001). UvrD-like and *E. coli* Rep DNA helicase belong to SF1, SF2 contains all the DEAD and DEAH box helicases and Rho and *E. coli* DnaB belong to SF4 and SF5. DEAD box helicases are commonly named as DDX and all other helicases are named as DHX (Abdelhaleem et al., 2003). The detailed analysis of helicases of *B. malayi* was done using the information on the website (http://www.dexhd.org) and the results are presented in the following sections.

3.2. DEAD box helicases

The family of DEAD box helicases is the best and widely characterized. These helicases belong to SF2 and are involved in almost all the aspects of RNA metabolism such as pre mRNA splicing, ribosome biogenesis, nucleocytoplasmic transport, nuclear transcription, organellar gene expression, translation and RNA decay (Linder, 2006). The DEAD-box proteins core reveals no substrate specificity for nucleic acid binding. It is most likely that the N and C-terminal regions and the intervening sequences provide help and specificity for binding to other proteins (Cordin et al., 2006; Rocak and Linder, 2004). The *B. malayi* genome was investigated using DEAD-box helicase as query and 14 hits were found (Supplementary Table 2). According to the information on human and yeast homologues (http://www.dexhd.org/eukdead) these helicases have been further classified (Table 1) and described in detail in the following sections.

a. Bm1_11355. It is a homologue of human DDX1 (Table 1) and is almost similar in size. The sequence of all the domains and the comparison with the human counterpart show that except domain Ia all the other domains contain highly conserved similar sequences (Fig. 1, 1a and 1b).
b. Bm1_16185. It is a homologue of human DDX3Y (Table 1) and is slightly larger in size and contains longer N-terminal region (Supplementary Fig. 1 and 1a). It is located on the Y chromosome in human and is highly homologous to DDX3X, which is present on the X chromosome. The sequence comparison of all the domains with human counterpart indicates that except domain Ia all the other domains are highly similar (Supplementary Fig. 1 and 1a and 1b).
c. Bm1_51245. It is a homologue of human DDX4 (Table 1) and is slightly smaller in size and has a smaller N-terminal region as compared to its human counterpart (Supplementary Fig. 1 and 2a and 2c). The sequence of all the domains and the comparison with the human counterpart show that except domains I, II, III and V, all the other domains contain some variation in the conserved sequence (Supplementary Fig. 1 and 2a and 2c).
d. Bm1_08350. It is also a homologue of human DDX4 (Table 1) and is slightly larger in size and has a longer N-terminal region as compared to its human counterpart (Supplementary Fig. 1 and 2b and 2c). The sequence of all the domains and the comparison with the human counterpart show that except domains I, II, III and VI, all the other domains contain some variation in the conserved sequence (Supplementary Fig. 1 and 2b and 2c).
e. Bm1_52675. It is a homologue of human DDX5 (Table 1) and is slightly smaller in size but has a longer N-terminal and a shorter C-terminal as compared to its human counterpart (Supplementary Fig. 1 and 3a and 3b). The sequence of all the domains and the comparison with the human counterpart