Biotechnology Advances 27 (2009) 122–132

Contents lists available at ScienceDirect

Biotechnology Advances

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

Research review paper

Improved insights into the transcriptomes of the human hookworm *Necator americanus* — Fundamental and biotechnological implications☆

Elida M. Rabelo[a,b,*], Ross S. Hall[a], Alex Loukas[c], Leanne Cooper[c], Min Hu[a], Shoba Ranganathan[d,e], Robin B. Gasser[a,*]

[a] Department of Veterinary Science, The University of Melbourne, 250 Princes Highway, Werribee, Victoria 3030, Australia
[b] Departamento de Parasitologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Avenida Presidente Antônio Carlos 6627, Campus Pampulha, CEP 31270-901 Belo Horizonte, Minas Gerais, Brazil
[c] Division of Infectious Diseases, Queensland Institute of Medical Research, Brisbane, Queensland 4006, Australia
[d] Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia
[e] Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119260

ARTICLE INFO

Article history:
Received 28 August 2008
Received in revised form 4 October 2008
Accepted 6 October 2008
Available online 18 October 2008

Keywords:
Human hookworm
Nematoda
Strongylida
Transcriptomics
Genomics
Bioinformatics

ABSTRACT

Hookworms of humans are blood-feeding parasitic nematodes of major socio-economic significance in a wide range of countries. They cause a neglected tropical disease (NTD) called "hookworm disease" (=necatoriasis and/or ancylostomiasis). *Necator americanus* is the most widely distributed hookworm of humans and is a leading cause of iron deficiency anaemia, which can cause physical and mental retardation and deaths in children as well as adverse maternal–foetal outcomes. Currently, there is a significant focus on the development of new approaches for the prevention and control of hookworms in humans. Technological advances are underpinning the discovery of drug and vaccine targets through insights into the molecular biology and genomics of these parasites and their relationship with the human host. In spite of the widespread socio-economic impacts of human necatoriasis, molecular datasets for *N. americanus* are scant, limiting progress in molecular research. The present article explores all currently available EST datasets for adult and larval stages of *N. americanus* using a semi-automated bioinformatic pipeline. In the current repertoire of molecules now available, some have been or are being considered as candidate vaccines against *N. americanus*. Among others, the most abundant sets of molecules relate to the pathogenesis-related protein (PRP) superfamily, comprising various members, such as the *Ancylostoma*-secreted or activation-associated proteins (ASPs) and the kunitz-type proteins, both of which are inferred to play key roles in the interplay between *N. americanus* and the human host. Understanding the molecular biology of these and other novel molecules discovered could have important implications for finding new ways of disrupting the pathways that they are involved in, and should facilitate the identification of new drug and vaccine targets. Also, the bioinformatic prediction of the essentiality of genes and gene products as well as molecular network connectivity of nematode-specific genes, together with sequencing by 454 technology, are likely to assist in the genomic discovery efforts in the very near future, to also underpin fundamental, molecular research of hookworms.

© 2008 Elsevier Inc. All rights reserved.

Contents

1. Introduction . . . 123
2. Expressed sequence tag (EST) datasets and bioinformatics. . . . 123
3. Bioinformatic exploration of new EST data derived from the adult stage of *N. americanus* . . . 124
   3.1. Initial bioinformatic processing and analysis of the EST dataset . . . 124
   3.2. Gene ontology (GO) classifications and pathway mapping . . . 126
   3.3. Predictions of gene functions based on 'loss-of-function' phenotypes in *C. elegans* . . . 126
4. Comparative analysis of publicly available EST datasets representing the adult or L3 stage of *N. americanus* . . . 126

☆ Note: New nucleotide sequence data produced for this paper are available in the GenBank database under accession numbers GE624935–GE626596.

* Corresponding authors. Department of Veterinary Science, The University of Melbourne, 250 Princes Highway, Werribee, Victoria 3030, Australia.
E-mail addresses: elidam.rabelo@gmail.com (E.M. Rabelo), robinbg@unimelb.edu.au (R.B. Gasser).

0734-9750/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.biotechadv.2008.10.002

5. Combined analysis of all currently available EST data for *N. americanus* . . . 129
6. Biotechnological implications and concluding remarks . . . 129
7. Authors' contributions . . . 131
Acknowledgements . . . 131
References . . . 131

1. Introduction

Soil-transmitted helminths (STHs) are of major socio-economic and human health importance in a wide range of countries. Many of them cause neglected tropical diseases (NTDs) (Hotez, 2007). Estimates indicate that more than 1 billion people are infected with these helminths (Hotez et al., 2005; Bethony et al., 2006), and that up to 135,000 deaths occur annually. Morbidity from nematodes is substantial and surpasses lung cancer and diabetes in disability adjusted life years (DALY) (Hotez et al., 2006). Hookworms, including *Necator americanus* and *Ancylostoma duodenale*, are destructive parasites (WHO, 2005). They are dioecious nematodes which inhabit the small intestine of their human host. The infective, third-stage (filariform) larvae (L3) penetrate human skin and migrate *via* the circulatory system and lung to finally reside as adults usually in the duodenum. The adults attach *via* their buccal capsule to the intestinal mucosa, rupture capillaries and feed on blood. The pathogenesis of hookworm infection is mainly a consequence of the blood loss, which occurs during attachment and feeding. In a large number of developing countries, hookworms are the leading cause of iron deficiency anaemia, which, in heavy infections, can cause physical and mental retardation and sometimes deaths in children as well as adverse maternal–foetal outcomes (Hotez et al., 2005; WHO, 2005).

Presently, the control of hookworms relies mainly on the treatment of humans with anthelmintic drugs, such as albendazole, mebendazole, pyrantel pamoate and/or levamisole (see Bethony et al., 2006). There is a potential risk of human hookworms developing genetic resistance to these drugs (if they are used excessively) and, based on the experience with nematodes of livestock (Wolstenholme et al., 2004), there has been a significant focus on the development of a hookworm vaccine (Hotez et al., 2005, 2006; Loukas et al., 2006) through the auspices of the Human Hookworm Vaccine Initiative (Hotez et al., 2003). Also advances in the development of genomic and bioinformatic tools (e.g., Mitreva et al., 2007; Moser et al., 2005; Ranjit et al., 2006; Nagaraj et al., 2007a,b) are underpinning the discovery of key molecules in hookworms, which should provide insights into fundamental aspects of their development, reproduction and their relationship with the host. Interestingly, more genomic studies of hookworms of animals, such as *A. caninum*, have been performed (Mitreva et al., 2005a; Moser et al., 2005; Abubucker et al., 2008; Datu et al., 2008) than of human hookworms. For instance, more than 80,000 expressed sequence tags (ESTs) of *A. caninum* are now available in public gene databases. In contrast, in spite of the socio-economic impact specifically of human hookworm disease (Bethony et al., 2006; Loukas et al., 2006), genomic datasets available for *A. duodenale*, *A. ceylanicum* and *N. americanus* are limited (cf. Daub et al., 2000; Parkinson et al., 2004; Ranjit et al., 2006), and they have not been analysed extensively. Therefore, in the present article, we extend previous studies by generating an additional EST dataset for the adult stage of *N. americanus*, subject these data to detailed bioinformatic exploration and undertake comparative analyses of all publicly available datasets for this parasitic nematode.

2. Expressed sequence tag (EST) datasets and bioinformatics

ESTs were determined from a cDNA library constructed from adult *N. americanus*. An experimental line ("Shanghai strain") of *N. americanus* (originally isolated in Shanghai, China), kindly provided by Drs Bin Zhan and Peter Hotez, was maintained in hamsters at the George Washington University. Total RNA from adult *N. americanus* (n = 12 worms; six of each sex) was extracted using Trizol (Invitrogen), following the manufacturer's instructions. Ten µg of this RNA was used as a template for the synthesis of double-stranded DNA using the SMART cDNA kit (BD Bioscience), after which the cDNA was modified with adapters and cloned into the *Sfi* I site of the pTriplEx2 plasmid (BD Bioscience) and packaged into λ arms. The titre and percentage of recombinant phages in the library were determined using the recommended protocols. *Escherichia coli* (strain BM25.8) was transduced with recombinant phage, from which the excision of the pTriplEx2 phagemid library was accomplished. Colonies obtained from the excised library were screened by the polymerase chain reaction (PCR) to verify that plasmids had inserts. The excised library was then sequenced, employing a 3730xl DNA analyzer (Applied Biosystems). The TempliPhi DNA Sequencing Template Amplification system (GE Healthcare) was used to sequence each clone using the 5′ λ TriplEx2 sequencing primer. The EST data were initially analysed and annotated using the semi-automated platform ESTExplorer (Nagaraj et al., 2007a), available at http://estexplorer.els.mq.edu.au/estexplorer/index.php. This pipeline includes a suite of programs that treat sequences in an ordered way. In phase I, all ESTs were pre-processed (SeqClean, without RepeatMasker), aligned/clustered using the Contig Assembly Program CAP3, employing a minimum sequence overlap length "cut-off" of 30 bases and an identity threshold of 95% for the removal of flanking vector and adapter sequences, followed by assembly. In phase II, gene ontologies (GO) were inferred (at the DNA-level) using BLAST2GO (Conesa et al., 2005). In phase III, the sequences (contigs or singletons) representing each cluster of ESTs were then conceptually translated into peptides using ESTScan. Following InterProScan (to define protein domains or motifs), the peptides were mapped to respective pathways in *Caenorhabditis elegans* using KEGG (release 40) Orthology-Based Annotation System (KOBAS). Inferred proteins were subjected to homology searches in the National Center for Biotechnology Information (NCBI) using the program BLASTX. Homologies with an *e*-value of ≤1e−05 were recorded as significant.

In addition, four other *N. americanus* EST datasets from previous studies were available from GenBank, including 272 from the adult stage (Daub et al., 2000; Parkinson et al., 2004) and 266 derived from adult gut tissue (Ranjit et al., 2006), 345 from fourth-stage larvae and 4149 from third-stage infective larvae (Parkinson et al., 2004). These data were analysed separately or combined with the new dataset generated in the present study. Sequences were queried (at the amino acid level) against three different databases containing protein sequences from different organisms. Data were compared with protein sequences available for (i) *C. elegans* (from WORMPEP v.167 *via* Wormbase), (ii) parasitic nematodes (peptide sequences from conceptually translated ESTs) and (iii) organisms other than nematodes (from the NCBI non-redundant protein database) (Benson et al., 2007). Sequences were classified as '*N. americanus* adult' (=NaAd) or '*N. americanus* L3' (=NaL3), followed by the contig or singleton number. For the L3 singletons, the original NCBI numbers were maintained. Peptides (representing clusters of ESTs) with no homology following BLASTX analysis were then subjected to analysis using the programs SignalP 3.0 (Bendtsen et al., 2004) and TMHMM (a membrane topology prediction program; Krogh et al., 2001) to predict signal and transmembrane peptides, respectively. These same peptides were also subjected to a homology search against the "dbEST — others" database (www.ncbi.nlm.nih.gov/dbEST/) using TBLASTX.

Download English Version:

https://daneshyari.com/en/article/14406

Download Persian Version:

https://daneshyari.com/article/14406

Daneshyari.com