Characterization and phylogenetic analysis of NDR domain-containing protein from *Strongyloides ratti* using Degenerate Primers-Based Polymerase Chain Reaction

Abbas Jolodar*

Department of Basic Sciences, Biochemistry and Molecular Biology Section, Faculty of Veterinary Medicine, Shahid Chamran University of Ahvaz, Ahvaz, Iran

Running title: Characterization of NDR domain-containing protein from *Strongyloides ratti*

* Corresponding author: Dr. Abbas Jolodar, Associate Professor in Faculty of Veterinary Medicine, Shahid Chamran University of Ahvaz, 61355-145, Ahvaz, Iran.
Phone: +98-61-33330073; Fax: +98-61-33360807; jolodara@scu.ac.ir

Abstract

The intestinal nematode *Strongyloides ratti* is important as an optimal model the human pathogen, *S. stercoralis*. A pair of degenerate primers was designed to amplify the target gene associated to the NDR domain-containing proteins based on the two conserved regions using the related nematode protein sequences. The putative NDR domain-containing gene fragment of 249 bp from the *S. ratti* was amplified using Touchdown RT-PCR. The results showed that the amplified fragment between these two motifs from *S. ratti* with the C-terminus part of reference protein has 66.3% similarities. By searching the EST (Expressed Sequence Tags) GenBank database, an overlapping 1082 bp cDNA fragment named SrNDR was compiled and found to contain a 972-nucleotide open reading frame that encodes a protein of 324 amino acids of 23.4 kDa expected molecular mass and a calculated pI of 7.64. The phylogram of SrNDR revealed that the gene variants of *S. ratti* SrNDR were grouped together with a strong bootstrap score of 87. Domain search analysis, showed the e-values of 1.05e-100 with the conserved NDR1 domain-containing of α/β hydrolase superfamily protein (pfam03096) between amino acid residues 9 to 287. The core domain of SrNDR folded with 12 alpha-helical conformation surrounded by 8 stranded antiparallel-beta-sheet. The three-dimensional structure model of SrNDR (residues 31-319) showed 100% similarity with human protein NDRG1 in 248 amino acids with 77% coverage. By using this program, we found a prediction of mostly *a*-helical (44.86%), Strand (1.08%) and random coils (54.05%) for the coding sequence of SrNDR. Based on the alignment of this protein with the corresponding nematodes, it can play a vital role in the survival of the parasite in the host cells.

Key words: *NDR domain-containing protein, Touchdown RT-PCR, Degenerate primers*

1. Introduction

The intestinal nematode *Strongyloides ratti*, a common parasite of rats is parasitic nematodes with a unique life-cycle characterized by obligate parasitic and facultative free-living generations (Bethony et al., 2006; Buonfrate et al., 2020). This parasite is important as a suitable model parasite which is genetically very similar to the human pathogen, *S. stercoralis*.

The N-myc downstream regulated gene (NDRG) protein family consists of 4 members, NDRG1-4 in human with 57-65% amino acid homology (Okuda and Kondoh, 1999). They are belongs to α/β hydrolase superfamily (ABHD) which is well conserved through evolution. It has been shown that the excretory and secretory (ES) products of parasitic nematodes such as *Haemonchus contortus* (Wang et al., 2019), *Heligmosomoides polygyrus* (Moreno et al., 2011), and *Mesocestoides corti* (Vendelova et al., 2016) were rich in ABHD proteins. Similar expressions of ABHD proteins/homologs were also demonstrated in free-living and parasitic parasites such as *Caenorhabditis elegans* ABHD5 (Xie and Roy, 2015), *Cryptosporidium parvum* Type II thioesterase (CpTEII) (Guo et al., 2019), and *Schistosoma japonicum* lysophospholipase (Fan et al., 2008). Apart from the expression of ABHD proteins in a wide range of mammals, the expression of a small amount of this proteins and its homologs in plants such as *Arabidopsis thaliana* ABHD11 and ABHD5 (Vijayakumar et al., 2016; Park et al., 2013), and *Saccharomyces cerevisiae* ABHD5 homologs (Ghosh et al., 2008) has also been reported. Members of this family are also found in wide variety of multicellular eukaryotes, including an Ndr1 type protein in *Helianthus annuus* (Common sunflower), known as Sf21.

Although the exact function of NDR domain-containing proteins has not been clearly explained, the recent evidence shows that mutation in these genes are associated with different neurological and physiological complications. It was predicted that this protein is a cytoplasmic protein associated in energy metabolism, stress responses, hormone responses, cell signaling, growth and differentiation (Qu et al., 2002). As indicated that NDRG1 is primarily localized in the cytoplasm, followed by its localization in the nucleus and mitochondrion, at probabilities of 47.8%, 26.1% and 8.7%, respectively NDRG1 is translocated from the cytoplasm to the nucleus in response to DNA damage and hypoxia (Bae et al., 2013). As reviewed by Fang et al., NDRG1 is also involved in embryogenesis and development, cell growth and differentiation, lipid biosynthesis and myelination, stress responses, immunity, DNA repair and cell adhesion among other functions. As shown in the literature review (Kovacevic and Richardson, 2006; Sahni et al., 2019), NDRG1 has been shown to be an effective inhibitor of metastasis signaling in a number of invasive cancers such as cancer prostate, pancreatic, breast, and colon cancers.

In this study, we aimed to do a preliminary characterization of a NDR domain-containing protein from *S. ratti* to gain a better understanding of parasite-host interaction.

2. Materials and Methods

2.1 RNA isolation and cDNA synthesis

Total RNAs from larvae *S. ratti* were prepared using the RNX plus solution (CinnaGen, Iran) according to the manufacturer's instructions. The extracted total RNA was calculated by absorbance at 260 nm and utilized directly or kept at -80°C. Briefly, 12 µl (2 μg each) of total RNA was incubated with 0.5 μg of ModT (modified oligodT) (5´-GGGTCTAGAGCTCGAGTCACTTTTTTTTTTTTTTT-3´) primer at 70°C, for 10 min.

The reaction was located on ice before adding 1 µl RNasin (CinnaGen, Iran), 1 µl dNTP mixture (120 mM of each nucleotide), 2.5µl of 5X enzyme buffer and 1 µl (200 U) of Moloney Murine Leukemia Virus (M-MulV) reverse transcriptase (CinnaGen, Iran). The reaction was incubated at 42°C for 1 h before inactivation at 70°C. Reverse transcriptase was

2.2 Sequences obtained from GenBank database

Sequences associated to the NDR domain-containing proteins were obtained from the GenBank database, after an extensive study related to target genes in several nematode species. The assemblies from different species included in the family of *Strongyloides* were retrieved, to make a database of proteins associated with *S. ratti*. The maximum approved e-value in this study was 8e-07; sequence scores less than this value were rejected.

2.3 RT-PCR amplification using degenerate primers

On the basis of the preserved sequences, two degenerate primers were made. To generate a homologous DNA fragment for the isolation of the *S. ratti* NDR domain-containing protein, a fragment of the gene was amplified from cDNA in the polymerase chain reaction (PCR) using degenerate primers (Jolodar, 2019). Primer sequences were: sense primer sequence NDR-F (5′-GCGgaattcGGNGCNTGGGAYTA), related to N-terminal region of protein (GAWDY) and antisense primer sequence NDR-R (5′-GCGaagcttCCRCANCCYTTRCA), and related to C-terminal region of protein (CKGCG). (N, R, and Y represent A/G/C/T, A/G and C/T, respectively). To facilitate the directional cloning, *Eco*RI and *Bam*HI were introduced to the 5´ends of the forward and reverse primers, respectively. Semi-Nested RT-PCR reactions were done for cDNA by use of touchdown PCR (TD-PCR) conditions. In order to facilitate Touchdown Semi-Nested RT-PCR, cDNA was produced using the modified oligo(dT) (Mod-T). This technique let us the amplification of the right transcripts in a two-round PCR protocol. For performing the first round of amplification, ModT-R primer (5´-CCCAGATCTCGAGCTCAGTG) was designed. This was complemented to the 5'-end tail of the modified oligodT (ModT) primer. TD-PCR reactions were performed the following conditions: 94°C×3 min for one cycle; 94°C × 30s, 42°C × 50s, 72°C × 1 min for 15 cycles by increasing 1°C per cycle; 94°C × 30s, 57°C × 50s, 72°C× 40s for 20 cycles. The first round of Touchdown PCR (TD-PCR) was done by ModT-R and NDR-F primers and the original cDNA as template. The second round of PCR was done by using NDR-F and NDR-R primers and the diluted first PCR (one-tenth) as templates. The PCR conditions for second rounds were 30 cycles with denaturation at 94°C for 0.5 min, annealing at 57°C for 0.5 min, and extension at 72°C 1.0 min and final extension at 72ºC for 5 min. PCR reaction was performed using cDNA from *S. ratti* as template with the mix: Tris-HCl 20mM, KCl 50mM, MgCl2 1.4mM, dNTPs 0.2mM of each, 0.5 U *Taq* Polymerase, each primer at 0.4 µM, and 5 µl of cDNA. Analytical agarose gel containing 50 µg/ml of safe stain was prepared. PCR products were fractionated on 1% agarose gel electrophoresis and detected by UV transilluminatore.

2.4 DNA sequence analysis

The amplified cDNA fragments were subjected to sequencing in both strands using a dideoxy termination method and run on an Applied Biosystems 373 DNA sequencer. The complete cDNA was determined by using overlapping fragments. The primer was designed using Primer3 program. The sequence was compared with sequences in the database using the BLAST algorithm from the NCBI website (ncbi.nlm.nih.gov). To evaluate the evolutionary relationships between NDR domain-containing genes, we retrieved target sequences of the other species from the NCBI GenBank. The putative signal peptides were analyzed using SignalP software (Nielsen et al., 1997) (cbs.dtu.dk/services/SignalP). The multiple alignments were made using the CLUSTAL_W program (Thompson et al., 1994) and edited with the BOXSHADE software (ch.embnet.org/software/BOX_form.html). The CDD-Search software from the NCBI site was used to determine the conserved domains (Marchler-Bauer et al, 2017). The molecular weight (MW) and theoretical isoelectric point (pI) of the deduced amino acid sequences were analyzed using the export protein analysis program (ca.expasy.org/tools). The secondary structure of protein was predicted using the PSIPRED Protein Sequence Analysis Workbench (bioinf.cs.ucl.ac.uk/psipred) (Buchan and Jones, 2019). It was used DAS-Transmembrane Prediction server (tmdos.bioinfo.se). The 3D structure prediction was performed by the Phyre2 program (Kelley and Sternberg, 2009). Phylogenetic analysis and genetic distance was carried out by the “neighbor-joining” method bootstrap tests 1000 using MEGA11 software (Tamura et al., 2007).

3. Results

3.1 Semi-Nested RT-PCR amplification using degenerate primers

NCBI GenBank protein database were searched using keyword "NDR domain-containing protein nematodes". Twelve protein sequences including *Wuchereria bancrofti* (EJW87670.1), *Brugia malayi* (XP_001899587.2), *Litomosoides sigmodontis* (VDM91312.1), *Loa loa* (XP_020302625.1)], *Acanthocheilonema viteae* (VBB26456.1)], *Onchocerca flexuosa* (VDO39025.1), *Litomosoides sigmodontis* (VDM91312.1), *Ditylenchus destructor* (KAI1725449.1), *Caenorhabditis elegans* (NP_510634.1), *Haemonchus contortus* (CDJ84079.1), *Nippostrongylus brasiliensis* (WKY14995.1), *Oesophagostomum dentatum* (KHJ98162.1) and *Cylicocyclus nassatus* (CAJ0609661.1) were retrieved and used in the alignment. Based on the multiple alignments using sequence data of reference sequences shows two blocks of highly preserved protein sequences GAWDY (for sense primer) and CKGCG (for antisense primer). Degenerate primers were chosen from regions of amino acid positions around 111-115 for the sense primers and around 223-227 for the antisense primers (Figure 1).

Those two blocks are translated using IUPAC coding system. Some amino acids are coded for by more triplet codon possibilities than others. The fold of degeneracy for each primer is estimated by multiplying the degenerate values of different amino acids using the IUPAC coding system. The fold of degeneracy for both primers was calculated 32. Starting with 0.5 g of larvae sample, 4 µg of total RNA were extracted and cDNA was synthesized. The TD-PCR starts with a low annealing temperature (42°C) for 15 cycles, followed by 20 cycles at a high annealing temperature (57°C). The low annealing situations permit the short conserved primer regions to hybridize to their complementary strand. By changing to a high annealing temperature, specificity can be increased.

3.2 Identification of NDR domain-containing gene

The degenerate primers were amplified the putative NDR domain-containing gene fragment of 340 bp from the *S. ratti* cDNA template in the second round of PCR. The amplified NDR-containing partial gene from *S. ratti* was confirmed by sequence analysis. The amplified DNA contained an open reading frame of 116 amino acids with sequence identity to other NDR domain-containing protein. This is in concordance with the expected amplicons predicted by manual calculation. Sequence comparison was done with GenBank database using the BLAST program from NCBI GenBank. In order to compare the nucleotide sequence of SrNDR domain-containing gene with the sequences available in the GenBank database, the BLASTn program available in the NCBI website was applied using the "Highly similar sequences (megablast)" program. Comparisons between the amplified cDNA fragment sequences with BLASTn program showed that this sequence is similar (99%) to the only two sequences from *S. ratti including* a 975 bp partial mRNA sequence (XM_024650741.1) and a 513 bp genomic assembly (LN609396.1). These two sequences are almost identical. Searching the EST (Expressed sequence tag) showed a few overlapping sequences named BI323694.1, FC811605.1 and FC817168.1 with the size of 496, 666 and 652 bp, respectively. An overlapping 1082 bp cDNA fragment named SrNDR was compiled and found to contain a 972-nucleotide open reading frame that encodes a protein of 324 amino acids of 23.4 kDa expected molecular mass and a calculated pI of 7.64 (Figure 2). Based on sequence alignments, it includes C-terminus part of protein. Domain database searching showed the e-values of 1.05e-100 with the conserved NDR1 containing domain of alpha/beta hydrolase superfamily protein (pfam03096) between amino acid residues 9 to 287 (Marchler-Bauer *et al.* 2017).

The assembled nucleotide sequence showed 99-84% identity with the 5 sequences of NDR-containing partial mRNA genes. Taxonomic report of SrNDR showed similarity with 5 hits from the family *Strongyloididae*, and a total of 4 hits belonged to the genus *Strongyloides*. It was revealed that *S. ratti* with 2 hits has the most similarity in this family, while only one hit was related to *S. papillosus* species (Table 1). *Parastrongyloides trichosuri* showed the least similarity with one hit.

3.3 Secondary structure and amino acid characterization

The PSIPRED protein sequence analysis method predicted the secondary structure of SrNDR protein. Each residue was assigned values for alpha helix, beta sheet, and coils using a window of 7 residues (Figure 3A). The use of these parameters, the probability of a given residue calculated, and the conformation with the largest confidence were predicted. Based on the secondary structure data and computational three dimensional structure model (Figure 3B), the protein was folded into 12 alpha-helical conformation and 8 stranded antiparallel-beta-sheet. The three-dimensional structure folded core domain of SrNDR (residues 31-319) showed 100% similarity with human protein NDRG1 in 248 amino acids with 77% coverage. The core of SrNDR is composed of an alpha/beta hydrolase fold with a central eight-stranded β sheet surrounded by alpha helices.

By using this program, we found a prediction of mostly *a*-helical (44.86%), Strand (1.08%) and random coils (54.05%) for the coding sequence of SrNDR (Figure 3A). The three-dimensional (3D) structure of SrNDR protein (Figure 3B) was modeled using 181 amino acids (98% of the entire coding sequence) with 100% confidence the single highest scoring

template (Kelley and Sternberg, 2009). DAS-Transmembrane analysis of the SrNDR amino acid sequence identified three putative transmembrane domains. The putative transmembrane domains are located between amino acids (65-69), (131-140) and (277-285). This analysis suggests that SrNDR may be a transmembrane protein, however; the subcellular location of this protein is not known.

3.4 Phylogenetic analysis

The phylogeny tree of SrNDR nucleotide sequence was performed by the Neighbor-Joining method using MEGA11 software with all 5 homolog sequences that were observed in the GenBank. Phylogenetic analysis clearly showed an evolutionary relationship of SrNDR with the only 2 homolog sequences from *S. ratti* (one partial mRNA, and other a genome assembly), revealing that this sequence was closely related to *Strongyloides* genus but, divergent from other parasitic nematode (*Parastrongyloides trichosuri*) orthologs (Figure 4). Among the 5 nucleotide sequences given in (Table 1), SrNDR and two sequences from *S. ratti* (XM 024650741.1 and LN609396.1) with high similarity are in the same cluster. However, the sequence from *S. stercoralis* (LL999077.1) with bootstrap value 65 is separated from both of them. The other sequence *Strongyloides papillosus* (LM525602.1) was placed in a separate cluster. *Parastrongyloides trichosuri (*LM523192.1) sequence is used as an out group.

In order to analyze SrNDR in the larger context of NDR domain-containing proteins, phylogenetic analysis was conducted based on the deduced amino acid sequences from *S. ratti* with 18 sequences from parasitic nematode family, as shown in Figure 5. The constructed phylogram showed two distinct clusters (A and B). In cluster A, the gene variant of *S. ratti* SrNDR is grouped together with mostly filarial and nematode parasite with a bootstrap score of 87. Within this cluster, SrNDR formed a separated branch. As it is expected SrNDR used in this study is perfectly arranged in cluster A showing higher genetic similarity with different NDR domain-containing protein homologs. All intestinal nematodes parasites belonging to Cluster B are placed together with a strong bootstrap score of 98. *Cephalotus follicularis* (GAV68301.1) is used as an out group.

The genetic distance of SrNDR was calculated with NDR-domain containing nucleotide variant sequences using MEGA11 software. According to Table 1, among the five NDR-containing nucleotide sequences available in GenBank, SrNDR had the highest and lowest genetic distance with *S. papillosus* (LM525602.1) (10.6%) and sequences from *S. ratti* (XM 024650741.1 and LN609396.1) (0.3%), respectively.

4. Discussion

Members of the NDRG1 family all possess an NDR domain-containing of an esterases, lipases, proteases, peroxidases, dehalogenases, and epoxide hydrolases which are belongs to alpha/beta hydrolase superfamily (ABHD) proteins that are highly conserved and ubiquitously distributed throughout the organisms (Carr and Ollis, 2009). It is suggested that this gene may act as a stress response or potentially as a transcription factor (Sun et al., 2013;

Fang et al., 2014). However, the precise molecular and cellular function of these family members is still unknown.

Parasitic helminthes have evolved sophisticated and highly complex mechanisms for cohabitant with hosts in different environments (Harris, 2017). They could release Excretory/Secretory (ES) proteins into the host environment to suppress host immune response so as to ensure their survival (Harnett, 2014). It has been reported that one of the components of ES isolated from *Haemonchus contortus* was a member of the NDR containing-domain protein that interacted with host T cells (Shapiro et al., 2001). Currently, to our knowledge, no structural characterization of this protein has been done from *S. ratti*. Since NDR domain-containing proteins are relatively conservative with each other within the ABHD superfamily (Carr and Ollis, 2009), we decided to isolate the counterpart of this protein from *S. ratti*.

Degenerate primers are commonly used for isolation of unidentified gene sequences in related organisms. This allows us to discover unknown sequences of new members of gene families with unknown molecular functions. We designed degenerate primers from several known sequences associated with NDR domain-containing protein in different nematode species related to target genes in the *S. ratti*. The alignments of the selected sequences revealed two highly conserve motif regions of amino acids that enabled us to design degenerate primers. By synthesizing the degenerate primers, the preserved sequence gene fragments from the related organisms were isolated using Degenerate primers-based polymerase chain reaction technique. At first, a primer pair in the normal PCR was used to isolate the gene fragment, which was not successful. Since it was speculated that this lack of success might be due to the low level of expression of this gene, the amount of cDNA template was increased in each reaction. Again, no product was produced. Therefore, we decided to perform a TD-PCR technique. In this technique, the temperature selected for the annealing step is initially set 5°C lower than the calculated melting point of the primers. Annealing under conditions of low stringency allows the formation of more primer-template hybrids.

Multiple alignment and phylogenetic analysis of SrNDR at nucleotide and amino acid sequences revealed that the gene was the NDR domain-containing protein. Based on the phylogenetic analysis with helminthes homologs, SrNDR was supposed to belong to ABHD subfamily. SrNDR possessed no signal peptide, suggesting that this reported gene is incomplete in the N-terminus part, or the protein as an intracellular protein perhaps not be secreted from cells.

The function of NDR family proteins in the cell has not yet been well defined. However, it is believed that the proteins of this family of putative signaling molecules have an essential role in the host-parasite relationship. The role of NDR1 as a representative of a large group of ABHD family proteins whose functions is involved in the Control of programmed cell death in *Arabidopsis* has been reported (Shapiro et al., 2001). By comparing the role of HcABHD protein in *Haemonchus contortus*, particularly its involvement in cell proliferation and apoptosis, SrNDR could also be one of the proteins that may play a vital control over the cell cycle and the survival of key host cells.

Acknowledgments
This study was financially supported by a research grant from the Vice President of Research Affairs Office at the Shahid Chamran University of Ahvaz, Ahvaz, Iran.

Authors' Contribution
Study concept and design: Jolodar A.
Acquisition of data: Jolodar A.
Analysis and interpretation of data: Jolodar A.
Drafting of the manuscript: Jolodar A.
Critical revision of the manuscript: Jolodar A.
Administrative, technical, and material support: Jolodar A.

Ethics
The authors have observed all ethical points including non-plagiarism, double publication, data distortion and data manipulation in this article.

Conflict of Interest
The authors declare that they have no conflicts of interest.

Data Availability
The data that support the findings of this study are available on request from the corresponding author.

Reference

Bae DH, Jansson PC, Huang ML, Kovacevic Z, Kalinowski D, Lee, CS, et al. The role of NDRG1 in the pathology and potential treatment of human cancers. Journal of Clinical Pathology. 2013;66:911-917. [DOI:10.1136/jclinpath-2013-201692]

Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, David Diemert D, et al. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. Lancet. 2006;367:1521-1532. [DOI:10.1016/S0140-6736(06)68653-4]

Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Research. 2019;47:402-407. [DOI:10.1093/nar/gkz297]

Buonfrate D, Bisanzio D, Giorli G, Odermatt P, Fürst T, Greenaway, C. The global prevalence of *Strongyloides stercoralis* infection. Pathogens. 2020;9:468. [DOI:10.3390/pathogens9060468]

Carr PD, Ollis DL. α/β Hydrolase fold: an update. *Protein and Peptide Letters*. 2009;16:1137-48. [DOI:10.2174/092986609789071298]

Fan J, Yang W, Brindley PJ. Lysophospholipase from the human blood fluke, Schistosoma japonicum. International Journal of Infectious Diseases. 2008;12:143-51. [DOI: 10.1016/j.ijid.2007.05.005]

Fang BA, Kovacevic Z, Park KC, Kalinowski DS, Jansson PJ, Lane DJ, et al. Molecular functions of the iron-regulated metastasis suppressor, NDRG1, and its potential as a molecular target for cancer therapy. Biochimica et Biophysica Acta. 2014;1845:1-19. [DOI:10.1016/j.bbcan.2013.11.002]

Ghosh AK, Ramakrishnan G, Chandramohan C, Rajasekharan R, 2008. CGI-58, the causative gene for Chanarin-Dorfman syndrome, mediates acylation of lysophosphatidic acid. Journal of Biological Chemistry. 2008; 283: 24525-33. [DOI: 10.1074/jbc.M801783200]

Guo F, Zhang H, Eltahan R, Zhu G. Molecular and biochemical characterization of a type II thioesterase from the zoonotic protozoan parasite Cryptosporidium parvum. Frontier in cellular and infection microbiology. 2019; 9:199. [DOI: 10.3389/fcimb.2019.00199]

Harnett W. Secretory products of helminth parasites as immunomodulators. Molecular *and* Biochemical Parasitology. 2014;195:130-6. [DOI:10.1016/j.molbiopara.2014.03.007]

Harris, N.L., 2017. Recent advances in type-2-cell-mediated immunity: insights from helminth infection. *Immunity* 47:1024-36. [DOI:10.1016/j.immuni.2017.11.015]

Jolodar A. Designing of Degenerate Primers-Based Polymerase Chain Reaction (PCR) for Amplification of WD40 Repeat-Containing Proteins Using Local Allignment Search Method. Journal of Sciences. 2019;30(2):105-111. [DOI:10.22059/jsciences.2019.271719.1007349]

Kelley LA, Sternberg MJE. Protein structure prediction on the Web: a case study using the Phyre server, Nature protocols. 2009;4(3):363-71. [DOI:10.1038/nprot.2009.2] [PMID:19247286]

Kovacevic Z, Richardson DR. "The metastasis suppressor, Ndrg-1: a new ally in the fight against cancer". Carcinogenesis. 2006; 27(12):2355-66. [DOI:10.1093/carcin/bgl146] [PMID 16920733]

Lu M, Tian X, Tian AL, Li C, Yan R, Xu L, et al. A Novel α/β Hydrolase Domain Protein Derived From Haemonchus contortus Acts at the Parasite-Host Interface. Frontiers in Immunology. 2020;11:1388. [DOI: 10.3389/fimmu.2020.01388]

Marchler-Bauer A. *"CDD/SPARCLE: functional classification of proteins via subfamily domain architectures."*, Nucleic Acids Research. 2017;45:200-203. [DOI:10.1093/nar/gkw1129]

Moreno Y, Gros PP, Tam M, Segura M, Valanparambil R, Geary TG, et al. Proteomic analysis of excretory-secretory products of Heligmosomoides polygyrus assessed with next-

generation sequencing transcriptomic information. PLOS Neglected Tropical Diseases. 2011;5:e1370. [DOI: 10.1371/journal.pntd.0001370]

Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering. 1997;10:1-6. [DOI:10.1093/protein/10.1.1]

Okuda T, Kondoh H. Identification of new genes ndr2 and ndr3 which are related to Ndr1/RTP/Drg1 but show distinct tissue specificity and response to N-myc. Biochemical and Biophysical Research Communications. 1999;266(1):208-15. [DOI:10.1006/bbrc.1999.1780]

Park S, Gidda SK, James CN, Horn PJ, Khuu N, Seay, DC, et al. The a/b hydrolase CGI-58 and peroxisomal transport protein PXA1 coregulate lipid homeostasis and signaling in Arabidopsis. Plant Cell. 2013;25:1726-39. [DOI: 10.1105/tpc.113.111898]

Qu X, Zhai Y, Wei H, Zhang C, Xing G, Yu Y et al. Characterization and expression of three novel differentiation-related genes belong to the human NDRG gene family. Molecular and Cell Biochemistry. 2002;229**:**35-44. [DOI:10.1023/a:1017934810825]

Shapiro AD, Zhang C. The Role of *NDR1* in Avirulence Gene-Directed Signaling and Control of Programmed Cell Death in Arabidopsis. Plant Physiology. 2001;127:1089-1101. [DOI:10.1104/pp.010096]

Sahni S, Park KC, Kovacevic Z, Richardson DR. "Two mechanisms involving the autophagic and proteasomal pathways process the metastasis suppressor protein, N-myc downstream regulated gene 1". Biochimica et Biophysica Acta. 2019; (6):1361-1378. [DOI:10.1016/j.bbadis.2019.02.008]

Sun J, Zhang D, Bae DH, Sahni S., Jansson P, Zheng Y, et al. Metastasis suppressor, NDRG1, mediates its activity through signaling pathways and molecular motors. Carcinogenesis. 2013;34:1943-1954. [DOI:10.1093/carcin/bgt163. PMID 23671130]

Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution. 2007;24:1596-1599. [DOI:10.1093/molbev/msm092]

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research. 1994;22:4673-4680. [DOI:10.1093/nar/22.22.4673]

Vendelova E, de Lima JC, Lorenzatto KR, Monteiro KM, Mueller T, Veepaschit J, et al. Proteomic analysis of excretory-secretory products of Mesocestoides corti metacestodes reveals potential suppressors of dendritic cell functions. PLOS Neglected Tropical Diseases. 2016;10:e0005061. [DOI: 10.1371/journal.pntd.0005061]

Vijayakumar A, Vijayaraj P, Vijayakumar, AK, Rajasekharan R. The Arabidopsis ABHD11 mutant accumulates polar lipids in leaves as a consequence of absent acylhydrolase activity. Plant Physiology. 2016;170:180-93. [DOI: 10.1104/pp.15.01615]

Wang T, Ma, G, Ang CS, Korhonen PK, Xu R, Nie S, et al. Somatic proteome of Haemonchus contortus. International Journal of Parasitology. 2019;49:311-20. [DOI: 10.1016/j.ijpara.2018.12.003]

Xie M, Roy R. 2015. The causative gene in chanarian dorfman syndrome regulates lipid droplet homeostasis in C. elegans. PLoS Genetics. 2015;11:e1005284. [DOI: 10.1371/journal.pgen.1005284]

Legends

Figure 1. Alignments of nematode NDR domain-containing proteins related to *S. ratti* for designing degenerate primers. Shading indicates identity (black) or conservative substitutions (grey). The conserved amino acids for designing degenerate primers are indicated by arrows.

Figure 2. A schematic representation of SrNDR domain.

Figure 3. A schematic representation of the secondary structure (A) and three dimensional model (B) of SrNDR.

Figure 4: Phylogenetic analysis of SrNDR nucleotide sequences with homolog sequences using the Neighbor-joining method. Bootstrap value is based on 1000 replicates. The numbers in front of the species are the accession numbers of the related gene variants in the GenBank. The numbers above the lines indicate the relationship between the groups.

Figure 5: Phylogenetic tree constructed from the amino acid sequences of SrNDR with the related sequences using Neighbor-joining analysis. Bootstrap numbers are based on 1000 replicates. The numbers in front of the species are the accession numbers of the related genes in the GenBank. The numbers above the lines indicate the relationship between the groups.

Table 1. Taxonomy report of SrNDR sequence based on blastn program.

Table 2. The genetic pairwise distances of SrNDR compared to the related nematode variant nucleotide sequences.

Teble 1:

Organism	Blast Name	Score	# of Hits
Strongyloididae	nematodes		5
. Strongyloides	nematodes		4
. . Strongyloides ratti BcDNA.GH02439 (SRAE_X000159500), partial mRNA	nematodes	1786	2
. . Strongyloides stercoralis genome assembly S_ratti_ED321, scaffold srae_chrx_scaffold0000002	nematodes	743	1
. . Strongyloides papillosus genome assembly S_papillosus_LIN, scaffold SPAL _ scaffold0000034	nematodes	616	1
. Parastrongyloides trichosuri genome assembly P_trichosuri_KNP, scaffold PTRK _ contig0000003	nematodes	505	1

Table 2:

	1	2	3	4	5	6
1 -SR-NDR_Strongyloides_ratti						
2- XM_024650741.1_Strongyloides_ratti	0.003					
- LN609396.1_Strongyloides_ratti	0.003	0.000				
4- LL999077.1_Strongyloides_stercoralis	0.071	0.068	0.068			
5- LM525602.1_Strongyloides_papillosus	0.106	0.104	0.104	0.106		
6-LM523192.1_Parastrongyloides_trichosuri	0.367	0.365	0.365	0.395	0.395	