Cloning and characterization of two neuropeptide genes from cereal cyst nematode, *Heterodera avanae* from India

Prasoon Thakur[1], Amita Sharma[1], Sashi B Rao[2], Mukesh Kumar[1], Nagavara Prasad G[1], Nidhi Tyagi[1], Divya Kamaraju[1], Pradeep Papolu[1], Prakash Banakar[1] & Uma Rao[1]*

[1]Division of Nematology, Indian Agricultural Research Institute (IARI), New Delhi-110012; [2]The Ella Foundation, Hyderabad, India-500039; Uma Rao – Email: umarao@iari.res.in; Phone: 011-25842721; *Corresponding author

Abstract:
The cereal cyst nematode, *Heterodera avenae (*Wollenweber, *1924)* is one of the most important plant parasitic nematodes of cereals. It is an obligate sedentary endo parasite causing considerable crop losses in wheat, barley and oats worldwide. FMRFamide-like peptides (FLPs) play critical role as neurotransmitters or neuromodulators in the nervous system and proposed as one of the important targets for the plant parasitic nematode management. Therefore, for the first time we have cloned and characterized two neuropeptide genes (*flp-12* and *flp-16*) from the cDNA library of feeding female of *H. avenae*. Sequence analysis of FLPs revealed that both the neuropeptides are closely related with the parasitic as well as free-living nematodes. The *flp-12* contains putative *22* residue long signal peptide at N-terminal suggesting its association with extra-cellular functions, while *flp-16* does not contain signal peptide. Besides this, we have found highly conserved motif KFEFIRF in *flp-12* and RFGK motif in *flp-16*. These two *flp* genes could be interesting and potential targets for functional validation to explore their utility for designing management strategies.

Background:
The Cereal cyst nematode (CCN) is a composite group of 12 species that infects graminaceous cereals and grasses resulting in high crop losses and distributed worldwide [1]. Among them, the most economically important species is *H. avenae (*Wollenweber, 1924) that can resist harsh environmental conditions and spread through the means by which soil is disseminated [2, 3]. Although chemical nematicides are the most reliable means of controlling CCN; their use is consistently being discontinued due to their toxicity to humans and the environmental safety. Novel and specific targets are thus needed to develop new strategies against these CCNs [4].

Neuromuscular system plays critical role in the location and penetration of host roots, secretory activities, migration, alimentation, and reproduction of plant parasitic nematodes (PPNs). Neuropeptide signaling system has been proposed to be the potential target for the management of plant parasitic

nematodes [5]. FMRFamide-like peptides (FLPs) have shown to be widely expressed in the nervous system of the cyst-forming nematodes, *H. glycines*, *Globodera pallida*, and *G. rostochiensis* [6]. The largest families of neuropeptides in nematodes are the FLPs, which possess a C-terminal Arg-Phe-NH2 signature and play a central role in motor activities. There are 28 *flp* genes reported in *Caenorhabditis elegans* encoding at least 72 distinct peptides [7]. In addition, 21 FLPs have been identified in another very economically important species of the root knot nematode, *Meloidogyne incognita* [8]. Functions of *flp* genes are regulated by MicroRNAs (miRNAs) and *flp-12* and *flp-16* were identified as potential targets in pine wood nematode, *Bursaphelenchus xylophilus* [9]. Although studies on role of FLPs on behavior of plant parasitic nematodes are limited, a series of RNA interference studies of *flp* genes in *G. pallida* revealed aberrant behavioral phenotypes and migrational abilities [10]. *flp-12* is predicted to be crucial for the normal muscular function in *G. pallida* and *M. incognita* as it has interfered with

the nematode migration in a sand column in response to the root defusates [11].

In *C. elegans*, some *flp* gene knockouts have induced multiple aberrant behavioral phenotypes [12] and each *flp* gene is expressed in a unique set of neurons [13] consistent with FLPs roles in a wide array of different motor behaviors. While expression patterns of FLPs in plant parasitic nematodes (PPN) do not appear to reflect those seen in *C. elegans*, a similar restricted and unique neuronal expression pattern was reported for five *flp* genes examined in *G. pallida* [14]. CCN genomic information is completely nonexistent and in the present study for the first time, we have cloned and characterized two genes, *flp*-12 and *flp*-16 from Indian isolates of *H. avenae* using orthologous sequences present in *M. incognita*. These two neuropeptides of *H. avenae* could be useful to design strategies for its management.
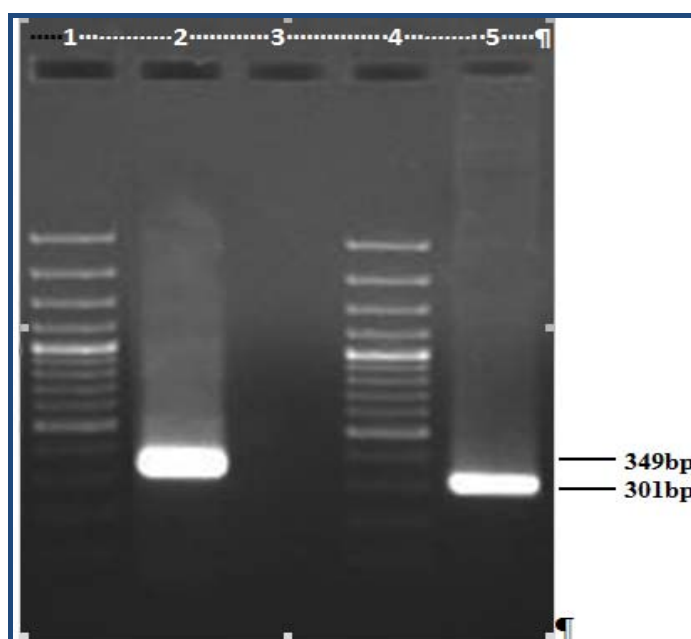
Figure 1: PCR amplification of *flp-12* and *flp-16* genes from cDNA library of *Heterodera avenae*. Lane 1:100 bp DNA marker; 2- *flp12:* 4- 100 bp DNA marker; 5- *flp16*

Methodology:
Nematode collection and multiplication on wheat plants
Cysts of CCN were collected from wheat roots growing in the fields of Indian Agricultural Research Institute, New Delhi, India. Genetic homogeneity of the population was characterized by sequencing and PCR - RFLP of ITS (Internal transcribed spacers) of rDNA (ribosomal DNA). Genetically homogenous populations were multiplied on wheat plants growing in pots under green house conditions and mature brown cysts were collected. Wheat plants were infected with juveniles and after 4 – 8 weeks, feeding females protruding out of the wheat roots were collected, frozen and used for RNA extraction.

cDNA library construction, PCR amplification, cloning and sequencing of flp genes of H. avenae
Total RNA was extracted from approximately 500 mg of frozen feeding females of *H. avenae.* Poly (A+) mRNA was purified and a cDNA library was constructed using Clontech's smart cDNA synthesis protocol as per manufacturer's instruction. *Flp-12* and

flp-16 were PCR amplified from the cDNA library using the primers designed based on the published nucleotide sequence of *flp-12* and *flp-16* genes from *M. incognita*, (Accession No: AY804187 and Accession No: EU549831 respectively). The sequences of the primer sets were forward primer: 5'-CCCAAGTTTGAGCTCTAAAAACAC 3'; reverse primer: 5'-TCATCGTCCAAATCGAATGA3' for *flp-12* and forward primer: 5'ATTTGATTGCCTCCTTAACACG3'; reverse primer: 5'- GTTTTGAATCTCTCCCAAATCG - 3' for *flp-16*.

PCR amplification reactions were performed in 25 μl reaction volumes containing 2.5 μl 10 x assay buffer (100mM Tris-HCl (pH 8.8 at 25^{0}C), 500 mM KCl, 15 mM MgCl$_2$, 0.1% gelatin, 0.05% Tween 20 and 0.05% NP-40), 200 μM each of dATP, dCTP, dGTP and dTTP (Fermentas), 16 ng primer, 1 unit of Taq polymerase (MBI Fermentas, Genetix, India) and 1 μl cDNA library. The PCR cycles consisted of initial denaturation at 94^{0}C for 4 min, followed by 35 cycles of amplification, denaturation at 94^0 C for 60 s, annealing at 55^{0}C for 30 s and extension at 72^{0}C for 1 min with a final extension at 72^{0}C for 10 min. The amplified product was run on 1.0% Agarose gel electrophoresis to confirm the size of the amplified product.

The fresh PCR product was cloned into pGEM-T cloning vector (Promega, USA) using standard protocol. Freshly prepared competent cells of *Escherichia coli DH5a* were transformed with the recombinant plasmids. Positive clones were selected by using ampicillin, blue white colonies and colony PCR. Inserts were confirmed by restriction digestion with *EcoRI*. The positive clones were custom sequenced by ABI solid sequencing platform.
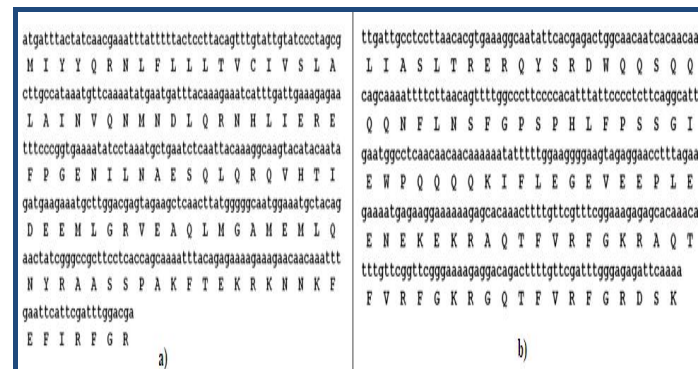

Figure 2: cDNA sequence in lower case of *flp-12* **(a)** and *flp-16*; **(b)** with the translated open reading frame. The deduced amino acid sequence is shown in one-letter symbols.

Computational Sequence Analysis
The amino acid sequences were deduced from the corresponding nucleotide sequences by using ExPASy translates available at http://expasy.org/. In order to find the homology with the plant parasitic nematodes, we performed the protein BLAST for the protein sequence and nucleotide BLAST for the nucleotide sequences using default parameters [15]. Comparison with the homologous sequences was done with ClustalW [16]. Theoretical isoelectric point (pI) and molecular weight (mw) for the conceptually translated protein sequences were calculated by the Expasy Protparam tool. Gene ontology term was assigned through AmiGO BLAST. Signal peptides were predicted by the SignalP server [17].

Phylogenetic tree was plotted by the maximum likelihood method based on the Jones-Taylor-Thornton model of amino acid change using Bioedit software package [18].
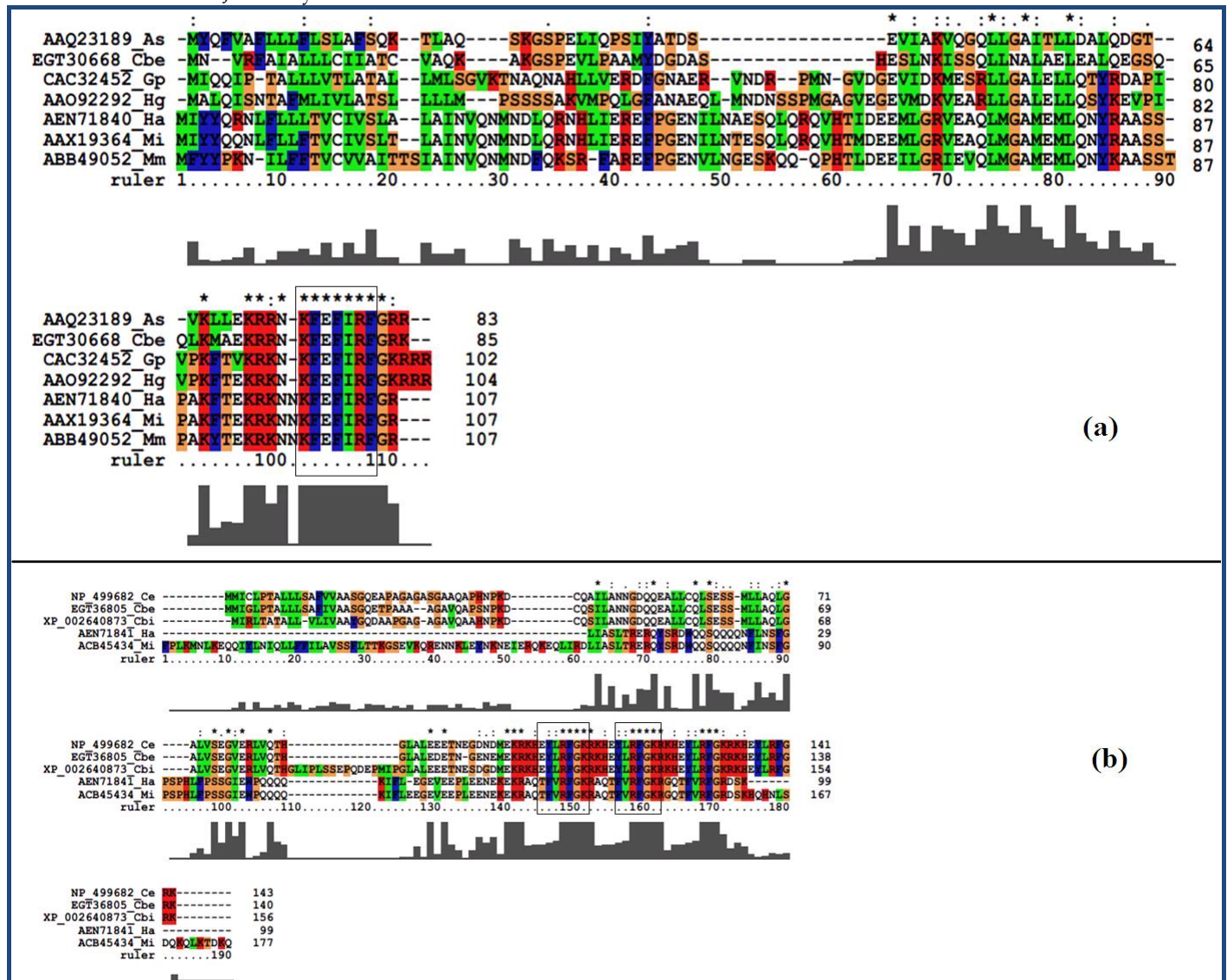
Figure 3: MSA of the predicted amino acid sequence of the *flp-12* **(a)** and *flp-16* **(b)** from *H. avenae* with *flp* sequences from other nematodes. The sequences are denoted by their Genbank identifier followed by the species abbreviation. Species abbreviations: Ha, *Heterodera avenae*; Mi, *Meloidogyne incognita*; Mm, *Meloidogyne minor*; Gp, *Globodera pallida*; Hg, *Heterodera glycines*; As, *Ascaris suum*; Cbe, *Caenorhabditis brenneri*; Ce, *Caenorhabditis elegans*; Cbi, *Caenorhabditis briggsae*.

Discussion:

Cloning of flp genes from cDNA of H.avenae

The coding sequences of the two FLPs (*flp-12* and *flp-16*) were amplified from the cDNA library of feeding female stage of *H. avenae*. PCR amplification gave 349 bp and 310 bp for *flp-12* and *flp-16* respectively **(Figure 1).** Amplified *flp* genes were cloned into pGEM-T TA cloning vector and recombinant plasmids were confirmed by colony PCR using specific gene primers. Further the plasmids were prepared for the positive recombinants and inserts were confirmed by restriction digestion with *EcoRI*. The cloned genes were sequenced using ABI solid sequencing platform and sequences are given in **(Figure 2a & 2b)**.

Characterization of the sequences

Sequence of the partial cDNA of *flp-12* (349 bp) and *flp-16* (301 bp) obtained from cDNA library of *H. avenae* were submitted to

Genbank sequence database (Accession Nos: JF437713 and JF437714). The percentage of the GC content was 35.2% and 39.2% for *flp-12* and *flp-16* respectively which were similar to the previously reported GC content for other plant parasitic nematodes **[8]**. BLASTN was carried out to determine the homology against non-redundant Genbank database. *Flp-12* and *flp-16* showed 96% and 95 % identity respectively with that of *M. incognita*. The nucleotide sequences conceptually translated into the corresponding amino acids resulting in 107 (*flp-12*) and 99 (*flp-16*) amino acids were submitted to Genbank (Accession no AEN71840 -FLP-12) and AEN71841 -FLP-16). Further, Protparam results showed a molecular weight (mw) of 12562.5 daltons (Da) and a Theoretical pI of 7.92 for *flp-12* while, mw for *flp-16* was 11765.1 Da and pI was 9.86. SignalP server was used to identify the signal peptide in both the FLPs. SignalP result of *flp-12* predicted 22 residue long signal peptide at N-terminal (1-22) and no significant hits were obtained for

BIOINFORMATION

flp-16. This indicates that *flp-12* could be involved in extra cellular function like signal transduction.

Sequence comparison
Six in case of *flp-12* and four for the *flp-16* protein sequences of closely related species were retrieved from Genbank database using protein Blast Search. Best hits obtained with the *M. incognita* for both FLPs were aligned pairwise to find the similarity with *H. avenae*. Pairwise alignment of the *flp-12* between *H. avenae* and *M. incognita* revealed the substitution of Q-R,F-L,T-A,T-A and M-I at 6, 12, 20, 49 and 60th positions respectively. However, in case of *flp-16* the substitution of I to K at 25th amino acid position and deletion of E at the 52nd position was observed in *H. avenae* sequence compared to *M. incognita*. Further, the homologous sequences aligned using multiple sequence alignment (MSA) suggested that *flp-12* is comparatively more conserved than *flp-16* among the closely related species. In addition, COOH terminal of *flp-12* is found to be highly conserved with the motif, KFEFIRF **(Figure 2a)**. Whereas, MSA of *flp-16* sequences revealed that [Y/F] [L/V] RFGK motif has two repeats at COOH Terminal. However, FVRFGK motif is conserved in the *flp-16* of both *M. incognita* and *H. avenae*, whereas YLRFGK motif is conserved in the *flp-14* of free living nematodes (*C. brenneri*, *C. elegans* and *C. briggsae*) **(Figure 3b).** This strongly suggests that these conserved motifs are essential in signal transduction mediated by GPCR (G-protein coupled receptor) **[19, 20]**. Analysis of *flp-12* and *flp-16* of *H. avenae* suggested that *flp-12* to be more conserved across parasitic nematodes with higher similarity to root knot nematode species rather than the cyst nematodes. It has diverged from the free living nematodes suggesting its vital role for parasitism. Even though *flp-16* is found to be homologous with free living counterparts, it is again highly similar to *flp-16* of *M. incognita*. Further, AmiGO BLAST hit confirmed *flp-12* involvement in receptor signalling pathway resulting in axon guidance (GO: 0007411) and negative chemotaxis (GO: 0050919) whereas, *flp-16* is involved in negative regulation of striated muscle contraction (GO: 0045988).

nematodes. The same homologous sequences as in **Figure 3** were used in this analysis. Abbreviation: PPNs-Plant parasitic nematodes.

Phylogenetic analysis:
Extending this study, we carried out phylogenetic analysis among closely related species of *H. avenae*. We constructed the phylogenetic tree of both *flp-12* and *flp-16* gene with parasitic as well as free living nematodes. Phylogenetic tree grouped *flp-12* of *H. avenae* with all PPNs. However very interestingly, it showed more closeness to the two root knot nematode species, *M. incognita* and *M. minor* than the cyst nematode species, *G. pallida* and *H. glycines* confirming the results of BLAST. However, *Ascaris suum* and *C. brenneri* formed two independent out groups suggesting their evolutionary distance **(Figure 4a).** Phylogenetic analysis of *flp-16* indicated that *H. avenae* and *M. incognita* were grouped together with high level of similarity. Additionally, it was also related to *flp-14* of three free living nematodes (*C. elegans*, *C. briggsei* and *C. brenneri*) **(Figure 4b)**. Significantly, phylogenetic analysis is in excellent agreement with the MSA result confirming the similarity.

Conclusion:
Present results indicate that FMRFamide like peptides are important for both plant parasitic and free living nematodes since they are conserved across different genera in spite of evolutionary pressure for diversification. The observed results confirm the phylogenetic relationship of these two genes determined by MSA thus, enhancing the confidence. However, *flp-12* is more conserved between the parasitic nematodes and diverged from free living nematodes indicating their possible role for parasitism. The application of RNAi to the identified neuronal genes in *H. avenae* could open the door to unraveling the underlying biology of these two *flp* genes in CCN. Work is in progress on functional validation of these two neuro peptides-encoding *flp-12* and *16* genes of *H. avenae* by RNAi through soaking in double stranded RNA solution under *in-vitro* conditions.

References:
[1] Yan G *et al. Phytopathology.* 2010 **100**: 216 [PMID: 20128694]
[2] Abidou H *et al. J Nematol.* 2005 **37**: 146 [PMID: 19262854].
[3] Mokabli A *et al. Nematology.* 2001 **3**: 171
[4] Andres MF *et al. Plant Pathol.* 2001 **50**: 270
[5] Maule AG *et al. Curr Top Med Chem*. 2002 **2**: 733 [PMID: 12052188]
[6] Kimber MJ *et al. Mol Biochem Parasit.* 2001 **116**: 199 [PMID: 11522352]
[7] Husson SJ *et al. Commun Agric Appl Biol Sci.* 2005 **70**: 153 [PMID: 16366296]
[8] Abad P *et al. Nat Biotechnol.* 2008 **26**: 909 [PMID: 18660804]
[9] Huang QX *et al. PLoS One.* 2010 **5**: 13271 [PMID: 20967259]
[10] Kimber MJ *et al. Faseb J.* 2007 **21**: 1233 [PMID: 17200420]
[11] Dalzell JJ *et al. Int J Parasitol.* 2010 **40**: 91 [PMID: 19651131]
[12] Nelson LS *et al. Science*. 1998 **281**: 1686 [PMID: 9733518
[13] Kim, K *et al. J Comp Neurol*. 2004 **475**: 540 [PMID: 15236235]
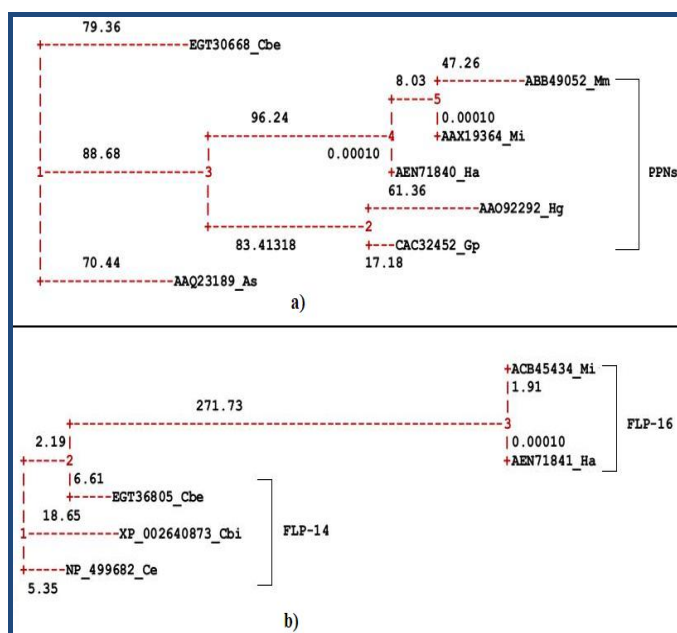
Figure 4: Phylogenetic relationships of the predicted *flp-12* **(a)** and *flp-16* **(b)** from *H. avenae* and FLP sequences from other

BIOINFORMATION

[14] Kimber MJ *et al. Int J Parasitol.* 2002 **32**: 1095 [PMID: 12117492]
[15] Altschul SF *et al. J Mol Biol.* 1990 **215**: 403 [PMID: 2231712]
[16] Larkin MA *et al. Bioinformatics.* 2007 **23:** 2947 [PMID: 17846036]
[17] Petersen TN *et al. Nat Methods.* 2011 **8:** 785 [PMID: 21959131]

[18] Hall TA, *Nucleic Acids Symp Ser.* 1999 **41:** 95
[19] GearyTG *et al. Trends Pharmacol Sci.* 2005 **26**: 56 [PMID: 15681019]
[20] Kubiak TM *et al. Biochem Biophys Res Commun.* 2003 **301**: 456 [PMID: 12565883]