Published in final edited form as:

Int J Parasitol. 2013 June ; 43(7): 515–520. doi:10.1016/j.ijpara.2013.01.008.

***Strongyloides stercoralis* daf-2 encodes a divergent ortholog of *Caenorhabditis elegans* DAF-2★**

Holman C. Massey Jr.^{a,*}, Najju Ranjit^{a,b}, Jonathan D. Stoltzfus^a, and James B. Lok^a

^aDepartment of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

We hypothesize that developmental arrest in infectious larvae of parasitic nematodes is regulated by signaling pathways homologous to *Caenorhabditis elegans* DAF (dauer formation) pathways. Alignment of *Strongyloides stercoralis* (*Ss*) DAF-2 with DAF-2 of *C. elegans* and homologs of other species shows that most structural motifs in these insulin-like receptors are conserved. However, the catalytic domain of *Ss*-DAF-2 contains two substitutions (Q1242 and Q1256), that would result in constitutive dauer formation in *C. elegans* or diabetes in vertebrate animals. *Ss*-*daf-2* also shows two alternately spliced isoforms, the constitutively expressed *Ss*-*daf-2a*, and *Ss*-*daf-2b*, which is only expressed in stages leading to parasitism.

Keywords

Strongyloides stercoralis; Insulin receptor; *Ss*-*daf-2*; DAF-2; Alternate splicing; Dauer

In contrast to free-living nematodes, which develop to reproductive maturity under optimal growth conditions and arrest development as dauer L3s under conditions of stress, most parasitic nematodes invariably develop to developmentally arrested L3s (L3i). L3i resemble the dauer L3s of free-living species in that they are non-feeding, resistant to environmental stress and have a similar pharyngeal morphology (Hotez et al., 1993). It is at this stage that host seeking occurs, either actively in the case of geohelminths, which develop to infective L3i in the external environment, or passively in the case of vector-borne L3i (e.g. - the filariae). Similar to dauers of free-living nematodes, where conditions supporting growth cause release from the dauer state, parasitic L3i resume development upon acquisition of a host. Thus, following on earlier work on parasite activation involving host signaling (Petronijevic et al., 1986) and gene activity (Petronijevic and Rogers, 1983), the “dauer hypothesis,” or “daf-c paradigm” (Hotez et al., 1993) was proposed, postulating that the L3i of parasitic nematodes are functional homologs of dauer L3s of free-living nematodes such as *Caenorhabditis elegans*. Dauer development is regulated by several interacting signal

★Note: Nucleotide sequence data reported in this paper are available in GenBank under the accession numbers: **JX171292** (pPV238), **KC312931** (pPV472) and **JX416330** (*Ss*-*daf-2* genomic region).

© 2013 Australian Society for Parasitology. Published by Elsevier Ltd. All rights reserved.

*Corresponding author. Tel.: +1-267-974-7242, +1-215-573-4756; fax: +1-215-573-7023. hmassey@vet.upenn.edu (H. C. Massey, Jr.).

^bPresent address: Department of Biology, Queensland Institute of Medical Research, Herston, Qld, 4029, Australia

Note: Supplementary data are associated with this article

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

transduction pathways in *C. elegans* (reviewed in: http://www.wormbook.org/chapters/www_dauer/dauer.html). In this study, we have focused on one of these pathways, the insulin-like pathway, that is central to regulation of entry and exit from the dauer state.

Components of an insulin-like signaling pathway have been found in several species of parasitic nematode (Massey et al., 2003; Gao et al., 2009; Hu et al., 2010; Stoltzfus et al., 2012b). The insulin-like pathway in *C. elegans* is regulated by as many as 40 insulin-like peptides, some of which act as signaling agonists and some as antagonists (Pierce et al., 2001), operating through the single insulin-like receptor protein tyrosine kinase, DAF-2 (Kimura et al., 1997). Insulin-like receptor protein tyrosine kinases such as DAF-2 form an ancient family of proteins found in all metazoan taxa (Renteria et al., 2008). These proteins play key roles in development, energy metabolism and regulation of lifespan (Garofalo, 2002; Bartke, 2008). DAF-2 and its homologs act through a signaling cascade, resulting in the phosphorylation and export from the cell nucleus of forkhead transcription factors such as DAF-16A and DAF-16B (Cahill et al., 2001), which are the terminal signaling molecules in this pathway (Ogg et al., 1997). Our laboratory and others have characterized the *daf-16* orthologs of *Strongyloides stercoralis* (Massey et al., 2003) and other parasites (Gao et al., 2009), and shown that these genes can partially complement *C. elegans* (*Ce*)-*daf-16* null mutants (Massey et al., 2006; Hu et al., 2010). Likewise, we have discovered *Ss-age-1*, which encodes an ortholog of the insulin-regulated PI3-kinase AGE-1 in *C. elegans*, and have demonstrated that administration of the PI3-kinase inhibitor, LY294002, inhibits initiation of feeding by L3i of *S. stercoralis* (Stoltzfus et al., 2012a), a functional indicator of resumption of development (Ashton et al., 2007). These results strongly implicate the insulin-like signaling pathway in regulation of parasitic and free-living development of *S. stercoralis* in a manner similar to dauer and reproductive development in *C. elegans*.

To further investigate the developmental implications of insulin-like signaling in *S. stercoralis*, we have characterized the *Ss-daf-2* insulin-like receptor gene and its putative protein product. The complete genomic region around *Ss-daf-2* including the flanking genes and the complete *Ss-daf-2a* cDNA were cloned and sequenced as described in the legend to Fig. 1. The *Ss-daf-2* gene is compact, spanning 4,536 bp, in contrast to the *Ce-daf-2* gene which contains 16 introns and spans some 33 kb of the genome (Kimura et al., 1997). *Ss-daf-2* includes a 2,231 bp exon 1, a 43 bp intron 1, a 114 bp exon 2, a 101 bp intron 2, and a terminal 2,047 bp exon 3 (Fig. 1A). An *S. stercoralis* expressed transcript database, derived from de novo assembled transcripts of seven developmental stages (Stoltzfus et al., 2012b), was searched for *Ss-daf-2* transcripts. No assembled transcripts for *Ss-daf-2* contained SL1 or SL2 spliced leader sequences (Blaxter and Liu, 1996). In this database, the longest *Ss-daf-2* 5' untranslated region (UTR) sequence ended 32 bases upstream of the start codon, indicating that the *Ss-daf-2* message is not normally trans-spliced to an SL1-like leader as are many nematode messages (Blaxter and Liu, 1996). In the predominantly expressed isoform, *Ss-daf-2a*, alternate splicing causes exon 2 to be skipped, creating a 258 bp intron composed of intron 1, exon 2 and intron 2 (designated intron 1+2). The minor isoform, composed of exons 1, 2 and 3, is designated *Ss-daf-2b* (Fig. 1A, C; Supplementary Figs. S1, S2). Sequence alignment of the *S. stercoralis* and *Strongyloides ratti* (*Sr*)-*daf-2* genomic sequences shows conservation of these alternative splice sites, suggesting similar processing of the *Sr-daf-2* gene (Fig. 1B; Supplementary Fig. S1). The 4,278 bp *Ss-daf-2a* coding sequence encodes a protein of 1,426 amino acid residues, and the 4,392 bp *Ss-daf-2b* coding sequence encodes a protein of 1,464 amino acid residues compared with 1,843 amino acid residues for the complete *Ce*-DAF-2.

Beyond the 5' end of *Ss-daf-2*, there are only 553 bp between the start codon and the *Ss-smad-1* gene, which is transcribed in the opposite direction (Fig. 1A). The protein encoded by *Ss-smad-1* is closely related to *Ce*-DAF-14, but the general term 'SMAD' is used here

because the *Ss-smad-1* gene belongs to a multicopy family of DAF-14-like genes found in the *S. stercoralis* genome (Stoltzfus et al., 2012b), whose functions have not been investigated. Whether separate promoters for *Ss-daf-2* and *Ss-smad-1* are found in the small space between the coding sequences, or whether the two genes are controlled by a common bi-directional promoter in this region has not been determined; however, the latter is suggested by the expression profile of the two genes, which is very similar for both messages, with constitutive expression detected in all seven life stages surveyed by RNAseq and having a modest maximum in parasitic females (Stoltzfus et al., 2012b). The *S. ratti* orthologs of these two genes, *Sr-daf-2* and *Sr-smad-1*, are encoded in the same relative positions on contig 3 of Chromosome 1 in the second draft of the *S. ratti* genome (Sanger Center, <http://www.sanger.ac.uk/resources/downloads/helminths/strongyloides-ratti.html>) (Fig. 1B). The *Ss-daf-2* 3' UTR contains a canonical AATAAA polyadenylation signal starting at position 4,392 of the *Ss-daf-2a* cDNA (numbered from the start codon) beginning 18 bp upstream of the site of polyadenylation of the mRNA transcript as determined by 3' rapid amplification of cDNA ends (RACE). The terminating codon of *Ss-daf-2* is separated by 116 bp from the terminating codon of a putative downstream gene reading in the opposite direction. This intronless gene encodes a serpin-like peptidase inhibitor (designated here as *Ss-srpn-1*). The terminating codon of *Ss-srpn-1* overlaps the polyadenylation site of the *Ss-daf-2* message. In the *S. ratti* genome, there is no nearby gene downstream of *daf-2* and the most likely ortholog of *Ss-srpn-1* is found on a separate contig (Chromosome 1 contig 4), indicating that at least one genomic rearrangement occurred in this region subsequent to the divergence of these two species.

Alignment of the predicted *Ss-DAF-2* protein sequence with homologs from other nematode species, the insect *Drosophila melanogaster* insulin-like receptor, the human insulin receptor isoform B (IR-B) and the human insulin-like growth factor 1 receptor (IGF1R), also supports strong conservation of major protein domains and structural motifs within this family of proteins (Fig. 1D; Supplementary Fig. S2). From the alignment, we predict *Ss-DAF-2* possesses a secretory signal sequence (SIG) followed by two receptor leucine-rich domains (RL1 and RL2) separated by a cysteine-rich furin-like domain (CR) containing 10 highly conserved internal disulfide bonds. These structures are supported by a stem consisting of three fibronectin type III domains (FN1-3). The middle, FN2, domain is interrupted by an insert domain which is divided into two parts (Ina and Inb) by protease cleavage. The length of the insert domain is affected by alternate splicing of the message. In *Ss-DAF-2A*, the junction between exon 1 and exon 3 (excluding exon 2) creates the first codon of a consensus basic protease cleavage (PC) site, RIKR, where the alpha and beta subunits of the receptor protein are predicted to be separated during receptor maturation. In the *Ss-DAF-2B* subunit, inclusion of exon 2 destroys this codon but includes two motifs, KNKK and RNKK, encoded by exon 2 that could serve the same purpose (Supplementary Fig. S1). Depending on which of the predicted cleavage sites in exon 2 is utilized, the C-terminal end of the alpha subunit of *Ss-DAF-2B* will be longer than *Ss-DAF-2A* by 22 or 28 amino acids, and the N-terminal end of the beta subunit of *Ss-DAF-2B* will be longer by 17 or 11 amino acids, respectively.

By homology to other insulin-like receptors, the alpha and beta subunits of the mature *Ss-DAF-2* receptor are predicted to be joined by a single disulfide bond between Cys 662 in the insert domain of the alpha subunit (ina) and Cys 944 in the FN3 domain of the beta subunit, and the linked alpha/beta half-receptor is predicted to be joined to a second identical half-receptor by three interchain disulfide bonds arising in the insert domains and a fourth in the FN1 domains of the two half-receptors (Supplementary Fig. S2; Fig. 1D). The beta subunits of the receptor molecule are predicted to cross into the cytoplasm at the transmembrane domain (TM). Just inside the cell membrane, the predicted juxtamembrane (JM) domain of each beta subunit contains the highly conserved NPEY motif (residues 1,049-1,052) which,

by homology, is predicted to be important as a site of tyrosine (Y1052) auto-phosphorylation and docking for *Ss*-IST-1 and *Ss*-IST-2 (Stoltzfus et al., 2012b), the nematode homolog of the mammalian insulin receptor substrates (IRS1, 2, 3 and 4), and for *Ss*-AAP-1 (Stoltzfus et al., 2012a), the homolog of the PI3 kinase p50/p55 subunit in *C. elegans* (Wolkow et al., 2002). The predicted protein tyrosine kinase catalytic domain (PYK), which follows, contains the activation loop (residues 1,227–1,248) which contains the presumably autophosphorylated tyrosine residues Y1236, Y1237, Y1240 and Y1241, corresponding to the human Y1158, Y1162 and Y1163, and *C. elegans* Y1412, Y1415 and Y1416. The position of the first autophosphorylated tyrosine in *Ce*-DAF-2 is offset by one position compared with human insulin receptor in the alignment, and both of these positions are occupied by tyrosines in *Ss*-DAF-2, so that the latter receptor has four tyrosines in the activation loop compared with three in the other species (Fig. 2, Supplementary Fig. S2). Autophosphorylation of these residues upon insulin binding in the human receptor induces a conformational change that exposes the active site, activating the kinase domain of the receptor molecule (Hubbard, 1997).

Cladistic analysis of an alignment of insulin-like receptor tyrosine kinases from representative phyla, using the distantly related *S. stercoralis* and *C. elegans let-23A* sequences as the outgroup, confirms the relatedness of the insulin-like receptor molecules to each other (Fig. 1E). Nevertheless, well supported divisions between insulin-like receptors of diverse nematode groups (supported by 100% bootstrapping scores) are in agreement with modern phylogenetic analysis of these organisms (Blaxter et al., 1998), so it would not be surprising if divergent characteristics should arise among these groups.

Interestingly, two residues highly conserved in the activation loop and the adjacent P+1 loop (Hubbard, 1997) of insulin receptors of all taxa included in our alignment, including free-living and parasitic nematodes, are substituted in *S. stercoralis*, *S. ratti* and the related species *Parastrongyloides trichosuri* (GenBank Accession Number **ADN44512**) (Fig. 2; Supplementary Fig. S2). *Strongyloides stercoralis* Q1242 is substituted for the basic residues K or R in all other taxa. This residue is adjacent to Y1241 (human Y1163), which may be involved in interactions with downstream signaling proteins (Hubbard, 1997). When the corresponding mutation, R1164Q, is found in the human insulin receptor, non-insulin-dependent diabetes mellitus (NIDDM) results (Cocoza et al., 1992). Likewise, *Ss* Q1256 corresponds in position to the dauer constitutive mutation, *daf-2(e1391)* (P1431L), in *C. elegans* (Kimura et al., 1997) and a (P1178L) mutation in the human insulin receptor which results in insulin resistance (Kim et al., 1992). Q1256 is located in the P+1 loop implicated in binding to a peptide model of the human insulin receptor substrate (IRS) (Hubbard, 1997). The two “Q” substitutions, Q1242 and Q1256, are found exclusively in the genera *Strongyloides* and *Parastrongyloides* (Fig. 2).

Expression reporter plasmids were made in which the promoter for either *Ce-daf-2* or *Ss-daf-2* was fused to the *gfp* gene with species-specific transcriptional terminators (pPV238 and pPV472, respectively; Supplementary Fig. S3). These constructs were used to transform their cognate species by gonadal microinjection. Expression patterns were examined in L1s of a stable transformed line of *C. elegans* and in transgenic larval progeny of microinjected *S. stercoralis* free-living females (http://www.wormbook.org/chapters/www_strongyloides/strongyloides.html). GFP expression from the *Ss-daf-2* transcriptional reporter was very low but detectable above background fluorescence in non-transformed control worms (Fig. 3B). Signal from the *Ss-daf-2*-based reporter was predominantly seen in the intestines of larval *S. stercoralis* (Fig. 3A), which was consistent with the pattern seen in larval *C. elegans* expressing a similar reporter construct based on the *Ce-daf-2* promoter (Fig. 3C) and with reports for *C. elegans* in Wormbase (http://www.wormbase.org/species/expr_pattern/Expr7058#021--10) where fluorescence of a DAF-2::GFP fusion protein was seen

predominantly in the intestine, head neurons and elsewhere in the nervous system. However, expression patterns reported in Wormbase varied depending on the method of detection. Neuronal expression as reported for *Ce*-DAF-2 (Kimura et al., 2011) was not seen in *S. stercoralis* transformed with the *Ss-daf-2* reporter construct.

To gain insight into the roles *Ss*-DAF2A and *Ss*-DAF-2B play in development, abundances of transcripts encoding these receptors were determined at seven different developmental stages. *Ss-daf-2*-specific transcripts were detected in all life stages of *S. stercoralis* surveyed (Fig. 3D, E). The overall pattern of transcript abundance was largely a function of the quantitative pattern of *Ss-daf-2a*, the more abundant of the two *Ss-daf-2* isoforms (Fig. 3D). *Ss-daf-2a* transcripts were detected at similar levels in all life stages except the parasitic female, in which they were significantly upregulated relative to levels in the in vivo activated L3is (L3+). It is noteworthy that similar levels of *Ss-daf-2a* were found in post-parasitic L1s and L3s, which develop predominantly to free-living adults, and post free-living L1s and L3i, which develop exclusively to parasitic females. A greater degree of regulation was seen in the less abundant isoform *Ss-daf-2b* (Fig. 3E). *Ss-daf-2b* was virtually undetectable in free-living females. Levels of this transcript increased as development progressed from post free-living L1 to post free-living L3 (both L3i and L3+) where transcript levels were significantly elevated over all other stages assayed. Levels of *Ss-daf-2b* transcripts declined following the L3+ and approached the level of detectability in post-parasitic L3s. Thus, *Ss-daf-2b* transcripts are found predominantly in life stages leading up to and including the reproductive parasitic female and decline to the limits of detectability in post-parasitic L3s and free-living females, suggesting a possible distinct role for *Ss*-DAF-2B in parasitic development.

It may be noted that similar alternative splice forms are not found in other parasitic nematodes, but have arisen independently in human evolution, where the two isoforms serve distinct roles in development. The C-terminal end of the alpha subunit of the human insulin receptor is part of the ligand binding site (Ward and Lawrence, 2009), and similar to exon 2 of *S. stercoralis*, the human exon 11 encodes a 12 amino acid extension near the C-terminal end of the alpha subunit of the human insulin receptor. The human fetal isoform, IR-A, in which exon 11 is skipped, binds IGF-II and has enhanced affinity for insulin. The adult isoform, IR-B, which includes exon 11, does not bind IGF-II and has reduced affinity for insulin (Belfiore et al., 2009).

Thus, whereas the DAF-2 insulin-like receptors of *S. stercoralis* and other *Strongyloides* and *Parastrongyloides* spp. bear close resemblance to insulin-like receptors highly conserved in all nematodes and other metazoa, they show distinct divergence in both the ligand binding and catalytic domains, and these differences may reflect the unusual life histories of this distinctive group of parasites, which include both free-living and parasitic generations. This is in stark contrast to other parasitic nematodes whose larvae invariably develop into parasitic adults (http://www.wormbook.org/chapters/www_genomesStrongyloides/genomesStrongyloides.html). Indeed, it may be significant that alternate splicing of the *Ss-daf-2b* transcript inserts a 38 amino acid peptide at the point of peptide cleavage during receptor maturation, resulting in the insertion of a short peptide at the C-terminal end of the alpha subunit (Fig. 1D; Supplementary Fig. S1). Since it has been shown that the C-terminal residues of the alpha subunit are integral to the ligand binding site of the human insulin receptor (Ward and Lawrence, 2009; Smith et al., 2010), and that the alternative spliced forms show distinct ligand binding patterns, the question is raised as to whether *Ss*-DAF-2A and *Ss*-DAF-2B respond in the same way to endogenous insulin-like signaling molecules, or whether they show differential binding patterns reflecting free-living and parasitic developmental patterns of the parasitic L4 and adult stages in *Strongyloides* and *Parastrongyloides*. Developmental patterns distinctly related to parasitism such as growth

and reproduction in the filariform (morphologically dauer-like) parasitic L4s and adults are not predicted by the *C. elegans* insulin-like signaling paradigm and fall outside the scope of a strict “dauer hypothesis”. Thus, since their life cycles contain complete generations of both free-living and parasitic development, *Strongyloides* spp. constitute unique models to study contrasting regulation of development in the parasitic and free-living forms.

The “dauer hypothesis” (Hotez et al., 1993) was formulated before dauer signaling pathways (reviewed in: http://www.wormbook.org/chapters/www_dauer/dauer.html) were elucidated at the molecular level, but it has nevertheless served well to focus studies of parasitic nematode development on the roles played by these pathways. Research in a number of laboratories including our own has shown that components of insulin signaling (Massey et al., 2003; Castelletto et al., 2009; Gao et al., 2009; Hu et al., 2010; Stoltzfus et al., 2012b) function in parasites in ways similar to their homologous pathways in *C. elegans*. The present study of the insulin-like receptor DAF-2 in *S. stercoralis*, while suggestive of function similar to that in *C. elegans*, highlights some divergences in the insulin-like receptor, DAF-2, that highlight the need to study mechanisms of developmental regulation in vivo in the parasites themselves. Such studies would greatly benefit from robust methods for transgenesis, gene knockout and targeted gene modification in species such as *S. stercoralis* and *S. ratti*. Significant strides have been made in transgenesis *Strongyloides* spp. (http://www.wormbook.org/chapters/www_strongyloides/strongyloides.html), and stable transposition and expression of transgenes using the *piggyBac* transposon has been achieved in *S. ratti* (Shao et al., 2012), but full utilization of this technique will require methods for wild type gene knockout, replacement or silencing, which have yet to be accomplished. Unequivocal testing of the dauer hypothesis awaits development of the technical capacity to experimentally manipulate the expression or function of *Ss-daf-2* in the parasites themselves, and thereby deduce its role in the biology of *Strongyloides* spp.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grants from the National Institutes of Health (NIH), USA (AI-50688 to J.B. Lok and RR02512 to M. Haskins). NIH training grant AI007532 helped support J.D. Stoltzfus. We thank Beth Maguire, Dr. Thomas Nolan and Dr. Michelle Castelletto for scientific and technical assistance; Drs. Matthew Berriman, Nancy Holroyd and Alan Tracey at the Sanger Center, Hinxton, UK, for access to *S. stercoralis* and *S. ratti* genomic sequencing data; and the *Caenorhabditis* Genetics Center, University of Minnesota, USA which is funded by the NIH National Center for Research Resources, for providing *C. elegans* strains.

References

- Ashton FT, Zhu X, Boston R, Lok JB, Schad GA. *Strongyloides stercoralis*: Amphidial neuron pair ASJ triggers significant resumption of development by infective larvae under host-mimicking in vitro conditions. *Exp Parasitol*. 2007; 115:92–97. [PubMed: 17067579]
- Bartke A. Insulin and aging. *Cell Cycle*. 2008; 7:3338–3343. [PubMed: 18948730]
- Belfiore A, Frasca F, Pandini G, Sciacca L, Vigneri R. Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. *Endocr Rev*. 2009; 30:586–623. [PubMed: 19752219]
- Blaxter M, Liu L. Nematode spliced leaders--ubiquity, evolution and utility. *Int J Parasitol*. 1996; 26:1025–1033. [PubMed: 8982784]
- Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, Vierstraete A, Vanfleteren JR, Mackey LY, Dorris M, Frisse LM, Vida JT, Thomas WK. A molecular evolutionary framework for the phylum Nematoda. *Nature*. 1998; 392:71–75. [PubMed: 9510248]

- Cahill CM, Tzivion G, Nasrin N, Ogg S, Dore J, Ruvkun G, Alexander-Bridges M. Phosphatidylinositol 3-kinase signaling inhibits DAF-16 DNA binding and function via 14-3-3-dependent and 14-3-3-independent pathways. *J Biol Chem*. 2001; 276:13402–13410. [PubMed: 11124266]
- Castelletto ML, Massey HC Jr, Lok JB. Morphogenesis of *Strongyloides stercoralis* infective larvae requires the DAF-16 ortholog FKTF-1. *PLoS Pathog*. 2009; 5:e1000370. [PubMed: 19360119]
- Cocozza S, Porcellini A, Riccardi G, Monticelli A, Condorelli G, Ferrara A, Pianese L, Miele C, Capaldo B, Beguinot F, et al. NIDDM associated with mutation in tyrosine kinase domain of insulin receptor gene. *Diabetes*. 1992; 41:521–526. [PubMed: 1607076]
- Gao X, Frank D, Hawdon JM. Molecular cloning and DNA binding characterization of DAF-16 orthologs from *Ancylostoma* hookworms. *Int J Parasitol*. 2009; 39:407–415. [PubMed: 18930062]
- Garofalo RS. Genetic analysis of insulin signaling in *Drosophila*. *Trends Endocrinol Metab*. 2002; 13:156–162. [PubMed: 11943559]
- Hotez P, Hawdon JM, Schad GA. Hookworm larval infectivity, arrest and amphiparatenesis: the *Caenorhabditis elegans* daf-c paradigm. *Parasitol Today*. 1993; 9:23–26. [PubMed: 15463660]
- Hu M, Lok JB, Ranjit N, Massey HC Jr, Sternberg PW, Gasser RB. Structural and functional characterisation of the fork head transcription factor- encoding gene, *Hc-daf-16*, from the parasitic nematode *Haemonchus contortus* (Strongylida). *Int J Parasitol*. 2010; 40:405–415. [PubMed: 19796644]
- Hubbard SR. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *Embo J*. 1997; 16:5572–5581. [PubMed: 9312016]
- Kim H, Kadowaki H, Sakura H, Odawara M, Momomura K, Takahashi Y, Miyazaki Y, Ohtani T, Akanuma Y, Yazaki Y, et al. Detection of mutations in the insulin receptor gene in patients with insulin resistance by analysis of single-stranded conformational polymorphisms. *Diabetologia*. 1992; 35:261–266. [PubMed: 1563582]
- Kimura KD, Riddle DL, Ruvkun G. The *C. elegans* DAF-2 insulin-like receptor is abundantly expressed in the nervous system and regulated by nutritional status. *Cold Spring Harb Symp Quant Biol*. 2011; 76:113–120. [PubMed: 22123849]
- Kimura KD, Tissenbaum HA, Liu Y, Ruvkun G. *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science*. 1997; 277:942–946. [PubMed: 9252323]
- Massey HC, Bhopale MK, Li X, Castelletto M, Lok JB. The fork head transcription factor FKTF-1b from *Strongyloides stercoralis* restores DAF-16 developmental function to mutant *Caenorhabditis elegans*. *Int J Parasitol*. 2006; 36:347–352. [PubMed: 16442538]
- Massey HC Jr, Ball CC, Lok JB. PCR amplification of putative *gpa-2* and *gpa-3* orthologs from the (A +T)-rich genome of *Strongyloides stercoralis*. *Int J Parasitol*. 2001; 31:377–383. [PubMed: 11306115]
- Massey HC Jr, Nishi M, Chaudhary K, Pakpour N, Lok JB. Structure and developmental expression of *Strongyloides stercoralis* *fkf-1*, a proposed ortholog of *daf-16* in *Caenorhabditis elegans*. *Int J Parasitol*. 2003; 33:1537–1544. [PubMed: 14572516]
- Ogg S, Paradis S, Gottlieb S, Patterson GI, Lee L, Tissenbaum HA, Ruvkun G. The fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. *Nature*. 1997; 389:994–999. [PubMed: 9353126]
- Petronijevic T, Rogers WP. Gene activity and the development of early parasitic stages of nematodes. *Int J Parasitol*. 1983; 13:197–199. [PubMed: 6853019]
- Petronijevic T, Rogers WP, Sommerville RI. Organic and inorganic acids as the stimulus for exsheathment of infective juveniles of nematodes. *Int J Parasitol*. 1986; 16:163–168. [PubMed: 3721706]
- Pierce SB, Costa M, Wisotzkey R, Devadhar S, Homburger SA, Buchman AR, Ferguson KC, Heller J, Platt DM, Pasquinelli AA, Liu LX, Doberstein SK, Ruvkun G. Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev*. 2001; 15:672–686. [PubMed: 11274053]

- Renteria ME, Gandhi NS, Vinuesa P, Helmerhorst E, Mancera RL. A comparative structural bioinformatics analysis of the insulin receptor family ectodomain based on phylogenetic information. *PLoS One*. 2008; 3:e3667. [PubMed: 18989367]
- Shao H, Li X, Nolan TJ, Massey HC Jr, Pearce EJ, Lok JB. Transposon-mediated chromosomal integration of transgenes in the parasitic nematode *Strongyloides ratti* and establishment of stable transgenic lines. *PLoS Pathog*. 2012; 8:e1002871. [PubMed: 22912584]
- Smith BJ, Huang K, Kong G, Chan SJ, Nakagawa S, Menting JG, Hu SQ, Whittaker J, Steiner DF, Katsoyannis PG, Ward CW, Weiss MA, Lawrence MC. Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists. *Proc Natl Acad Sci U S A*. 2010; 107:6771–6776. [PubMed: 20348418]
- Stoltzfus JD, Massey HC Jr, Nolan TJ, Griffith SD, Lok JB. *Strongyloides stercoralis age-1*: a potential regulator of infective larval development in a parasitic nematode. *PLoS One*. 2012a; 7:e38587. [PubMed: 22701676]
- Stoltzfus JD, Minot S, Berriman M, Nolan TJ, Lok JB. RNAseq Analysis of the Parasitic Nematode *Strongyloides stercoralis* Reveals Divergent Regulation of Canonical Dauer Pathways. *PLoS Negl Trop Dis*. 2012b; 6:e1854. [PubMed: 23145190]
- Ward CW, Lawrence MC. Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor. *Bioessays*. 2009; 31:422–434. [PubMed: 19274663]
- Wolkow CA, Munoz MJ, Riddle DL, Ruvkun G. Insulin receptor substrate and p55 orthologous adaptor proteins function in the *Caenorhabditis elegans daf-2*/insulin-like signaling pathway. *J Biol Chem*. 2002; 277:49591–49597. [PubMed: 12393910]

Highlights

Strongyloides stercoralis daf-2 encodes a divergent insulin-like receptor tyrosine kinase.

Amino acid changes Q1242 and Q1256 in *Ss*-DAF-2 would be deleterious in other species.

The *Ss-daf-2* is expressed in two splice isoforms, *daf-2a* (exons 1 + 3) and *daf-2b* (exons 1 + 2 + 3).

Exon-2 in *Ss-daf-2b* is located in a position that may alter ligand binding specificity or affinity.

Ss-daf-2a is expressed at all life stages; *daf-2b* is only expressed in stages leading to parasitism.

Fig. 1.

Gene and protein structure of *Strongyloides stercoralis* (*Ss*)-DAF-2 and phylogeny. (A) The *Ss-smad-1*, *Ss-daf-2* and *Ss-srpn-1* gene region. (B) The *Strongyloides ratti* (*Sr*)-*daf-2* and *smad-1* genes are organized as in *S. stercoralis*. *Sr-srpn-1* is not downstream of *daf-2*, but elsewhere in the genome. (C) The *Ss-daf-2* transcript is predicted to encode a signal sequence (SIG), the alpha fragment and the beta fragment, or alpha and beta subunits of the processed receptor molecule. The intron separating the predicted alpha and beta subunits is alternately spliced to include exon 2 in the *Ss-daf-2b* isoform according to our RNAseq data. The consensus protease cleavage (PC) sites are found at the exon 1/exon 3 junction in *Ss-DAF2A* and within exon 2 in *Ss-DAF-2B* (See Supplementary Fig. S1 for details). (D) The *Ss-DAF-2*, *Caenorhabditis elegans* (*Ce*)-DAF-2 and human insulin receptor (*Hs-InsR*) proteins are shown. Note the predicted alpha and beta subunit extensions due to alternate splicing to include alternately spliced exons 2 and 11 are indicated in the *S. stercoralis* and human receptors, respectively. Proteolytically processed monomers are represented by bars representing the polypeptides with intrachain disulfide bonds represented as small gray bars beneath the receptor leucine-rich (RL)1, cysteine-rich (CR) and RL2 domains. The intra-subunit disulfide bond joining the beta subunit fibronectin type III (FN3) domain with alpha subunit insert domain is indicated by a gray loop beneath the domains. The cysteines participating in disulfide bonds joining the two processed monomers to form the holoreceptor are shown as vertical gray lines above the bars. The positions of the two “Q substitutions” (R>Q and P>Q) in the catalytic PYK domain are indicated above the bar representing the *Ss-DAF-2* molecule. Prom, promoter region; UTR, untranslated region;

TM, transmembrane; JM, juxtamembrane. (E) Phylogeny of selected insulin-like receptors and DAF-2 molecules with the more distantly related *Ss*- and *Ce*-LET-23A (GenBank Accession Number **CAA93882**) receptor tyrosine kinases included as an outgroup. Clade numbers are noted in Roman numerals. *Strongyloide stercoralis* (*Ss*) and *Parastrongyloides trichosuri* (*Pt*) are in Clade IVa, *Ascaris suum* (*As*) and *Loa loa* (*Ll*) are in Clade III, *Caenorhabditis elegans* (*Ce*) and *Caenorhabditis briggsae* (*Cb*) are in Clade V and *Trichinella spiralis* (*Ts*) is in Clade I of a recent phylogenetic analysis of nematodes (Blaxter et al., 1998); *Drosophila melanogaster* insulin-like receptor (*Dm*-INR), human insulin receptor isoform B (*Hs*-IR-B), and the human insulin-like growth factor 1 receptor (*Hs*-IGF1R). For accession numbers of the sequences, see Fig. 2. For *Strongyloides* culture, DNA and RNA extraction, PCR, reverse transcription-PCR (RT-PCR), rapid amplification of cDNA ends (RACE) and inverse PCR methods, see http://www.wormbook.org/chapters/www_strongyloides/strongyloides.html. Degenerate primers, D2F4 (ACaGTtAAAATTGGAGATTTTGG) and D2R6 (tCATTActTAAWCCWAtATATGGTTG), where W is the degenerate nucleotide (A + T), were used to amplify a conserved 215 bp fragment of *Ss*-DAF-2. Lowercase underlined bases signify mismatches with the authentic *Ss*-*daf-2* sequence. Complete genomic and cDNA (*Ss*-*daf-2a* only) sequences were cloned using inverse PCR and 3' and 5' RACE methods as previously described (Massey et al., 2001) and completely sequenced by multiple forward and reverse reads. These data were subsequently confirmed by complete genomic sequencing at the Sanger Center, UK (data not shown). The sequence of the alternatively spliced *Ss*-*daf-2b* isoform message was obtained by RNAseq at a depth of over 1,000 reads overlapping *Ss*-*daf-2b* exon2 (Stoltzfus et al., 2012b). Alignment and phylogenetic analysis to construct a neighbor-joining tree with 1,000 iterations of bootstrapping for the DAF-2 and LET-23 receptors of *S. stercoralis* and other species was done using Geneious Pro version 5.5.6 (<http://www.geneious.com/>) using the Clustal W algorithm.

Ss-DAF-2 residue	+1242	+1256
Parasitic nematodes		
<u>Ss-DAF-2</u>	GMARDIYYHEYY <u>Q</u> PTGKRLIPVRWMA <u>Q</u> ES	
<u>Sr-DAF-2</u>	GMARDIYYHEYY <u>Q</u> PAGKRLIPVRWMA <u>Q</u> ES	
<u>Pt-DAF-2 (ADN44512)</u>	GMARDIYYHEYY <u>Q</u> PTGKRLIPVRWMA <u>Q</u> ES	
<u>Mi-(Minc08945)*</u>	GMARDIYYHEYYKPAGKRLMPVRWMA <u>P</u> ES	
<u>Dv-(Dviv_isotig22935)**</u>	GMARDIYYHEYY <u>K</u> PNGKRLMPVRWMA <u>P</u> ES	
<u>As-INSR (ADY40225)</u>	GMARDIYYHEYY <u>K</u> PTGKRLMPVRWMA <u>P</u> ES	
<u>Bm-DAF-2 (AAW50597)</u>	GMARDIYYHEYY <u>K</u> PAGKRLMPVRWMA <u>P</u> ES	
<u>Ll-INSR (XP_003144691)</u>	GMARDIYYHEYY <u>K</u> PAGKRLMPVRWMA <u>P</u> ES	
<u>Ts-INSR (XP_003371555)</u>	GMARDVYVKDY <u>Y</u> RPQGRRLLMPVRWMA <u>P</u> EA	
Free-living nematodes		
<u>Ce-DAF-2 (NP_497650)</u>	GMARDLFYHDY <u>Y</u> KPSGKRMPVRWMS <u>P</u> ES	
<u>Cb-DAF-2 (AAQ90014)</u>	GMARDLFYHDY <u>Y</u> KPSGKRMPVRWMS <u>P</u> ES	
Insects		
<u>Dm-INSR (CAY93025)</u>	GMTRDIY <u>Y</u> ETDY <u>Y</u> RKGGKGLLPVRWMS <u>P</u> ES	
Vertebrates		
<u>Hs-IGF1R (NP_000866)</u>	GMTRDIY <u>Y</u> ETDY <u>Y</u> RKGGKGLLPVRWMS <u>P</u> ES	
<u>Hs-INSR (P06213)</u>	GMTRDIY <u>Y</u> ETDY <u>Y</u> RKGGKGLLPVRWMA <u>P</u> ES	

Fig. 2.

The activation and P+1 loops of insulin-like receptors found in nematodes, insects and humans. Auto-phosphorylated tyrosine residues in the activation loop are highlighted in gray. The “Q substitutions” at positions 1,242 and 1,256 are highlighted in black with the corresponding basic and proline residues in other species underlined. Species represented: *Ss*, *Strongyloides stercoralis*; *Sr*, *Strongyloides ratti*; *Pt*, *Parastrongyloides trichosuri*; *Mi*, *Meloidogyne incognita**; *Dv*, *Dictyocaulus viviparus***; *As*, *Ascaris suum*; *Bm*, *Brugia malayi*; *Ll*, *Loa loa*; *Ts*, *Trichinella spiralis*; *Ce*, *Caenorhabditis elegans*; *Cb*, *Caenorhabditis briggsae*; *Dm*, *Drosophila melanogaster*; *Hs*, *Homo sapiens*. * *Meloidogyne incognita* (root knot nematode) http://www.inra.fr/meloidogyne_incognita. ** *Dictyocaulus viviparus* (bovine lungworm) - <http://genome.wustl.edu/genomes/detail/dictyocaulus-viviparus/>.

Fig. 3.

Developmental expression patterns of *Strongyloides stercoralis* (*Ss*)-*daf-2*. (A) Localization of GFP expression driven by the *Ss-daf-2* promoter in transgenic post free-living (PFL) larvae (L1-L2) of *S. stercoralis* (pPV472 transgene: *Ss-daf-2* promoter::*gfp*(S65C)::*Ss-era-1* terminator). (B) *Strongyloides stercoralis* control post free-living L1 (no transgene; image exposure identical to A). (C) Localization of reporter transgene expression driven by the *daf-2* promoter in *Caenorhabditis elegans* L1 (pV238 transgene: *Ce-daf-2* promoter::*gfp*(S65C)::*Ce-unc-54* terminator). Scale bars in A–C = 100 μm. DIC, differential interference contrast image. (D, E) Levels of *Ss-daf-2* transcript isoforms, *Ss-daf-2a* (D) and *Ss-daf-2b* (E), as determined by RNAseq of polyadenylated RNA from seven developmental stages each in biological triplicate: free-living females (FL Female), infectious L3s (L3i), in vivo activated L3s (L3+), parasitic females (P Female), predominantly (>95%) heterogonically developing post-parasitic L1s (PP L1), and post-parasitic approximately L3s heterogonically developing to free-living adults (PP L3). Mean transcript abundances $\pm$ the 95% confidence limits (error bars) for the three biological replicates were calculated using Cufflinks version 2.0.0 (<http://cufflinks.cbc.umd.edu/>), expressed as fragments per kb of coding exon per million mapped reads (FPKM) and log transformed as described elsewhere (Stoltzfus et al., 2012b). Transcript abundances were analyzed for developmentally related differences using Cuffdiff version 2.0.2. Statistical algorithms employed by Cuffdiff are described at <http://cufflinks.cbc.umd.edu/howitworks.html>. For both *Ss-daf-2a* and *Ss-daf-2b* pairwise comparisons were conducted between abundances in L3+ and the abundance in each of the other stages. Statistical probabilities ≤ 0.05 were considered significant. Bracket in (D) denotes the only significant comparison result by this criterion and bracket in (E) denotes the only non-significant result. For all other comparisons in (E), $P < 0.001$.