Biochemical and molecular characterization of two cytidine deaminases in the nematode *Caenorhabditis elegans*

Fiona J. THOMPSON*[1], Collette BRITTON†, Isla WHEATLEY*, Kirsty MAITLAND*, Glenda WALKER*, Shrikant ANANT‡, Nicholas O. DAVIDSON‡ and Eileen DEVANEY*[1]

*Department of Veterinary Parasitology, University of Glasgow, Bearsden Road, Glasgow G61 1QH, Scotland, U.K., †Wellcome Centre for Molecular Parasitology, Anderson College, Dumbarton Road, Glasgow G11 6NU, Scotland, U.K., and ‡Department of Internal Medicine, Washington University School of Medicine, St Louis, Missouri 63110, U.S.A.

Two cytidine deaminases (CDDs) from the free-living nematode *Caenorhabditis elegans* have been cloned and characterized. Both *Ce*-CDD-1 and *Ce*-CDD-2 are authentic deaminases and both exhibit RNA-binding activity towards AU-rich templates. In order to study their temporal and spatial expression patterns in the worm, reporter gene constructs were made using approx. 2 kb of upstream sequence. Transfection of *C. elegans* revealed that both genes localized to the cells of the intestine, although their temporal expression patterns were different. Expression of *Ce-cdd-1* peaked in the early larval stages, whereas *Ce-cdd-2* was expressed in all life cycle stages examined. RNA-interference (RNAi) assays were performed for both genes, either alone or in combination, but only *cdd-2* RNAi produced a consistent visible phenotype. A proportion of eggs laid from these worms were swollen and distorted in shape.

Key words: developmental regulation, intestinal localization, reporter gene.

INTRODUCTION

Cytidine deaminases (CDDs) belong to a family of enzymes that catalyse the deamination of cytidine to uridine and, as such, have an important role in the pyrimidine salvage pathway in many cell types. In mammals, one member of the family (apobec-1) has acquired an additional role in the editing of apolipoprotein B (apoB) mRNA [1]. In this process, a cytidine (C6666) is deaminated to a uridine, resulting in the production of two proteins (apoB100 and apoB48) with differing functions from a single mRNA. In most mammals, apoB editing is tissue-specific, occurring in the small intestine, but not the liver. The catalytic subunit responsible for the deamination reaction is a CDD (apobec-1). C → U RNA editing is dependent upon several additional proteins [2] that are only now being characterized at a molecular level [3–5]. Previous studies have documented the presence of apobec-1 in cells and tissues that do not express apoB, leading to the hypothesis that additional templates for editing might exist. However, despite many efforts, only a few other likely substrates for CDD-catalysed RNA editing have been described (reviewed in [6,7]). Transfection of yeast cells with apoB and apobec-1 results in the editing of apoB C6666, suggesting that the necessary complementation factors are present in yeast [8]. More recently, an endogenous CDD from yeast was shown to have C → U editing activity on a reporter mRNA, but the native substrate has not been identified [9].

We have described previously [10] a cDNA encoding a CDD from the parasitic nematode *Brugia pahangi* (*Bp-cdd*) that is highly up-regulated as the parasite transfers from the mosquito vector to the mammalian definitive host. Because genes that are triggered by the transition from vector to mammal might have important roles in the establishment of the infection, *Bp-cdd* has been the focus of extensive characterization [11,12]. Although phylogenetic analysis grouped the *Brugia* CDD with the cytosine nucleoside deaminases of prokaryotes [11,13] that neither bind nor edit RNA, the temporal expression pattern of the *Brugia* molecule was not consistent with a role in the salvage pathway. In addition, the recombinant *Brugia* CDD was shown to possess RNA-binding activity with an affinity for AU-rich RNA templates [11], in contrast with the *Escherichia coli* and *Bacillus subtilis* molecules, which fail to bind RNA. However, *Bp*-CDD displayed no editing activity towards an apoB substrate. The ability of some CDDs to bind AU-rich sequences has led to the hypothesis that some members of the family might have acquired a more general role in RNA metabolism, in a manner analogous to that described for other RNA-binding proteins that affect RNA stability by interacting with AU-rich elements in the 3′ untranslated region of target genes [14]. Recent studies [15] with apobec-1 supported this hypothesis, and demonstrated that transfection of a mammalian cell line with apobec-1 resulted in a significant increase in the half life of c-*myc* RNA.

Because of the difficulties inherent in defining gene function in parasitic nematodes such as *Brugia*, in which it is not possible to carry out transfection or to knock out genes, we sought to identify CDDs in the genetically tractable free-living nematode *Caenorhabditis elegans*. Screening of the *C. elegans* database, ACeDB, identified two genes annotated as CDDs, which we named *cdd-1* and *cdd-2*. In the present study, we report the biochemical and molecular characterization of *Ce-cdd-1* and *Ce-cdd-2*.

Abbreviations used: apoB, apolipoprotein B; ACF, apobec-1 complementation factor; 5-AU, a radiolabelled cRNA template containing five tandem repeats of AUUUA; (*Ce-*/*Bp-*)CDD, cytidine deaminase (derived from *Caenorhabditis elegans*/*Brugia pahangi* respectively); GFP, green fluorescent protein; L1, L2 etc., first-stage larva, second-stage larva, etc.; MBP, maltose-binding protein; 5′-RACE, 5′-rapid amplification of cDNA ends; ORF, open reading frame; RNAi, RNA interference; RT, reverse transcription; THU, tetrahydrouridine.

[1] To whom correspondence should be addressed (e-mail e.devaney@vet.gla.ac.uk).

MATERIALS AND METHODS

PCR amplification of *Ce-cdd-1* and *Ce-cdd-2*

Cosmids C47D2 and F49E8 were supplied by the *C. elegans* genome consortium. Cosmid DNA was extracted by standard methods [16], and the genomic copy of each gene was generated by PCR. Reactions were performed using 100 μM of each gene-specific primer (*Ce-cdd-1* primers C and D, and *Ce-cdd-2* primers E and F; Table 1) on 100 ng of cosmid DNA using AmpliTaq™ (PerkinElmer, Norwalk, CT, U.S.A.). Cycle conditions were as follows: denaturation at 95 °C for 2 min; 30 cycles of 94 °C for 1 min, 55 °C for 1 min and 72 °C for 3 min; and a final extension for 10 min at 72 °C. The resulting products were cloned into the pCR™ 2.1 vector (Invitrogen, Carlsbad, CA, U.S.A.), and the sequence was confirmed using the Sequi-Therm EXCEL II™ Long-Read™ DNA sequencing Kit-LC (Epicentre Technologies, Cambio Ltd., Cambridge, U.K.) with fluorescently labelled (using the tricarbocyanine dye, IRD800) T7 and T3 primers on an automated sequencer (both from MWG-Biotech, Milton Keynes, U.K.). cDNA fragments were amplified by PCR (primers and conditions as above) using first-strand cDNA, generated from RNA isolated from mixed-stage worms. The cDNA fragments were similarly cloned into the pCR™ 2.1 vector (Invitrogen), and the sequence was confirmed.

Phylogenetic analysis

The catalytic domains of *Ce*-CDD-1 and *Ce*-CDD-2 were used to generate a distance matrix with the RNA-editing enzymes and other known cytosine nucleoside and nucleotide deaminases using the PHYLIP analysis program, as described previously [17]. The distance matrix was generated using the PROTDIST program, followed by the generation of a tree by the KITSCH programme, which was plotted using the DRAWTREE program [18].

Expression of recombinant proteins

A second set of gene-specific primers with suitable restriction sites to allow directional in-frame cloning of each cDNA molecule were designed (*Ce-cdd-1* primers G and H, and *Ce-cdd-2* primers I and J; Table 1). Each fragment was amplified by PCR (as described above), digested (using *Xba*I–*Pst*I or *Pst*I–*Bam*HI for *Ce-cdd-1* and *-2* respectively), and subcloned into pMAL (New England Biolabs, Beverly, MA, U.S.A.). Recombinant proteins [maltose-binding protein (MBP)–CDD-1 and MBP–CDD-2] were expressed following the manufacturer's protocol. The proteins were purified on amylose columns, eluted with 10 mM maltose, and dialysed into the appropriate buffer. Where required, the MBP portion was cleaved by Factor X digestion (New England Biolabs).

Biochemical analysis of recombinant proteins

Recombinant proteins were assayed for enzyme activity and RNA binding, as described previously [11]. MBP-fusion proteins were used in all experiments, and compared with MBP alone. Enzyme activities were measured using [^{3}H]deoxycytidine as a substrate, and analysed by TLC [19]. UV cross-linking of *Ce*-CDD-1 and *Ce*-CDD-2 was performed using a synthetic RNA template, as described previously [11].

Temporal expression pattern of the *cdd* mRNA in *C. elegans*

A panel of cDNAs was used to analyse the temporal expression pattern of *cdd-1* and *cdd-2* mRNA throughout the *C. elegans* life cycle, using RNA isolated from synchronous populations of worms harvested at 2 h intervals [20]. Gene-specific primers (*Ce-cdd-1*, primers A and B; *Ce-cdd-2*, primers E and F) were designed to span an intron, so that the resulting cDNA and genomic fragments could be clearly distinguished on 4 % (w/v) agarose gels (see Table 1 for sizes). Following 35 cycles of PCR, the products were blotted on to nylon membranes [Osmonics (formerly Micron Separation Incorporated), Westborough, MA, U.S.A.], probed with a [γ-^{32}P]ATP end-labelled primer [16] and subjected to autoradiography. The *ama-1* gene (for RNA polymerase II) was used to check the integrity of the mRNA in these experiments [20]. The resulting autoradiographs were analysed on a Kodak Digital Image Station (Rochester, NY, U.S.A.).

Table 1 Primers used in this study

For the A + B *Ce-cdd-1* PCR fragments the size of the genomic DNA is 285 bp, and the cDNA is 170 bp; for the E + F *Ce-cdd-2* PCR fragments, the genomic DNA is 642 bp in size, and the cDNA is 500 bp.

Primer	Sequence (5′ → 3′)
A	CAAGTATAGCAACTTCAAAG
B	CAACAACAATGTACTTAAA
C	ATGACTACAACAAAGGCAAACC
D	CTATTCCTTATTGTTGTGATC
E	ATGGCCGCGAATAGTCTTCCTC
F	TTAAGAGACGACTGTAACGTTATG
G	GCCGTCTAGAATGACTACAACAAAGGC
H	GCCGCTGCAGCTATTCCTTATTGTTGTG
I	GCCGGGATCCATGGCCGCGAATAGTCTTCCTC
J	GCCGTCTAGATTAAGAGACGACTGTAACGTTATG
K	AGGCGCGCCTTTGGAAAAAGTCAAACCAGAAA
L	AGGCGCGCCTTGACTATCCGATTCATTTGCTT
M	AGGCGCGCCGGAATGGCGGGAAATGTGGAAAGG
N	AGGCGCGCCCACTTTGCATGCGTTCCACAACGC
P	CTATTCCTTATTGTTGTGATC
Q	GCGCTTCTCTCTGCACAAATCG
R	CACCAGCGCAGCTCCAACTTTG
S	TTAAGAGACGACTGTAACGTTATG
T	GCACTACGCTCAGCGCAGATAG
U	GAGAAGCGCTGCACCAACTGG
V	GCCGCTGCAGCCTGCGAAGAACTTC
W	GCCGGGATCCAGCGCAGCTCCAAC
X	GCCGCTGCAGCCTATTTGTCTATACG
Y	GCCGGGATCCCTGTAGTTAACTATGAC
Z	GCCGGGATCCTGAACTAGTTCCACATC

5′-Rapid amplification of cDNA ends (5′-RACE)

5′ RACE was performed exactly as described in the manufacturer's instructions (5′ RACE System Version 2; Gibco BRL, Gaithersburg, MD, U.S.A.). Briefly, first-strand cDNA was generated from mixed-stage RNA using gene-specific primers (*Ce-cdd-1* primer P and *Ce-cdd-2* primer S). These products were then tailed and used in nested PCR reactions (*cdd-1* primers Q and R, and *cdd-2* primers T and U) to amplify the 5′ end of each molecule. The products were cloned into pCR™ 2.1, and sequenced as described above.

Promoter analysis

For the *Ce-cdd-1* gene, the region amplified included approx. 2 kb of upstream sequence and extended into the second exon of the *Ce-cdd-1* open reading frame [(ORF); 83 bases of the first exon, 46 bases of the first intron and 33 bases of the second exon, 2167 bases in total, primers V and W; Table 1). Primers were designed to amplify either approx. 2 kb upstream of the *Ce-cdd-2* gene (primers X and Y), or 2 kb of upstream sequence and 40

```
L27943    hCDA          GICAERTAIQKAVSEGYKDF...MQDDFISPCGACRQVM
U20865    scCDD         CICAERSAMIQVLMAGHRSG...WEDQCVSPCGVCRQFI  CYTIDINE DEAMINASE
M60916    ECCDA         TVHAEQSAISHAWLSGEKAL...AITVNYTPCGHCRQFM

U61949    Ce-CDD-2      TICAERSAIVSAVSQGYTKF...ELSEPASPCGLCRQFL
U64861    Ce-CDD-1      TICAERSAIVTALTKGHRKF...ELEAPCSPCGVCRQVL CYTOSINE NUCLEOSIDE DEAMINASE
X91065    Bp-CDD        TICAERSAVTRAVAEGYREF...TPAEPTAPCGLCRQFL

L26234    hAPOBEC-1     TNHVEVNFIKKFTSERDFHP...TWFLSWSPCWECSQAI
L07114    rAPOBEC-1     NKHVEVNFIEKFTTERYFCP...TWFLSWSPCGECSRAI
NM020661  hAID          GCHVELLFLRYISDW.DLDP...TWFTSWSPCYDCARHV C-TO-U RNA DEAMINASE
AF161698  hAPOBEC-2     AAHAEEAFFNTILPAFDPAL...TWYVSSSPCAACADRI
                          *                     *  *
```

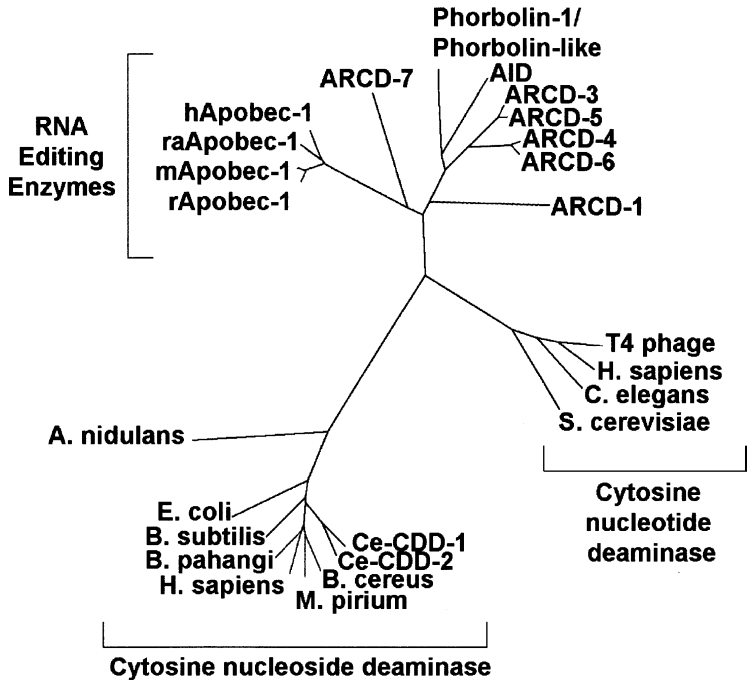

Figure 1 Alignment of the active site of nematode CDDs with the RNA editing deaminases and representative members of the CDD family (upper panel), and phylogenetic relationship of the CDD family (lower panel)

Upper panel: sequences were aligned using the ClustalW program of MacVector. Identical amino acids are boxed and shown in bold; similar amino acids are merely emboldened. Gaps have been introduced to maximize the alignment. Zinc binding amino acids are indicated with asterisks. Lower panel: the catalytic domains of *Ce*-CDD-1 and *Ce*-CDD-2 were used to generate a distance matrix with the RNA-editing enzymes and other known cytosine nucleoside and nucleotide deaminases using the PHYLIP analysis program. The distance matrix was generated using the PROTDIST program, followed by the generation of a tree by the KITSCH program, which was plotted using the DRAWTREE program. *A. nidulans*, *Aspergillus nidulans*; ARCD, apobec-1-related CDD; AID, activation-induced deaminase; *B. cereus*, *Bacillus cereus*; *B. subtilis*, *Bacillus subtilis*; *H. sapiens*, *Homo sapiens*; hApobec-1, human Apobec-1; *M. pirium*, *Mycobacterium pirium*; mApobec-1, mouse Apobec-1; raApobec-1, rat Apobec-1; *S. cerevisiae*, *Saccharomyces cerevisiae*.

bases of *Ce-cdd-2* ORF (primers X and Z). The resulting PCR products were digested (with *Pst*I and *Bam*HI), gel-purified and cloned into the plasmid pPD96.04, which contains both green fluorescent protein (GFP) and *LacZ* reporter genes, and a nuclear localization signal [21]. Cloning sites and orientation were confirmed by ABI sequencing. Each construct was used to transform *C. elegans* N2, essentially as described previously [22]. In brief, purified plasmid DNA was injected into the distal arm of the hermaphrodite gonad at a concentration of 50 μg/ml, together with plasmid pRF4 at 100 μg/ml (containing a dominant mutant allele of the *rol-6* gene) [23]. Lines in which F2 and subsequent generations showed a right roller phenotype were analysed for GFP expression by UV microscopy. Alternatively, worms were fixed and stained for *LacZ* expression exactly as described previously [24]. At least two separate lines were analysed for each construct. Images were recorded using an Olympus BX60 fluorescence microscope.

RNA interference (RNAi) by injection

Both cDNAs were excised from pCR™2.1 (*Bam*HI and *Eco*RV) and subcloned into the pBSc vector (Stratagene, La Jolla, CA, U.S.A.). Qiagen plasmid preparations were generated and linearized as described previously ([25]; also see http://www.

ciwemb.edu), then single-stranded RNA was generated from the T7 and T3 promoters as directed in the manufacturer's protocol (Promega, Madison, WI, U.S.A.). The complementary RNA strands were annealed, purified and re-suspended at 1 mg/ml ([25]; also see http://www.ciwemb.edu), and used to transform *C. elegans* N2 strain by micro-injection. Controls were injected with water alone. In some experiments, RNA was synthesized using the MEGAscript T3 and T7 kits (Ambion, Austin, TX, U.S.A.), the resulting RNA was cleaned on Qiagen RNeasy columns (Qiagen, Chatsworth, CA, U.S.A.), and then annealed and used for injection exactly as described above. Worms were left for 24 h at 20 °C, before moving on to new OP50-seeded plates, and then progeny were examined for any RNAi mutant phenotype.

RNAi by feeding

Fragments of both *Ce-cdd-1* and *-2* were amplified by PCR (conditions as above) from genomic DNA using gene-specific primers containing *Asc*I sites (*Ce-cdd-1* K and L, *Ce-cdd-2* M and N; Table 1). For *Ce-cdd-1*, the fragment was 2250 bp in size; for *Ce-cdd-2*, the fragment was 827 bp. PCR fragments were digested with *Asc*I, gel-purified [16] and cloned into the *Mlu*I site of the L4440 vector [26]. Plasmids were maintained in DH5α cells (Gibco BRL). Qiagen Mini-preps of each plasmid were made, and used to transform competent HT115 cells [26]. The RNAi feeding protocol used was exactly as described in [27]. In brief, bacteria were seeded on to plates containing 1 mM isopropyl β-D-thiogalactoside ('IPTG'), induced overnight at approx. 22 °C, and then worms at a specific stage of development were placed on to plates containing L4440 without an insert (negative control), plates with L4440 containing the appropriate insert, or L4440 containing an embryonic lethal double-stranded RNA (kindly provided by B. Roberts, University of Glasgow). Plates were incubated at three different temperatures (16, 21 and 25 °C), and worms were monitored at frequent intervals.

RESULTS

Cloning *Ce-cdd-1* and *Ce-cdd-2*

The *C. elegans* database, ACeDB, was screened for homologies with *Bp-cdd* and two potential CDDs were identified. These were named *Ce-cdd-1* (cosmid position C47D2.2; X chromosome) and *Ce-cdd-2* (F49E8.4; chromosome IV). Both genes were amplified by reverse transcription (RT)-PCR on mixed-stage *C. elegans* RNA or by PCR on cosmid DNA using gene-specific primers and oligo (dT). The resulting PCR products were cloned into pCR™ 2.1, and the sequence of each gene was confirmed. The two *C. elegans* genes are approx. 60% identical with each other and with *Bp-cdd* at the DNA level, and 55% identical at the amino acid level. The ORF for each of the *C. elegans* genes predicts a protein of approx. 17.5 kDa. Sequence analysis of the active site of the *C. elegans* enzymes demonstrated that both amino acid sequences contain the motif for the zinc-binding region of CDDs and deoxycytidylate deaminases (Figure 1, upper panel), and a number of sites that might function in post-translational modification. Phylogenetic analysis using the catalytic domain of the enzymes indicates that the *C. elegans* CDDs are clustered with the cytosine nucleoside deaminases (Figure 1, lower panel).

Ce-CDD-1 and *Ce*-CDD-2 are authentic deaminases

The ORF of both *Ce-cdd-1* and *Ce-cdd-2* were expressed as fusion proteins in an MBP vector, purified by amylose-affinity chromatography and used as fusion proteins in a variety of assays. Both fusion proteins demonstrated saturable CDD activity (Figure 2A), with a K_m of 1.5×10^{-5} M for both enzymes,

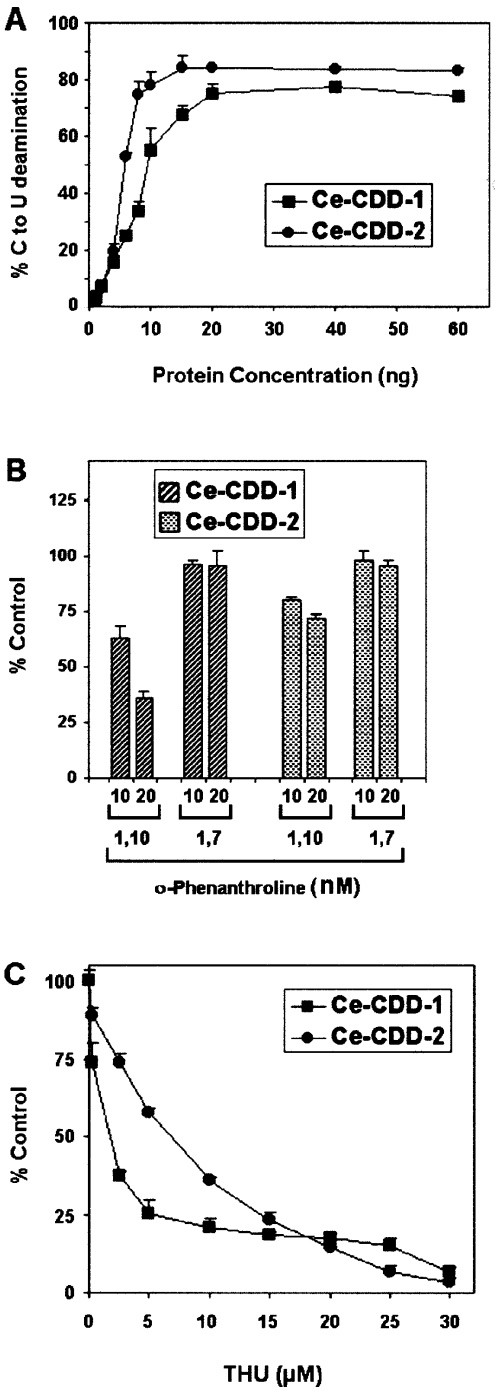

Figure 2 Enzymic activity of *Ce*-CDDs

CDD activity was determined by *in vitro* incubation of purified *Ce*-CDD-1 and *Ce*-CDD-2 at 37 °C using [^{3}H]deoxycytidine as a substrate. The percentage conversion of C into U was determined by TLC. (**A**) Deaminase activity with increasing amounts of protein (0–60 ng, 4 h incubation). (**B**) Inhibition of CDD activity by zinc chelation using increasing concentrations (10, 20 nM) of 1,10-*o*-phenanthroline (1, 10) and its inactive isomer, 1,7-*o*-phenanthroline (1, 7). Control CDD reactions lacking the inhibitors were also performed, and the value was taken as 100% (results not shown). (**C**) Inhibition of CDD activity with increasing concentrations of THU. Incubations were performed with 20 ng of either *Ce*-CDD-1 or *Ce*-CDD-2 for 4 h, with increasing concentrations (0–30 μM) of THU.

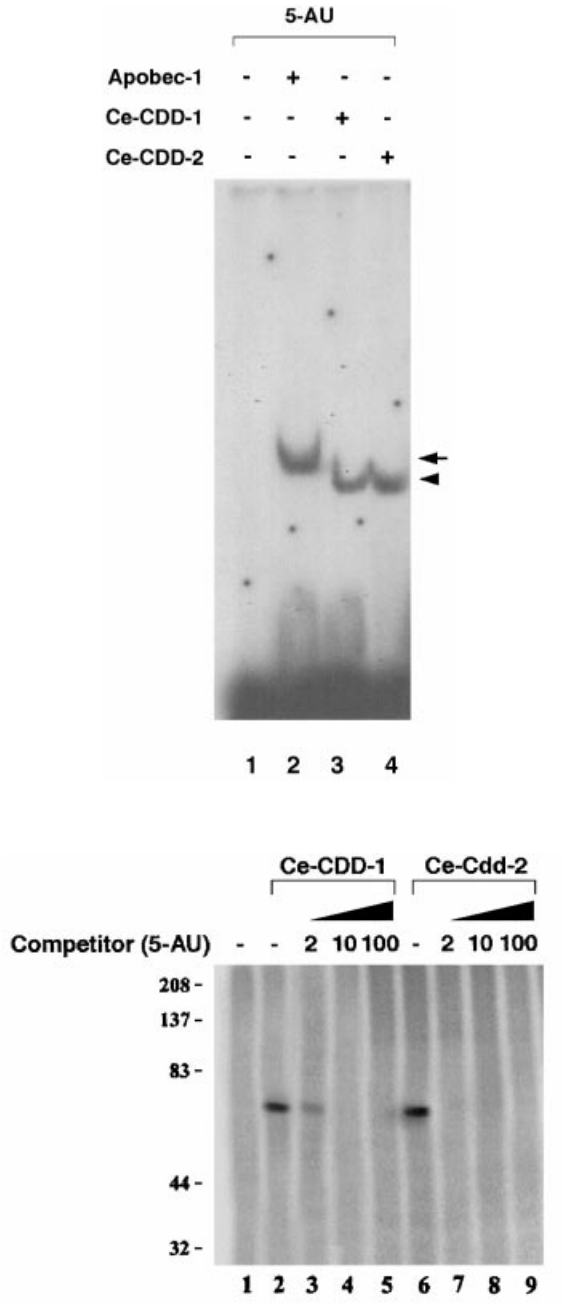

Figure 3 *Ce*-CDDs are AU-rich RNA-binding proteins

Upper panel: electromobility shift assay. A radiolabelled cRNA (5-AU) was incubated with 250 ng of recombinant protein, as indicated. The complex was separated on a 5% (w/v) native polyacrylamide gel, and subjected to autoradiography. The location of shift by apobec-1 is shown by the arrow, that of the *Ce*-CDDs is shown by the arrowhead. Lower panel: RNA-binding competition assay. A UV-cross-linking assay was performed with the radiolabelled 5-AU template and *Ce*-CDDs. Where indicated, competition for binding to the radiolabelled template was performed by addition of 2–10-fold-higher levels of non-labelled 5-AU RNA. The complex was separated by SDS/PAGE (10% gels), and subjected to autoradiography. Migration of molecular-mass standards in the gel is shown to the left.

and a V_{max} of 30 units/mg for *Ce*-CDD-1 and 15 units/mg for *Ce*-CDD-2. CDDs are Zn^{2+}-dependent enzymes, and the activity of both of the *C. elegans* enzymes could be inhibited by the Zn^{2+} chelator 1,10-*o*-phenanthroline, but not the inactive analogue, 1,7-*o*-phenanthroline (Figure 2B), with CDD-1 being more sensitive to inhibition than CDD-2 ($P < 0.001$). Addition of increasing amounts of tetrahydrouridine (THU), a competitive inhibitor of CDDs, inhibited both enzymes (Figure 2C). Again, CDD-1 was more sensitive to inhibition than CDD-2, with 50% inhibition achieved at 2.5 μM THU for CDD-1 and 7.5 μM THU for CDD-2 ($P < 0.001$).

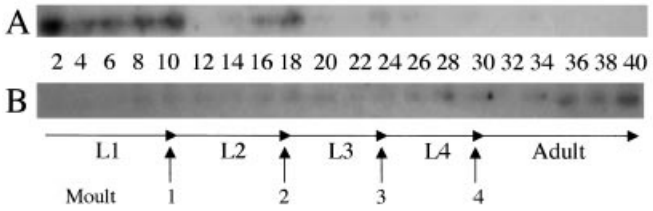

Figure 4 RT-PCR analysis of *Ce-cdd-1* (A) and *Ce-cdd-2* (B) expression throughout the *C. elegans* life cycle

First-strand cDNA was synthesized from RNA isolated from synchronous cultures of *C. elegans* at 2 h intervals. PCR was performed using gene-specific primers, as described in the Materials and methods section. Products were run on 4% (w/v) agarose gels, blotted, probed with ^{32}P-labelled probes and then subjected to autoradiography.

Ce-CDD-1 and *Ce*-CDD-2 are RNA-binding molecules

Previous studies using UV-cross-linking assays have indicated that both apobec-1 and *Bp*-CDD are RNA-binding molecules with an affinity for AU-rich templates. To investigate whether the *C. elegans* CDDs demonstrated similar activity, UV-cross-linking was undertaken using a radiolabelled cRNA template containing five tandem repeats of AUUUA (5-AU). As shown in the upper panel of Figure 3, both *C. elegans* recombinant proteins bound the radiolabelled template, as did the positive control, apobec-1. Furthermore, the binding of both recombinant proteins to the 5-AU template could be competed using a 10-fold excess of non-labelled 5-AU template (Figure 3, lower panel).

Ce-cdd-1 and *Ce-cdd-2* are differentially expressed throughout the lifecycle

Bp-cdd is first expressed following the transfer of third-stage larvae (L3) from the mosquito vector to the mammalian host [12] and thereafter is highly expressed throughout the post-infective L3 and fourth (L4) stages only. To investigate the transcription pattern of the *C. elegans* genes, RT-PCR was performed using gene-specific primers on a panel of cDNAs prepared from synchronized worms at 2 h intervals throughout the life cycle [20]. In our hands, both mRNAs were undetectable by ethidium bromide staining of PCR products, and could only be detected by Southern blotting of a 35-cycle PCR, as described in the Materials and methods section. In these experiments, *Ce-cdd-1* and *Ce-cdd-2* showed very different patterns of expression: *Ce-cdd-1* was detected in the L1 stage, between 2 and 10 h of development, with levels of transcription decreasing in the L2 stage, and no signal apparent thereafter (Figure 4A). In contrast, *Ce-cdd-2* first appeared in the 8 h sample (late L1), and then was more or less constitutively expressed up to the 34 h sample (young adult), after which expression increased in the 36–40 h samples (Figure 4B). The integrity of the RNA in each sample was assessed by RT-PCR using primers for the *ama-1* gene [20], which was visualized by ethidium bromide staining (results not shown).

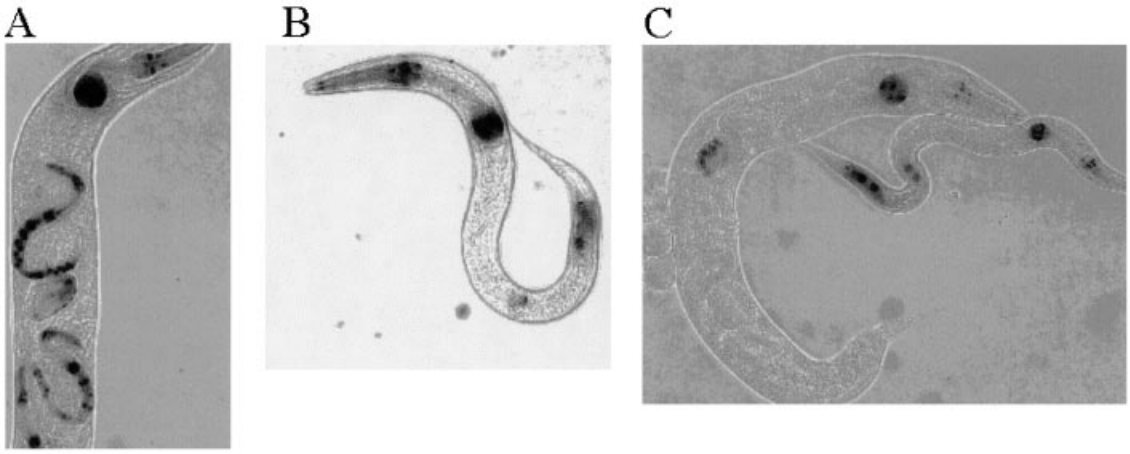

Figure 5 The *Ce-cdd-1* reporter construct is expressed predominantly in the early larval stages

N2 worms were transfected with the *Ce-cdd-1* translational fusion reporter gene construct, and stained for *LacZ* expression for 10 min. Worms were viewed on an Olympus BX60 fluorescence microscope. (**A**) Bagged adults containing eggs and hatched L1 (magnification ×14); (**B**) L3 (magnification ×28); and (**C**) L4 and adult (magnification ×14). Especially noteworthy is staining of the intestinal cells, and intense staining of the pharyngeal bulbs.

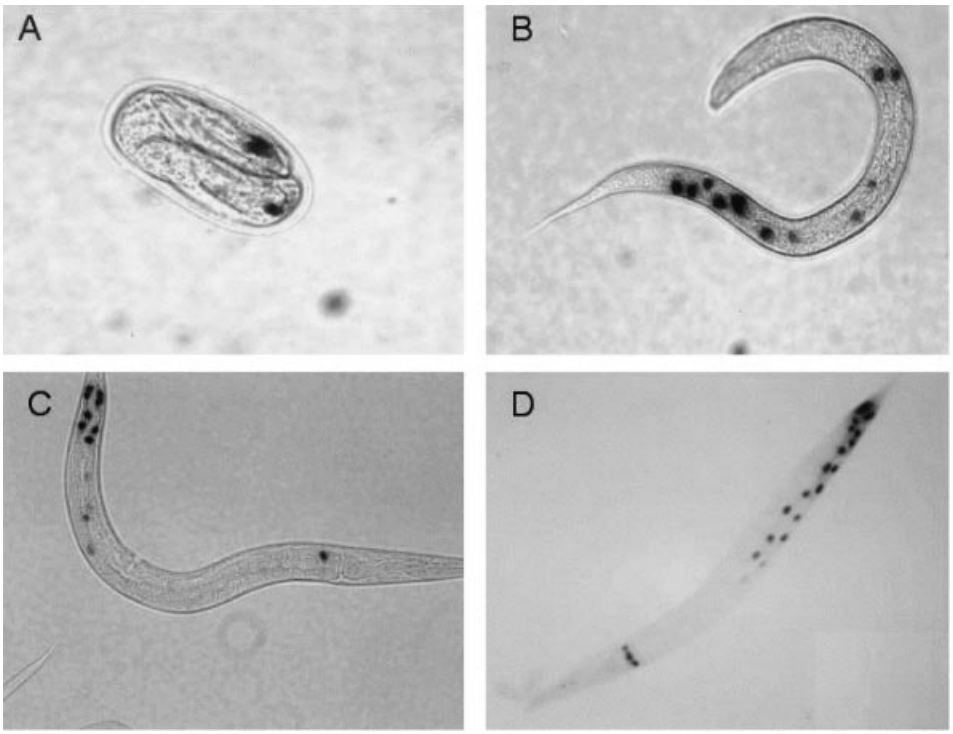

Figure 6 The *Ce-cdd-2* reporter construct is expressed throughout the life cycle

N2 worms were transfected with the *Ce-cdd-2* translational fusion reporter gene construct and stained for *LacZ* expression for 4.5 h. Worms were viewed on an Olympus BX60 fluorescence microscope. (**A**) Egg, ×20; (**B**) L2, ×20; (**C**) L4, ×10; (**D**) adult, ×10. Note staining is confined to the intestinal cells.

Ce-cdd-1 and *Ce-cdd-2* are expressed in the cells of the gut

In order to generate reporter constructs containing the appropriate signals for authentic gene expression, the transcriptional start site for each gene was determined using 5′-RACE. Subsequent sequencing of the products revealed that both transcripts were *trans*-spliced to the conserved 22 nt nematode spliced leader sequence SL-1. There was a single base between the 3′ end of SL-1 and the ATG in *Ce-cdd-1*, but no untranslated region (‘UTR’) in *Ce-cdd-2*. These results, together with sequence information in ACeDB, were used to design primers to amplify approx. 2 kb of upstream sequence. In the case of *Ce-cdd-1*, a translational fusion construct was generated containing 2 kb of

upstream sequence and 162 bases of gene sequence. With *Ce-cdd-2*, a transcriptional fusion containing only upstream sequence and a second translational fusion were analysed. Considering that both gave similar results, only the translational fusion is discussed. These fragments were cloned into the plasmid pPD96.04 and used to transform *C. elegans* in order to define the temporal and spatial expression patterns dictated by the putative promoter elements.

In lines transformed with the *cdd-1* reporter gene construct, faint staining was first detected in the egg, but the most intense staining was observed in the L1 and L2 stages (Figure 5). In the adult worm, transfection with the *cdd-1* construct resulted in a higher incidence of 'bagging' (eggs hatching within the parent), and in these animals intense staining of the newly hatched L1 *in utero* was observed (Figure 5A). This phenotype was only observed in transformed worms. Staining in the L1 and L2 stages was confined to the gut cells and to the pharyngeal bulbs. Between the L2 and the L3 stages, most of the intestinal cells ceased to express the reporter gene, except for a few cells towards the tail of the worm (Figure 5B). In later life cycle stages (L4 and adult), staining was largely confined to the pharynx and a cluster of cells at the tail of the worm (Figure 5C). For the *cdd-2* construct, reporter gene expression was confined to the cells of the intestine with no staining in the pharynx. The temporal expression pattern of *cdd-2* also differed from that of *cdd-1*: staining for the *cdd-2* reporter construct was first observed in the egg, but the number of positive cells increased through successive larval stages and was maintained into the adult worm (Figure 6). In the Figures shown, the *LacZ* staining was allowed to develop for 10 min for *cdd-1*, and for 4.5 h for *cdd-2*. The pattern of GFP expression for both reporter gene constructs was very similar to that described for *LacZ*.

RNAi with *Ce-cdd-1* and *Ce-cdd-2* results in a minimal phenotype

Sense and anti-sense single-stranded RNAs were produced for *Ce-cdd-1* and *Ce-cdd-2*, the corresponding strands were annealed, and the double-stranded RNA molecules were injected into the worms. In initial experiments no visible phenotype was observed, with the exception that the transformed worms appeared to be slightly longer and thinner than the wild-type ones. Subsequent experiments produced variable results: in two out of three experiments in which double-stranded RNA for *cdd-2* was injected, a significant proportion of injected worms died, whereas the water controls or worms injected with *cdd-1* double-stranded RNA were viable. The surviving P_0 and F_1 worms injected with *cdd-2* double-stranded RNA laid eggs, some of which were visibly abnormal. These eggs appeared to have very fragile eggshells, and were swollen and distorted in appearance at all temperatures. Micro-injection was also performed using both double-stranded RNAs for *cdd-1* and *cdd-2* together, but no additive effect was observed (results not shown).

Several different experiments were then performed using the RNAi feeding vector. In these experiments, worms at different stages of development were placed on plates expressing the appropriate construct or on control plates with *E. coli* containing L4440 vector alone. Plates were incubated at 16 °C, 21 °C or 25 °C, and observed at frequent intervals. No consistent phenotype in terms of gross morphology or development of the worm was observed at any temperature with either *cdd-1* or *cdd*-2, and RNAi had no effect upon viability or fertility of worms. However, in several experiments it was noted that eggs laid from F1 worms that had been fed L4440 containing the *cdd-2* insert were unable to hatch properly. These eggs looked very similar to those described above from worms micro-injected with *cdd-2* double-stranded RNA (Figure 7). No such phenotype was observed with the *cdd-1* construct at any temperature.

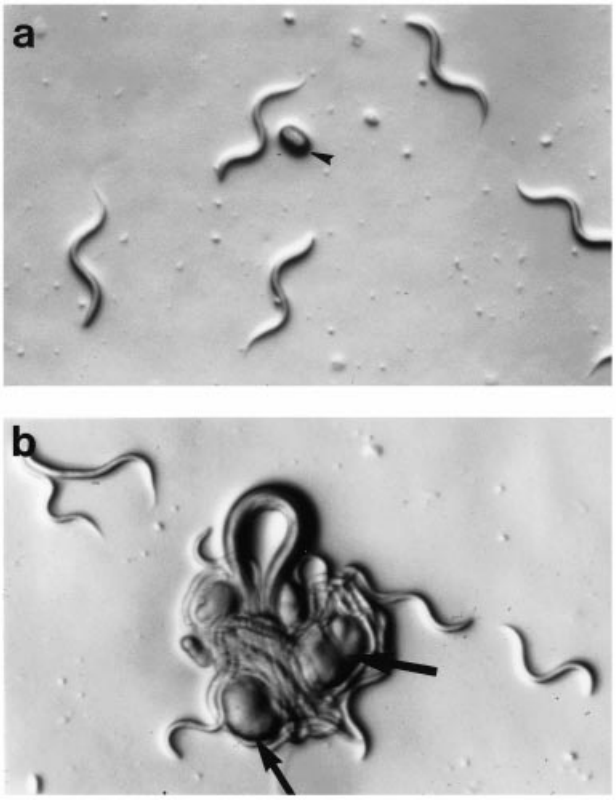

Figure 7 RNAi with *Ce-cdd-2* results in abnormal eggs

(**A**) Worms grown on a control plate seeded with HT115 cells containing the L4440 plasmid. The arrowhead shows a normal egg. (**B**) Worms grown on plates seeded with HT115 cells containing L4440 with the *cdd-2* insert. Eggs are from the F3 generation. Note the swollen appearance of the eggs (shown by an arrow). Both panels were photographed at a $\times 35$ magnification.

DISCUSSION

In this paper we describe the characterization of two CDDs from the nematode *C. elegans*. Biochemical analysis of the *C. elegans* recombinant proteins demonstrated that these were active deaminases that could be inhibited by Zn^{2+} chelation or with the reversible inhibitor THU. Even though all the nematode CDDs analysed to date cluster with the cytosine nucleoside deaminases rather than the RNA-editing family of enzymes, they do display RNA-binding activity towards AU-rich RNA templates, in contrast with other members of the cytosine nucleoside cluster. The significance of this observation in the worm is the focus of ongoing study. Neither of the *C. elegans* recombinant proteins bound to an apoB RNA template (results not shown), and therefore it is unlikely that either protein has RNA-editing activity, at least towards apoB. We attempted to define further the role of these proteins in *C. elegans* using double-stranded RNAi assays [25,28–30], a powerful tool for defining gene function in *C. elegans*. However, RNAi failed to give a definitive phenotype for either *cdd*-1 or *cdd*-2, or for both genes together. In both the feeding experiments and by micro-injection, it was clear that a proportion of eggs laid from adult worms exposed to

the *cdd-2* construct were abnormal and failed to hatch. This phenotype was specific to *cdd-2*. In two out of three microinjection experiments, there was a higher rate of mortality in worms injected with *cdd*-2 double-stranded RNA compared with those injected with either *cdd-1* or water alone, but no visible phenotype was observed in the surviving worms or in their progeny, making it difficult to conclude that this effect is related to inhibition of *cdd-2*. It is known that not all genes tested by RNAi have visible phenotypes when inhibited [27], and a *cdd* knock-out gene might fall into this category. In the mouse, deletion of apobec-1 did not provide any gross phenotype, except for the absence of apoB48 in the serum [31]. Recent studies on mammalian CDDs have revealed the existence of a family of related proteins [6], and in the *C. elegans* genome there is at least one other predicted protein that could potentially compensate for any phenotype obtained by RNAi with *Ce-cdd-1* and *-2*. This gene, *Ce-cdd-3*, is located on chromosome II (cosmid position C29F5.3), and was detected initially by scanning ACeDB with a PROSITE CDD signature motif [13]. In addition, there are three other predicted proteins in Worm Pep that share the active site motif of CDDs and deoxycytidylate deaminases.

Given the tight temporal and spatial expression of the *Bp*-CDD [12], we anticipated that the *C. elegans* CDDs might be similarly expressed at defined points in the life cycle and perhaps also expressed in the hypodermis. RT-PCR analysis demonstrated that *Ce-cdd-1* was most abundantly expressed in the early larval stages, whereas *Ce-cdd-2* was more highly expressed in later life cycle stages. These expression patterns broadly confirm results from microarray experiments, in which *Ce-cdd-2* was reported to be more highly expressed in adult worm mRNA compared with mixed-stage mRNA, and to be more abundant in hermaphrodite than in male worms [32]. *Ce-cdd-1* has also been analysed in whole genome microarray studies [33], and was shown to be expressed at extremely low levels in all stages analysed. The temporal expression patterns obtained with the reporter gene constructs confirmed these observations.

For both reporter constructs, approx. 2 kb of upstream sequence was sufficient to drive expression in the cells of the intestine. Consistent with the results of RT-PCR, *Ce-cdd-1* was most abundant in the early larval stages, while in the L3 and later stages, only the pharynx and a cluster of cells at the posterior of the gut were stained for reporter gene expression. In contrast, with the *cdd-2* reporter the intestinal cells were still strongly positive in the adult worm. Analysis of the upstream sequence of both genes revealed several potential control elements, which might direct gut-specific expression, including GATA elements, Skn-1 elements and, in *Ce-cdd-2*, an E-box domain [34–37]. Further analysis will be necessary to identify the functional elements that define tissue-specific expression in the intestine of *C. elegans*.

The localization of the CDDs to the intestinal cells of *C. elegans* is an interesting observation, particularly since apobec-1 is restricted to epithelial cells lining the gastrointestinal tract in humans [38]. The gut in *C. elegans* is an extremely active tissue, and the presence of CDDs in the cells of the intestine might reflect the requirement for nucleic acid synthesis in this compartment, and a role for CDD in the salvage pathway. However, gut function in *C. elegans* is not restricted to the digestion of food. For example, the major secretory products of the *C. elegans* intestine are the vitellogenins, or yolk proteins, that are secreted into the pseudocolem for uptake into the developing oocyte [39,40]. Intriguingly, vitellogenins are distantly related to apoB, and it is generally accepted that apoB arose from a vitellogenin precursor after the nematode/vertebrate divergence [41,42]. Thus the co-localization of the ancestral apoB (vitellogenin) and apobec-1 (CDD) to the intestine of *C. elegans* might hold clues for understanding the mechanisms by which apoB editing evolved in higher organisms. However, it is relevant to note that there is no evidence to date that the vitellogenins of *C. elegans* are either edited or provide a substrate for binding of CDD. Preliminary experiments in which RNAi was carried out in a transgenic reporter line that expressed a yolk protein (yp170) fused to GFP [40] showed no visible phenotype or differences in fertility.

Although we have yet to provide evidence that RNA editing catalysed by a CDD occurs in nematodes, it is interesting to note that the related family of editing enzymes, adenosine deaminases active on RNA ('ADARs') are found in *C. elegans* and have known editing substrates [43,44]. Recent studies in the mammalian editing field have identified an essential co-factor for apobec-1 activity, namely apobec-1 complementation factor (ACF). ACF contains multiple RNA-binding domains [3,4] and, *in vitro*, apobec-1 and ACF are sufficient for editing. A gene homologous with that of ACF is present in the *C. elegans* genome, and it will be of interest to characterize possible interactions between the *C. elegans* CDDs and the ACF in future studies.

This study was funded by a grant from the Wellcome Trust. N.O.D. was supported by grants HL-38180 and DK 56260 from the National Institutes of Health. We would like to thank Sarah Hunter for help with the expression of recombinant proteins, Iain Johnstone and co-workers for the provision of the *C. elegans* cDNA panel, Brett Roberts for the positive control for RNAi (all at the University of Glasgow), Barth Grant (Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, NY, U.S.A.) for provision of the *yp170::GFP* reporter strain, and John Gilleard (University of Glasgow) for many helpful discussions.

REFERENCES

1 Teng, B., Burant, C. F. and Davidson, N. O. (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein. Science **260**, 1816–1818
2 Navaratnam, N., Shah, R., Patel, D., Fay, V. and Scott, J. (1992) Apolipoprotein B mRNA editing is assosciated with UV crosslinking of proteins to the editing site. Proc. Natl. Acad. Sci. U.S.A. **90**, 222–226
3 Mehta, A., Kinter, M. T., Sherman, N. E. and Driscoll, D. M. (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. Mol. Cell. Biol. **20**, 1846–1854
4 Lellek, H., Kirsten, R., Diehl, I., Apostel, F., Buck, F. and Greeve, J. J. (2000) Purification and molecular cloning of a novel essential component of the apolipoprotein B mRNA editing enzyme-complex. J. Biol. Chem. **275**, 19848–19856
5 Blanc, V., Navaratnam, N., Henderson, J. O., Anant, S., Kennedy, S., Jarmuz, A., Scott, J. and Davidson, N. O. (2001) Identification of GRY-RBP as an apolipoprotein B RNA-binding protein that interacts with both apobec-1 and apobec-1 complementation factor to modulate C to U editing. J. Biol. Chem. **276**, 10272–10283
6 Gerber, A. P. and Keller, W. (2001) RNA editing by base deamination: more enzymes, more targets, new mysteries. Trends Biochem. Sci. **26**, 376–384
7 Anant, S. and Davidson, N. O. (2001) Molecular mechanisms of apolipoprotein B mRNA editing. Curr. Opin. Lipidol. **12**, 159–165
8 Dance, G. S., Sowden, M. P., Yang, Y. and Smith, H. C. (2000) APOBEC-1 dependent cytidine to uridine editing of apolipoprotein B RNA in yeast. Nucleic Acids Res. **28**, 424–429
9 Dance, G. S., Beemiller, P., Yang, Y., Mater, D. V., Mian, I. S. and Smith, H. C. (2001) Identification of the yeast cytidine deaminase CDD1 as an orphan C → U RNA editase. Nucleic Acids Res. **29**, 1772–1780
10 Martin, S. A. M., Hunter, S. J., Thompson, F. J. and Devaney, E. (1996) Stage specific gene expression in the post-infective L3 of the filarial nematode, *Brugia pahangi*. Mol. Biol. Parasitol. **79**, 109–112
11 Anant, S., Martin, S. A., Yu, H., MacGinnitie, A. J., Devaney, E. and Davidson, N. O. (1997) A cytidine deaminase expressed in the post-infective L3 stage of the filarial nematode, *Brugia pahangi*, has a novel RNA-binding activity. Mol. Biochem. Parasitol. **88**, 105–114

12 Hunter, S. J., Thompson, F. J., Tetley, L. and Devaney, E. (2001) Temperature is a cue for gene expression in the L3 of the filarial nematode *Brugia pahangi*. Mol. Biochem. Parasitol. **112**, 1–9
13 Mian, I. S., Moser, M. J., Holley, W. R. and Chatterjee, A. (1998) Statistical modelling and phylogenetic analysis of a deaminase domain. Comput. Biol. **5**, 57–72
14 Zaidi, S. H. E., Denman, R. and Malter, J. S. (1994) Multiple proteins interact at a unique *cis*-element in the 3′-untranslated region of amyloid precursor protein mRNA. J. Biol. Chem. **269**, 24000–24006
15 Anant, S. and Davidson, N. O. (2000) An AU-rich sequence element (UUUN[A/U]U) downstream of the edited C in apolipoprotein B mRNA is a high-affinity binding site for Apobec-1: binding of Apobec-1 to this motif in the 3′ untranslated region of c-myc increases mRNA stability. Mol. Cell. Biol. **20**, 1982–1992
16 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
17 Retief, J. D. (2000) Phylogenetic analysis using PHYLIP. Methods Mol. Biol. **132**, 243–258
18 Anant, S., Mukhopadhyay, D., Sankaranand, V., Kennedy, S., Henderson, J. O. and Davidson, N. O. (2001) ARCD-1, an apobec-1-related cytidine deaminase, exerts a dominant negative effect on C to URNA editing. Am. J. Physiol. Cell. Physiol. **281**, C1904–C1916
19 MacGinnitie, A. J., Anant, S. and Davidson, N. O. (1995) Mutagenesis of apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, reveals distinct domains that mediate cytosine nucleoside deaminase, RNA binding, and RNA editing activity. J. Biol. Chem. **270**, 14768–14775
20 Johnstone, I. L. and Barry, J. D. (1996) Temporal reiteration of a precise gene expression pattern during nematode development. EMBO J. **15**, 3633–3639
21 Fire, A., White-Harrison, S. and Dixon, D. (1990) A modular set of *lac Z* fusion vectors for studying gene expression. Gene **93**, 189–198
22 Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991) Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. **10**, 3959–3970
23 Kramer, J. M., French, R. P., Park, E. C. and Johnson, J. J. (1990) The *Caenorhabditis elegans* rol-6 gene, which interacts with the sqt-1 collagen gene to determine organismal morphology, encodes a collagen. Mol. Cell. Biol. **10**, 2081–2089
24 Mounsey, A., Molin, L. and Hope, I. A. (1999) Gene Expression Patterns. In *C. elegans*: A Practical Approach (Hope, I. A., ed.), pp. 181–200, Oxford University Press, Oxford
25 Fire, A., Xu, S., Montgomery, M., Kostas, S., Driver, S. and Mello, C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature (London) **391**, 806–811
26 Timmons, L., Court, D. L. and Fire, A. (2001) Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. Gene **263**, 103–112
27 Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M. and Ahringer, J. (2000) Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. Nature (London) **408**, 325–330
28 Barstead, R. (2001) Genome-wide RNAi. Curr. Opin. Chem. Biol. **5**, 63–66
29 Gonczy, P., Echeverri, G., Oegema, K., Coulson, A., Jones, S. J., Copley, R. R., Duperon, J., Oegema, J., Brehm, M., Cassin, E. et al. (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature (London) **408**, 331–336
30 Kuwabara, P. E. and Coulson, A. (2000) RNAi-prospects for a general technique for determining gene function. Parasitol. Today **16**, 347–349
31 Hirano, K., Young, S. G., Farese, R. V., Ng, J., Sande, E., Warburton, C., Powell-Braxton, L. M. and Davidson, N. O. (1996) Targeted disruption of the mouse apobec-1 gene abolishes apolipoprotein B mRNA editing and eliminates apolipoprotein B48. J. Biol. Chem. **271**, 9887–9890
32 Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V. and Kim, S. K. (2001) Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. U.S.A. **98**, 218–223
33 Hill, A. A., Hunter, C. P., Tsung, B. T., Tucker-Kellogg, G. and Brown, E. L. (2000) Genomic analysis of gene expression in *C. elegans*. Science **290**, 809–812
34 Britton, C, McKerrow, J. H. and Johnstone, I. L. (1998) Regulation of the *Caenorhabditis elegans* gut cysteine protease gene *cpr*-1: requirement for GATA motifs. J. Mol. Biol. **283**, 15–27
35 Britton, C., Redmond, D. L., Knox, D. P., McKerrow, J. H. and Barry, J. D. (1999) Identification of promoter elements of parasite nematode genes in transgenic *Caenorhabditis elegans*. Mol. Biochem. Parasitol. **103**, 171–181
36 Fukushige, T., Schroeder, D. F., Allen, F. L., Goszczynski, B. and McGhee, J. D. (1996) Modulation of gene expression in the embryonic digestive tract of *C. elegans*. Dev. Biol. **178**, 276–288
37 Stroeher, V. L., Kennedy, B. P., Millen, K. J., Schroeder, D. F., Hawkins, M. G., Goszczynski, B. and McGhee, J. D. (1994) DNA–protein interactions in the *Caenorhabditis elegans* embryo: oocyte and embryonic factors that bind to the promoter of the gut-specific ges-1 gene. Dev. Biol. **163**, 367–380
38 Hadjiagapiou, C., Giannoni, F., Funahashi, T., Skarosi, S. F. and Davidson, N. O. (1994) Molecular cloning of a human small intestinal apolipoprotein B mRNA editing protein. Nucleic Acids Res. **22**, 1874–1879
39 Kimble, J. and Sharrock, W. J. (1983) Tissue-specific synthesis of yolk proteins in *Caenorhabditis elegans*. Dev. Biol. **96**, 189–196
40 Grant, B. and Hirsh, D. (1999) Receptor-mediated endocytosis in the *Caenorhabditis elegans* oocyte. Mol. Biol. Cell **10**, 4311–4326
41 Baker, M. E. (1998) Is vitellogenin an ancestor of apolipoprotein B-100 of human low-density lipoprotein and human lipoprotein lipase? Biochem. J. **255**, 1057–1060
42 Chen, J. S., Sappington, T. W. and Raikhel, A. S. (1997) Extensive sequence conservation among insect, nematode, and vertebrate vitellogenins reveals ancient common ancestry. J. Mol. Evol. **44**, 440–451
43 Morse, D. P. and Bass, B. L. (1999) Long RNA hairpins that contain inosine are present in *Caenorhabditis elegans* poly(A)$^+$ RNA. Proc. Natl. Acad. Sci. U.S.A. **96**, 6048–6053
44 Hough, R. F., Lingam, A. T. and Bass, B. L. (1999) *Caenorhabditis elegans* mRNAs that encode a protein similar to ADARs derive from an operon containing six genes. Nucleic Acids Res. **27**, 3424–3432

Received 11 December 2001/12 February 2002; accepted 18 April 2002