Volume 13 Number 14 1985 Nucleic Acids Research

The *C. elegans* vitellogenin genes: short sequence repeats in the promoter regions and homology to the vertebrate genes

John Spieth, Karen Denison, Sarah Kirtland, Jerome Cane and Thomas Blumenthal

Program in Molecular, Cellular and Developmental Biology, and Department of Biology, Indiana University, Bloomington, IN 47405, USA

Received 7 May 1985; Accepted 24 June 1985

ABSTRACT

The nematode Caenorhabditis elegans contains a small family of vitellogenin genes which is expressed abundantly, but only in the intestine of the adult hermaphrodite worm. In order to identify possible regulatory elements, we have sequenced the DNA surrounding the 5' ends of five of the six genes. Contained within regions which have largely diverged from one another, two different heptameric sequences are found repeated within the first 200 bp upstream of each of the genes. The first sequence, TGTCAAT, is present as a perfect heptamer at least once upstream of each gene. It is repeated in both orientations four to six times in each 5' flanking region, allowing a one-base mismatch. The second sequence, CTGATAA, is also present as a perfect heptamer in a restricted region upstream of each gene. These two sequence elements may be involved in regulation of the vitellogenin genes. Remarkably, the CTGATAA sequence is present in a similar location in the promoter regions of vertebrate vitellogenin genes. In fact, our data reveal a surprising degree of similarity between the nematode and vertebrate vitellogenins.

INTRODUCTION

Oocytes of the nematode, Caenorhabditis elegans, accumulate four major yolk proteins: two that are about 170 kdaltons and two smaller proteins of 115 and 88 kdaltons (1, 2). The vitellogenins, precursors of the yolk proteins, are encoded by a small family of genes whose expression is limited by stage, sex, and tissue. They are synthesized in the adult hermaphrodite intestine, cotranslationally secreted into the body cavity, and subsequently taken up by the gonad (3, 4). The vitellogenin gene family is composed of six genes, each about 5.1 kb in length including a few short introns (4, 5). The relationships between the genes are summarized in Fig. 1. The 170 kd polypeptide, yp170A, is encoded by three very closely related genes: vit-3, vit-4, and vit-5. (The "vit" terminology, for "vitellogenin," is used in place of "YP" in order to conform with standard C. elegans genetic nomenclature.) Of the three, only vit-5 is known to be expressed. vit-3 and vit-4 are located in tandem on the X chromosome and are nearly identical to each other in the regions sequenced. The other 170 kd polypeptide, yp170B, is encoded by vit-2.

vit-1 is a pseudogene (see below). The two smaller yolk proteins arise by cleavage of a precursor of 180,000 daltons (6) which is encoded by vit-6, a distant member of the vit-1-vit-5 gene family (5).

Although transcription of the vitellogenin genes is apparently restricted to the 32 cells of the adult hermaphrodite intestine, the mRNAs accumulate to a level sufficient to make them among the most abundant mRNAs in a mixed population of larval and adult worms (4). We have determined the nucleotide sequences surrounding the 5' ends of five of the six vitellogenin genes. We have found that although the sequences of the promoter regions are largely diverged, certain small repeated elements are found upstream of all the members of this gene family. One of these is reminiscent of those found in several groups of co-regulated yeast genes: a short sequence repeated several times in both orientations upstream of each gene. In addition, there is another very highly conserved sequence found in a limited region of each promoter.

Walker et al. (7) have previously noted a relatively high degree of sequence conservation in the 5' untranslated region and the signal sequences of chicken and Xenopus vitellogenin genes. We show here that a surprising amount of homology also exists between the vertebrate and C. elegans genes in these regions. Furthermore, one of the presumptive regulatory sequences that we have identified is also present in the promoter regions of the vertebrate vitellogenin genes.

RESULTS AND DISCUSSION

Locating and Sequencing the 5' Ends

The approximate locations of the 5' ends of the six vitellogenin genes were determined by hybridization of ^{32}P-labeled cDNA, primed with calf-thymus DNA fragments, to restriction enzyme-digested λ clones containing the genes (4, 5). Some of the 5' ends were more precisely located by S1 and mung bean nuclease mapping (4). Restriction fragments (from 1.2 to 1.7 kb in length) containing the 5' regions of vit-1, vit-2, vit-4, vit-5, and vit-6 were subcloned into pUC9. These clones were extensively mapped using restriction enzymes which recognize 4 bp sequences. The nematode insert DNA was then extracted from the plasmids, digested with Sau3a, cloned into m13mp18 and 19, and sequenced by the dideoxy method (8). Fragments were aligned by comparison of the restriction maps with the restriction sites predicted by the sequences of the Sau3a fragments. The alignments were subsequently confirmed by sequen-

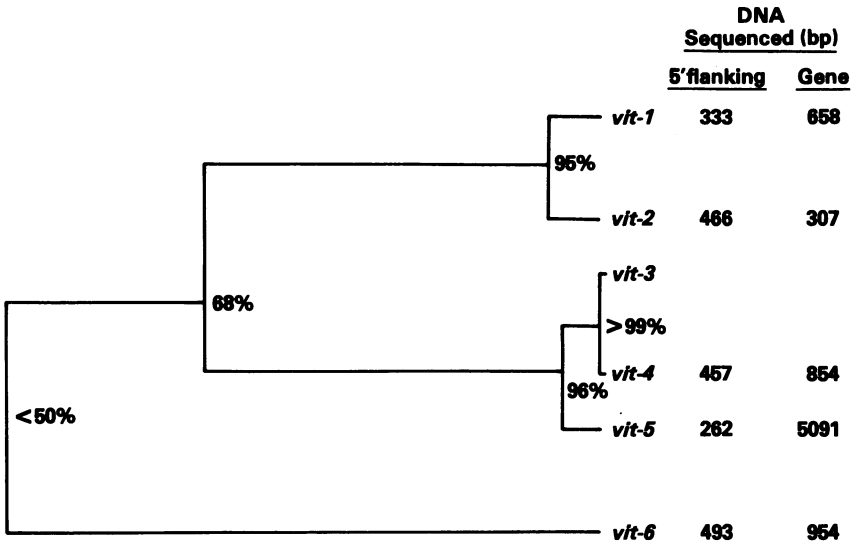

Figure 1. Schematic representation of the C. elegans vitellogenin gene family. Numbers at the divergence points represent percent homologies based on sequences of the following numbers of bases throughout the coding regions of each gene: vit-1, 2658; vit-2, 2507; vit-3, 650; vit-4, 1294; vit-5, 5091; vit-6, 954. The number of base pairs of contiguous DNA at the 5' end for which we currently have sequence data is shown on the right.

cing overlapping fragments cloned from restriction sites discovered by the previous sequencing.

The precise locations of the 5' ends of vit-2, vit-4, vit-5, and vit-6 mRNAs were determined by primer extension, accompanied by dideoxy sequencing, using synthesized fifteen base primers homologous to sequences near the 5' ends (not shown). Each gene showed a single, unambiguous 5' end: an A surrounded by pyrimidines (see Fig. 2).

We have determined the nucleotide sequence of portions of all genes except vit-3 (see Fig. 1 for the amount of contiguous DNA sequenced). In this communication we present only the sequences from -250 to about +60 (+1 indicates the first base in the mRNA). Figure 2 shows the coding sequences aligned by the first AUG in the mRNA, and the non-coding sequences aligned by the 5' ends of the mRNAs, as determined by primer extension experiments. The vit-1 sequence is aligned with the others by homology.

vit-1 is a pseudogene

Sequencing of vit-1 revealed a single base pair deletion, at position

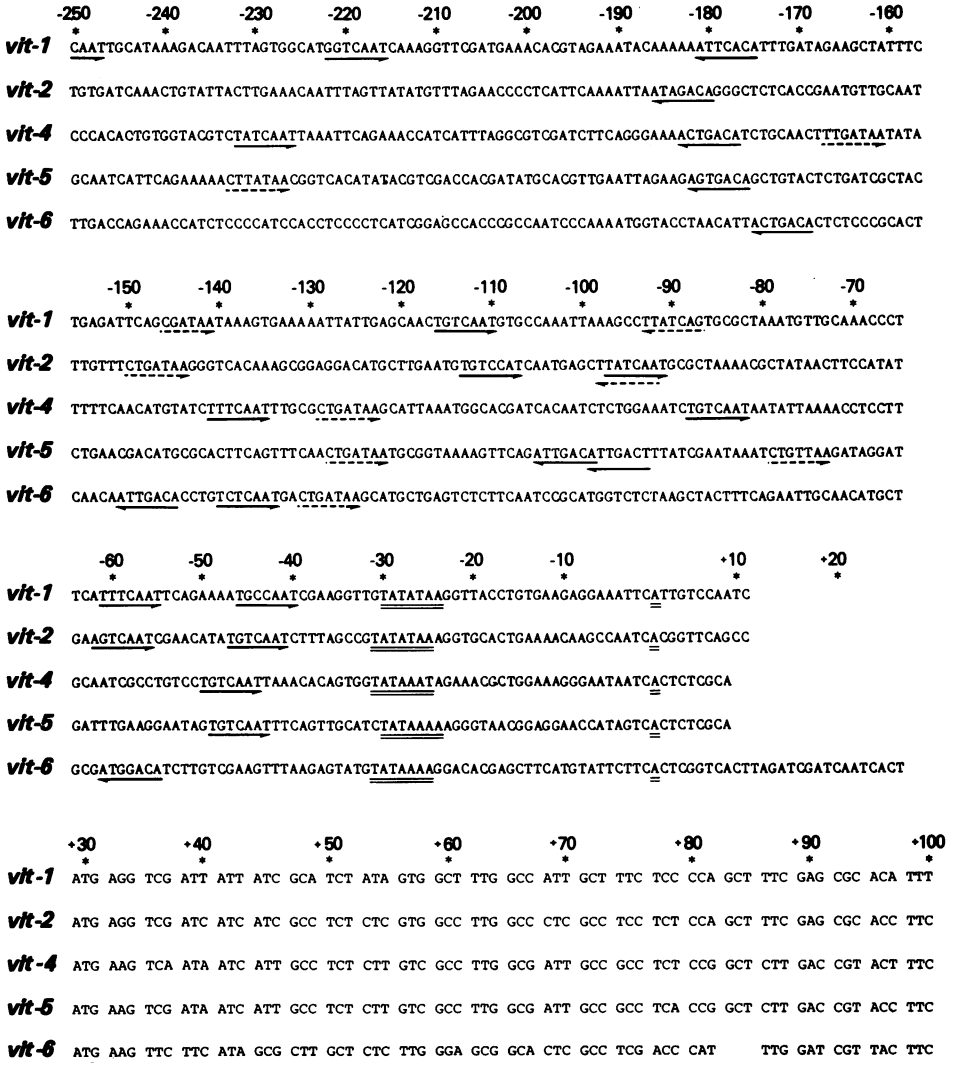

Figure 2. DNA sequences surrounding the 5' ends of the vitellogenin genes. DNA sequences of the 5' flanking and untranslated regions, determined by the methods of Sanger et al. (8) are aligned by the 5' ends of the mRNAs (numbered +1) as determined by primer extension. The TATA boxes are double-underlined. TGTCAAT-related sequences: solid arrows. CTGATAA-related sequences: broken arrows. DNA sequences of the coding regions are aligned by the first AUG in the mRNA and numbered from the first base in the mRNA of vit-6.

+184, with respect to the other vitellogenin genes. (Because this was the only such alteration we found, we checked the sequence in this region of vit-1 by sequencing both strands with the dideoxy as well as the chemical sequencing technique.) The deletion results in a stop codon in frame. Furthermore, multiple stop codons are also present in the other two reading frames. Thus, vit-1 can not be translated, but we do not know if it is transcribed. We performed a primer extension experiment with a primer specific for the predicted sequence of vit-1 mRNA and no extension product was detected. In a control experiment the same oligonucleotide was able to prime synthesis from the appropriate vit-1 M13 genomic clone (E. Zucker and J. Cane, unpublished). We conclude that vit-1 mRNA does not accumulate, a result consistent with either lack of transcription or mRNA instability due to premature translation termination (9).

The Coding and 5' Untranslated Regions

We have determined the nucleotide sequence of at least the first 307 bp of the coding region of each of the genes (not including vit-3). In this region vit-1 through vit-5 are all highly homologous. vit-6 is much more diverged, but nevertheless clearly related to the other genes. vit-2 is 82% homologous to vit-1, 67% to vit-5, and 50% to vit-6. vit-4 is 95% homologous to vit-5.

vit-1 through vit-5 have exceptionally short 5' untranslated sequences: from 9 to 11 bases in length (Fig. 2), making them among the shortest 5' untranslated sequences known (10). Interestingly, the 5' untranslated sequences of the vertebrate vitellogenin genes are also exceptionally short: only 13 bases long (7).

The 5' Flanking Sequences

The strong sequence conservation present within the coding regions does not extend into the 5' flanking sequences (Fig. 2). The five genes do have TATA boxes which are present at position -30 or -31 in each promoter. In addition, there are sequence elements which have been retained in the first 250 bp upstream of all five genes. (Since no additional resemblances between sequences further upstream could be detected, only the first 250 upstream bp are included in Fig. 2.)

A Multiply Repeated CAAT Box-like Element

The sequence, TGTCAAT, is present eight times as a perfect heptamer upstream of the five genes, at least once in every presumptive promoter region. If a 1 bp mismatch is allowed, the sequence is present a total of 27 times in the upstream regions, between four and six times in the 300 bp

Table 1
TGTCAAT-related Sequences in vit-1-vit-6

Gene	Orientation[+]		Consensus Sequence TGTCAAT	Location[‡]
vit-1	→		TATCAAT	-253
	→		GGTCAAT	-222
	←		TGTGAAT	-181
	→	*	TGTCAAT	-116
	→		TTTCAAT	-61
	→		TGCCAAT	-46
vit-2	←		TGTCTAT	-186
	→		TGTCCAT	-113
	→		TATCAAT	-97
	→		AGTCAAT	-62
	→	*	TGTCAAT	-47
vit-4	←	*	TGTCAAT	-302
	→		TATCAAT	-232
	←		TGTCAGT	-183
	→		TTTCAAT	-141
	→	*	TGTCAAT	-88
	→	*	TGTCAAT	-50
vit-5	←		TGTCACT	-182
	←	*	TGTCAAT	-105
	←		AGTCAAT	-99
	→	*	TGTCAAT	-49
vit-6	→		TGTCAAA	-272
	←		TGACAAT	-261
	←		TGTCAGT	-175
	←	*	TGTCAAT	-151
	→		TCTCAAT	-140
	←		TGTCCAT	-61

A	2	3	1	0	24	24	1
C	0	1	1	26	2	1	0
G	1	21	0	1	0	2	0
T	24	2	25	0	1	0	26
	T	G	T	C	A	A	T

*Indicates perfect consensus sequence.
[+]→ reads TGTCAAT, ← reads ATTGACA, in strand shown in Fig. 2.
[‡]Distance from first base in mRNA.

closest to each 5' end (Table 1). It is also present two additional times in the short intron located at position +123 in vit-6 (not shown). It can be present in either orientation. Furthermore, the heptamer is located preferentially in two positions: it is always present in one orientation at about

Table 2
Similarity Between TGTCAAT and Other Sequences Implicated in Gene Expression

Consensus	Sequence
Vitellogenins	TGTCAAT
CAAT box	GGT(C)CAATCT
Immunoglobulin octamer	ATGCAAAT
Opposite Orientation	
Vitellogenins	ATTGACA
Bacterial -35 region	TGTTGACAATTT
Yeast general amino acid control	AA(T)GTGACTC
Adenovirus enhancer core	A(C)GGAAGTGA(C/A)

position -180, and in vit-1 through vit-5, again at about position -45 in the opposite orientation. Within the sequence, not all of the seven bases are equally variable: the T's at position 3 and 7 and the C in position 4 are particularly well conserved.

The sequence bears striking homology to certain other short sequences previously implicated in gene expression (Table 2). In one orientation, TGTCAAT is identical to the CAAT box at six of its seven positions (11). However, it differs from this sequence both by being nearer to the TATA box than in previously observed cases and by being multiply repeated in each promoter region in both orientations. It is also quite similar to the octamer found in either orientation at position -90 upstream of immunoglobulin genes (12). In the opposite orientation, the sequence (ATTGACA) shares six contiguous bases with the bacterial -35 element, a part of the RNA polymerase recognition site (13). It also shares five bases with the yeast upstream activating sequence for general amino acid control (14) and with a consensus sequence for several enhancer elements, called the "adenovirus core enhancer" (15). TGTCAAT is not generally found upstream of C. elegans genes. We have searched the published promoter regions of two collagen genes (16), four actin genes (17), a myosin gene (18), and a gene for the major sperm protein (19). The sequence is present only in the latter, at about position -85.

Many regulatory elements, including those controlling genes from sources

as diverse as humans and yeast, have been found to be repeated in promoter regions (14, 20, 21). We believe the high degree of repetition of the TGTCAAT sequence element, specifically in the vitellogenin gene promoter regions, strongly implicates it in regulation of these C. elegans genes. The fact that TGTCAAT is located in about the same position in each gene adds further support to this idea. Like TGTCAAT, several previously identified eucaryotic regulatory elements, including the heavy metal response element of the metallothionein genes (20, 21), and the heat shock response element (22), have been found to occupy a position just upstream of the TATA box. Finally, the observation that the sequence is so closely related to those already implicated in RNA polymerase binding, to upstream activating sites and to enhancers suggests that these cis-acting elements may be functionally more similar to each other than previously suspected.

A Sequence Found Upstream of Nematode and Vertebrate Vitellogenin Genes

The sequence CTGATAA is present as a perfect heptamer between positions -90 and -150 in each vitellogenin promoter region, again in both orientations (Table 3). Allowing a 1 bp mismatch, the sequence appears a total of ten times in the five upstream regions (Fig. 2). In contrast, it is missing from all of the other published C. elegans promoter regions. A perfect CTGATAA heptamer is present in the same orientation between -128 and -150 in every gene but vit-1. In vit-1, this region contains the same sequence with the first T deleted. We speculate that this deletion could result in lack of transcription of vit-1. Surprisingly, CTGATAA also appears in both the chicken vitellogenin and Xenopus vitellogenin A gene promoters between positions -88 and -97 (23). It is present an additional time upstream of the chicken gene. A closely related sequence TTGATAA is present at about position -200 in the Xenopus B vitellogenin genes; and another closely related sequence, CTGATGA, is present upstream of the two sequenced Drosophila yolk protein genes, YP1 and YP2 (24-26). Of the 26 examples of this sequence listed in Table 3, the T at position 2 and the A at position 7 are invariant while the G at position 3 and the T at position 5 vary only once each. The high degree of conservation and the restricted location of this sequence suggest that it may play an important role in regulation of vitellogenin gene expression.

Other Sequences

We also noticed the presence of another potentially important sequence, consensus TGCGCA/T, between positions -100 and -150 of each vitellogenin gene (Fig. 2). This sequence is usually immediately adjacent to the CTGATAA sequence. It is very closely related to the heavy metal response element of

Table 3
CTGATAA-related Sequences

Gene	Orientation[+]		Consensus Sequence CTGATAA	Location[‡]
vit-1	←	*	CTGATAA	-93
vit-2	→	*	CTGATAA	-150
	←		TTGATAA	-98
vit-4	←		ATGATAA	-446
	→		TTGATAA	-167
	→	*	CTGATAA	-129
vit-5	→		CTTATAA	-233
	→	*	CTGATAA	-128
	→		CTGTTAA	-79
vit-6	→	*	CTGATAA	-131
Chicken Vitellogenin II	←	*	CTGATAA	-176
	→	*	CTGATAA	-88
Xenopus Vitellogenin A1	→		ATGATAA	-527
	←		CTGACAA	-234
	→	*	CTGATAA	-88
A2	→		CTGATCA	-319
	→	*	CTGATAA	-97
B1	←		ATGATAA	-311
	→		TTGATAA	-201
	←		CTGGTAA	-77
B2	←		ATGATAA	-314
	→		TTGATAA	-202
Drosophila YP1	←		CTGATTA	-257
	→		CTGCTAA	-236
	←		CTGATGA	-228
Drosophila YP2	←		CTGATGA	-161

A	4	0	0	23	0	22	26
C	18	0	0	1	1	1	0
G	0	0	25	1	0	2	0
T	4	26	1	1	25	1	0
	C	T	G	A	T	A	A

*Indicates perfect consensus sequence.
[+]→ reads CTGATAA, ← reads TTATCAG, in strand shown in Fig. 2.
[‡]Distance from first base in mRNA.

the metallotheinin genes (20, 21).

A symmetrical sequence recently reported to be repeated upstream of vertebrate vitellogenin genes as well as another estrogen-regulated liver-specific gene, VLDLII, is not present in the portions of the C. elegans genes we have sequenced (23). Since it was found within the first 375 upstream bp of the vertebrate genes, we should have found it in at least vit-2, vit-4, and vit-6 if it occupied an homologous position in the C. elegans genes.

In comparing the sequences of promoter regions of estrogen-regulated liver specific genes of chicken, the vitellogenin gene II and VLDLII, Burch discovered four short sequence elements present in both (27). One of these, CTGGTCAAT, bears a striking resemblance to both of the sequence elements found upstream of the nematode vitellogenin genes. It shares six contiguous bases with TGTCAAT and five out of seven bases with CTGATAA. Determination of the functional significance of these homologies awaits molecular dissection of the vitellogenin promoters.

Comparison of Nematode and Vertebrate Vitellogenins

Figure 3 shows the predicted amino acid sequences of the first 71 residues of the vitellogenins encoded by vit-1-vit-6, beginning translation from the first AUG in the mRNA. The vit-2-encoded polypeptide is 91% homologous to vit-1, 75% to vit-5 and 31% to vit-6 in the region shown in Fig. 3. The vit-4- and vit-5-encoded proteins are 97% homologous in this region. All of the encoded proteins begin with a hydrophobic sequence which is presumably a signal sequence responsible for co-translational secretion.

The similarities in size between the nematode and vertebrate vitellogenins and between the 5' untranslated regions of their mRNAs, as well as the high degree of sequence conservation between the chicken and frog genes (7), led us to compare the predicted amino acid sequences of the nematode and vertebrate genes. Figure 3 shows the predicted amino acid sequences of the first 71 residues (encoded by the first three exons) of five vertebrate vitellogenins: one from chick (7, 27, 28) and four from Xenopus (7, 29). Amino acid residues found in at least one vertebrate and one nematode protein are shaded. The results reveal a striking degree of sequence homology. The first exon of the vertebrate genes, which codes for the signal sequence, shows the highest degree of conservation with the corresponding region of the nematode genes. For instance, in this region of the Xenopus vitellogenin B1 gene (Fig. 3, up to first triangle), 9 out of 13 amino acid residues are identical with the C. elegans vit-2 protein. [Comparison of signal sequences of unrelated proteins does not generally reveal homologous signal sequences (30)]. The seven amino acid sequence encoded by the second exon (Fig. 2, between triangles) is the

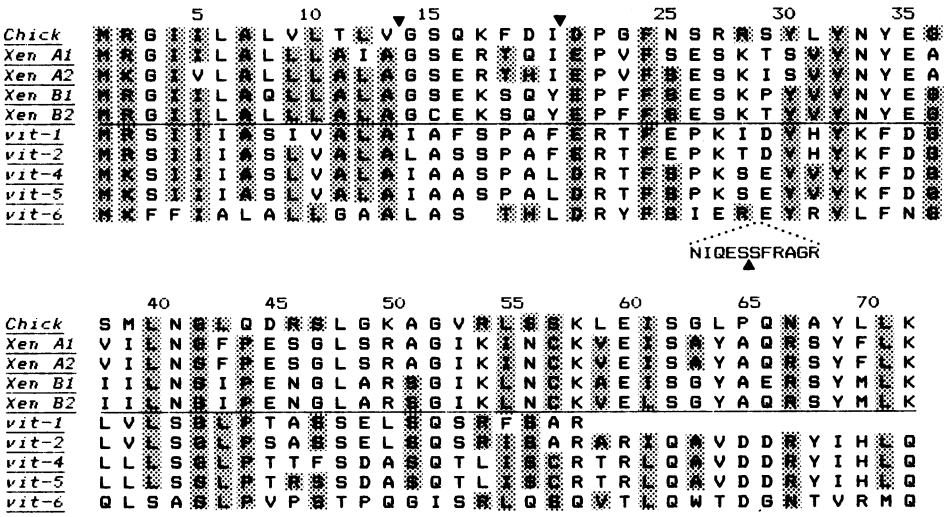

Figure 3. Predicted amino acid sequences of the amino termini of vertebrate and nematode vitellogenins. The chicken and Xenopus vitellogenins (7, 27-29) are aligned with the products of the vit-1-vit-5 nematode genes without the introduction of gaps. The vit-6 polypeptide was aligned with the other nematode vitellogenins by the introduction of one single-amino acid gap and an 11-amino acid insertion. Inverted triangles mark the position of introns present in all vertebrate genes. The position of an intron present only in vit-6 is marked by a triangle. Residues found in identical locations in at least one vertebrate and one nematode vitellogenin are shaded. The standard single-letter code for amino acids is used.

most diverged within the vertebrate lines (29) and shows no detectable homology with the corresponding portions of the predicted nematode proteins, although the number of amino acids between regions of homology has been maintained. In the part of the protein encoded by the third exon, the Xenopus B1 sequence is identical to the C. elegans vit-2 sequence at 13 of 51 positions. Overall, 35 of the 71 positions for which sequence data is available contain identical amino acids between at least one vertebrate and one nematode vitellogenin. The relationship between the vitellogenins encoded by the Xenopus B1 gene and vit-2 was verified by computer analysis of the amino acid sequences using the mutation data matrix and the ALIGN program (31). There are 22 identities out of 71 possible matches between residues. The alignment score is 14.08 SD, which is highly significant (31). The above comparisons were made without inserting gaps within the vertebrate genes or vit-1-vit-5. In order to align vit-6 with the other nematode genes, however, it was necessary to pl-

ace a one-amino acid gap and an 11-amino acid (plus intron) insertion into vit-6.

In the regions of the vertebrate genes sequenced to date, the vit-1-vit-5 family is more closely related to the Xenopus gene family than it is to the chicken gene. All of the vertebrate genes are more closely related to the vit- 1-vit-5 family than they are to vit-6. This observation is consistent with the fact that the products of both the vertebrate and vit-1-vit-5 genes are taken up intact by the oocytes while the vit-6 vitellogenin is processed before uptake.

In contrast to the surprisingly large similarity between nematode and vertebrate vitellogenins, we could detect no similarity between the nematode and Drosophila vitellogenins (24-26).

In sum, our data indicates that the vitellogenin genes of nematodes and vertebrates are distant members of a single gene family. In the hundreds of millions of years since the separation of the nematode and vertebrate lines several dramatic differences have arisen. The multiple large introns present in the vertebrate genes are absent from the nematode genes. Furthermore, while the vertebrate gene products are processed to much smaller polypeptides in the oocyte, the nematode gene products remain nearly or entirely intact.

Regulation of Vitellogenin Gene Transcription

The genes whose promoter regions have been characterized are strictly regulated during C. elegans development. Vitellogenin gene expression is restricted in three ways: by tissue, by sex, and by stage: the mRNA is present in abundance, but only in the adult hermaphrodite intestine (4). We hypothesize that the sequences we have identified are cis-acting elements involved in this developmental regulation. The highly repeated nature of one of the elements, TGTCAAT, could be responsible for the abundance of vitellogenin gene expression in the adult hermaphrodite intestine. The presence of the other sequence, CTGATAA, in a restricted region of both nematode and vertebrate vitellogenin promoters suggests that it also plays a role in regulation of this gene family. Experiments to test these ideas will require developmentally regulated expression in nematodes transformed with the vitellogenin promoter regions fused to a gene coding for an assayable product.

ACKNOWLEDGEMENTS

We are grateful to W. Klein, B. Polisky, and S. Strome for comments on the manuscript, and to W. Sharrock for much helpful advice. This work was

supported by research grant GM30870 from the National Institute of General Medical Sciences.

REFERENCES

1. Klass, M.R., Wolf, N., and Hirsh, D. (1979) Develop. Biol. 69, 329-335.
2. Sharrock, W.J. (1983) Develop. Biol. 96, 182-188.
3. Kimble, J. and Sharrock, W.J. (1983) Develop. Biol. 96, 189-196.
4. Blumenthal, T., Squire, M., Kirtland, S., Cane, J., Donegan, M., Spieth, J., and Sharrock, W.J. (1984) J. Mol. Biol. 174, 1-18.
5. Spieth, J. and Blumenthal, T. (1985) Mol. Cell. Biol. In press.
6. Sharrock, W.J. (1984) J. Mol. Biol. 174, 419-431.
7. Walker, P., Brown-Luedi, M., Germond, J-E., Wahli, W., Meijlink, F., Vanhet, P.W., Schip, A.D., Roelink, H., Gruber, M., and Ab, G. (1983) EMBO J. 2, 2271-2279.
8. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
9. Boss, J.M., Darrow, M.D., and Zitomer, R.S. (1980) J. Biol. Chem. 255, 8623-8628.
10. Kozak, M. (1984) Nucl. Acids Res. 12, 857-873.
11. Benoist, C., O'Hare, K., Breathnach, R., and Chambon, P. (1980) Nucl. Acids Res. 8, 127-142.
12. Parslow, T.G., Blair, D.L., Murphy, W.J., and Granner, D.K. (1984) Proc. Natl. Acad. Sci. USA 81, 2650-2654.
13. Rosenberg, M. and Court, D. (1979) Ann. Rev. Genet. 13, 319-354.
14. Hinnebusch, A.G. and Fink, G.R. (1983) J. Biol. Chem. 258, 5238-5247.
15. Hearing, P. and Shenk, T. (1983) Cell 33, 695-703.
16. Kramer, J.M., Cox, G., and Hirsh, D. (1985) J. Biol. Chem., in press.
17. Files, J.G., Carr, S., and Hirsh, D. (1983) J. Mol. Biol. 164, 355-375.
18. Karn, J., Brenner, S., and Barnett, L. (1983) Proc. Natl. Acad. Sci. USA 80, 4253-4257.
19. Klass, M.R., Kinsley, S., and Lopez, L.C. (1984) Mol. Cell. Biol. 4, 529-537.
20. Carter, A.D., Felber, B.K., Walling, M., Jubier, M-F., Schmidt, C.J., and Hamer, D.H. (1984) Proc. Natl. Acad. Sci. USA 81, 7392-7396.
21. Stuart, G.W., Searle, P.F., Chen, H.Y., Brinster, R.L., and Palmiter, R.D. (1984) Proc. Natl. Acad. Sci. USA 81, 7318-7322.
22. Pelham, H.R.B. (1982) Cell 30, 517-528.
23. Walker, P., Germond, J-E., Brown-Luedi, M., Givel, F., and Wahli, W. (1984) Nucl. Acids Res. 12, 8611-8626.
24. Hung, M-C. and Wensink, P. (1983) J. Mol. Biol. 164, 481-492.
25. Hung, M-C. and Wensink, P. (1981) Nucl. Acids Res. 9, 6407-6419.
26. Hovemann, B. and Galler, R. (1982) Nucl. Acids Res. 10, 2261-2274.
27. Burch, J.B.E. (1984) Nucl. Acids Res. 12, 1117-1135.
28. Geiser, M., Mattaj, I.W., Wilks, A.F., Seldran, M., and Jost, J-P. (1983) J. Biol. Chem. 258, 9024-9030.
29. Germond, J-E., Walker, P., Heggeler, B., Brown-Luedi, M., de Bony, E., and Wahli, W. (1984) Nucl. Acids Res. 12, 8595-8609.
30. Von Heinje, G. (1983) D.J. Biochem. 133, 17-21.
31. Barker, W.C., Ketcham, L.K., and Dayhoff, M.D. (1978) in Atlas of Protein Sequence and Structure, ed. Dayhoff, M.O. (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 3, pp. 359-362.