HAL
open science

Molecular Evolution of Ankyrin: Gain of Function in Vertebrates by Acquisition of an Obscurin/Titin-Binding-related Domain (OTBD).

Alexander A. Hopitzan, Anthony J. Baines, Ekaterini Kordeli

To cite this version:

Alexander A. Hopitzan, Anthony J. Baines, Ekaterini Kordeli. Molecular Evolution of Ankyrin: Gain of Function in Vertebrates by Acquisition of an Obscurin/Titin-Binding-related Domain (OTBD).. Molecular Biology and Evolution, 2006, 23, pp.46-55. 10.1093/molbev/msj004 . hal-00008681

HAL Id: hal-00008681

https://hal.science/hal-00008681v1

Submitted on 13 Sep 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

This is a pre-copy-editing, author-produced PDF of an article accepted for publication in "Molecular Biology and Evolution" following peer review.
The definitive publisher-authenticated version is available online at: http://dx.doi.org/10.1093/molbev/msj004

Molecular Evolution of Ankyrin: Gain of Function in Vertebrates by Acquisition of an Obscurin/Titin-Binding-related Domain (OTBD).

(Research Article)

Alexander A. Hopitzan[1], Anthony J. Baines[2], and Ekaterini Kordeli[1*]

[1]Institut Jacques Monod/CNRS/Universités Paris VI et VII; 2 Place Jussieu, F-75251 Paris cedex 05, France

[2]Department of Biosciences, University of Kent, Canterbury, Kent, CT2 7NJ, England

*Corresponding author: Ekaterini Kordeli

Email: kordeli@ijm.jussieu.fr
Telephone: +33 (1) 44 27 42 18
Fax: +33 (1) 44 27 59 94.

Running Title: Molecular Evolution of Ankyrins

Keywords: Ankyrin; Ank_{G107}; Striated Muscle; Module; Titin; Obscurin.

Abstract

Ankyrins form a family of modular adaptor proteins that link between integral membrane proteins and the cytoskeleton. They evolved within the metazoa as an adaptation for organizing membrane microstructure and directing membrane traffic. Molecular cloning has identified one *Caenorhabditis elegans* (*unc-44*), two *Drosophila* (*Dank1*, *Dank2*) and three mammalian (*Ank1*, *Ank2*, *Ank3*) genes. We have previously identified a 76 amino acid alternatively spliced sequence that is present in muscle polypeptides encoded by the rat *Ank3* gene. A closely related sequence in a muscle *Ank*1 product binds the cytoskeletal muscle proteins obscurin and titin. This Obscurin/Titin-Binding-related Domain (OTBD) contains repeated modules of 18 amino acids: three are encoded by *Ank1* and *Ank2*, two by *Ank3*; this pattern is conserved throughout vertebrate ankyrin genes. The *Caenorhabditis elegans* ankyrin, UNC-44, contains one 18 amino acid module, as does the ankyrin gene in the urochordate *Ciona intestinalis,* but the insect ankyrins contain none. Our data indicate that an ancestral ankyrin acquired a 18 amino acid module which was preserved in the ecdysozoa/deuterostome divide, but it was subsequently lost from arthropods. Successive duplications of the module led to a gain of function in vertebrates as it acquired obscurin/titin binding activity. We suggest that the OTBD represents an adaptation of the cytoskeleton that confers muscle cells with resilience to the forces associated with vertebrate life.

Introduction

Ankyrins mediate linkage of integral proteins to the spectrin-based cytoskeleton. Gene knock-out experiments, siRNA depletion, and natural mutations point to a crucial role of ankyrins in organizing membrane domains and in delivering ion channels and cell adhesion proteins to requisite membrane sites (Bennett and Baines 2001; Mohler et al. 2003; Kizhatil and Bennett 2004; Mohler et al. 2004).

Three ankyrin genes in mammals (*Ank1, Ank2* and *Ank3*), two in *D. melanogaster* (*Dank1*, *Dank2*) and one in *C. elegans* (*unc-44*) are known to date (Lambert et al. 1990; Lux, John, and Bennett 1990; Otto et al. 1991; Dubreuil and Yu 1994; Kordeli, Lambert, and Bennett 1995; Otsuka et al. 1995; Peters et al. 1995; Bouley et al. 2000). So far, no ankyrins have been found in plants, fungi, yeast, and bacteria. Thus, ankyrins are likely to have evolved early in evolution of metazoans to meet the requirement of animal cells for generation and maintenance of complex membrane structures (Mohler, Gramolini, and Bennett 2002b).

Ankyrins are modular proteins that contain a series of domains that are highly conserved between gene products (the N-terminal membrane-binding Ank-repeats, the spectrin-binding ZU5 domain, and death domain). Additionally, the different genes encode unique regions in the C-terminal regions. Mammals generate functionally specialized isoforms during development and in different tissues by complex patterns of mRNA splicing (Bennett and Baines 2001) including by splicing the C-terminal regions (Hall and Bennett 1987; Davis, Davis, and Bennett 1992; Mohler, Gramolini, and Bennett 2002a).

We previously identified a unique 76-residue insertion (the "76aa" insert) in the C-terminal region of Ank_{G107}. Ank_{G107} is a muscle isoform of ankyrin-G (as products of the *Ank3* gene are known) which is targeted to the sarcolemma (Gagelin et al. 2002). Elucidation of the

expression pattern of *Ank3* in muscle (Hopitzan et al. 2005) strongly suggested that 76aa is present in all muscle ankyrins-G, but is not expressed outside muscle.

The C-terminal regions of *Ank1* and *Ank2* gene products (ankyrins-R and ankyrins-B, respectively) contain a sequence closely similar to 76aa, with highest conservation in sAnk1, an *Ank1* gene product. sAnk1 is a muscle-specific, integral, truncated (~25 kDa) ankyrin which localizes to M and Z lines of sarcomeres and is thought to link the sarcoplasmic reticulum to myofibrils (Zhou et al. 1997; Birkenmeier et al. 1998; Gallagher and Forget 1998). Two giant myofibrillar proteins, titin and obscurin, interact with the 76aa regions of sAnk1; obscurin also binds the 76aa region of an *Ank2* gene product (Bagnato et al. 2003; Kontrogianni-Konstantopoulos and Bloch 2003; Kontrogianni-Konstantopoulos et al. 2003). Titin and obscurin belong to the same family of modular proteins of vertebrate striated muscle. Both proteins have been suggested to play crucial roles in myofibrillogenesis (Young, Ehler, and Gautel 2001; Granzier and Labeit 2002; Russell et al. 2002; Tskhovrebova and Trinick 2003; Kontrogianni-Konstantopoulos et al. 2004; Tskhovrebova and Trinick 2004). Vertebrate titin forms flexible filaments of more than 1µm in length that span half a sarcomere from the Z- to the M-line in vertebrate striated muscle. The two most N-terminal domains of titin (two Ig domains, ZIg1 and ZIg2) are involved in binding sAnk1 at the Z-line (Kontrogianni-Konstantopoulos and Bloch 2003). Obscurin colocalizes with sAnk1at the level of both, Z- and M-lines. The last C-terminal 400 residues were shown to bind sAnk1 (Bagnato et al. 2003; Kontrogianni-Konstantopoulos et al. 2003).

Titin and obscurin homologs (Tskhovrebova and Trinick 2003; Hooper and Thuma 2005), as well as ankyrin (*C. elegans* body wall muscles; Chen, Ong, and Bennett 2001), have been also reported in invertebrate muscle. Interestingly, an ankyrin-like protein has been localized in the myoplasm and muscle cells of ascidian eggs and embryos, and seems to play a role in muscle

development, possibly coordinating the linkage between membrane cytoskeleton, sarcoplasmic reticulum, and sarcomeres (Jeffery and Swalla 1993).

These findings raise a number of intriguing questions about the origin of the 76aa-related sequences, as well as the functional consequences of their presence in ankyrin molecules. Here we report that 76aa homologous sequences are present in all paralogous vertebrate ankyrin genes and show a modular architecture. Intriguingly, one of these 18 amino acid modules is present in both the ecdysozoan and deuterostome invertebrate lineages. We conclude that an early metazoan acquired such a module, which was adapted in vertebrate evolution by successive duplication to interact with muscle proteins, including obscurin/titin. These data indicate a gain of function in vertebrate ankyrin evolution.

Materials and Methods

Sequence retrieval

Protein sequences homologous to the *Ank3* 76aa insert were retrieved by PSI-BLAST search (Altschul et al. 1997) querying the NCBI BLAST server (http://www.ncbi.nlm.nih.gov/BLAST/) against the non-redundant and RefSeq databases. Conserved sequences in the *C. elegans* ankyrin UNC-44 AO13 were identified using the Wise2 server (http://www.ebi.ac.uk/Wise2/). Textual Wise2 outputs were graphically displayed and processed with the genome annotation tool Artemis release 6.0 (http://www.sanger.ac.uk/Software/Artemis/v6/).
Nucleotide sequences homologous to the 76aa insert were searched by BLAT (http://genome.ucsc.edu) and by the tBLASTn procedure (Altschul et al. 1990) against non-redundant, EST, RefSeq, and whole genome databases on following sites: http://www.ncbi.nlm.nih.gov; http://www.ensembl.org/ (H. sapiens v25.34e.1, M. musculus v25.33a.1, R. norvegicus v25.3c.1, D. rerio v25.4.1, G. gallus v25.1b.1 , A. gambiae v 25.2b.1, F. rubripes v25.2c.1, D. melanogaster v25.3b.1, P. troglodytes v25.1.1 , A. mellifera v25.1.1, T. nigroviridis v 25.1.1, C. familiaris vBROADD1, C. elegans v25.116a.1, B. taurus vBtau_1.0); http://genome.jgi-psf.org/Xentr3/Xentr3.home.html, http://www.sanger.ac.uk/cgi-bin/blast/submitblast/x_tropicalis (X. tropicalis assembly3); http://genome.jgi-psf.org/ciona4/ciona4.home.html (C. intestinalis assembly1); http://atlas.cnio.es/Caenorhabditis_briggsae/ (C. briggsae v25.25.1). Predicted matches (e.g. exon-intron boarders) were confirmed by eye and in the case of less conserved fish sequences refined by annotating the 76aa insert to genomic sequence using Wise2. For hidden Markov model analysis, the program HMMer was used (Eddy 1998). Protein sequence was deduced from extracted exonic sequences using the Lasergene (DNASTAR) software package.

Nucleotide sequences corresponding to the ZU5 domain (smart entry: http://smart.embl-heidelberg.de/smart/do_annotation.pl?DOMAIN=ZU5&BLAST=DUMMY, pfam entry: http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00791) were retrieved as described above. All database searches were performed before December 2004; nucleotide sequence queries against non-redundant, EST, and RefSeq databases have been repeated before March 2005. Results are available online as supplementary data.

Sequence Alignments

All multiple nucleotide and protein sequence alignments were performed using ClustalX (Thompson et al. 1997) and were manually optimized using the BioEdit software (Hall 1999). Sequence repeats within the OTBD were initially identified using the programs REPRO (http://ibivu.cs.vu.nl/) and RADAR (http://www.ebi.ac.uk/Radar/). The "sequence logo" representing aligned module sequences was created using WebLogo (Crooks et al. 2004). Multiple nucleotide sequence alignments are available online as supplementary data.

Structure predictions

To predict the structure of the OTBD, we used the metaserver at Genesilico (http://genesilico.pl; Kurowski and Bujnicki 2003). This server is a gateway to multiple high quality methods of secondary structure prediction and fold recognition. Additionally, we used servers for PONDR (http://www.pondr.com) and Globplot (http://globplot.embl.de; Linding et al. 2003) to analyse the potential for disordered structure.

Phylogenetic Analysis

Phylogenetic analysis was conducted on 48 nucleic acid sequences corresponding to the ZU5 domain (for sequence alignment see supplementary data). The construction of the maximum

likelihood (ML) tree was performed with the program Treefinder (www.treefinder.de) using the general time reversible substitution model for nucleotides (GTR) with 4 rate categories and default parameters (Jobb, von Haeseler, and Strimmer 2004). Robustness of the topology was assessed using 1000 bootstrap replicates. The tree was rerooted using *C. elegans* and *C. briggsae* as outgroups.

Results and Discussion

1. The muscle-specific 76aa insert of ankyrins-G defines a novel domain conserved among vertebrate ankyrins

Figure 1 shows a schematic diagram of the muscle ankyrin Ank_{G107}, encoded by the rat *Ank3* gene. In the C-terminal region is a 76 amino acid insert encoded by three consecutive exons (Gagelin et al. 2002; Hopitzan et al. 2005). We previously reported that sequences homologous to the 76aa are encoded by all three mammalian ankyrin genes (Gagelin et al. 2002). To gain further insight into the nature of the 76aa, we performed a multiple sequence alignment of C-terminal domains of ankyrins from human, mouse and rat (fig. 2*A*).

In previous alignments using available ankyrin-B isoforms, the *Ank2* gene products displayed conservation over only about half of the 76aa sequence (Gagelin et al. 2002). However, in database searches, we found sequences in which homology extends further: these sequences include human partial protein product CAD98033 (derived from the uterus EST DKFZp686M09125) and rat RefSeq entry XP_227735. Fig. 2*A* includes alignment of these sequences, and indicates that homology extends to the whole region corresponding to the 76aa. *Ank2* mRNAs can evidently be spliced to include or exclude the first exon of the 76aa homologous sequence.

PSI-BLAST searches revealed that no proteins other than ankyrins contain the 76aa sequence. Surprisingly, in the *C. elegans* ankyrin UNC-44, a short stretch of about 20 residues shows homology with mammalian 76aa sequence. The alignment is with the central region of the 76aa. By contrast, no homology was found with either of the *Drosophila melanogaster* ankyrins (*Dank1* and *Dank2*).

Further analysis of the 76aa region using the programs REPRO and RADAR revealed two internal sequence repetitions which we termed modules I and II (fig. 2*A*). Both modules are

present within all three vertebrate gene products as confirmed by a multiple sequence alignment (fig. 2*B*). A less conserved third module (module III) was identified only in the C-terminal domains of *Ank1* and *Ank2* transcripts. The pattern of sequence conservation among modules is shown in fig. 2*C*. It is noteworthy that the sequence conserved in *C. elegans* corresponds to module II. All modules, except module III of *Ank2*, are made of 18 residues. The only known protein-protein interactions occurring within this region are between the *Ank1* encoded sAnk1 and the two muscle proteins titin and obscurin (Bagnato et al. 2003; Kontrogianni-Konstantopoulos and Bloch 2003; Kontrogianni-Konstantopoulos et al. 2003). The sites of interaction reside, at least partially, within modules I and II, respectively (fig. 2*A*).

Based on these observations, we propose that the conserved sequences corresponding to the ankyrin-G 76aa should be designated "**O**bscurin/**T**itin **B**inding-related **D**omain" (OTBD). Having noted the presence of a module II-like sequence in *C. elegans* UNC-44, we wondered if vertebrate and worm sequences have common origins. If this is the case, related sequences might be in simple chordates too. The genome of the urochrodate *Ciona intestinalis* is now available (Dehal et al. 2002). We searched the genome scaffolds initially by tBLASTn, using rat ankyrin-G OTBD as a query sequence. This yielded no significant hits. As a more sensitive alternative, we constructed a hidden Markov model of the aligned 18 amino acid modules from human, mouse, rat and *C. elegans*. This was used to search all possible translations of the *C. intestinalis* ankyrin gene. A single hit was obtained. This sequence is shown in fig. 2*B* aligned with other 18 amino acid modules. To establish if the *C. intestinalis* sequence is likely to be expressed as part of an ankyrin protein, we used the predicted amino acid sequence to query the EST database by tBLASTn. This revealed 16 ESTs with 100% identity, thus this single module represents part of an expressed protein. The ESTs are part of Unigene cluster Cin.5740: assembly of this cluster using CAP3 reveals a sequence encoding part of a conventional ankyrin with ZU5 and death

domains, as well as the module. The HMM detects the 18 amino acid module in this peptide with an E value of 0.005.

Since the HMM for the 18 amino acid module was a sensitive means for detecting related sequences in the urochordate, we wondered if our initial Blast searches had failed to detect such a module in insect ankyrins. Re-analysis of the *D. melanogaster* ankyrins revealed no significant hits. We also used this HMM to search the ENSEMBL databases of *Anopheles gambiae* and *Apis mellifera* peptides: again, no significant hits were found. We conclude that the 18 amino acid module is represented in both ecdysozoan and deuterostomal organisms, but that it is not preserved in insect ankyrins.

We further investigated the possibility that the OTBD represents a known protein structure. We used the 18 amino acid HMM to probe sequences represented in the Protein Databank (PDB). This search revealed no significant hits. Direct BLASTP analysis of sequences in the PDB using the OTBD sequence of human ankyrin-R gave no hits with expectation values better than 0.3, so we cannot assign a fold to the OTBD with confidence. We also submitted the OTBD sequence of human ankyrin-R to the Genesilico metaserver for structure prediction. The algorithms INGBU, 3DPSSM, MGENTHREADER, FFAS failed to give a confident or consistent prediction of fold. However, the secondary structure prediction methods PSIPRED, JNET, SABLE, PROF, JUFO and PROFSEC indicated possible secondary structure elements in the OTBD sequence. The output of the JUFO program is typical of the predictions obtained; its α-helix and β-strand predictions are annotated on fig. 2*A*. To probe the sequences for disordered structure, the programs DISEMBL, DISOPRED, GLOBPLOT and PONDR were used. All these programs predicted elements of disordered structure between the modules. To illustrate this, fig. 2*A* also shows the PONDR predictions for disordered sequence. Note that the second half of each 18 amino acid module is predicted to be α-helical and the modules are connected by disordered

polypeptide. Kyte-Doolittle (Kyte and Doolittle 1982) hydropathy analysis indicates that the sequence is relatively hydrophilic, and JNET (Cuff and Barton 2000) solvation analysis indicates that most of the OTBD is likely to be accessible to water. It seems possible that the OTBD is flexible, with structured regions contributed by 18 amino acid modules joined by flexible linkers.

2. Exons encoding the OTBD: expression and organization

To investigate the conservation of the OTBD among species we extended the multiple protein sequence alignment to all so far available ankyrin data, including genomic sequences (fig. 3). Protein sequences shown in fig. 2*A* were annotated to genomes to localize OTBD-encoding exons. Nucleotide sequences were subsequently extracted, translated, and aligned (fig. 3*A*).

In mammals, the OTBD is encoded by three consecutive exons corresponding to rat *Ank3* exons 43 to 45 (Hopitzan et al. 2005), and human *ANK1* exons 39a to 41 (Gallagher et al. 1997). Exon 41 has three internal splice sites giving rise to four segments. In muscle-specific sAnk1 isoforms, alternative splicing joins segments 1 and 4 (Birkenmeier et al. 1993; Birkenmeier et al. 1998). It is of interest that segment 4 encodes module III. We found that *ANK2* exons 44 and 45, as defined by Mohler and colleagues (Mohler et al. 2003) corresponded to the last two OTBD exons. The first exon was here identified by annotating the partial protein sequence shown in fig. 2*A* to the *ANK2* gene locus.

Several lines of evidence point to a muscle-restricted expression of the first OTBD exon. (1) Expression of exon 39a is restricted to muscle under the control of an alternative promoter (Birkenmeier et al. 1998; Gallagher and Forget 1998). (2) All three OTBD exons of *Ank3* have been found so far only in muscle tissue (Gagelin et al. 2002; Hopitzan et al. 2005). (3) The first OTBD exon encodes the entire module I, which contains half of the region that binds muscle protein titin (figs. 2*A*, 3*A*). It will be of importance to determine the critical residues implicated in

titin binding as has been done for obscurin. (4) BLAST analysis of the EST databases reveals no expression outside muscle-rich tissues. (5) Although multiple full-length cDNAs have been obtained for ankyrin-B splice variants, none have yet been obtained from muscle; correspondingly none contain the first exon of the OTBD.

Based on these data, we propose that module I bestows ankyrins with important muscle specific-functions, such as providing binding sites for muscle proteins. The last two OTBD exons are not restricted to muscle tissues; they have been also found in transcripts from reticulocytes, liver, bone marrow, and brain (Otto et al. 1991; Gallagher et al. 1997). It is noteworthy that both exons are needed to constitute module II. Crucial residues for binding obscurin are localized within both last OTBD exons; this requirement could apply to other yet unidentified protein interactions and could be the reason why these two exons have always been found co-expressed.

Our results strongly suggest the presence of the three OTBD exons in vertebrates in general, including birds, amphibian, and fishes (fig. 3*A*).

To further study the conservation of the OTBD among vertebrates, we aligned nucleotide sequences corresponding to the most conserved middle exon (fig. 3*B*). A striking observation was that both exon structure and nucleotide sequence are conserved among all paralogous ankyrin genes from fish to man.

3. Phylogentic relationships among ankyrin genes

To understand more of the origin and evolution of the OTBD we conducted phylogenetic analyses on available ankyrin sequences. A defining characteristic of ankyrins is the ZU5 domain. This domain binds spectrin and is highly conserved in all known ankyrins. A ML tree was constructed from 48 nucleotide sequences encoding ankyrin ZU5 domains, retrieved from

genomic databases. The tree defines two major clusters corresponding to vertebrate and invertebrate ankyrins.

In previous analysis of ankyrins, Bouley et al. (2000) examined the phylogeny of ankyrins by comparing the sequences of ANK-repeats of a small number of ankyrins. They concluded that *C. elegans* and *D. melanogaster* ankyrins do not represent direct orthologs of any vertebrate ankyrins, and that duplication events that led to fruitfly ankyrins 1 and 2 occurred independently of the expansion of the vertebrate genes. In these respects, their data are consistent with the proposed divergence of bilateran metazoa into ecdysozoa and deuterostomes (Aguinaldo et al. 1997; Halanych 2004).

Figure 4 shows the results of our more extensive analysis of the nucleotide sequences of the spectrin-binding ZU5 domain. We have taken advantage of recent genome and EST sequencing to analyze a much wider range of organisms that was available to Bouley et al. (2000). Like Bouley et al. (2000), our data support independent gene expansion events in the arthropod and vertebrate lineages. The sequences of all insects available have two ankyrin genes, compared to single genes in available nematode sequences.

Vertebrate sequences all fall into three categories that we name *Ank1*, *2*, and *3* in accordance with mammalian gene nomenclature. Interestingly, our data reveal six ankyrins in teleost fish, with two in each of the three ankyrin groups. This is consistent with the recently reported whole genome duplication event in the ray-finned fish lineage (Van de Peer 2004; Volff 2005). An interesting question is whether the three paralog ankyrin gene pairs are functional or whether one copy of each has evolved as a pseudogene. To the extent that the Ensembl annotations reveal that the key functional domains are preserved in each predicted protein, any degeneration is limited.

The single ankyrin gene in the urochordate *C. intestinalis* appears to be the ortholog of all

vertebrate ankyrins. This echoes the pattern in most other vertebrate superfamilies, which have a single *C.intestinalis* ortholog (Leveugle et al. 2004).

4. Origin and Function of the OTBD; relation to titin and obscurin

Based on all the data above, we propose that the full OBTD represents a vertebrate adaptation. We looked for further arguments in favor of this hypothesis by analyzing the OTBD-binding sequences of titin and obscurin. The first two N-terminal Ig domains of human titin, which bind the sAnk1 OTBD, were used in a PSI-BLAST search. Retrieved sequences revealed very high conservation among vertebrates (94%, 93% and 82% identity when compared to *M. musculus*, *G. gallus*, and *T. nigroviridis*, respectively). Homologous invertebrate sequences were much less conserved: *C. elegans* titin showed 40% identity; the most related insect titin sequence (an *A. mellifera* kettin isoform; Lakey et al. 1993) showed only 36% of identity.

Regarding obscurin, we performed a PSI-BLAST search using the C-terminal 400 residues of the human sequence. Unlike titin, overall conservation of this sequence among vertebrates is rather low: 64%, 38% and 30% identity when compared to *R. norvegicus*, *T. nigroviridis*, and *G. gallus*, respectively. However, highly conserved, short sequence stretches observed within this region could be important for ankyrin binding. No significant hit with an invertebrate sequence was found, and no obscurin is annotated in either Wormbase or Flybase.

These results strongly suggest that protein-protein interactions between the cytoskeleton and muscle proteins titin and obscurin have evolved differently in vertebrates and invertebrates. We propose that vertebrate ankyrins gained specific functions in muscle by the acquisition of the OTBD, whereas any related functions in invertebrates are mediated by different protein interactions. This hypothesis is supported by the diverse molecular architecture of invertebrate

sarcomeres and by the strong relationship between titin isoforms and sarcomere properties (Tskhovrebova and Trinick 2003).

Conclusions

Here we characterize a novel sequence from the C-terminal domain of ankyrins which is unique to vertebrate ankyrin genes. We describe this sequence as an Obscurin/Titin-Binding-related Domain (OTBD) since it is established that these two giant muscle proteins bind this sequence in a product of the mammalian *Ank1* gene.

The domain is encoded by three exons and is composed of up to three modules that potentially contain α-helical elements; the modules are likely to be joined to each other by flexible linkers. Amino acid and nucleotide sequences as well as splice sites corresponding to modules I and II are highly conserved from fish to man. Tissue expression information derived from sequence analysis of cloned isoforms and ESTs suggests that the first OTBD exon, which encodes the entire module I, is expressed exclusively in muscle tissues.

The presence of conserved sequence corresponding to module II in *C. elegans* and *C. intestinalis* suggests that the OTBD has evolved by successive duplications of an ancestor sequence that first arose before the divergence of ecdysozoa and deuterostomes.

Insect ankyrin genes duplicated independently and have lost module II. Comparison of titin and obscurin sequences implicated in binding to *Ank1*-encoded OTBD show conservation only among vertebrates. These data support the hypothesis that the OTBD confers specialized function(s) to vertebrate ankyrins.

One of the features of vertebrate life is the increased size that vertebrate organisms have compared to their precursors. One view of ankyrins is that they are a solution to the problems of independent motility in metazoa (Bennett and Baines 2001): for example, they contribute to membrane resilience to the forces of muscle contraction. By extension, the OTBD, by linking two giant cytoskeletal proteins to membrane structures, would consolidate the linkage of sarcomeres to the sarcoplasmic reticulum, and eventually to the sarcolemma. Identification of further protein

interactions of the OTBDs encoded by the other two ankyrin genes will greatly advance our understanding of the complex functions of the ankyrin family in vertebrate tissues.

Supplementary Material

- Table indicating the sources of nucleotide sequences used in this study (Hopitzan05-0255_ Table.pdf)
- Multiple nucleotide sequence alignment corresponding to ZU5 domains from which the phylogenetic tree (Figure 4) was built (Hopitzan05-0255_ ZU5.pdf and Hopitzan05-0255_ ZU5.fasta).
- Multiple nucleotide sequence alignment of conserved exonic sequences homologous to the 76aa insert from which the multiple protein sequence alignment (Figure 3) was deduced (Hopitzan05-0255_ OTBD.pdf and Hopitzan05-0255_ OTBD.fasta).

Acknowledgements

This work was supported by the Centre National de la Recherche Scientifique (CNRS), the Universities Paris 6 and 7, and by grants from EC (HPRN-CT-2000–00096) and from the Association Française contre les Myopathies (AFM). AJB was a recipient of a Leverhulme Trust Research Fellowship; AAH was supported in part by a fellowship from the Association Française contre les Myopathies (AFM).

Literature Cited

Aguinaldo, A. M., J. M. Turbeville, L. S. Linford, M. C. Rivera, J. R. Garey, R. A. Raff, and J. A. Lake. 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. Nature **387**:489-493.

Altschul, S. F., W. Gish, W. Miller, E. W. Meyers, and D. J. Lipman. 1990. Basic Local Alignment Search Tool. **215**:403.

Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Nucleic Acids Res. **25**:3389-3402.

Bagnato, P., V. Barone, E. Giacomello, D. Rossi, and V. Sorrentino. 2003. Binding of an ankyrin-1 isoform to obscurin suggests a molecular link between the sarcoplasmic reticulum and myofibrils in striated muscles. J. Cell Biol. **160**:245-253.

Bennett, V., and A. J. Baines. 2001. Spectrin and Ankyrin-Based Pathways: Metazoan Inventions for Integrating Cells Into Tissues. Physiol. Rev. **81**:1353-1392.

Birkenmeier, C. S., J. J. Sharp, E. J. Gifford, S. A. Deveau, and J. E. Barker. 1998. An Alternative First Exon in the Distal End of the Erythroid Ankyrin Gene Leads to Production of a Small Isoform Containing an NH2-Terminal Membrane Anchor. Genomics **50**:79-88.

Birkenmeier, C. S., R. A. White, L. L. Peters, E. J. Hall, S. E. Lux, and J. E. Barker. 1993. Complex patterns of sequence variation and multiple 5' and 3' ends are found among transcripts of the erythroid ankyrin gene. J. Biol. Chem. **268**:9533-9540.

Bouley, M., M.-Z. Tian, K. Paisley, Y.-C. Shen, J. D. Malhotra, and M. Hortsch. 2000. The L1-Type Cell Adhesion Molecule Neuroglian Influences the Stability of Neural Ankyrin in the Drosophila Embryo But Not Its Axonal Localization. J. Neurosci. **20**:4515-4523.

Chen, L., B. Ong, and V. Bennett. 2001. LAD-1, the Caenorhabditis elegans L1CAM homologue, participates in embryonic and gonadal morphogenesis and is a substrate for fibroblast growth factor receptor pathway-dependent phosphotyrosine-based signaling. J. Cell Biol. **154**:841-856.

Crooks, G. E., G. Hon, J.-M. Chandonia, and S. E. Brenner. 2004. WebLogo: A Sequence Logo Generator. Genome Res. **14**:1188-1190.

Cuff, J. A., and G. J. Barton. 2000. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. Proteins **40**:502-511.

Davis, L. H., J. Q. Davis, and V. Bennett. 1992. Ankyrin regulation: an alternatively spliced segment of the regulatory domain functions as an intramolecular modulator. J. Biol. Chem. **267**:18966-18972.

Dehal, P., Y. Satou, R. K. Campbell, J. Chapman, B. Degnan, A. De Tomaso, B. Davidson, A. Di Gregorio, M. Gelpke, D. M. Goodstein, N. Harafuji, K. E. M. Hastings, I. Ho, K. Hotta, W. Huang, T. Kawashima, P. Lemaire, D. Martinez, I. A. Meinertzhagen, S. Necula, M. Nonaka, N. Putnam, S. Rash, H. Saiga, M. Satake, A. Terry, L. Yamada, H.-G. Wang, S. Awazu, K. Azumi, J. Boore, M. Branno, S. Chin-bow, R. DeSantis, S. Doyle, P. Francino, D. N. Keys, S. Haga, H. Hayashi, K. Hino, K. S. Imai, K. Inaba, S. Kano, K. Kobayashi, M. Kobayashi, B.-I. Lee, K. W. Makabe, C. Manohar, G. Matassi, M. Medina, Y. Mochizuki, S. Mount, T. Morishita, S. Miura, A. Nakayama, S. Nishizaka, H. Nomoto, F. Ohta, K. Oishi, I. Rigoutsos, M. Sano, A. Sasaki, Y. Sasakura, E. Shoguchi, T. Shin-i, A. Spagnuolo, D. Stainier, M. M. Suzuki, O. Tassy, N. Takatori, M. Tokuoka, K. Yagi, F. Yoshizaki, S. Wada, C. Zhang, P. D. Hyatt, F. Larimer, C. Detter, N. Doggett, T. Glavina, T. Hawkins, P. Richardson, S. Lucas, Y. Kohara, M. Levine, N. Satoh, and D. S. Rokhsar.

2002. The Draft Genome of Ciona intestinalis: Insights into Chordate and Vertebrate Origins. Science **298**:2157-2167.

Dubreuil, R. R., and J. Yu. 1994. Ankyrin and beta-Spectrin Accumulate Independently of alpha-Spectrin in Drosophila. PNAS **91**:10285-10289.

Eddy, S. R. 1998. Profile hidden Markov models. Bioinformatics **14**:755-763.

Gagelin, C., B. Constantin, C. Deprette, M.-A. Ludosky, M. Recouvreur, J. Cartaud, C. Cognard, G. Raymond, and E. Kordeli. 2002. Identification of AnkG107, a Muscle-specific Ankyrin-G Isoform. J. Biol. Chem. **277**:12978-12987.

Gallagher, P. G., and B. G. Forget. 1998. An Alternate Promoter Directs Expression of a Truncated, Muscle-specific Isoform of the Human Ankyrin 1 Gene. J. Biol. Chem. **273**:1339-1348.

Gallagher, P. G., W. T. Tse, A. L. Scarpa, S. E. Lux, and B. G. Forget. 1997. Structure and Organization of the Human Ankyrin-1 Gene. Basis for complexity of pre-mRNA processing. J. Biol. Chem. **272**:19220-19228.

Granzier, H., and S. Labeit. 2002. Cardiac titin: an adjustable multi-functional spring. J. Physiol. **541**:335-342.

Halanych, K. M. 2004. The new view of animal phylogeny. Annual Review of Ecology, Evolution, and Systematics **35**:229-256.

Hall, T. A. 1999. BioEdit:a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. **41**:95-98.

Hall, T. G., and V. Bennett. 1987. Regulatory domains of erythrocyte ankyrin. J. Biol. Chem. **262**:10537-10545.

Hooper, S. L., and J. B. Thuma. 2005. Invertebrate Muscles: Muscle Specific Genes and Proteins. Physiol. Rev. **85**:1001-1060.

Hopitzan, A. A., A. J. Baines, M.-A. Ludosky, M. Recouvreur, and E. Kordeli. 2005. Ankyrin-G in skeletal muscle: Tissue-specific alternative splicing contributes to the complexity of the sarcolemmal cytoskeleton. Exp. Cell Res. Epub ahead of print. http://dx.doi.org/10.1016/j.yexcr.2005.1004.1013

Jeffery, W. R., and B. J. Swalla. 1993. An ankryin-like protein in ascidian eggs and its role in the evolution of direct development. Zygote **1**:197-208.

Jobb, G., A. von Haeseler, and K. Strimmer. 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evolutionary Biology **4**:18.

Kizhatil, K., and V. Bennett. 2004. Lateral Membrane Biogenesis in Human Bronchial Epithelial Cells Requires 190-kDa Ankyrin-G. J. Biol. Chem. **279**:16706-16714.

Kontrogianni-Konstantopoulos, A., and R. J. Bloch. 2003. The Hydrophilic Domain of Small Ankyrin-1 Interacts with the Two N-terminal Immunoglobulin Domains of Titin. J. Biol. Chem. **278**:3985-3991.

Kontrogianni-Konstantopoulos, A., D. H. Catino, J. C. Strong, W. R. Randall, and R. J. Bloch. 2004. Obscurin regulates the organization of myosin into A bands. Am J Physiol Cell Physiol **287**:C209-217.

Kontrogianni-Konstantopoulos, A., E. M. Jones, D. B. van Rossum, and R. J. Bloch. 2003. Obscurin Is a Ligand for Small Ankyrin 1 in Skeletal Muscle. Mol. Biol. Cell **14**:1138-1148.

Kordeli, E., S. Lambert, and V. Bennett. 1995. AnkyrinG. A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier. J. Biol. Chem. **270**:2352-2359.

Kurowski, M. A., and J. M. Bujnicki. 2003. GeneSilico protein structure prediction meta-server. Nucleic Acids Research **31**:3305-3307.

Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J Mol Biol **157**:105-132.

Lakey, A., S. Labeit, M. Gautel, C. Ferguson, D. P. Barlow, K. Leonard, and B. Bullard. 1993. Kettin, a large modular protein in the Z-disc of insect muscles. EMBO J. **12**:2863-2871.

Lambert, S., H. Yu, J. T. Prchal, J. Lawler, P. Ruff, D. Speicher, M. C. Cheung, Y. W. Kan, and J. Palek. 1990. cDNA Sequence for Human Erythrocyte Ankyrin. PNAS **87**:1730-1734.

Leveugle, M., K. Prat, C. Popovici, D. Birnbaum, and F. Coulier. 2004. Phylogenetic Analysis of Ciona intestinalis Gene Superfamilies Supports the Hypothesis of Successive Gene Expansions. J Mol Evol **58**:168.

Linding, R., R. B. Russell, V. Neduva, and T. J. Gibson. 2003. GlobPlot: exploring protein sequences for globularity and disorder. Nucleic Acids Research **31**:3701-3708.

Lux, S., K. M. John, and V. Bennett. 1990. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. Nature **344**:36-42.

Mohler, P. J., A. O. Gramolini, and V. Bennett. 2002a. The Ankyrin-B C-terminal Domain Determines Activity of Ankyrin-B/G Chimeras in Rescue of Abnormal Inositol 1,4,5-Trisphosphate and Ryanodine Receptor Distribution in Ankyrin-B (-/-) Neonatal Cardiomyocytes. J. Biol. Chem. **277**:10599-10607.

Mohler, P. J., A. O. Gramolini, and V. Bennett. 2002b. Ankyrins. J Cell Sci **115**:1565-1566.

Mohler, P. J., J.-J. Schott, A. O. Gramolini, K. W. Dilly, S. Guatimosim, W. H. duBell, L.-S. Song, K. Haurogne, F. Kyndt, M. E. Ali, T. B. Rogers, W. J. Lederer, D. Escande, H. L. Marec, and V. Bennett. 2003. Ankyrin-B mutation causes type 4 long-QT cardiac arrhythmia and sudden cardiac death. Nature **421**:634-639.

Mohler, P. J., I. Splawski, C. Napolitano, G. Bottelli, L. Sharpe, K. Timothy, S. G. Priori, M. T. Keating, and V. Bennett. 2004. A cardiac arrhythmia syndrome caused by loss of ankyrin-B function. PNAS **101**:9137-9142.

Otsuka, A. J., R. Franco, B. Yang, K. H. Shim, L. Z. Tang, Y. Y. Zhang, P. Boontrakulpoontawee, A. Jeyaprakash, E. Hedgecock, and V. I. Wheaton. 1995. An ankyrin-related gene (unc-44) is necessary for proper axonal guidance in Caenorhabditis elegans. J. Cell Biol. **129**:1081-1092.

Otto, E., M. Kunimoto, T. McLaughlin, and V. Bennett. 1991. Isolation and characterization of cDNAs encoding human brain ankyrins reveal a family of alternatively spliced genes. J. Cell Biol. **114**:241-253.

Peters, L. L., K. M. John, F. M. Lu, E. M. Eicher, A. Higgins, M. Yialamas, L. C. Turtzo, A. J. Otsuka, and S. E. Lux. 1995. Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene family and the major ankyrin in kidney, is expressed in alternatively spliced forms, including forms that lack the repeat domain. J. Cell Biol. **130**:313-330.

Russell, M. W., M. O. Raeker, K. A. Korytkowski, and K. J. Sonneman. 2002. Identification, tissue expression and chromosomal localization of human Obscurin-MLCK, a member of the titin and Dbl families of myosin light chain kinases. Gene **282**:237-246.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research **25**:4876.

Tskhovrebova, L., and J. Trinick. 2003. Titin: properties and family relationships. Nat Rev Mol Cell Biol **4**:679-689.

Tskhovrebova, L., and J. Trinick. 2004. Properties of Titin Immunoglobulin and Fibronectin-3 Domains. J. Biol. Chem. **279**:46351-46354.

Van de Peer, Y. 2004. Tetraodon genome confirms Takifugu findings: most fish are ancient polyploids. Genome Biology **5**:250.

Volff, J.-N. 2005. Genome evolution and biodiversity in teleost fish. Heredity **94**:280-294.

Young, P., E. Ehler, and M. Gautel. 2001. Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor protein involved in sarcomere assembly. J. Cell Biol. **154**:123-136.

Zhou, D., C. S. Birkenmeier, M. W. Williams, J. J. Sharp, J. E. Barker, and R. J. Bloch. 1997. Small, Membrane-bound, Alternatively Spliced Forms of Ankyrin 1 Associated with the Sarcoplasmic Reticulum of Mammalian Skeletal Muscle. J. Cell Biol. **136**:621-631.

Figure legends:

Figure 1. Schematic representation of Ank_{G107} containing the muscle-specific 76-residue insert (76aa). Ankyrin domains are designated as follows: SpBd, spectrin-binding; DD, death domain; C-ter, C-terminal. Amino acid sequence of 76aa and encoding *Ank3* exons are shown on top and below, respectively.

Figure 2. Multiple alignment of C-terminal ankyrin sequences corresponding to the rat ankyrin-G 76aa insert.

(A) Shown are the protein sequences encoded by *Ank1* (*H. sapiens* AnkR, NP_065211; *M. musculus* AnkR, AAC24156), *Ank2* (*H. sapiens* AnkB, CAD98033; *R. norvegicus* AnkB, XP_227735), *Ank3* (*R. norvegicus* AnkG, AJ428573), and *unc-44* (*C. elegans* UNC-44, AAB41827).

Amino acid residues of human *Ank1* gene product (AnkR, NP_065211) implicated in binding titin and obscurin are indicated on top of the alignment (double underlined and in bold). Underneath the sequences are structure predictions, generated as described in the text. Predicted α-helices are shown as cylinders; β-strand as an arrow; disorder as an open rectangle. Note that the β-strand predicted by JUFO is suggested by PONDR to be disordered, so both predictions are indicated in the figure.

(B) Identification of three internal sequence repetitions (modules I-III) within ankyrin sequences corresponding to the 76aa insert by optimized ClustalX alignment. Amino acid sequence of the *C. intestinalis* module was deduced from EST clone BW435037 (nucleotides 556 to 621). Modules are underlined (B) and indicated by open boxes in (A). Conserved amino acids in more than 80% (A) or 60% (B) of the sequences are highlighted in black (identical residues) or grey background (similar residues according to Blosum62 residue weight table).

(C) Graphical representation of the aligned module sequences shown in (B) as “sequence logo”.

Figure 3. Multiple alignment of OTBD amino acid and nucleotide sequences as deduced from vertebrate genomes.

(A) Compared sequences correspond to the most conserved modules I and II. Human *ANK1* OTBD exons 39a to 41 are indicated on top. Conserved amino acids in 80% or more of the sequences are highlighted in black (identical residues) or grey background (similar residues according to Blosum62 residue weight table).

(B) Multiple alignment of nucleotide sequences corresponding to human *Ank1* exon 40. Coding sequence is in uppercase, splice acceptor and donor sites (ag/gt) are in lower case. Exons are represented by boxes on top of the alignment; shown below is the protein sequence of human sAnk1 (NP_065211). Residues implicated in binding of titin and obscurin are in bold and double underlined. Introns are not drawn to scale. Identical nucleotides in 80% or more of the sequences are highlighted in black background.

Figure 4. Phylogenetic tree of ankyrins.

The maximum likelihood tree shows the relation between ankyrins based on sequences corresponding to the ZU5 domain; *C. elegans* and *C. briggsae* served as outgroups. Bootstrap values calculated from 1000 replicates are indicated.

Figure 5. Proposed model of evolutionary events leading to OTBD in present-day ankyrins.

The figure shows a simplified summary of our data. An early metazoan acquired a sequence module (shown in black at the root of the tree), which was preserved after the ecdysozoa/deuterostome split. Nematodes retain this in present-day UNC-44 ankyrin, but it has

been lost from arthropod ankyrins. DNA closely related to this module is present in the ankyrin gene of the urochordate *C. intestinalis*. In vertebrates, the module was preserved in the coding sequence of ankyrin genes, underwent successive duplications to form modules I, II and III, and thus gained function as an obscurin/titin-binding domain. The ankyrin gene itself also underwent successive duplications giving rise to *Ank1, 2,* and *3*; module III has been lost from *Ank3*. The ray-finned fish lineage went through a round of tetraploidization/rediploidization, the legacy of which is that each of *Ank1, 2,* and *3* is duplicated in present-day teleosts, and retains modules of the OTBD.

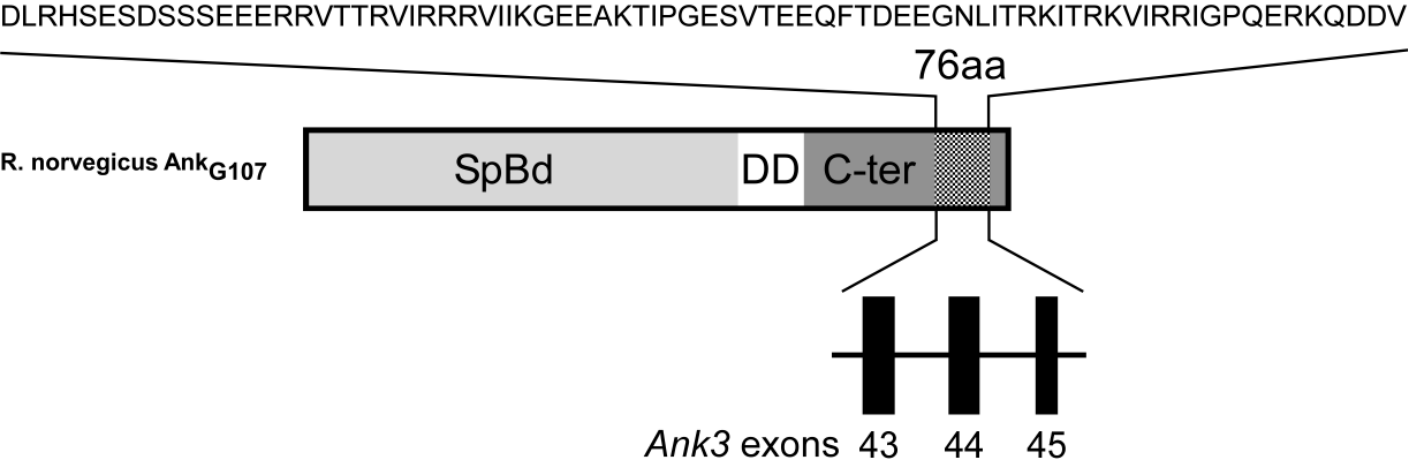

Figure 1

Hopitzan et al.

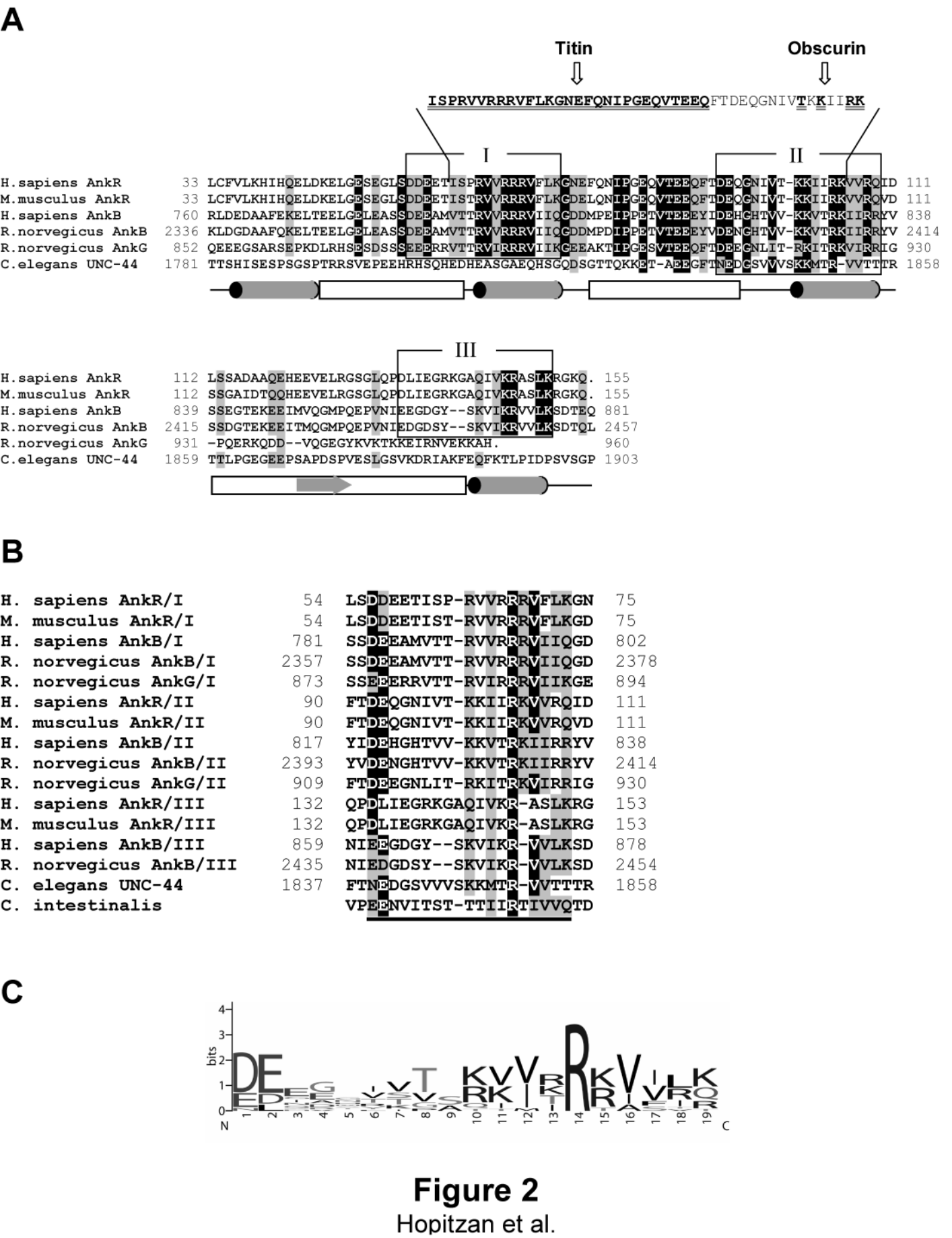

Figure 2

Hopitzan et al.

```
                                exon 39a                                  exon 40                                   exon 41
H. sapiens Ank1        ..ESEGLSDDEETISTRVVRRRVFLK GNEFQNIPGEQVTEEQFTDEQGNIVTKK IIRKVVRQIDLSSA..

P. troglody Ank1       ..ESEGLSDDEETISTRVVRRRVFLK GNEFQNIPGEQVTEEQFTDEQGNIVTKK IIRKVVRQIDLSSA..
M. musculus Ank1       ..ESEGLSDDEETISTRVVRRRVFLK GDELQNIPGEQVTEEQFTDEQGNIVTKK IIRKVVRQVDSSGA..
R. norvegicus Ank1     ..ESGGLSDDEETISTRVVRRRVFLK GDELQNIPGEQVTEEQFTDEQGNIVTKK IIRKVVRQVDSSGA..
C. familiaris Ank1     ..ESEGVSDDEETISTRVVRRRVFLK GNEFQNIPGEQVTEEEFTDEQGNIVTKK IIRKVVRQIDSSGA..
G. gallus Ank1         ..ESQGLADDEESLSTRVVRRRVLVK GNEVLHLPGEHVTEEQFTDEQGNIITKK VIRKVVHQLDL-DT..
X. tropicalis Ank1     ..EGGG-EDDEETLTTKVVRRRVIVK GNEVPDIPGEQVTEEEFTDEQGNIVTKK IVRKVLRRVGPPGT..
D. rerio Ank1a         ..ENEGVADEEENVTTRVVRRRVILK GDEAEDLPGEQVSEEQFTDEHGNIVTKK IVRKVVRRGK--GE..
T. nigroviridis Ank1b  ..EVEGVSDEEENVTTRVVRRRVILK GEEVEDLPGEHVSEEQFMDEHGNIVTKK IVRKVVRRGKGAGE..
F. rubripes Ank1x      ..EVEGVSDEEENVTTRVVRRRVILK GDEVGDLPGEHVSEEQFMDEHGNIVTKK IVRKVVRRGKGSGE..
H. sapiens Ank2        ..ELEASSDEEAMVTTRVVRRRVIIQ GDDMPEIPPETVTEEYIDEHGHTVVKK VTRKIIRRYVSSEG..
P. troglody Ank2       ..ELEASSDEEAMVTTRVVRRRVIIQ GDDMPEIPPETVTEEYIDEHGHTVVKK VTRKIIRRYVSSEG..
M. musculus Ank2       ..ELEASSDEEAMVTTRVVRRRVIIQ GDDMPDIPPETVTEEYVDENGHTVVKK VTRKIIRRYVSSDG..
R. norvegicus Ank2     ..ELEASSDEEAMVTTRVVRRRVIIQ GDDMPDIPPETVTEEYVDENGHTVVKK VTRKIIRRYVSSDG..
C. familiaris Ank2     ..ELEASSDEEAMVTTRVVRRRVIIQ GDDMPDIPPETVTEEYIDEHGHTVVKK VTRKIIRRYVSSDG..
B. Taurus Ank2         ..ELEASSDEEAMEATRVVRRRLIIQ GDDMPDIPPETVTEEYIDEHGHTVVKK VTRKIIRRFISSDG..
G. gallus Ank2         ..ELENSSDEDEMVTTRVVRRRVIIQ ADSMPEMPPETVTEEQYTDEHGHTVVKK VTRKIIRRYVSPDG..
X. tropicalis Ank2     ..ELETTSDEDEMVTTRVVRRRVIIQ ADDMPDILPETVTEEQYTDEHGHIVVKK VTRKIIRRYVLPDG..
D. rerio Ank2a         ..VLS-DSSDEDVLTTRIVRRRVIIQ GDDVPDISPQSVTEEQYTDAQGNLVVKK VTRKVIRRCVSSDG..
D. rerio Ank2b         ..LLS-DSSEEEVLTTRVVRRRVVIQ GDDMPEIPPQTVTEEQYTDEYGNMVVKK ITRKVIRKYVSADG..
T. nigroviridis Ank2a  ..SLSSDSSEEEVLTTRVVRRRVIIQ ADDLPELPPQSVTEEAYMDQHGNMVVKK VTRKVIKKYVSTDG..
T. nigroviridis Ank2b  ..SMS-DSSDDDFFTSRIVRRRFVIQ ADDFPEFATQSVTEEKYRDENGHIVVKK ITRKIIRKCVSAEA..
F. rubripes Ank2a      ..ALSSDSSEEEVLTTRVVRRRVIIQ ADNLPDLPPQTVTEEKYTDEHGNMVVKK ITRKVIHKYVSADG..
H. sapiens Ank3        ..ESDSSSEEERRVTTRVIRRRLIIK GEEAKNIPGESVTEEQFTDEEGNLITRK ITRKVLRRIVIPQE..
P. troglody Ank3       ..ESDSSSEEERRVTTRVIRRRLIIK GEEAKNIPGESVTEEQFTDEEGNLITRK ITRKVLRRIVIPQE..
M. musculus Ank3       ..ESDSSSEEEQRITTRVIRRRVIIK GEEAKNIPGESVTEEQFTDEEGNLITRK ITRKVIRRIG-PQE..
R. norvegicus Ank3     ..ESDSSSEEERRVTTRVIRRRVIIK GEEAKTIPGESVTEEQFTDEEGNLITRK ITRKVIRRIG-PQE..
C. familiaris Ank3     ..ESDSSSEEERRVTTRVIRRRLIIK GEEAKSIPGESVTEEQFTDEEGNLITRK ITRKVIRRIAIPQE..
G. gallus Ank3         ..ESDSSSGEEQRITTRVYRRRLILK GEEAKNIPGESVTEEQFTDEEGNVITRK ITRKVVRRVVIPHE..
X. tropicalis Ank3     ..ESVSSSDEAERITTKVTRRRLILK GEQAKNIPGESVTEEQYTDEEGNIITRK VTRKVLKRIVVPHE..
D. rerio Ank3a         ..H.HRGLDEERTVTTKVFRRRLILK GEEARNIPAESVSEEQFIDEDGNIITRK VIRKVICRVDSPTP..
T. nigroviridis Ank3a  ..SLDSGSEEEQTVTTRIFRRRLILK GEEARNLPGESVTEEHYLDRDGNLVSRK VIRKVIRRVSTPTP..
T. nigroviridis Ank3b  ..EPSSSEEEEPSVTTRVFRRRLVVK GEEARNLPRESVTEDQFIDEDGNLVTRK VIRKVVRRVVGTEE..
F. rubripes Ank3a      ..SVDSGSEEEQTVTTRIFRRRLILK GEEAKNIPGESMTEEHYMDHDGNLISRK VIRKVIRRVSTPTP..
                                   I                                                   II
```

Figure 3A

Hopitzan et al.

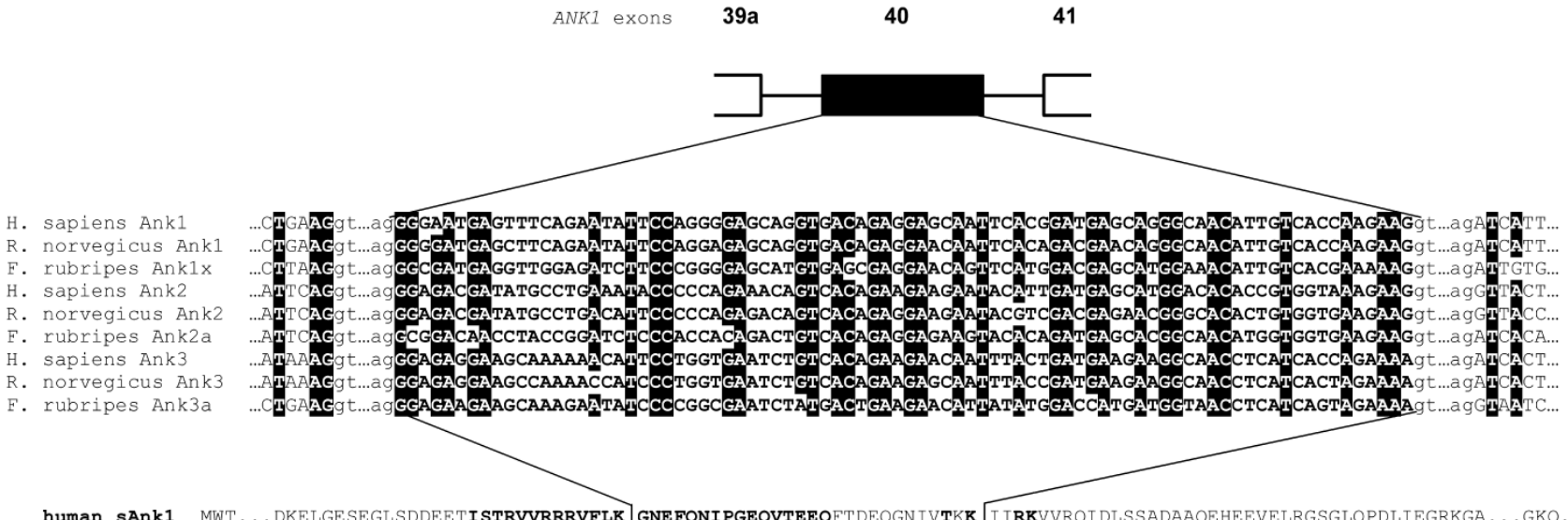

Figure 3B

Hopitzan et al.

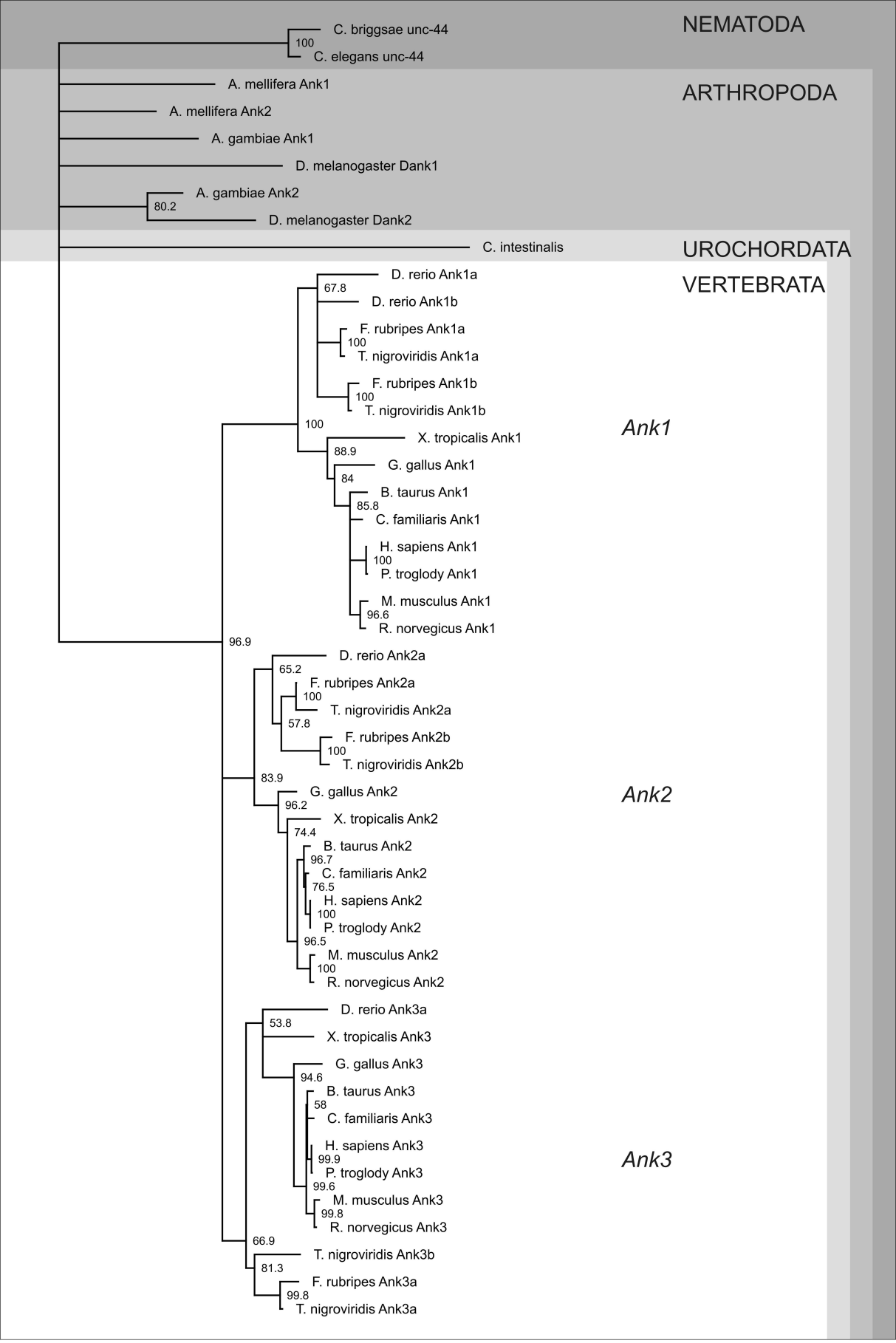

Figure 4

Hopitzan et al.

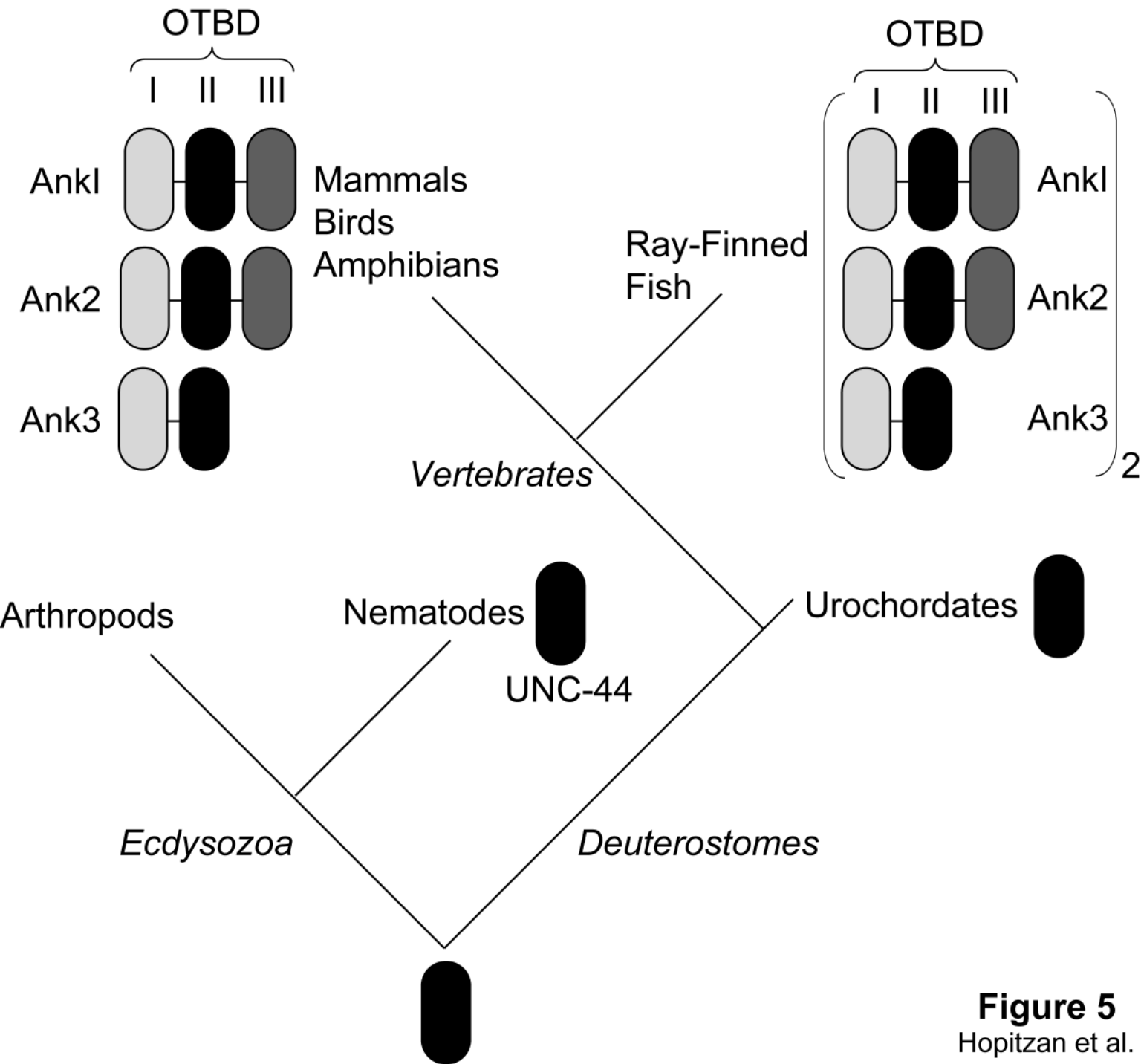

Figure 5
Hopitzan et al.