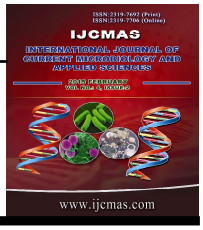
Int.J.Curr.Microbiol.App.Sci (2015) 4(2): 687-700

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 4 Number 2 (2015) pp. 687-700**
http://www.ijcmas.com



Original Research Article

Molecular cloning, characterization and phylogenetic analysis of an actin gene from the marine mollusk *Rapana venosa* (class *Gastropoda*)

Martin Ivanov[1]*, Elena Todorovska[1], Mariana Radkova[1], Oleg Georgiev[2], Aleksandar Dolashki[3] and Pavlina Dolashka[3]

[1]AgroBioInstitute, Agricultural Academy, 8 Dragan Tsankov blvd., 1164 Sofia, Bulgaria
[2]Institute of Molecular Life Sciences, University Zurich-Irchel, 190 Winterthurer Str., CH-8057 Zurich, Switzerland
[3]Institute of Organic Chemistry with Centre of Phytochemistry, BAS, 9 G. Bonchev str., 1113 Sofia, Bulgaria
**Corresponding author*

Keywords

Rapana venosa, Actin, Cloning, Sequence analysis, Phylogenetics

A B S T R A C T

A full length transcript encoding actin from the veined rapa whelk *Rapana venosa (RvAct1)* was isolated using RT-PCR and 5'/3' RACE. The cDNA spans 1565 bp and contains a long 3' untranslated region of 293 bp and 111 bp of 5' untranslated sequence. The open reading frame encodes a polypeptide of 376 amino acids. Sequence comparison showed that *RvAct1* is more closely related to non-muscle type actins (β or γ) rather than muscular (α) actins of human. Phylogenetic analysis revealed high homology of *RvAct1* with the cytoplasmic and non-muscle actins of species belonging to the same taxonomy group (phylum *Molluska*) and with evolutionary higher invertebrates (*Arthropoda*) and vertebrates. The study presented information for the relationships among species from different taxa and added proof to the conservative nature of the actin gene family. It provided new information for the prevalence of a specific amino acid (Cys) at position 272 in the cytoplasmic actins which could be used as a potential phylogenetic mark for discrimination of mollusks from the more evolutionary complex invertebrate and vertebrates. The presented actin sequence will extend the knowledge on the molecular evolution of class *Gastropoda* and will help further development of universal actin isoform classification in invertebrates.

Introduction

Actin is a member of a super family of ATPases that includes the ATP-binding domain of the Hsp70-related proteins, sugar kinases, actin-related proteins and prokaryotic actin homologs that despite the absence of amino acid (AA) sequence similarity, share a common fold (Hurley, 1996; Goodson and Hawse, 2002). The common core architecture consists of two large domains connected by a hinge and a nucleotide binding site located in the cleft between the domains.

As a part of the housekeeping gene families the actin is highly conserved in a wide variety of taxonomy groups from single cell organisms to human. In eukaryotes, the actin is a globular contractile protein with an approximate size of 42-kDa which is involved in the muscle contraction by forming muscle filaments. As a major cytoskeletal component actin plays a vital role in many different cell functions such as cell motility, movement of organelles, phagocytosis and signal transduction (Valentijn *et al.,* 1999; Rogers and Gelfand, 2000; Doussau and Augustine, 2000; Munn, 2001). Because of its constant expression in the cell, the actin is widely used in gene expression (Noonan *et al.,* 1990; Oleksiewicz *et al.,* 2001) and transgenic research (Cadoret *et al.,* 1999). The conserved nature of the actin genes among different species made them a potential candidate in the phylogenetic studies (Baldauf *et al.,* 2000).

In invertebrates at present there is no approved and unified classification of the actin isoforms mainly due to the small number of reported actin genes, number even smaller in the *Mollusca* taxon. Class *Gastropoda* represents the majority of the specimens of the phylum *Mollusca* with over 70,000 different organisms. The other two major branches includes the classes *Bivalvia* (clams) with over 20,000 and *Cephalopoda* (squid, octopus) with around 1000 of the most evolutionary evolved organisms from the phylum *Mollusca*. The representatives of phylum *Mollusca* are valuable sources of peptides possessing immunostimulatory functions (Lammers *et al.,* 2012) with large application in pharmaceutical industry and medicine for health disease prevention (Yamashita *et al.,* 2012; Flamini *et al.,* 2009; Toshkova *et al.,* 2009; Dolashka *et al.,* 2014). They are also sources of housekeeping genes such as those encoding actins, mostly with application in genomics and phylogenetic research.

Rapana venosa has been discovered by Valenciennes in 1846 and was classified as a member of *Muricidae* family, class *Gastropoda*, phylum *Mollusca*. It`s a native organism in the Sea of Japan, Yellow and East China Sea. Because of it's high tolerance to different ecological factors, *Rapana venosa* became a widespread marine organism and at present it is a dominant species of the benthic ecosystems in the Black Sea (Kos'yan, 2013).

The actins isolated from various representatives of the three major *Mollusca* classes have high amino acid homology and almost constant coding sequence length (Hongming *et al.,* 2007; Sin *et al.,* 2007; Adema, 2002;). The mollusc's actins showed homology with both muscle and cytoplasmic actin isoforms from vertebrates which makes their absolute classification impossible.

In contrast, six defined isoforms of actin have been identified in vertebrates based on their primary structure and isoelectric point. In terms of their tissue localization these isoforms are split into two major categories, muscle actin isoforms (α-skeletal, α-vascular, α-cardiac and γ-enteric) and non-muscle cell actin isoforms (β- and γ-cytoplasmic) (Vandekerckhove and Weber, 1979).

The insufficient gene information is one of the major problems that researchers face when dealing with molecular evolution in phylum *Mollusca*. Here we report the successful isolation and characterization of actin encoding transcript from marine organism *Rapana venosa* and it`s phylogenetic relationships with muscle and non-muscle actins in a wide number of taxonomic groups. This study extends the existing EST and protein databases and

contributes to the taxonomical classification of the different actin isoforms in invertebrates. The actin sequence could be used as a valuable tool in future transgenic research projects, phylogenetic analysis and gene expression studies related to proteins or peptides with antimicrobial, antiviral and antitumor activity.

Materials and Methods

Animal material

Veined rapa whelk (*Rapana venosa*) animals with average shell size of 12 cm were collected from the Black sea. Muscle tissue from two different animals was separately dissected and cut down to smaller pieces of 0.5cm each. The material was grinded down with liquid nitrogen and stored at -70 °C until RNA isolation.

Total RNA isolation

Total RNA was extracted according to TRIzol® Reagent protocol (Life Technologies: 15596-026) separately from each animal. The integrity and quality of the extracted total RNA was tested on 1.2% agarose gel. The residual traces of genomic DNA were eliminated by treatment of total RNA with DNase I, RNase free (Thermo scientific: #EN0521). The quantity and the quality of the purified RNA were determined by UV spectrophotometer and agarose gel electrophoresis.

Degenerate primers, first strand DNA synthesis and PCR amplification

For successful amplification of *Rapana venosa* actin cDNA, two degenerate primers 488up (5`-CGGAGATGGTGTCACCCACAC-3`) and 842dw (5`-TGTATGTGGTCTCGTGGATACC-3`) (Hongming *et al.*, 2007) were used to amplify the internal actin specific fragment. Both primers define a highly conservative region with a total length of 374 bp. First strand actin cDNA was synthesized using RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific #K1631) and amplified with the degenerate primers under the following PCR conditions: 94°C for 1 min and 30 cycles of 30 s at 94°C, 30 s at 60ºC, and 45 s at 72°C, followed by a 5 min final extension at 72°C. All incubations and PCR reactions were performed on a Verity TM 9902 96-Well Termal cycler (Applied Biosystems).

5`RACE

First strand cDNA was synthesized using Roche 5`/3` RACE, 2nd Generation kit (Roche, # 03353621001). The synthesis was perform with 1,8 µg total RNA, 842dw as a specific primer (SP1) and Reverse Transcriptase. cDNA homopolymeric A-tail addition was done according to the manufacturer instructions, followed by purification with Illustra GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare, # 28903470). The 5`end of the actin cDNA was amplified using primers 842dw and Oligo d(T)-anchor, 2 µl of cDNA, 1x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs and 2.5U High Fidelity Taq polymerase (Roche Applied Science #1732650001) in a total reaction volume of 50 µl. Cycling parameters were as follows: 94°C for 2 min and 35 cycles of 30 s at 94°C, 30 s at 62ºC, and 1.30 min at 72°C with a final extension at 72°C for 8 min.

3`RACE

The 3’ end of actin coding DNA sequence in *R. venosa* was obtained using 5’/3’ RACE Kit, 2nd Generation (Roche). 1,8 µg total RNA was reverse transcribed according to

the manufacturer instructions and the obtained cDNA was amplified with PCR anchor primer and the degenerate sense primer 488up. PCR reaction was carried out with 2 µl cDNA product, 1.5 mM $MgCl_2$, 10 pM primer 488up, 0.2 mM dNTPs and 2.5U High Fidelity Taq polymerase in a total reaction volume of 50 µl. Cycling parameters were as follows: 94°C for 2 min and 32 cycles of 30 s at 94°C, 30 s at 60ºC, and 2 min at 72°C with a final 8 min extension at 72°C.

Cloning and sequencing

PCR products were assessed on 1% agarose gel. Clear bands were cut out and purified with Illustra GFX™ PCR DNA and Gel Band Purification Kit, and then ligated to a plasmid vector pTZ57R/T InsTAclone PCR Cloning Kit (Thermo Scientific - #K1214). The ligation product was then used to transform *E. coli DH5 α* cells. Plasmid DNAs of 7-8 positive colonies/fragment were subjected to sequencing on a capillary sequencer.

Bioinformatic analysis

BLAST search programs (*Blastn* and *Blastp)* (Altschul *et al.,* 1990) were used to compare the nucleotide and deduced amino acid sequences of *Rapana venosa* actin cDNA with the available actin sequences in the NCBI database (www.ncbi.nlm.nih.gov). For phylogenetic analyses, additional actin sequences from *Molluska* and several representatives of other taxonomic groups with homology of above 85% with the *Rapana venosa* transcript were retrieved from the GenBank (www.ncbi.nlm.nih. gov/genbank). Multiple alignment was conducted by the neighbor-joining (NJ) method (Saitou and Nei, 1987) with Clustal W software (Thompson *et al.,* 1994). The phylogenetic tree was constructed with the chosen aligned sequences. The degree of support for internal branches of the tree was assessed in 1 000 bootstrap replicates.

Result and Discussion

Structural characteristics of *Rapana venosa* actin cDNA

The full-length cDNA of the *Rapana venosa* actin gene, named *RvAct1* (AccNo KF410817), was isolated according to the strategy described in Fig. 1. The cDNA fragment, corresponding to the internal part of the actin gene (Fragment 1) was amplified by RT-PCR with primers 488up and 842dw. Using 3'RACE and 5'RACE, the 3' (Fragment 2) and 5' (Fragment 3) ends of the cDNA were amplified, cloned and subsequently sequenced. The sequences corresponding to the three fragments were further assembled using software Vector NTI (Invitrogen Advance 11.5). A contig covering the full length of the actin cDNA was obtained based on the correct overlapping of the fragments (Fig. 1).

The resulting transcript with a total length of 1565 bp includes 111 bp long 5'UTR, an open reading frame of 1131 bp and 3' untranslated region of 323 bp together with a 30 bp poly A tail. A polyadenylation signal (AATAAA) was identified 21 bp upstream of the polyA tail. The open reading frame encodes a protein of 376 amino acids (**Fig. 2**), a feature that is specific for majority of the mollusk actins (Hongming *et al.,* 2007; Des Groseillers *et al.,* 1994).

Three synonymous nucleotide substitutions (SNPs) were detected between the cDNA of the two animals used in this study (data not shown). One of the SNPs is located in the coding region (position 771 bp - T > C) while the other two (T > C) and (A > T) were found in the 3'UTR region. The substitution in the coding region was located

in the third position of the codon for the amino acid arginine without changing the identity of this AA. This difference could be attributed to the intra population variability of this species.

Rapana venosa actin gene classification

Due to the absence of a correct differentiation of actin isoforms in invertebrate organisms, the classification of each newly isolated actin transcript is based on those available in the higher vertebrates such as *Homo sapiens* in which six isoforms of actin have been identified based on their primary structure, isoelectric point and tissue localization: muscle and non-muscle isoforms (Vandekerckhove and Weber, 1979). The main differences between vertebrate cytoplasmic actins and their muscle counterparts are 20 specific AA residues which are accepted as diagnostic in classification assays (Kusakabe *et al.*, 1997).

Multiple alignment between the *R. venosa* actin amino acid sequence (*RvAct1*) and the six available human actin isoforms: skeletal α-actin (Gunning *et al.*, 1984; *Taylor et al.*, 1988), cardiac α-actin (Gunning *et al.*, 1983; Hamada *et al.*, 1982; Minty, 1986), smooth muscle α-actin (Hamada *et al.*, 1989; Nakano, 1991; Ueyama *et al.*, 1984; 1990), cytoplasmic β-actin (Nakajima-Iijima *et al.*, 1985; Ng *et al.*, 1985; Ponte *et al.*, 1984), cytoplasmic γ-actin (Erba *et al.*, 1986; 1988), smooth muscle γ-actin (Miwa *et al.*, 1991) showed a total of 15 matches: 6 Ala, 10 Val, 16 Met, 17 Cys, 76 Val, 103 Val, 162 Thr, 176 Leu, 201 Thr, 225 Gln, 260 Ala, 267 Leu, 287 Val, 297 Thr, 365 Ser with the amino acid residuals diagnostic for vertebrate cytoplasmic actins. Three positions: 153 Met, 272 Cys, 279 Phe, were identical with the muscular (α) human actins while position 129 Ala shows high variability within the mollusk phylum and some lower vertebrate organisms and differ from both muscle (Val) and cytoplasmic human actins (Thr) (**Fig. 3**). The diagnostic AA for human cytoplasmic actins at position 5 (Ile) was also different (Val) in almost all mollusk's organisms. Based on the predominant AA matches with the human cytoplasmic actin isoform we safely classify the *RvAct1* as a cytoplasmic actin. Several studies (Fisher and Bode, 1989; Lardans *et al.*, 1997; Cadoret *et al.*, 1999; Adema, 2002) have already reported that invertebrate's actins are very similar to vertebrate cytoplasmic (β)-actins even when they have been isolated from muscle tissue (Patwary *et al.*, 1996; Miyamoto *et al.*, 2002).

The isolated *RvAct1* belongs to the sugar kinase/HSP70/actin superfamily and possess 2 domains (small and large) as have been described for the actin binding proteins (Domínguez and Holmes, 2011). A four-subdomain nomenclature has been adopted (Kabsch *et al.*, 1990). The small domain consists of two sub-domains. The sub-domain 1 is located between residues 1-32, 70-144 and 338-375 while the sub-domain 2 includes residues 33 to 69. The large domain is also divided into two sub-domains: 3(residues 145-180) and 4 (residues 181-269) (Kabsch and Vandekerckhove, 1992). The *R. venosa* actin possesses conserved residues in positions that affect the binding to ATP (residues Lys 214, Glu 215, Gly 303, Tyr 307 and Lys 337) and Ca^{+2} (residues Asp 12, Lys 19, Gln 138 and Asp 155). The DNase I binding region (residues Arg 40 - Gly 47, Lys 62 - Ile 65, Thr 204 and Glu 208) is also found. Furthermore, all residues required to form actin filaments (322–325, 286–289, 166–169, 375 and 110–112) are present in the actin of *R. venosa* (data not shown) (Kabsch and Vandekerckhove, 1992). Two nuclear export signals (NES) (NES-1, residues 170–181;

NES-2, residues 211–222) which are leucine-rich domains necessary for the transfer between nucleus and cytoplasm are also found. It is prominent that these sites and several flanking residues are conserved in the predicted amino acid sequence of RvAct1, suggesting that this protein is cytoplasmic. The nucleotide-binding site residues (residues 12-15; 17, 19; 138 and 155-158) which form the nucleotide-binding domain (NBD) are also found in the actin of *R. venosa*. The nucleotide sits in a deep cleft formed between the two lobes of the NBD. Substrate binding to sugar kinase/HSP70/actin superfamily members is associated with closure of this catalytic site upper cleft between sub-domains 2 and 4. The functional activities of several members of the superfamily, including hexokinases, actin, and HSP70s, are modulated by allosteric effectors, which may act on the cleft closure in the eukaryotic cell.

A common N-terminal sequence pattern in invertebrate actins: 1 Met, 2 Cys, followed by a stretch of three acidic amino acids of either Glu or Asp specific for vertebrate non-muscle actins (Hongming *et al.,* 2007) was also validated for RvAct1 (Fig. 4). However, our study showed that some non-muscle and cytoplasmic actins from invertebrates do not fully match the above mentioned pattern and have identical N-terminal sequence to those of the cytoplasmic actins of *Gallus gallus* and human which do not have a 2 Cys AA. Several non-muscle invertebrate actins have different AA residues at position 2 (Gly, Trp, Ala, Arg, Ser) instead of Cys. Therefore, 2 Cys is not a common mark for non-muscle and cytoplasmic actins of invertebrates, especially for mollusks. Additional comparison by multiple alignment of the published non-muscle and cytoplasmic actin sequences with high homology to the *RvAct1* protein (>85%) inside the phylum *Mollusca* showed no specific amino acid motifs or residuals that can be used as a phylogenetic mark for differentiation of the three mollusks classes: *Bivalvia, Gastropoda* and *Cephalopoda* (data not shown).

Our study revealed a new interesting feature. At position 272 the amino acid Cys which is a diagnostic for cytoplasmic actins of vertebrates (*Reptilia*, *Amphibia*, *Actinopterygii*, *Aves* and *Mammalia*) was observed in the majority (>80%) of the cytoplasmic isoforms of the randomly selected from us 89 representatives of evolutionary higher invertebrates from phylum *Artropoda* but not in the representatives of phylum *Molluska* (Fig. 4). Therefore, we can speculate that this feature is a specific mark for distinguishing the non-muscle and cytoplasmic actins of mollusks from some actin isoforms of the higher invertebrate and vertebrate organisms. It is also inferred that the cytoplasmic actins of invertebrate *Arthropoda* are more similar to the cytoplasmic actins of vertebrates rather than of mollusk representatives. This could be attributed to the specific function of these actins in the more complex organisms of phylum *Arthropoda*.

Phylogenetic analysis of *Rapana venosa* actin

The phylogenetic analysis attempts to determine the evolutionary relationships between actins of invertebrates and evolutionary higher vertebrates. In the last decade, molecular phylogenetics has significantly redesigned our view on species classification and evolution. Among the genes used in phylogenetic studies the actin genes are useful tool because they are highly conserved among different organisms. However, there are difficulties in constructing phylogenetic relationships in

many large groups of organisms due to insufficient data available from many orders and families. To fill this gap we performed a phylogenetic analysis that examine the evolutionary relationship of the actin isolated from *Rapana venosa* with actin sequences from invertebrates (worm, mollusks, artropods) and vertebrates (fish, frog, snake, chicken, human) retrieved from GenBank with sequence identity of 97-99% and 93-99%, respectively.

The constructed phylogenetic tree is composed of 2 major clades. The first clade includes only vertebrate's muscle (α) isoforms and the non-muscle actin of worm (*Caernohabditis elegans*, phylum *Nematoda*). The second is divided into 2 sub-clades. The first sub-clade includes the cytoplasmic actin isoforms of vertebrates while the second one encompasses cytoplasmic, non-muscle and one defined muscle actin isoforms of invertebrates (phyla *Arthropoda* and *Molluska*).

As a member of phylum *Mollusca* the actin from *Rapana venosa* groups with the cytoplasmic actins in the larger sub-clade of the second clade. This is an additional confirmation to the amino acid comparison analysis for the affiliation of *R. venosa`s* actin to cytoplasmic isoforms as well as for the higher similarity of the mollusc's actins to the cytoplasmic rather than muscle isoforms of vertebrates. However, based on the specific grouping displayed in the phylogentic tree few potential cytoplasmic isoforms and one putative muscle isoform could be distinguished within the phylum *Molluska*. The putative muscle isoform includes non-muscle mollusk's actins isolated from muscle tissue (*C. gigas*, *H. bleekeri*, *O. vulgaris*), non-muscle actins (*M. galloprovincialis*, *P. fucata*) and one defined muscle actin (*A. californica*). The actin from *R. venosa* together with the cytoplasmic ones from *D. pealeii*, *I. paradoxus* and *S. maindroni* groups more closely to the putative muscle actin isoform. The observed differentiation of the mollusk's actins in this sub-clade might suggest the existence of distinct cytoplasmic isoforms and a muscle one, obviously due to gene duplication events and divergence from the ancestral cytoplasmic gene.

The smaller sub-clade includes cytoplasmic actins from classes *Insecta*, *Arachnida*, *Crustacea* of the phylum *Arthropoda*. The grouping of the representatives of these classes within the phylogenetic tree clearly showed that the cytoplasmic actins of the higher invertebrates are positioned in close proximity to the vertebrate cytoplasmic actins and molluskan ones.

In summary, the actin from *Rapana venos*a shows high similarity to the known cytoplasmic actins of representatives of different taxonomic groups from invertebrates to vertebrates. However, additional studies are highly necessitated for isolation and more precise characterization and classification of different isoforms of actins from various tissues and developmental stages from large number of species of phylum *Molluska* using most recent developments in molecular biology, next-generation sequencing (NGS) and bioinformatics.

Figure.1 Schematic representation of *Rapana venosa* cDNA cloning strategy (left) and visualization of the three isolated PCR fragments on a agarose gel, according to Material and Methods (right)

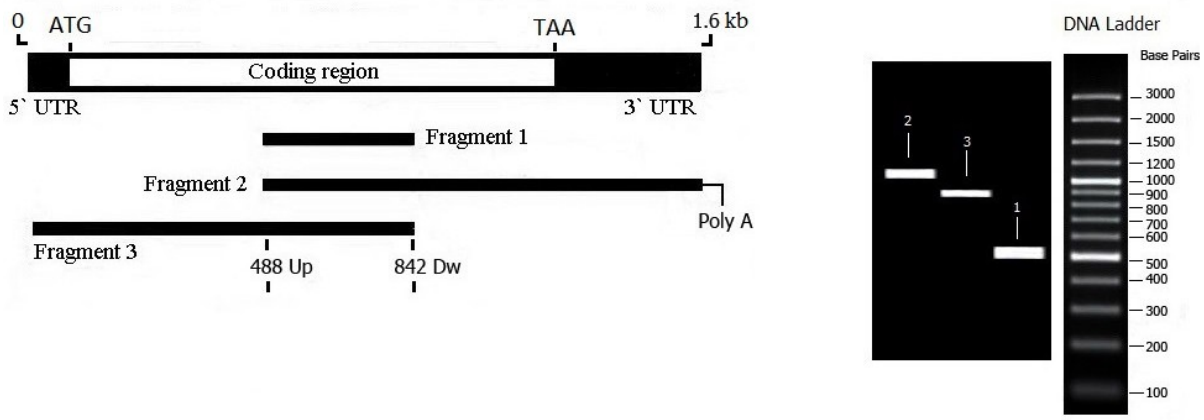


Figure.2 Nucleotide, amino acid sequence and ORF of actin cDNA from *Rapana venosa (RvAct1)*. The start codon (ATG), the stop codon (TAA) and the polyadenylation signal (AATAAA) are marked in bold

```
1     CGTCTTTCGATCCACTCTTCAGGCAGAGAAGGCACAAAGAGCCCTGTTCAAATTCTCTTCCAATTTTTAAGCACTTTTTTGCTTCCAGCCCTCCGACTC
                  M  C  D  D  E  V  A  A  L  V  V  D  N  G  S  G  M  C  K  A  G  F  A  G  D  D  A  P  R
100   TAACTTATCATCATGTGCGACGACGAAGTTGCTGCTCTTGTTGTGGACAATGGCTCCGGCATGTGCAAGGCCGGGTTTGCCGGGGACGACGCCCCCAGG
      A  V  F  P  S  I  V  G  R  P  R  H  Q  G  V  M  V  G  M  G  Q  K  D  S  Y  V  G  D  E  A  Q  S  K
199   GCTGTCTTCCCCTCCATCGTGGGTCGCCCCCGTCACCAGGGTGTGATGGTGGGCATGGGTCAGAAGGACAGCTACGTGGGAGACGAGGCCCAGAGCAAG
      R  G  I  L  T  L  K  Y  P  I  E  H  G  I  V  T  N  W  D  D  M  E  K  I  W  H  H  T  F  Y  N  E  L
298   CGTGGTATCCTGACCCTGAAGTACCCCATCGAGCACGGCATCGTCACCAACTGGGACGACATGGAGAAGATCTGGCACCACACCTTCTACAACGAGCTG
      R  V  A  P  E  E  H  P  V  L  L  T  E  A  P  L  N  P  K  A  N  R  E  K  M  T  Q  I  M  F  E  T  F
397   CGCGTGGCCCCCGAGGAGCACCCCGTCCTGCTGACCGAGGCCCCCCTCAACCCCAAGGCCAACCGTGAGAAGATGACCCAGATCATGTTCGAGACCTTC
      N  A  P  A  M  Y  V  A  I  Q  A  V  L  S  L  Y  A  S  G  R  T  T  G  I  V  L  D  S  G  D  G  V  T
496   AACGCCCCGGCCATGTACGTGGCCATCCAGGCCGTGCTGTCCCTGTACGCTTCCGGTCGTACCACCGGCATCGTGCTGGACTCCGGAGATGGTGTCACC
      H  T  V  P  I  Y  E  G  Y  A  L  P  H  A  I  L  R  L  D  L  A  G  R  D  L  T  D  Y  L  M  K  I  L
595   CACACTGTTCCCATCTACGAGGGTTACGCTCTTCCCCACGCCATCTTGAGGTTGGATCTGGCCGGCCGTGACCTGACAGACTACCTGATGAAGATCCTG
      T  E  R  G  Y  S  F  T  T  T  A  E  R  E  I  V  R  D  I  K  E  K  L  C  Y  V  A  L  D  F  E  Q  E
694   ACAGAGCGTGGCTACAGCTTCACCACCACAGCTGAACGAGAAATCGTCCGTGACATCAAGGAGAAGCTCTGCTACGTTGCCTTGGACTTTGAGCAGGAA
      M  Q  T  A  A  S  S  S  S  L  E  K  S  Y  E  L  P  D  G  Q  V  I  T  I  G  N  E  R  F  R  C  P  E
793   ATGCAGACAGCTGCATCCTCCTCCTCCCTGGAGAAGAGCTACGAGCTCCCTGACGGTCAGGTCATCACCATCGGCAACGAGCGATTCCGTTGTCCCGAG
      A  M  F  Q  P  S  F  L  G  M  E  S  A  G  I  H  E  T  T  Y  N  S  I  M  K  C  D  V  D  I  R  K  D
892   GCCATGTTCCAGCCATCTTTCCTGGGTATGGAATCTGCTGGTATCCACGAGACCACCTACAACTCCATCATGAAGTGTGACGTGGACATCCGTAAAGAC
      L  Y  A  N  T  V  L  S  G  G  S  T  M  Y  P  G  I  A  D  R  M  Q  K  E  I  S  A  L  A  P  P  T  M
991   TTGTACGCCAACACTGTGCTGTCTGGTGGCTCCACCATGTACCCTGGTATCGCCGACAGAATGCAGAAGGAAATCAGTGCCTTGGCTCCTCCCACCATG
      K  I  K  I  I  A  P  P  E  R  K  Y  S  V  W  I  G  G  S  I  L  A  S  L  S  T  F  Q  Q  M  W  I  S
1090  AAGATCAAGATCATCGCTCCCCCAGAACGCAAGTACTCTGTGTGGATCGGTGGCTCCATCCTGGCCTCTCTGTCCACCTTCCAGCAGATGTGGATCTCC
      K  Q  E  Y  D  E  S  G  P  S  I  V  H  R  K  C  F  *
1189  AAGCAGGAATACGATGAGTCTGGCCCCTCCATCGTCCACAGAAAGTGCTTCTAACTCAGGGCATCGTACAGACATCTTTCTCTTTCGTTAACACTTGT
1288  CTCCACACACTGCTGCCATACAAAAGATGGAGATTCAATATTTTGAAAAACCGACAAGTCATAATGCTTATTTCTTTTGTTTCTTGACTGCAAATGAGA
1387  CAACAACTTTCATGATCCCTTTTGTATTATTATTTTTGAACAATAGGCGGCTGTTTTATATCGGTTTGTTTTTTAGTTCAAATGCTTACATTGAATACT
1486  ATGTCCATACTTTTGTTACTTACAAACGAAATAAAATGGTTAACAATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Figure.3 Multiple alignment of *Rapana venosa* actin (AGZ87937.1) and the six actin isoforms found in *Homo sapiens*: cytoplasmic β-actin (AAH16045), cytoplasmic γ-actin (NP_001605.1), cardiac α-actin (NP_005150.1), skeletal α-actin (NP_001091.1), smooth muscle α-actin (NP_001604.1) and smooth muscle γ-actin (P63267.1). The diagnostic amino acids and position 272 are marked with asterisk (*)

Figure 4. Comparison of the N-terminal sequences and the diagnostic for vertebrate's amino acid at position 272 (Cys) of *Rapana venosa* actin (AGZ87937.1) and the actins from representatives of other taxonomic classes: *Mammalia* (NP_001605.1; AAH16045; NP_005150.1; NP_001091.1; NP_001604.1; P63267.1), *Aves* (CAA25004.1; NP_001007825.1), *Reptilia* (JAB54771.1), *Amphibia* (P15475.1), *Actinopterygii* (ACI23578.1), *Insecta* (AGR44827.1; AAY81972.1), *Arachnida* (ABV71392.1; BAE46505.1), *Malacostraca* (ADG45302.1; AAG16253.1), *Cephalopoda* (AAU11523.1; AEN25588.1), *Gastropoda* (AGZ87937.1; AAB49413.1; AAX19288.1), *Bivalvia* (AEK81538.1; ACD99707.1). The diagnostic for vertebrate's amino acid (Cys) at position 272 is marked with asterisk (*). A selected set of six of the studied 89 representatives of phylum *Arthropoda* are included in the multiple alignment for simplicity

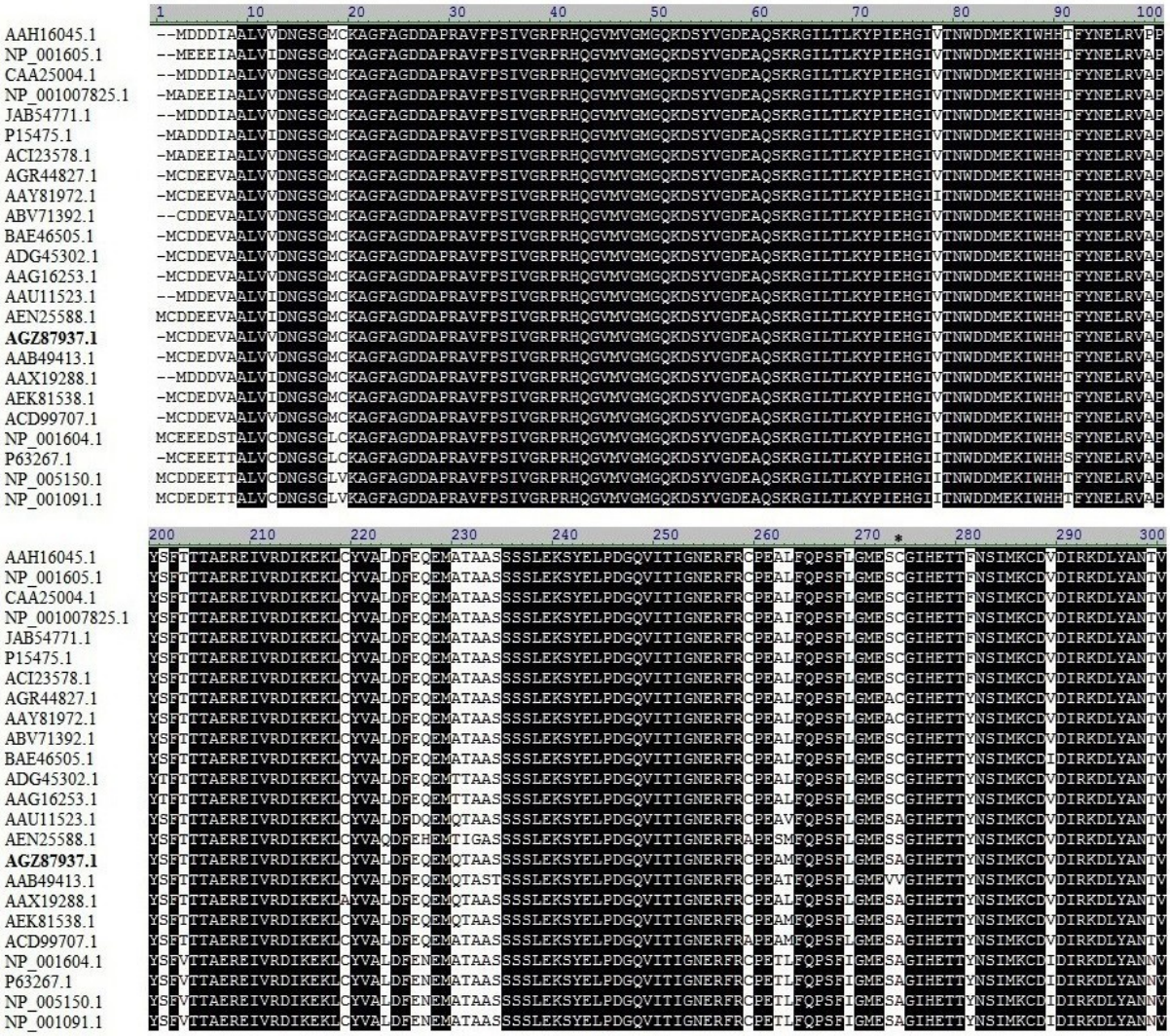
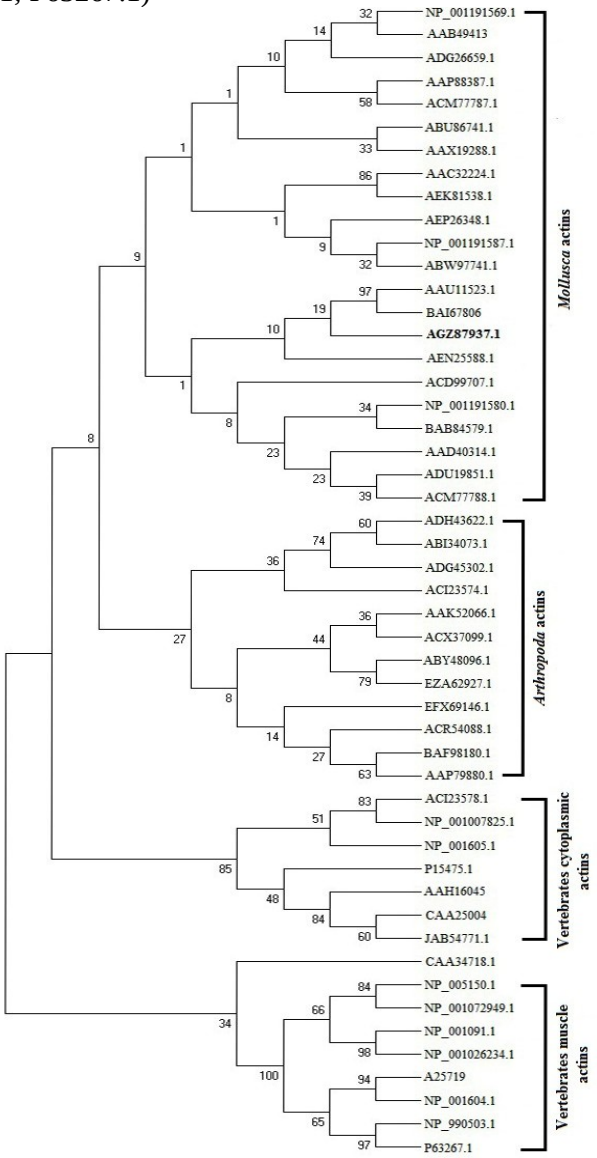


Figure.5 Phylogenetic tree of actin isoforms from *Rapana venosa* (AGZ87937.1) and different vertebrate and invertebrate species: *Aplysia californica* (NP_001191569.1; NP_001191587.1; NP_001191580.1), *Biomphalaria glabrata* (AAB49413), *Hyriopsis cumingii* (ADG26659.1), *Azumapecten farreri* (AAP88387.1), *Todarodes pacificus* (ACM77787.1), *Haliotis diversicolor* (ABU86741.1), *Haliotis iris* (AAX19288.1), *Dreissena polymorpha* (AAC32224.1), *Meretrix meretrix* (AEK81538.1), *Sinonovacula constricta* (AEP26348.1), *Crassostrea ariakensis* (ABW97741.1), *Doryteuthis pealeii* (AAU11523.1), *Idiosepius paradoxus* (BAI67806), *Sepiella maindroni* (AEN25588.1), *Pinctada fucata* (ACD99707.1), *Crassostrea gigas* (BAB84579.1), *Mytilus galloprovincialis* (AAD40314.1), *Heterololigo bleekeri* (ADU19851.1), *Octopus vulgaris* (ACM77788.1), *Eriocheir sinensis* (ADH43622.1), *Pacifastacus leniusculus* (ABI34073.1), *Marsupenaeus japonicas* (ADG45302.1), *Homarus americanus* (ACI23574.1), *Heliothis virescens* (AAK52066.1), *Polyrhachis vicina* (ACX37099.1), *Nilaparvata lugens* (ABY48096.1), *Cerapachys biroi* (EZA62927.1), *Daphnia pulex* (EFX69146.1), *Agelena silvatica* (ACR54088.1), *Ixodes persulcatus* (BAF98180.1), *Rhipicephalus microplus* (AAP79880.1), *Acipenser transmontanus* (ACI23578.1), *Xenopus borealis* (P15475.1), *Micrurus fulvius* (JAB54771.1), *Caenorhabditis elegans* (CAA34718.1), *Gallus gallus* (NP_001007825.1; CAA25004; NP_001072949.1; NP_001026234.1; A25719; NP_990503.1) and *Homo sapiens* (NP_001605.1; AAH16045; NP_005150.1; NP_001091.1; NP_001604.1; P63267.1)



Acknowledgments

This work was supported Bulgarian Ministry of Education, projects TK01-496/2009, ДНТС/01/6 and "Young researchers" DMU 03/26, grant №BG051PO001-3.3.06-0025, financed by the European Social Fund and Operational Programme Human Resources Development (2007 – 2013) and co-financed by Bulgarian Ministry of Education, Youth and Science, and Deutsche Forschungsgemeinstchaft (DFG-STE 1819/5-1/2012), Germany.

References

Adema, C.M., 2002. Comparative study of cytoplasmic actin DNA sequences from six species of Planorbidae Gastropoda: Basommatophora. *J. Molluscan Studies*, 68: 17-23.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.*, 215: 403-410.

Baldauf, S. L., Roger, A. J., Wenk-Siefert, I., Doolittle, W.F. 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. J. Sci. Washington, 972-977.

Cadoret, J.P., Debon, R., Cornuddela, L., Lardans, V., Morvan, A., Roch, P., and Boulo, V. 1999. Transient expression assays with the proximal promoter of a newly characterized actin gene from the oyster *Crassostrea gigas*. *FEBS Lett.*, 460: 81-85.

Des Groseillers, L., Auclair D., Wickham L., and Maalouf M. 1994. A novel actin cDNA is expressed in the neurons of *Aplysia californica*. *Biochim. Biophys. Acta*, 1217: 322–324

Dolashka, P., Nesterova, N., Zagorodnya, S., Dolashki A., Baranova, G., Golovan, A., Voelter, W. 2014. Antiviral Activity of hemocyanin *Rapana venosa* and its isoforms against Epstein-Barr virus. *Global Journal of Pharmacology*, 8(2): 206-212.

Dominguez R. and Holmes K.C. (2011). Actin Structure and Function. *Annu Rev Biophys.*, 40: 169–186.

Doussau, F., Augustine G. J. 2000. The actin cytoskeleton and neurotransmitter release: An overview. *Biochimie*, 82: 353-363.

Erba, H.P., Gunning, P., Kedes, L. 1986. Nucleotide sequence of the human gamma cytoskeletal actin mRNA: anomalous evolution of vertebrate non-muscle actin genes. *Nucleic Acids Res.*, 14(13): 5275-5294.

Erba, H.P., Eddy, R., Shows, T., Kedes, L., Gunning, P. 1988. Structure, chromosome location, and expression of the human gamma-actin gene: differential evolution, location, and expression of the cytoskeletal beta- and gamma-actin genes. *Mol. Cell. Biol.*, 8(4): 1775-1789.

Flamini, M.I., Fu, X.D., Sanchez, A.M., Giretti, M.S., Garibaldi, S., Goglia, L., Pisaneschi, S., Tosi, V., Genazzani, A.R., Simoncini, T. 2009. Effects of raloxifene on breast cancer cell migration and invasion through the actin cytoskeleton. *J. Cell Mol. Med.*, 13(8B): 2396-2407.

Fisher, D. A., Bode, H.R. 1989. Nucleotide sequence of an actin-encoding gene from *Hydra attenuata*: structural characteristics and evolutionary implications. Gene, 84: 55-64.

Goodson, H.V., Hawse, W.F. 2002. Molecular evolution of the actin family. *J. Cell Sci.*, 115: 2619-2622.

Gunning P., Ponte P., Blau H., Kedes, L. 1983. α-skeletal and α-cardiac actin genes are co-expressed in adult human skeletal muscle and heart. *Mol. Cell. Biol.*, 3: 1985-1995.

Gunning P., Ponte P., Kedes L., Eddy R., Shown, T. 1984. Chromosomal localization of the co-expressed human skeletal and cardiac actin genes. *Proc. Natl. Acad. Sci. USA*, 81:1813-1817.

Hamada H., Petrino, M.G, Kakunaga, T. 1982. Molecular structure and evolutionary origin of human cardiac muscle actin

gene. *Proc. Natl. Acad. Sci. USA*, 79: 5901-5905.

Hamada S., Nakano Y., Kakunaga, T. 1989. Structure of 3`-downstream segment of the human smooth muscle (aortic-type) α-actin-encoding gene and isolation of the specific DNA probe. *Gene*, 84: 455-462.

Hongming, M., Kangsen, M., Zhiguo, L., Wei, X. 2007. Cloning and characterization of an actin gene of *Chlamys farreri* and the phylogenetic analysis of mollusk actins. *Chinese J. Oceanology and Limnology*, 25: 304-309.

Hurley, J.H. 1996. The sugar kinase/heat shock protein 70/actin superfamily: implications of conserved structure for mechanism. *Annu. Rev. Biophys. Biomol. Struct*., 25: 137–162.

Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., Holmes, K.C. (1990). Atomic structure of the actin: DNase I complex. *Nature*, 347(6288): 37-44.

Kabsch, W., and Vandekerckhove, J., 1992. Structure and function of actin. *Annu. Rev. Biophys. Biomol.* 21: 49–76.

Kos'yan, A.R. 2013. Comparative analysis of *Rapana venosa* (Valenciennes, 1846) from different biotopes of the Black Sea based on its morphological characteristics. *Oceanology*, 53(1): 47-53.

Kusakabe, T., Araki, I., Satoh, N., Jeffery, W.R. 1997. Evolution of chordate actin genes: evidence from genomic organization and amino acid sequences. *J. Mol. Evol.*, 44: 289–298.

Lammers, R.J., Witjes, W.P., Janzing-Pastors, M.H., Caris, C.T., Witjes, J.A. 2012. Intracutaneous and intravesical immunotherapy with keyhole limpet hemocyanin compared with intravesical mitomycin in patients with non-muscle-invasive bladder cancer: results from a prospective randomized phase III trial. *J. Clin. Oncol*., 30(18): 2273-2279.

Lardans, V., Ringaut, V., Duclermortier P., Cadoret, J.P., Dissous C. 1997. Nucleotide and deduced amino acid sequences of *Biomphalaria glabrata* actin cDNA. *DNA Seq*., 7: 353-356.

Minty A., and Kedes L. 1986. Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. *Mol. Cell. Biol*., 6: 2125-2136.

Miwa, T., Manabe, Y., Kurokawa, K., Kamada, S., Kanda, N., Bruns, G., Ueyama, H., Kakunaga, T. 1991. Structure, chromosome location, and expression of the human smooth muscle (enteric type) gamma-actin gene: evolution of six human actin genes. *Mol. Cell Biol*., 11(6): 3296-306.

Miyamoto, H., Hamaguchi, M., Okoshi, K. 2002. Analysis of genes expressed in the mantle of oyster *Crassostrea gigas*. *Fisheries Sci*., 68: 651-658.

Munn, A.L. 2001. Molecular requirements for the internalisation step of endocytosis: insights from yeast. *Biochim Biophys Acta*, 1535: 236-257.

Nakajima-Iijima S., Hamada, H., Reddy, P., Kakunaga, T. 1985. Molecular structure of the human cytoplasmic β-actin gene: interspecies homology of sequences in the introns. *Proc. Natl. Acad. Sci. USA*, 82: 6133-6137.

Nakano, Y., Nishihara Y., Sasayama S., Miwa T., Hamada S., Kakunaga, T. 1991. Transcriptional regulatory elements in the 5' upstream and first intron regions of the human smooth muscle (aortic type) α-actin-encoding gene. *Gene*, 99: 285-289.

Ng, S.Y., Gunning, P., Eddy, R., Ponte, P., Leavitt, J., Shows, T., Kedes, L. 1985. Evolution of the functional human β-actin gene and its multi-pseudogene family: conservation of noncoding regions and chromosomal dispersion of pseudogenes. *Mol. Cell. Biol*., 5: 2720-2732.

Noonan, K.E., Beck, C., Holzmayer, T.A., Chin, J.E., Wunder, J.S., Andrulis, I.L., Gazdar, A.F., Willman, C.L, Griffith, B., Von Hoff, D.D., Roninson, I.B. 1990. Quantitative analysis of MDR1 (multidrug resistance) gene expression in human tumors by polymerase chain

reaction. *Proc. Natl. Acad. Sci*., 87: 7160-7164.

Oleksiewicz, M.B., Donaldson, A.I., Alexandersen, S. 2001. Development of a novel real-time RT-PCR assay for quantitation of foot-and-mouth disease virus in diverse porcine tissues. *J. Virol. Methods*, 92: 23-35.

Patwary, M.U., Reith, M., Kenchington, E.L. 1996. Isolation and characterization of cDNA encoding an actin gene from sea scallop (*Placopecten magellanicus*). *J. Shellfish Res*., 15: 265-270.

Ponte, P., Ng, S.Y., Engel, J., Gunning, P., Kedes, L. 1984. Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human beta-actin cDNA. *Nucleic Acids Res.*, 12: 1687-1696.

Rogers, S.L., Gelfand, V.I. 2000. Membrane trafficking, organelle transport, and the cytoskeleton. *Curr. Opin. Cell Biol*., 12: 57-62.

Saitou, N., Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol*., 4: 406-425.

Sin, Y.T., Bryant, M.J., Johnstone, A. 2007. *Zoological studies*, 46(6): 734-745.

Taylor, A., Erba, H.P., Muscat G.E., Kedes, L. 1988. Nucleotide sequence and expression of the human skeletal α-actin gene: evolution of functional regulatory domains. *Genomics*, 3: 323-336.

Thompson, J.D., Higgins, D.G., Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22: 4673-4680.

Toshkova, R., Ivanova, E., Hristova, R., Voelter, W., Dolashka-Angelova, P. 2009. Effect of Rapana venosa hemocyanin on antibody-dependent cell cytotoxicicity (ADCC) and mitogen responsibility of lymphocytes from hamsters with progressing myeloid tumors. *World Journal of Medical Sciences*, 4(2): 135-142.

Ueyama, H., Hamada H., Battula N., Kakunaga, T. 1984. Structure of a human smooth muscle actin gene (aortic type) with a unique intron site. *Mol. Cell. Biol.*, 4: 1073-1078.

Ueyama, H., Bruns, G., Kanda, N. 1990. Assignment of the vascular smooth muscle actin gene ACTSA to human chromosome 10. *Jinrui Idengaku Zasshi*., 35(2): 145-50.

Valentijn, K., Valentijn, J.A., Jamieson, J.D. 1999. Role of actin in regulated exocytosis and compensatory membrane retrieval: Insights from an old acquaintance. *Biochem. Biophys. Res. Commun*., 266: 652-661.

Vandekerckhove, J., Weber, K. 1979. The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine fast skeletal muscle, and rabbit slow skeletal muscle. A protein-chemical analysis of muscle actin differentiation. Differentiation, 14: 123-133.

Yamashita, M., Ogawa, T., Zhang, X., Hanamura, N., Kashikura, Y., Takamura, M., Yoneda, M., Shiraishi, T. 2012. Role of stromal myofibroblasts in invasive breast cancer: stromal expression of alpha-smooth muscle actin correlates with worse clinical outcome. *Breast Cancer*, 19(2): 170-176.