



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

Data on four apoptosis-related genes in the colonial tunicate *Botryllus schlosseri*

Nicola Franchi, Francesca Ballin, Lucia Manni\*, Filippo Schiavon, Lorian Ballarin

Department of Biology, University of Padova, Italy

## ARTICLE INFO

## Article history:

Received 13 April 2016

Received in revised form

3 May 2016

Accepted 11 May 2016

Available online 20 May 2016

## Keywords:

Sequence databases

Alignments

BLAST analysis

*Botryllus schlosseri*

Apoptosis-related transcripts

Bax

AIF1

PARP1

IAP7

## ABSTRACT

The data described are related to the article entitled “Recurrent phagocytosis-induced apoptosis in the cyclical generation change of the compound ascidian *Botryllus schlosseri*” (Franchi et al., 2016) [1]. Four apoptosis-related genes, showing high similarity with mammalian Bax (a member of the Bcl-2 protein family), AIF1 (apoptosis-inducing factor-1), PARP1 (poly ADP ribose polymerase-1) and IAP7 (inhibitor of apoptosis-7) were identified from the analysis of the transcriptome of *B. schlosseri*. They were named BsBax, BsAIF1, BsPARP1 and BsIAP7. Here, their deduced amino acid sequence were compared with known sequences of orthologous genes from other deuterostome species together with a study of their identity/similarity.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Specifications Table

Subject area	Biology
More specific subject area	Developmental Biology
Type of data	Tables, figures

DOI of original article: <http://dx.doi.org/10.1016/j.dci.2016.04.011>

\* Corresponding author at: Department of Biology, University of Padova via U. Bassi 58/B 35131 Padova Italy. Tel.: +39 049 8276252; fax: +39 049 8276199.

E-mail address: [lucia.manni@unipd.it](mailto:lucia.manni@unipd.it) (L. Manni).<http://dx.doi.org/10.1016/j.dib.2016.05.017>

2352-3409/© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

How data was acquired	Bioinformatic analysis, RACE
Data format	Raw and analysed data
Experimental factors	The partial transcripts present in the transcriptome, identified by BLAST analysis, were elongated through 5' and 3' RACE according to the 2nd generation of 5'/3' RACE kit
Experimental features	Analysis with BLAST, LALIGN, SMART, Clustal Omega
Data source location	Padova, Italy
Data accessibility	Data are available in this article and at GenBank via accession numbers GenBank: KU948200 for BsBAX, GenBank: KU948201 for BsPARP1, GenBank: KU948202 for BsAIF1, GenBank: KU948203 for BsIAP7.

---

### Value of the data

- The data provide the full-length sequences of four apoptosis-related transcripts from the colonial ascidian *B. schlosseri* useful to study the phylogeny trees of the corresponding proteins in chordates.
  - From the data, the protein primary structures can be deduced and, from that, three-dimensional models can be obtained, useful to compare the domain organization of the corresponding chordate proteins.
  - Expression studies, exploiting the present data, can contribute to elucidate the dynamics of the cyclical apoptosis, which characterizes the colonial blastogenetic cycle of the ascidian *B. schlosseri*.
- 

### 1. Data

The data reported include supporting information to the phylogenetic analyses of Franchi et al. [1]. They consist of transcript sequences, sequence alignments and comparisons of four apoptosis-related genes identified in the recently-obtained transcriptome of *B. schlosseri* [2]. The sequences show high similarity with mammalian transcripts for Bax, AIF1, PARP1 and IAP7 and were named BsBax, BsAIF1, BsPARP1 and BsIAP7, respectively. The expression of these genes was studied further in the above-reported paper [1].

### 2. Experimental design, materials and methods

Amplification and cloning of transcripts for BsBax, BsAIF1, BsPARP1 and BsIAP7 was achieved with specific primers designed on sequences found in our collection of transcriptomes [2]. In order to verify and complete the full length cDNA, PCR reactions were carried out with a denaturing step at 94 °C for 2 min, 40 cycles of 30 s at 94 °C, 40 s at 60 °C and 90 s at 72 °C, and a final extension at 72 °C for 10 min. Amplicons were separated using 1.5% agarose gel, purified, cloned and sequenced. The partial transcripts were elongated through 5' and 3' RACE according to the 2nd generation of 5'/3' RACE kit (Roche). [Supplementary Table 1](#) reports the specific primers used for amplicons production and their elongation through 5'- and 3'-RACE and for the *in situ* hybridisation experiments reported in [1].

The sequences from GenBank, reported in [Supplementary Tables 2–5](#), were used for alignments and sequence comparisons with the sequences of BsBax, BsAIF1, BsPARP1, BsIAP7, respectively. The



TGTTA  
 TAAGTTGATTGGCGATTGCTCTTCAGGCTCTCCTTGATTTCGAAGTTGTTAGTATATA  
 1 GGCACCGCGTATCTAACATATAGATATACGGTATACCATTCCATTTCTTGTGTGGCA M  
 1 ATG  
 2 S G G G D Q A K K E A G K I D D K N D E  
 4 TCGGGGGCGGTGATCAAGCGAAGAAGGAAGCAGGGGAAGATAGATGATAAAAAATGACGAG  
 22 D P R P S R P T R P R T A R P G R R T D  
 64 GATCCTCGACCATCTAGGCCAACCCGACCAGAGAACCAGGACAGGACAGGATAGAAAGCAGAT  
 42 P E S S L V Q G A V G G E G S Y T R H L  
 124 CCTGAATCGTCTTGGTCCAGGGAGCTGTTGGTGGGGAAGGTTTCATACACGAGACATCTA  
 62 T E E E R A V G N Q A R H L L T M F I S  
 184 ACTGAGGAGAAAAGAGCTGTTGGAAACCAAGCAAGCATCTTCTGACAATGTTTATCTCT  
 82 D R A E R D P D L S Q D K K G A V Q E T  
 244 GACAGAGCAGAACGTGATCCCACCTGTCGCAGGATAAGAAAGGAGCTGTACAGGAAAAA  
 102 V Q L C N K S S E S N A M S D S D A A L  
 304 GTTCAGTTATGTAACAAGAGCTCGGAAAGTAATGCGATGAGCGATAGCGATGCTGCTTTA  
 122 R E V S L T L R V I G D Q L N E D N D L  
 364 AGGGAGGTATCATTGACCTTGCGTGTATTGGAGATCAGTTGAATGAAGACAATGACCTTT  
 142 N S L I D R V L V Q P T K D I F M K V C  
 424 AACAGCCTTATTGATAGAGTTCTCGTTCAACCCACCAAGGACATCTTCATGAAAGTTTGT  
 162 K Q I F A D O N F N W G R V A A V F Y F  
 484 AAACAGATATTGCCGATCAAAATTTCAACTGGGGCAGAGTAGCCGCTGTATTCTACTTT  
 182 A Y R L I T K A V C T T M D E G I D W I  
 544 GCGTACAGACTTATAACAAAGGCTGTATGCACACCATTGGATGAAGGCATAGATTGGATT  
 202 K D I M S W A M D F L Y Q Y V L W W I V  
 604 AAAGATATTATGAGTTGGGCGATGGACTTCTTATATCAGTATGTTCTTTGGTGGATTGTC  
 222 Q R G G W G M I R E Y F G T P T A Y V V  
 664 CAAAGGGGAGGTTGGGGAATGATCAGAGAATACTTTGGAACACCCGACTGCGTATGTGGTT  
 242 S L S A A T I L T V L Y V W W K Q Q - 259  
 724 TCTCTGTCTGCCCAACAATACTTACTGTTTTGTATGTATGGTGAAGCAGCAATGA 850  
 TCT  
 TAAACTTTATCATTCTGTTTTTACCAAGATCACTGTCTTAGATATGAAATTTACCCAGT  
 GCCTTTTATAACGGCTACATCTGGTTTTATTTGTAGGAAGTTCAGATTTCTGTGGACATA  
 CTCTAGATCAGTAATCGGCAATCTATTTTGGCTCTGCAGACTGGTTCCAAAAATGCAACAC  
 ACAAAATTTCAATTAAGCGACTAACCTAATCTGTTTTATTAACATACAGTGATTGACCA  
 CTATAATTTCTAAAACCTTCACTTAACTGGCAAGATATATGAAATCCACGACAGCGTCA  
 CCGGTCTGCAAAACCGTGGGTTGCCGACCCTCCATATACATTCATTGTGAAATCGA  
 TCTAGATGTGATGAATATCAATGTTATGCATTGAAATTTTGCCAAACAAATGACAGGCCA  
 CTGAAATGGTTTTGAGATTTCTTAAAAATAAATGGAACGATGATTACATATCAGTTTT  
 GGCATAGCAACAAGGTGACATGCCATAGCAAATTTTCATTTGAAAAATATCAGGATTT  
 TGTGAGTGGGGTGAATCTCCGTCATTTGGTTGATGGACAAGCTGAATGTCGATACTTA  
 TATACGAGTATCCCCGTTTTGTCCAAAATAATGGTGTCTGCTTTGCATATGGTGTG  
 GTTTTTCCGGTACCAACAATGGATATGCTCCAGTTAGTCATGAGTTATACAAGGTGCTTA  
 TTTTCCCTCAAAATGAGCAAAATAATGATCCCTGAATGTCGCTCTCTATTGACATGCGATCTC  
 CTTTTTACCAGTAGACTTGAATTTCAATGGGTCTCGTTTTGTATTTTTGGCTCTATA  
 GATCAAGTACATTTTAAAGAAATGAAGCCTAATGATTTACTCATATTGCTCTGTTGCGGTG  
 TTGCTACTGTGTTCTAATCATGGGAGAAAAATCTCGCTCTCTACTGTGATATTTGA  
 GTCCGATGTGTTCTGACTTGACAATCCAGCATATATTCTCAGCGTTTTACGATCAATTT  
 ATTGCTCAATTTATGCTGATTTAATAAAAAATATCATGTAGTTTTTGTAGATAACCTACAAA  
 GATTGCTTACTCTGAAAAAAA

Fig. 2. cDNA and deduced amino acid sequence of BsBax. START and STOP codon are in green and red, respectively; 5' and 3' UTR are in grey. In yellow, the Bcl-2 domain and, within this, BH1, BH2 and BH3 domains are underlined. Light blue: transmembrane domain.

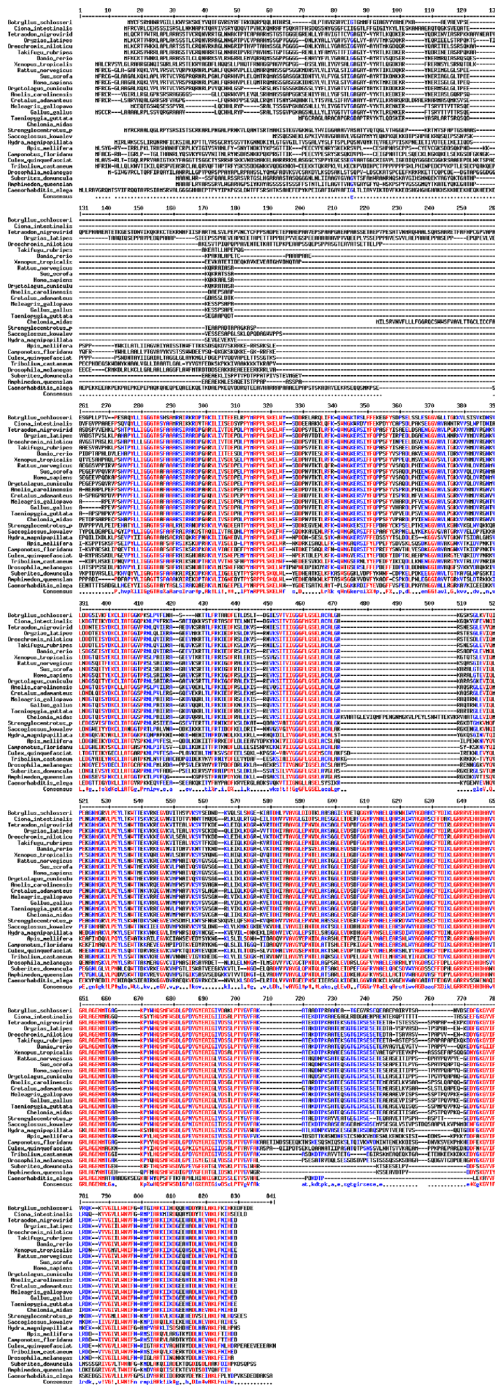


Fig. 3. Alignments of the deduced amino acid sequence of BSAF1 with known orthologous sequences from both vertebrate and invertebrates.

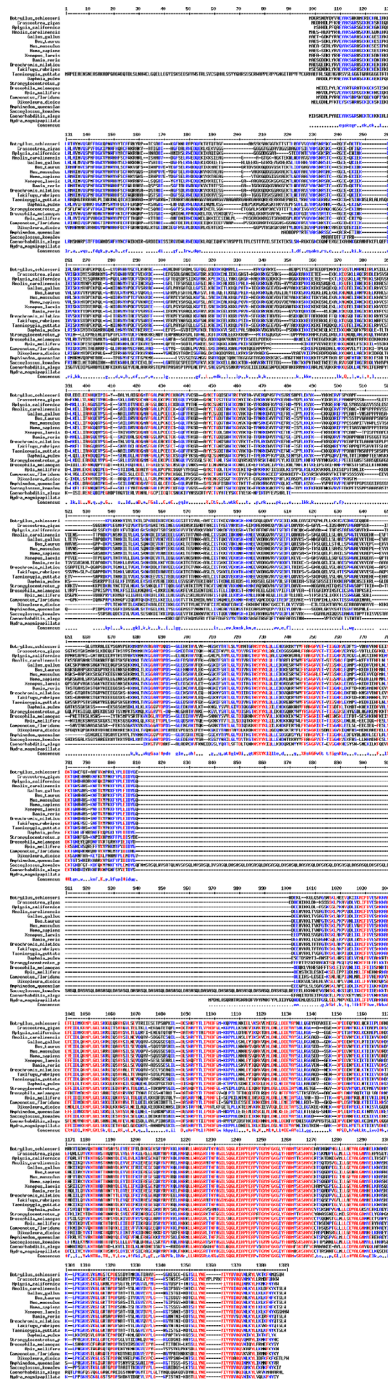
```

TTATAAGTATAGATTCAACTATAGAAGCTTCTCAATC
1
1 M Y C F S R M H
ATGATATTGTTTAGTCGGATGCAC
9 N R V G I L L K N V S K S K E Y V Q T F
25 AATCGGGTGGAACTACTGCTAAAGAAATGTGTCAAAAAGCAAAGAAATACGTTCAAAACATT
29 G V R G Y R F T R K D Q R P Q Q L H T H
85 GGGTCAGAGGATACAGGTTCACTAGAAAAGACCAACGACACAGCAGTTGCACAGCAT
49 R S L D L P T H V G S A V C I G T G M M
145 CGTCACTTGAACCTCCACGCATGTTGGGTCCTGCTGTTGCATTGGAACGGGATGATG
69 F F G I A G V Y Y A M E P K A A E V A E
205 TCTTTGGAATTGCTGGAGTTTATTATGCCATGGAGCCAAAAGCTGCTGAAGTCGCTGAA
89 V P S E E S G P L L P T V P E S A Q Y I
265 GTACCGTCGGAAGAACTCGCCCACTGTACCAACTGTCCCGAGTCTGCACAATATCTT
109 L I G G G T A S H S A M R A I R K R D P
325 TTGATCGGAGGACGAGCATCTCATTCCGCAATGCGCCCATCAGGAAAAGAGATCCA
129 T A K I L I I T F E E L R F Y M R P P I
385 ACTGCAAAAATTTAATCATCACGAAGAAGAACTTCGACCTTACATGGCACCCCGCTC
149 S K E L W Y S D D R E L A R O L I F K Q
445 TCAAAGGAACCTTGTACTCGGACGATCGGAGTTGGCGGCACTTATTTCAAACAA
169 W N G K T R S L F F E K E G F Y S D P S
505 TGGAAACGAAAACGAGGAGTTATTTTTGAGAAGGAGGATTTTATTCGGACCTTCA
189 E L S S L E S G G V G L L T G K K V L S
565 GAATTGCTGCTAGAAAAGCGGGGCTCGGACTACTTCTGGAAGAAAGGTTTGTGCG
209 I S V K D N S V T L D D G S I V K F E K
625 ATATCTGTAAGATAAATCCGTTACTCTCGATGACGGGTCGATTGTGAAATTTGAGAAA
229 C L I A T G G K P K S L P V F E N H S N
685 TGTCTCATCGCACCGGAAACCGAAAAGTTGCCAGTTTGTGAAATCACAGCAAC
249 K T T L F R T A A D F E T L D S L T D G
745 AAAACAACGCTTTTCGAACGGCGGCTTCGAGACTTTGGATTCTCTCACCGATGGA
269 I E S V T V I G G G P L G S E L A C A L
805 ATAGAACTGTACTGTGATAGGAGGAGATTCTTGGGAAGCGAATGCTTGC CGCTT
289 G H K G S K S G L K V T Q I F P E A G N
865 GGCCATAAAGCTCCAAGTCAGGCTTGAAGTGAAGCAATCTTCCGGAAGCCGGCAAC
309 M G R V L P E Y L T K W T T E K V R K E
925 ATGGCCCGCTTCCCGGAATATCTACAAAAGTGACGACGAAAGATTCGTAAGAG
329 G V D V I T H S V V K S V S E D N D K V
985 GGTGTCATGTGATCACTATTCCGTCGTAAGTCAGTGTGAGAGCAATGACAAAAGT
349 Q L S L S N G E K I A T D H L V V A V G
1045 CAGTGTCTCAGTACCGGAAAAGATGCGCACCGATCATCTGTTGTTGCCATAGT
369 L D I D T K L A A S A G I E V D E T Q G
1105 TTAGACATCGACGAACTGGCAGCTTCCGAGGTATCGAAGTGGATGAAACCGAAGCC
389 G Y R V N A E L Q A R N N I W V A G D A
1165 GGGTATCGGGTTAATGCCGAGTTGCAAGCGAAGAAACAATATCTGGTGGCAGGAGCGCA
409 A C F Y D I K L G R R R V E H H D H A V
1225 GCGTGTTCACGACATCAAGCTTGAAGGAGGAGGAGTGGAGCACATGACCATGCGGTG
429 V S G R L A G E N M T G A G K P Y W H Q
1285 GTCAGCGCAGATTGGCCGGAAAATATGACCGGTGCGGTAACCATCACTGGCACCAA
449 S M F W S D L G P D V G Y E A I G I V D
1345 TCCATGTTTTGGTCCGATCTGCTCCGATGTGGGATATGAAGCATTGAAATTTGGAC
469 A K L P T V G V F A K A T A A D T P R A
1405 GCCAAACTGCCTACTGTGCGAGTCTTTGTAAGCCACGGCAGCCGACAGCCACGCGCT
489 A A E A T G E G V R S E Q E A A E P A I
1465 GCAGCAGAACCCTGCGCAAGGCTACGCTGTAACGGAAGCGGCAACCTGCTATT
509 R V T S A N V D S E D F G K G V V F Y V
1525 CCGGTGACTTCAGCCAAATGTTACTCGGAAGACTTCGCAAGGTTGATGTTTCTACGTG
529 R Q K K V V G I L L W N I F G R T G I A
1585 CGACAAAAGAAAGTCTGGGACTGCTGTGGAACTTTCCGGCCACTGGAATTGCG
549 R K I I H D Q Q E H D D Y R E L A K L F
1645 AGAAAGTCAATCATGACCAACRAAGAACAGACGACTCCGAGAACTCCGAAACTGTTT
569 K I H K E D F E E - 578
1705 AAAATCACAAAGAAGATTTGAAGACGAA 1737
ATCCAACGTTTCGATCTCAGATGC
CAAATATGATATTGCACTATAAATACAGAAGTGTGTAATGCTGTTAAGTCTGT
ACCAAGAAGAAGCCGACCAACCTGATGAGGGAGTTGGAGCTTATACAGTCTTCCG
GGTAGGACCTTAACACTAGATAATTTCAACGTTTCCAAATATGATTAGAGTTCGAATT
CTAGCTACAGTTTTTTAAATAGACTAGAGTAGGAGATATTAATATAGAAATGACTG
AATTTTAAAAAAA

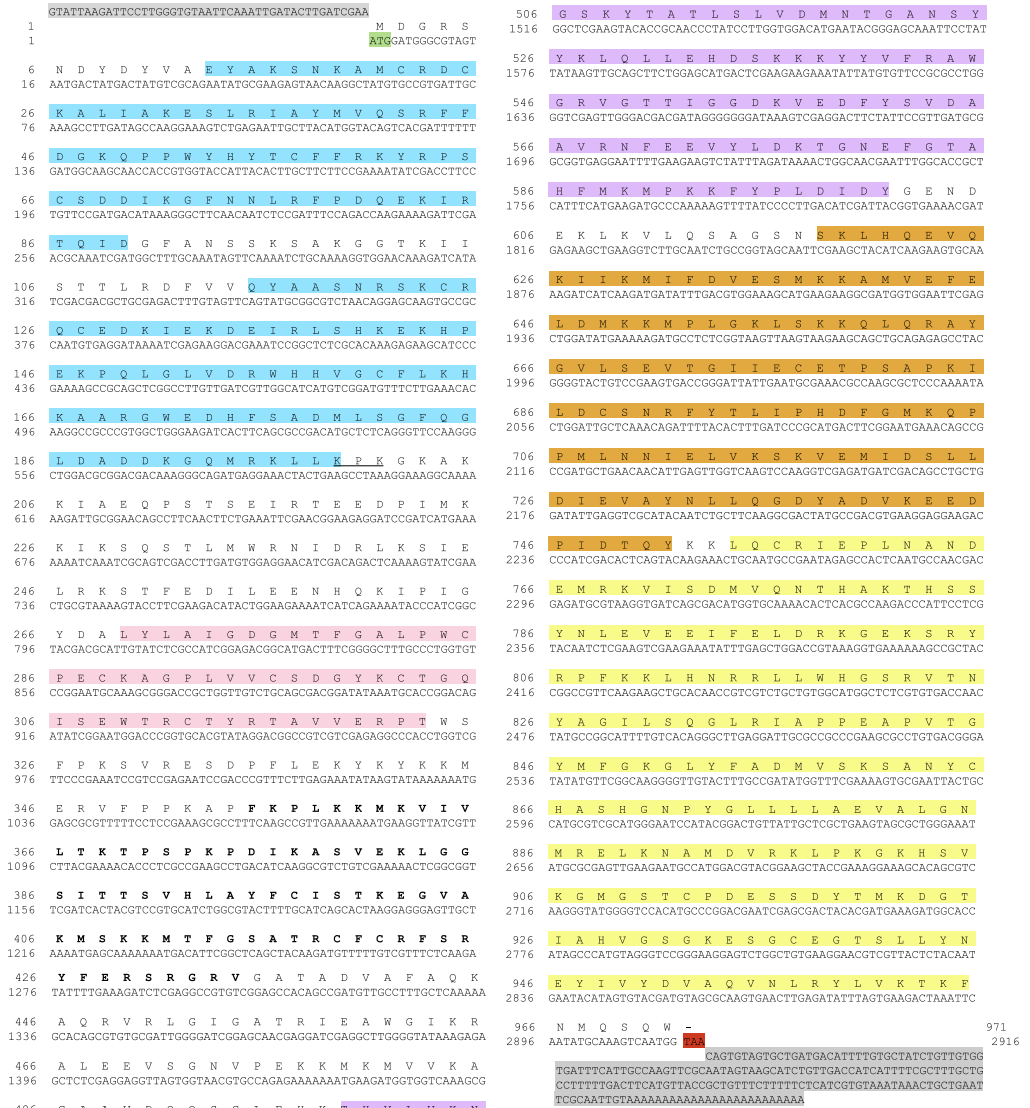
```

**Fig. 4.** cDNA and deduced amino acid sequence of BsAIF1. START and STOP codon are in green and red, respectively; 5' and 3' UTR are in grey. Orange: pyridine nucleotide-disulphide domain (PNDD); yellow: NADH-binding PNDD; light blue: C-terminal domain of dimerisation; light purple: nuclear localisation motif. Bold: the N-terminal transmembrane domain.





**Fig. 5.** Alignments of the deduced amino acid sequence of BsPARP1 with known orthologous sequences from both vertebrate and invertebrates.



**Fig. 6.** cDNA and deduced amino acid sequence of BsPARP1. START and STOP codon are in green and red, respectively; 5' and 3' UTR are in grey. Light blue: N-terminal zinc-finger domains; pink: PDR1 (poly(ADP-ribose)-synthase 1) domain; bold: BRCT (BRCA1 (breast cancer susceptibility protein C-terminus) domain; light purple: WGR (tryptophane-, glycine-, arginine-rich) motif; orange; C-terminal regulatory PARP domain; yellow: C-terminal catalytic PARP domain.



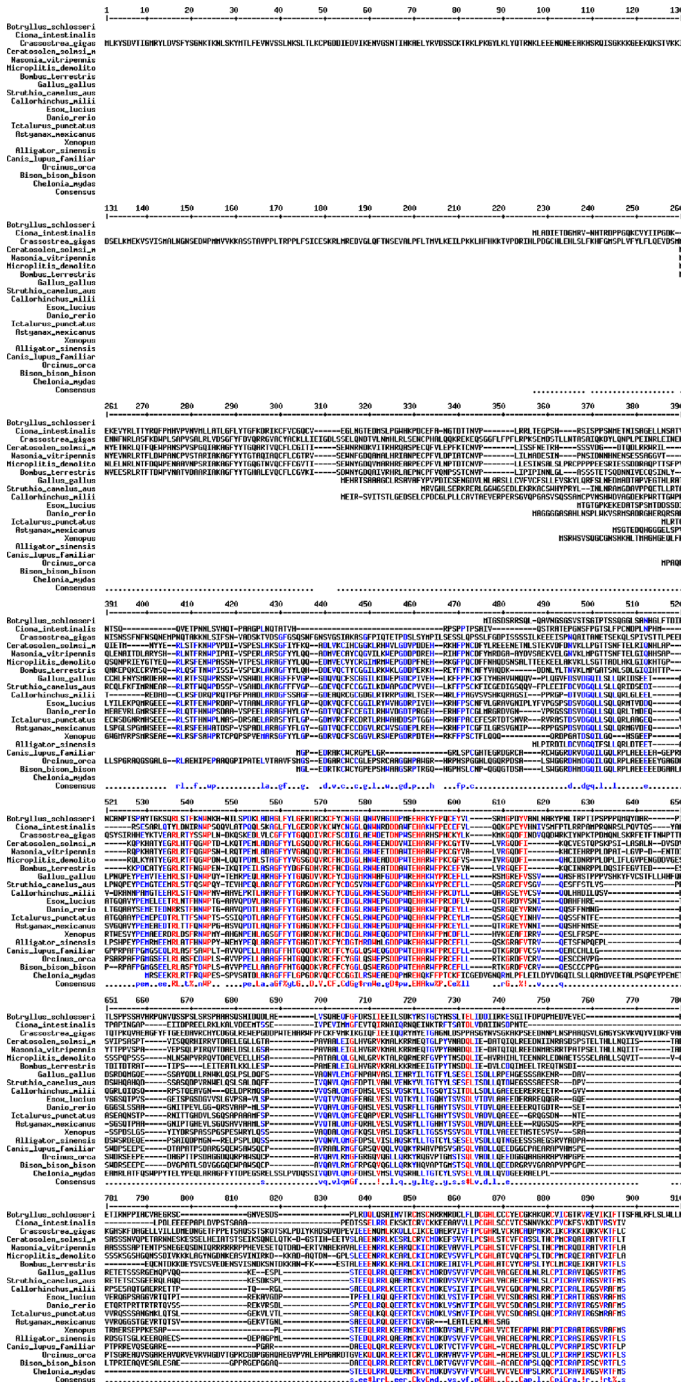


Fig. 7. Alignments of the deduced amino acid sequence of BslA7 with known orthologous sequences from both vertebrate and invertebrates.

```

AAGGTATTCTTTGTACCTATTATATAGTGTATATATGGAGTTGGGACGAGTAGA
ACCGCAGCAAAATACGTGCCAAGTTCTGTATGATAGTTCGCCTTGTGAAACTGGA

1                                     M I
1                                     ATGATT

      G S D S R R S Q L Q A V N G S G S V S T
GGATCGGACAGCAGGAGGTACAGCTTCAAGCTGTGAATGGAAGCGGAAGCGGTGCAACA

23  S G I P T S S Q G G L S A N H G L F T D
67  TCTGGTATACCTACCAGTTCTCAAGGCGGACTATCAGCAAACCATGGTCTCTTCACGGAC

43  I F N C A N P T S P A Y T G K S Q R L S
127 ATCTTCAACTGTGCAAAATCCAACCTTCTCCAGCATATACGGGGAAGTCGCAAAAGGCTTTCA

63  T F K N W N K H N I L S P D K L A D A G
187 ACCTTTAAAAACTGGAAACAAGCACAACTACTGAGCCAGATAAAATAGCCGACGCAGGA

      L F Y L G E R D R C K C F Y C N G G L Q
247 TTGTTTTATTGGGGCAAAGAGACAGATGTAATGCTTTTACTGCAATGGAGGATTGCAG

103 N W V A G D D P M E E H A K Y F P Q C E
327 AATTGGGTGGCAGGCGACGCCAATGGAGGAGCATGCCAAATATTTCCCAATGCGAA

123 Y V L S R M G P D Y V A N L N A R Y P N
387 TACGTTCTGTCAAGAATGGGACCCGATTACGTTGCAAACCTGAATGCAAGATATCCGAAT

143 L T R P T I P S P P P Q M Q Y D A R P I
447 TTAACGCGACCGACGATTCGGTCACCACCCCGCAAATGCAATATGACGCCCGGCAATT

163 T L S P P S S A V H R P Q N V Q S S P S
507 ACGCTGTCACCGCCAGCAGCGCTGTTACAGGCCACAAAACGTGCAATCATCGCCATCA

183 L S R S P A A A S Q S H I D Q D L A E L
567 CTCAGCCGGTCGCCAGCAGCTGCATCACAAGTCATATTGATCAAGACTTGGCGGAATTA

203 V S Q A E Q F G F D R S I I E E I L S D
627 GTATCGCAAGCAGAACAAATTCGGTTTTGATCGATCAATTATTGAGGAAATATTATCGGAT

223 K Y R S T G C Y H S S L T E L I D D I I
687 AAGTACAGTCAACGGGTTGTTATCACAGCAGCCTCAGAACCTATTGATGACATCAAT

243 R K E S G I T F D P Q P M E D V E V E C
747 AGAAAAGAAGTGGGATTACCTTCGATCCCCAGCCAATGGAAGACGTCGAAGTCGAATGT

263 H E T I R N P P I A C V A E G R S C G N
807 CATGAACAATACGGAATCCACCGATAGCATGCGTGGCAGAGGTAGATCCTGTGGTAA

283 V E S D S P L R D Q L Q S A I N V T R C
867 GTAGAAAGCGACTCGCCGCTACGCGATCAGCTGCAATCAGCAATAAACGTAACCCGTTGT

303 M S C N R R N R D C L F L D C G H L C C
927 ATGTCATGTAACAGGCGGAATCGTGATTGCTTATTCTTAGATTGCGGTCACTTGTGCTGC

323 C Y E C G K A K Q R C V I C G T R V R E
987 TGTTACGAATGCGGGAAAGCAAAGCAAAGATGCGTGATTGTGGGACCCGGGTTAGGGAA

343 V I K I F T T S F A L R F L S L W L L N
1047 GTGATAAAATATTTACGACATCTTTCGCATTAAGGTTCCCTTCACTGTGGCTTCTAAAC

363 Y C N -
1107 TACTGCAAT TAA
      ATATAGTTATCTTCAAGATCTTATTTTCGTTTACTATAATGAATAAA
      CATTCTGTGTCTTGATAAAAAAAA
    
```

Fig. 8. cDNA and deduced amino acid sequence of BslAP7. START and STOP codon are in green and red, respectively; 5' and 3' UTR are in grey. Orange: BIR domain; blue: RING (really interesting new gene) finger domain.

## Acknowledgements

This research was supported by grants from the CARIPARO Foundation (Excellence Project 2009) and the Italian M.I.U.R (P.R.I.N. grant 20109XZEPR to LB).

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.05.017>.

## Appendix B. Transparency Document

Transparency data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.05.017>.

## References

- [1] N. Franchi, F. Ballin, L. Manni, F. Schiavon, G. Basso, L. Ballarin, Recurrent phagocytosis-induced apoptosis in the cyclical generation change of the compound ascidian *Botryllus schlosseri*, *Dev. Comp. Immunol.* 62 (2016) 8–16.
- [2] D. Campagna, F. Gasparini, N. Franchi, N. Vitulo, F. Ballin, L. Manni, G. Valle, L. Ballarin, Transcriptome dynamics in the asexual cycle of the chordate *Botryllus schlosseri*, *BMC Genomics* 17 (2016) 275.
- [3] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.* 32 (2004) 1792–1797.
- [4] X. Huang, W. Miller, A time-efficient linear-space local similarity algorithm, *Adv. Appl. Math* 12 (1991) 337–357.