UC Irvine

UC Irvine Previously Published Works

Title

Characterization of a Cu/Zn superoxide dismutase-encoding gene region in *Drosophila willistoni*.

Permalink

<https://escholarship.org/uc/item/8f7515kv>

Journal

Gene, 147(2)

ISSN

0378-1119

Authors

Kwiatowski, J
Latorre, A
Skarecky, D
et al.

Publication Date

1994-09-01

DOI

10.1016/0378-1119(94)90085-x

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

GENE 08177

Characterization of a Cu/Zn superoxide dismutase-encoding gene region in *Drosophila willistoni**

(*Drosophila virilis*; gene homology; gene structure; polyadenylation signal; *Sophophora*)

Jan Kwiatowski^{a,b}, Amparo Latorre^{a,c}, Douglas Skarecky^a and Francisco J. Ayala^a

^aDepartment of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717, USA; ^bInstitute of Botany, Warsaw University, 00-478 Warsaw, Poland; and ^cDepartment of Genetics, University of Valencia, 46100 Burjassot, Spain

Received by J.A. Engler: 21 December 1993; Revised/Accepted: 1 March/2 March 1994; Received at publishers: 13 June 1994

SUMMARY

A Cu/Zn superoxide dismutase-encoding gene (*Sod*) from *Drosophila willistoni* was cloned and sequenced. The gene shows a typical structure for a fruit-fly *Sod* gene, with a coding region of 462 bp in two exons separated by a 417-bp intron. Comparison of the *Sod* sequences from *D. willistoni* and *D. melanogaster* suggests that these species are only remotely related. Downstream from the *Sod* gene, there is an ORF on the opposite strand that putatively encodes the last exon of an unidentified gene. The polyadenylation signals of the two genes are separated by only 61 bp in *D. willistoni*, conforming to the common picture of compact dipteran genomes.

A 1472-bp fragment containing Cu/Zn *Sod* was isolated from a genomic library of *D. willistoni* by cross-hybridization with Cu/Zn *Sod* cDNA from *D. melanogaster* and sequenced on both strands. The coding region (Fig. 1) of the gene consists of 462 bp, which is the same length previously obtained for other fruit-fly species (Kwiatowski et al., 1992a). The coding region is separated into two exons by a 417-bp intron, the shortest among the *Sod* genes sequenced so far in Diptera (Kwiatowski et al., 1992a). Downstream of the coding region a putative polyadenylation signal AATAAA starts at nt 1045. The distances between known *Drosophila Sod* sequences are consistent with the observation made on the basis of rRNA (Pelandakis et al., 1991) that *D. willistoni* is remotely related to *D. melanogaster*, and other species of

the *Sophophora* subgenus. The coding nt as well as aa differences between *D. willistoni* and *D. melanogaster* or *D. simulans* are larger than those between *D. willistoni* and *D. virilis*, a species of the subgenus *Drosophila* (Table I).

Comparison of a *D. willistoni Sod* gene region with a corresponding *D. virilis* sequence (Kwiatowski and Ayala, 1989) reveals a conserved fragment of 120 bp downstream from *Sod*, starting at nt 1258 (Fig. 1). This sequence can be identified as an ORF on the complementary strand to *Sod*. A search for similar sequences in the EMBL and PIR databases does not reveal any homologous genes or proteins. The unidentified sequence is flanked by a TAA stop triplet and an AG-3' consensus splicing site (Breathnach and Chambon, 1981), suggesting that it is the last exon of a gene. A putative polyadenylation signal is present on the complementary strand separated by only 61 nt from the one belonging to *Sod*. Since the poly(A) tail of mRNA usually begins 10–30 nt after the polyadenylation signal (Wahle, 1992), the transcribed portions of the two genes in *D. willistoni* must lay in a very close proximity, perhaps overlapping each other. In *D. virilis* the distance between the two genes is slightly larger, but

Correspondence to: Dr. J. Kwiatowski, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717, USA. Tel. (1-714) 856-8296; Fax (1-714) 725-2181; e-mail: jmkwiato@uci.edu
*On request, the authors will supply detailed experimental evidence for the conclusions reached in this Brief Note.

Abbreviations: aa, amino acid(s); bp, base pair(s); cDNA, DNA complementary to RNA; nt, nucleotide(s); ORF, open reading frame; rRNA, ribosomal RNA; Sod, superoxide dismutase(s); *Sod*, gene encoding Sod.

gatcaacaaaatgggtggttaaagctgtttgcgttattaatggcgatgccaaagggtaccgt 60
 m v v k a v c v i n g d a k g t v 17
 tttcttggagcaggaggaagtgaaataacagcaaatggggaataaatgtgaatgtgtt 120
 f f e q e 22
 ttaggaggcaaatgacgtttctcaaggctcagtcacaaagtacttagagcgagtgagagc 180
 cggcttcggcgggggcccttttccgtcggttggtccacattagaataatcaaatgagcc 240
 gagagcaattgctgcggtttcgttttgcgtgtcatcggttaaaagcagttatataagg 300
 cacatacatcac 360
 gattaggcatagacttttgggttatcatgaagcaccctccaccactgggtttactatc 420
 tggtttcttagctcccccttttacacatctcgtatctctctctctctctctctctctct 480
 catttacttttaggacaatggagctcccgctcaaaagtacccggcgaggtgacccggcctgg 540
 d n g a p v k v t g e v t g l g 38
 caaaggactgcatggtttccatgtccatgaatttggcgacacacacacacacacacacac 600
 k g l h g f h v h e f g d n t n g c m s 58
 ctctggccctcatttcaatccacatagcaggagcaggtggcgccctggagatgagaatcg 660
 s g p h f n p h s k e h g a p g d e n r 78
 tcatttggcgcatgttggcacaattgaggtcagtcagtcagtcagtcagtcagtcagtc 720
 h l g d l g n i e a s g s g p t a v n i 98
 caccgattccaaatcaccctagtcggtgccaatagatcattggagcagcaggtgtgtct 780
 t d s k l t l v g a n s i i g r t v v v 118
 ccattgcatcagcatgtatctggcgaagggtggacacagcagtcagtcagtcagtcag 840
 h a d p d d l g k g g h e l s k t t g n 138
 tgctgagctcgcattggttgcgtgtcattggcagtcagtcagtcagtcagtcagtcag 900
 a g a r i g c g v i g i a k l * 153
 ccctaagactatatacatatcatctatttctatatgtatcttctactgaattac 960
 acgatatacatatgtattgttatatcatcttctacttctgtgtctgttagttta 1020
 acaaaaaaaacctttaaggttgaatagtggttattgtgagtttttttttcaatac 1080
 tggctcctaacccgattgcagttgcgttttttttttttttttttttttttttttttt 1140
 aaatgtttgaaactatggtctctcgaactacgaacaaatgtgaaacgtaaatttttc 1200
 tacttgggtttttttcttagtatttgggtgcgttttttttttttttttttttttttt 1260
 ggtcttgtttaaatacagattgtagacaaatgccagtggaagaggtggcacaataatgt 1320
 t k n f y s q y v i g l s s t a f l t f
 ataacaccacaggtggcgaagcccaaaaatcctttgtatccctctgtgtgagaacac 1380
 y c w w t a f g l f g k y v e t s f v
 gaagagaataatgtagcagtgtagcaattttctattttgtcttctaaaggcatgtatct 1440
 acatttttcagactcgtaaatgaggcctgcag 1472

Fig. 1. The nt and deduced aa sequences of *Sod* and an unidentified gene. Canonical splicing pairs are underlined; polyadenylation signals are double underlined. Stop triplets are marked by asterisks. The sequence has been submitted to GenBank/EMBL under accession No. L13281.

TABLE I

The aa (below diagonal) and nt (above diagonal) percent identities between Cu/Zn superoxide Sod sequences of *Drosophila* species

Species ^a	Dw	Dm	Ds	Dv
Dw	—	79.9	80.1	81.0
Dm	87.6	—	97.6	81.4
Ds	87.6	100.0	—	81.0
Dv	90.2	86.9	86.9	—

^a Dw, *D. willistoni*; Dm, *D. melanogaster* (Kwiatowski et al., 1989b); Ds, *D. simulans* (Kwiatowski et al., 1989a); Dv, *D. virilis* (Kwiatowski and Ayala, 1989). *Sod* are 462 nt and Sod are 153-aa long.

still very short, since the putative polyadenylation signals are separated by only 80 bp.

A similar situation occurs in other fruit-fly species, where the distance between the AATAAA canonical

sequences, on opposite strands, are 55 bp in *Chymomyza amoena* (Kwiatowski et al., 1992b) and 93 bp in *Ceratitis capitata* (Kwiatowski et al., 1992a). In *Chymomyza* but not in *Ceratitis* there is an ORF homologous to the one found in *D. virilis* and *D. willistoni*. The polyadenylation signal in *Ceratitis* may thus come from another gene. The close proximity of *Sod* and the unidentified gene downstream on the complementary strand, together with the generally short introns encountered in dipterans (Hawkins, 1988; Kwiatowski et al., 1992a), suggest that there is evolutionary pressure towards eliminating unnecessary DNA load in these organisms.

ACKNOWLEDGEMENTS

We thank Drs. Cort Anderson and Jeff Powell (Yale University) for the genomic library of *D. willistoni*, and Dr. Gordon Tener (University of British Columbia) for the Cu/Zn *Sod* cDNA clone of *D. melanogaster*. Research supported by NIH grant GM42397.

REFERENCES

- Breathnach, R. and Chambon, P.: Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50 (1981) 349–383.
- Hawkins, J.D.: A survey on intron and exon lengths. *Nucleic Acids Res.* 16 (1988) 9893–9908.
- Kwiatowski, J. and Ayala, F.J.: *Drosophila virilis* Cu,Zn superoxide dismutase gene sequence. *Nucleic Acids Res.* 17 (1989) 2133.
- Kwiatowski, J., Gonzalez, F. and Ayala, F.J.: *Drosophila simulans* Cu,Zn superoxide dismutase gene sequence. *Nucleic Acids Res.* 17 (1989a) 6735.
- Kwiatowski, J., Patel, M. and Ayala, F.J.: *Drosophila melanogaster* Cu,Zn superoxide dismutase gene sequence. *Nucleic Acids Res.* 17 (1989b) 1264.
- Kwiatowski, J., Skarecky, D. and Ayala, F.J.: Structure and sequence of the Cu,Zn *Sod* gene in the Mediterranean fruit-fly *Ceratitis capitata*: intron insertion and deletion in the evolution of the SOD gene. *Mol. Phyl. Evol.* 1 (1992a) 72–82.
- Kwiatowski, J., Skarecky, D., Burgos, M. and Ayala, F.J.: Structure and sequence of the Cu,Zn superoxide dismutase gene of *Chymomyza amoena*: phylogeny of the genus and codon-use evolution. *Insect Mol. Biol.* 1 (1992b) 3–13.
- Pelandakis, M., Higgins, D.H. and Solignac, M.: Molecular phylogeny of the subgenus *Sophophora* of *Drosophila* derived from large subunit of ribosomal RNA sequences. *Genetica* 84 (1991) 87–94.
- Wahle, E.: The end of the message: 3'-end processing leading to polyadenylated messenger RNA. *BioEssays* 14 (1992) 113–118.