## 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, Jan Latorre, Amparo Skarecky, Douglas <u>et al.</u>

### **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 <a href="https://creativecommons.org/licenses/by/4.0/">https://creativecommons.org/licenses/by/4.0/</a>

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<sup>a,b</sup>, Amparo Latorre<sup>a,c</sup>, Douglas Skarecky<sup>a</sup> and Francisco J. Ayala<sup>a</sup>

<sup>a</sup>Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717, USA; <sup>b</sup>Institute of Botany, Warsaw University, 00-478 Warsaw, Poland: and <sup>c</sup>Department 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 crosshybridization 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.

$ \begin{array}{cccc} \texttt{gatcaacaaaatggtggttaaagctgtttgcgttattaatggcqatgccaagggtaccgt} \\ \texttt{m} & \texttt{v} & \texttt{k} & \texttt{a} & \texttt{v} & \texttt{v} & \texttt{i} & \texttt{n} & \texttt{g} & \texttt{a} & \texttt{k} & \texttt{g} & \texttt{t} & \texttt{v} \end{array} $	60 17
tttctttgagcaggag <u>gt</u> aàgtgaaataacagcaaatggggaataaatgigaaatgtgti f f e q e	120 22
ttaggaggcaaattgcagtttetcaaggtcagtccaaagtaettagaggcgagtgagagc	180
cggcttcggcggggccccttttccgtcggttggtcccacattagaaattcaaattgagcc	240
gagagcaattgctgcggtticgtttttgctgtgtcatcgttaaaagcagttatataaggg	300
cacatacatacacacacacatacaccgatgatgatgatgataaaaactacatagacataat	360
gattaggcatagacttttgttggcttatcatgaagcacccttcaccactggtttactatc	420
tgttttcttagetccccctittacacatcicgctatctctctctctctctctctctct	480
$ \begin{array}{c} {}_{catttacttitagqacaatggaqctcccgtcaaagttaccggcgaggtgaccggcctggg} \\ {}_{d \ \pi \ g \ a \ p \ v \ k \ v \ t \ g \ e \ v \ t \ g \ 1 \ g \ } \end{array} $	540 38
caaaggactgcatggtttccatgtccatgaatttggcgacaacaccaacggatgcatgtc $k = 1$ h g f h v h a f g d n t n g c m s	600 58
$\begin{array}{c} ctctggcccccatttcaatccacatgcaaggagcatggcgcccctggagatgagaatcg\\ s  g  p  h  f  n  p  h  s  k  e  h  g  a  p  g  d  e  n  r \end{array}$	660 78
tcatttgggcqatttgggcaacattgaggcatctggcagtggtccaactgctgttaatat $h \ 1 \ g \ d \ 1 \ g \ n \ i \ e \ a \ s \ g \ s \ g \ p \ t \ a \ v \ n \ i$	720 98
$\begin{array}{c} caccgattccaaaatcacccctagtcggtgccaatagtatcattggacgcaccgttgtcgt\\ \texttt{t \ \texttt{d} \ \texttt{s} \ \texttt{k} \ \texttt{i} \ \texttt{t} \ \texttt{l} \ \texttt{v} \ \texttt{g} \ \texttt{a} \ \texttt{n} \ \texttt{s} \ \texttt{i} \ \texttt{i} \ \texttt{g} \ \texttt{r} \ \texttt{t} \ \texttt{v} \ \texttt{v} \ \texttt{v} \end{array}$	780 118
ccatgccgatcccgatgatctgggcaagggtggacacgagctgagcaagacgactggtaa h a d p d d l g k g g h e l s k t t g n	840 138
tgctggagetcgcattggtgggggtgtcattggcattgccaagatctaagaatgaggagt $f$	900 153
ccctaagaciatatacatacatacatctaittctatatgiatcttatctactgaaattac	960
acgatatacacatatgtatigtatatcatitatatctactacttgtggtctgttagttta	1020
acaaaaaaaaacctttaaggttga <u>aataaa</u> tgggtatttgtgagtttttttttcaatat	1080
tggctcctaaccgattcgcagttgcgttttt <u>tttatt</u> acaaatttataaacaatatatac	1140
aaatgtttgaactatggcttctcgactacgaaacaatttgaaacgtaaaattttctaagt	1200
tacttgggtitttttctttågtatttggtgcgctttttaicatttgagtittgcttcttå	1260
ggtettgttaaaataegatigatagacaatgeecaaggagggggeaaataatgtaaa t k n f y s q y v i g l s s t a f l t f	1320
ataacaccaccaggtggcaaagcccaaaatcetttgtatacctetgtggagaacacc <u>tg</u> $y \in w w t a f g l f g k y v e t s f v$	1380
gaagagaataatgtagcagttagcaatttettattttggtettetaaaggeatgtatett	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*	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	

<sup>a</sup> 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 (Yalc 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.