Biologia (2023) 78:2517–2526
https://doi.org/10.1007/s11756-023-01389-4

ORIGINAL ARTICLE

Identification and characterization of *globin* gene from *Bombus terrestris* (Hymenoptera: Apocrita: Apidae)

Habeş Bilal Aydemir[1] · **Ertan Mahir Korkmaz**[2]

Received: 5 October 2022 / Accepted: 7 March 2023 / Published online: 29 March 2023

Abstract
Hemoglobins are mainly functional in the transport and storage of O_2, and are present in aerobic, lung respiratory metazoans. Based on recent studies, it has been determined that globin-like proteins are also found in invertebrate groups which have tracheal respiratory systems. Thus, it has been claimed that the globin molecules may have housekeeping functions beyond their known functions. In this study, the presence of a hemoglobin gene was investigated in a natural pollinator, *Bombus terrestris*. *Bombus terrestris* (Hymenoptera: Apidae: Bombinae) is an ecologically and economically important species, it has a wide distribution in high and low altitude habitats, can live in different localities in terms of altitude and temperature. The expression of *globin* gene was determined using real time qPCR and RACE system, and confirmed by SRA data. The potential *globin* ORF sequence was 516 bp nucleotide and 171 bp amino acid long. This data was deposited in the GenBank database under accession number OP312897. On the other hand, expression profiles were investigated in different hypoxic conditions at different developmental stages (larva, pupae and adult) and in different tissue types (abdomen, thorax, head and legs). It was determined that the *globin* gene is transcribed in all developmental stages, but only in abdomen; and that globin expression decreases with decreasing amount of oxygen.

Keywords *Bombus terrestris* · qPCR · RACE system · *Globin* · Gene expression

Introduction

Many aerobic organisms need a continuous and adequate supply of oxygen (O_2). For most animals, a successful gas exchange mechanism is critical to their survival, requiring adequate O_2 to be transported to the internal tissues. Hemoglobins, small heme-proteins, function in the transport and storage of O_2, and contribute to aerobic metabolism. Hemoglobins are either monomeric or polymeric proteins, and they typically consist of a subunit of approximately 150 amino acids containing eight α-helical segments called A-H helices, forming the 3-over-3 α-helical sandwich structure (Weatherall 1983). Binding of O_2 is mediated by a porphyrin group and an iron ion coordinated with the proximal histidine located at helix position F8 (i.e. amino acid 8 of Helix).

Although hemoglobin is common among metazoan taxa, these proteins have been considered unnecessary in insects, mainly due to their use of the tracheal system (Hardison 1996; Weber and Vinogradov 2001). Only a few insect species living in low oxygen conditions, as well as in aquatic environments, have been reported to have hemoglobins (Weber and Vinogradov 2001). In later studies, three hemoglobin-like genes (glob1, glob2 and glob3) have been characterized in *Drosophila melanogaster* (Fallén, 1823) (Burmester and Hankeln 1999; Hankeln et al. 2002; Burmester et al. 2006). The presence of multiple *Hb* genes in *Drosophila* suggests that this type of respiratory proteins may arise and be functional in insects even though these animals do not encounter markedly severe hypoxic (oxygen-poor) conditions during their life cycle. In addition, the presence of an Hb orthologous to *D. melanogaster glob1* and its expression in similar tissues have also been detected in the honey bee

✉ Habeş Bilal Aydemir
hbilalaydemir@hotmail.com

[1] Faculty of Science and Letters, Department of Molecular Biology and Genetics, Tokat Gaziosmanpaşa University, 60250 Tokat, Turkey

[2] Faculty of Science, Department of Molecular Biology and Genetics, Sivas Cumhuriyet University, 58140 Sivas, Turkey

Apis mellifera (Linnaeus, 1758) (Hankeln et al. 2006). This finding also suggests that Hbs may be more widely distributed among insects than previously thought. Remarkably, Chironomid midges, the only free-living insect group, can possess extracellular monomeric and dimeric hemoglobins (Nath 2018). Hemoglobin-like genes have also been characterized in some other insect species including *Anopheles gambiae* (Giles, 1902), *Aedes aegypti* (Linnaeus, 1762), *Bombyx mori* (Linnaeus, 1758), and *Samia cynthia* (Drury, 1773) (Burmester and Hankeln 2007; Kawaoka et al. 2009). Although the amino acid sequences of these hemoglobin genes are very diverse, the amino acids that are essential for heme and oxygen binding are conserved. Recently identified insect hemoglobins suggest that they are intracellular proteins and may be involved in respiration in relation to the tracheal system, and these findings point that *Hb* genes actually belong to the standard repertoire genes of insects and may have conserved functions (Burmester and Hankeln 2007; Burmester 2015).

In many vertebrate and invertebrate species, globins mediate the transport of O_2 in the blood and also increase the amount of O_2 available to muscle and other tissues (Weatherall 1983; Weber and Vinogradov 2001; Burmester and Hankeln 2014). However, some globins may also have other or additional functions: they can detoxify reactive oxygen species (ROS) (Flogel et al. 2004; Koch and Burmester 2016), can adjust nitric oxide (NO) turnover (Flogel et al. 2001; Hendgen-Cotta et al. 2008), may regulate O_2 homeostasis (Yadav et al. 2015; Gleixner et al. 2016) or play a role in cellular signal transduction (De Henau et al. 2015).

Bombus terrestris (Linnaeus, 1758) (Hymenoptera: Apidae: Bombinae) is an ecologically and economically important species, it has a wide distribution in high and low altitude habitats, can live in different localities in terms of altitude and temperature. In this study, we investigated the presence of a functional *globin* gene in the mechanism of oxygen transport and storage in this natural pollinator, *B. terrestris*.

Material and method

Specimen collections and RNA isolation

A colony of *Bombus terrestris* was obtained from Koppert Biological Systems (Antalya, Turkey) The colony was incubated until the number of individuals doubled, and then divided into three main groups as larvae, pupae and adults (Fig. 1). After 15 additional incubation days, adults were exposed to the hypoxic conditions for 1 to 10 h, gradiently. They were then separated into different tissue groups (abdomen, thorax, head and legs) and stored in RNAlater at -80 °C. Total RNA was extracted from each sample by using RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Each RNA pellet was dissolved in DEPC-treated H_2O for further analysis. Quantity and quality of total RNA were checked by Nanodrop (MaestroGen Inc.) and agarose gel electrophoresis, respectively. The RNA samples were stored at -80° C until use in RT-PCR.

RT-PCR, real-time PCR and RACE-system analysis

cDNA fragments were obtained via RT-PCR experiments using SuperScript™ IV First-Strand Synthesis System (Invitrogen) according to the manufacturer's instructions. Specific primers were designed manually according to the predicted *B. terrestris globin* (XM_003396782) and *Apis mellifera globin* (NM_001077823) references (Online Resource 1). All amplicons obtained were sequenced by a commercial sequencing service (Macrogen, Netherlands). Sequences were individually assembled, aligned using Geneious R9 and checked via BLAST homology search to detect *globin* sequences (Kearse et al. 2012). Real-time PCR was carried out with 2X SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad), and the amplification was detected using an CFX Connect™ Real-Time PCR Detection according to the manufacturer's protocol, with GLBRTF1 and GLBRTR4 primers (Online Resource 1). Housekeeping GAPDH gene was used for validation of real-time reactions. The 5` and 3` ends of the *B. terrestris* globin-like cDNA were obtained using the FirstChoice® RLM-RACE Kit (Invitrogen), and sequenced after cloning into the pCR®4-TOPO® vector (Invitrogen). All sequences from RACE were assembled and annotated using Geneious R9 and checked via BLAST (Kearse et al. 2012).

In addition to Real-Time PCR, the expression profile of *globin* gene was investigated with RNA-seq transcriptome datasets (SRR5614828, SRR5614826, SRR5125109, SRR5068604, SRR3929964, SRR329387) using "map to reference" and "calculate expression level" options in Geneious R9 (Kearse et al. 2012).

Characterization of putative *globin* gene and protein

Nucleotide sequence of candidate, putative *globin* gene of *B. terrestris* was arranged using RACE and transcriptome analysis results. The open reading frames were identified by the ORF Finder (www.ncbi.nlm.nih.gov/gorf/gorf.html). Amino acid composition was determined, and relatively synonymous codon usage bias (RSCU) was calculated using Geneious R9 (Kearse et al. 2012). The control/ regulator regions were determined using Sequence Manipulation Suite (Stothard 2000) and Methprimer (Li and Dahiya 2002).

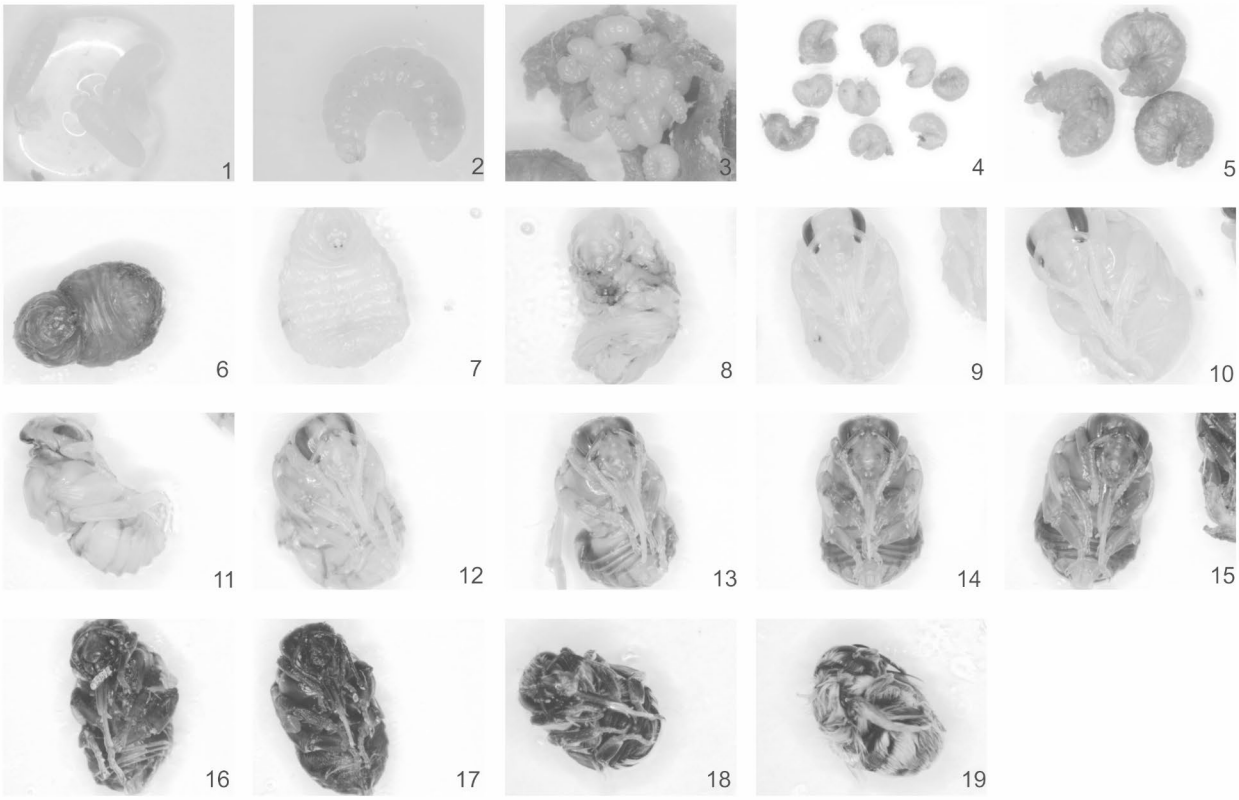


Fig. 1 Developmental stages of *B. terrestris* used in this study. Images 1–7 show the larval stage at 2-day intervals, images 8–18 show the pupal stage at 2-day intervals, image 19 shows the adult stage

The putative signal peptide sequences and their localization were determined via PredictProtein and Signal- BLAST/ Signal- 3 L programs, respectively. Cellular localization of putative protein was determined via PROTTER (Omasits et al. 2014) and DeepLoc 2.0 (Thumuluri et al. 2022). 1D Protein Structure Prediction Server (Chen and Kurgan 2007) and PSIPRED v3.3 (Buchan et al. 2013) programs were used to determine whether the protein folding reaction can occur spontaneously and to predict secondary structures of proteins (Online Resource 2). Among the program estimates, those with the lowest Gibbs free energy values (ΔG) were accepted as the candidate *globin* protein structures. Molecular weight, isoelectric point (pI) and Grand average of hydropathicity index (GRAVY value) of putative protein were calculated via Sequence Manipulation Suite (Stothard 2000). Tertiary structure of the protein was determined by Cn3D v4.3.1 (Wang et al. 2000) using defined globin proteins of *D. melanogaster* (NP_001287343) and *A. mellifera* (NP_001071291) as reference structures. Residual interactions of tertiary structure were estimated by ISIS (interaction sites identified from sequence) method in PredictProtein (Ofran and Rost 2007).

Comparative analysis of *globin* genes in insects

A sampling dataset of *globin* gene was manually retrieved in order to compare and to be used as reference sequence. *D. melanogaster* (NM_001300414), *A. mellifera* (NM_001077823), *A. gambiae* (AM182453), *Chironomus thummi thummi* (Kieffer, 1973) (M17691) sequences were downloaded from GenBank (http://www.ncbi.nlm.nih.gov) database. Amino acid sequences were aligned using CLUSTAL W with IUB weight matrix to determine homology (Fig. 2). Amino acid sequences were aligned to determine conserved amino acid positions with "Grantham distances" and "D value" estimates (Grantham 1974). If the D value is less than 100, the amino acid changes of the protein were considered as conserved, otherwise as a radical change.

Gene saturation was determined based on the values of Iss and Iss.c using DAMBE (Xia 2001), whereby Iss ≥ Iss.c indicates saturation and vice versa. Transitional and transversional mutation rates of gene were calculated in GTR + G model.

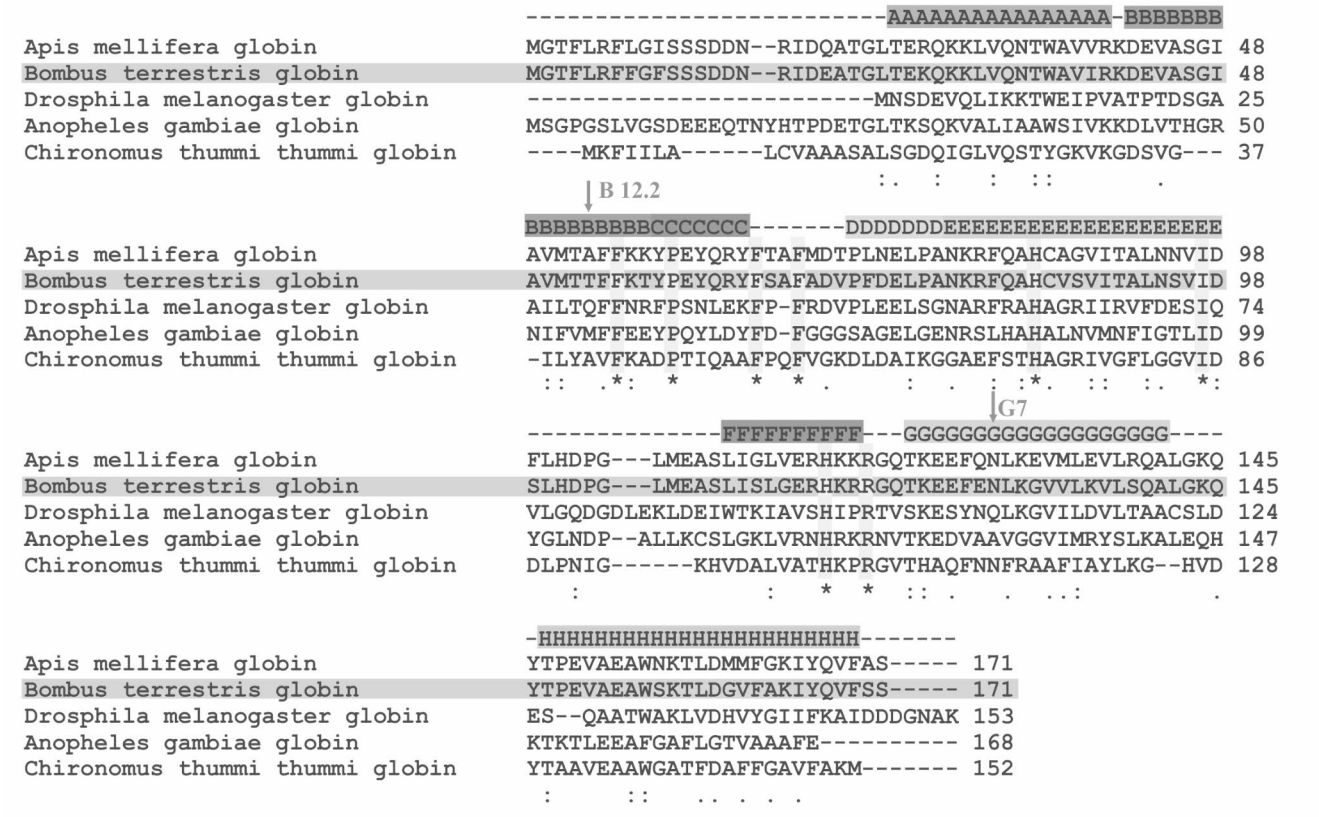


Fig. 2 Alignment of amino acid sequences. Conserved amino acids were marked

Results and discussion

cDNA sequences of potential *globin* gene were found to exhibit 99% homology with *B. terrestris* cytoglobin-2-like sequence (XM_003396785). Thus, it was determined that the *globin* gene is transcribed in *B. terrestris*, and in terms of developmental stages and body parts, the gene is transcribed in all developmental stages (larva, pupa and adult) but only in abdomen (Fig. 3). This is consistent with the previous studies showing that it is expressed in tissues such as the malpighian tube, testis, and a large part of the tracheal system located in the abdomen (Kawaoka et al. 2009) and it supports the existence of an orthologous *globin* gene. *B. terrestris* RNASeq data obtained from the SRA database also showed that the *globin* gene is expressed and has a significant RPKM value. Although globin transcript was not detected in the head samples in our study, in RNASeq analysis the gene seems to be expressed in head in the data with accession numbers SRR5125109 and SRR329387. This suggests that there may be neuroglobin-like proteins in *B. terrestris*.

It was determined in real- time PCR analysis that the *globin* gene is expressed albeit relatively low levels in all stress conditions. It may be explained by the fact that the globin protein in cytoplasm is present in hexaform which provides long half-life for proteins (Trent and Hargrove 2002). In addition, since the hexaform increases the thermal and pH stability of the protein structure, globin protein can have high durability in extreme conditions (Hamdane et al. 2003; Picotti et al. 2009). On the other hand, in expression studies performed in hypoxia environments, it was determined that globin expression decreases with decreasing amounts of oxygen (Fig. 4), similar to that observed in *C. tentans* and *D. melanogaster*. These findings suggest that globin protein may play a role in thermoregulation and/or oxygen transport and storage in this species (Lee et al. 2006; Gleixner et al. 2008; Herhold et al. 2020).

As a result of following RACE system analysis, a total of 1055 bp nucleotide sequences of the *globin* gene were obtained, 300 bp of which were obtained from 5` RACE and 755 bp were obtained from 3` RACE (Fig. 5). This data was deposited in the GenBank database under accession number OP312897. There was a potential transcription signal with high GC% content (≥ 50%) from 41st to 191st position. This site was evaluated for the presence of potential CpG islands. ORF was determined in the positions of 231st to 744th. This

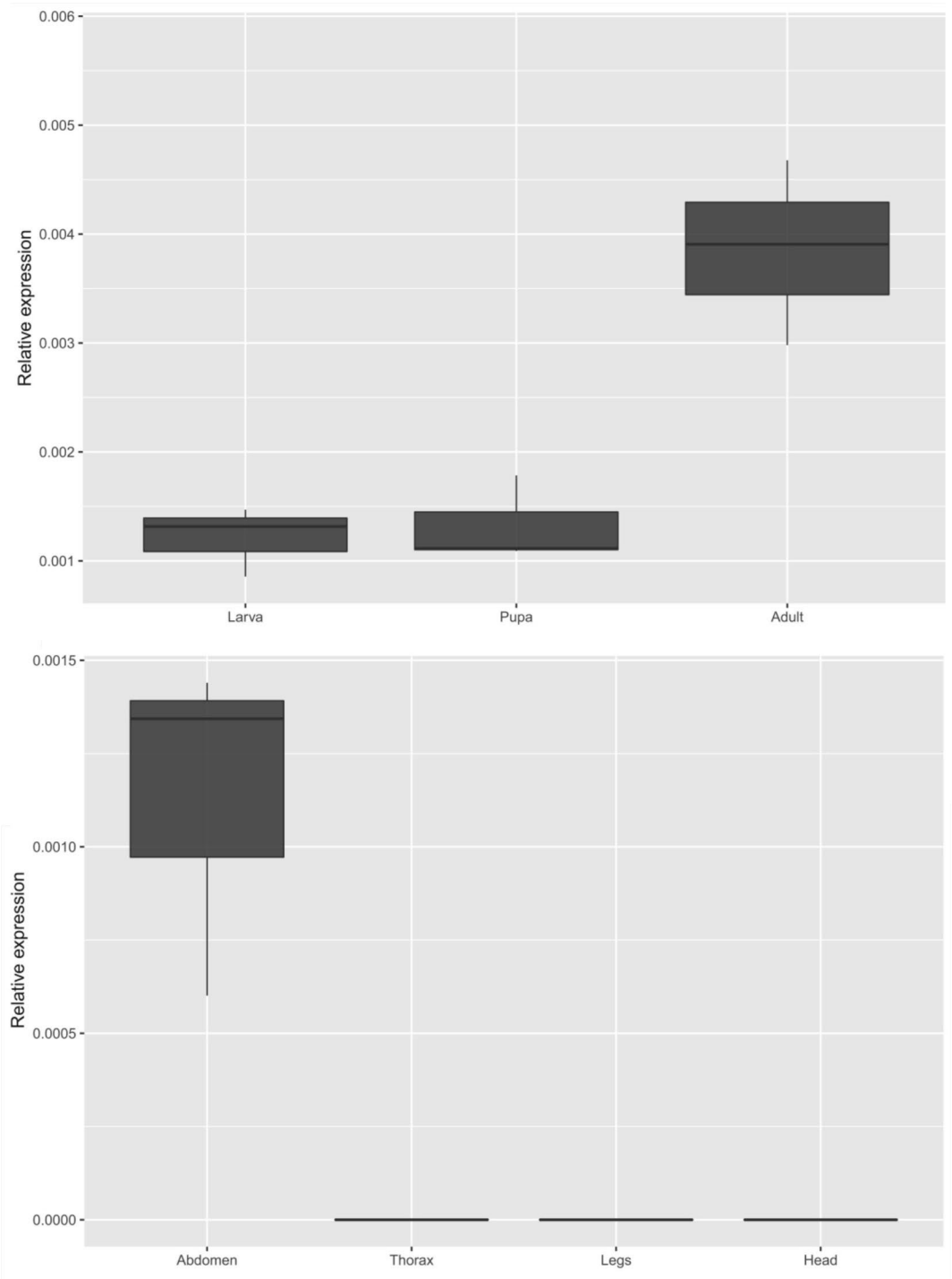


Fig. 3 The pattern of *globin* gene expression in developmental stages and body parts of *B. terrestris*

potential globin ORF sequence was 516 bp nucleotide and 171 bp amino acid long, similar to that in *A. mellifera* (Hankeln et al. 2006). The ORF started with AUG (methionine) and stopped with UGA codons. Alanine, valine, leucine and serine were the most represented amino acids (8.19%) in potential protein sequences; on the other hand, cysteine and tryptophan amino acids were the least common ones with the average of 0.58% and 1.17% percent, respectively. RSCU table of putative protein was given in Table 1. Thus, the presence of 5` and 3` regulatory sequences in the *globin* gene indicates the functionality of this gene in *B. terrestris*. However, further studies with protein-based analyses are needed to support our findings.

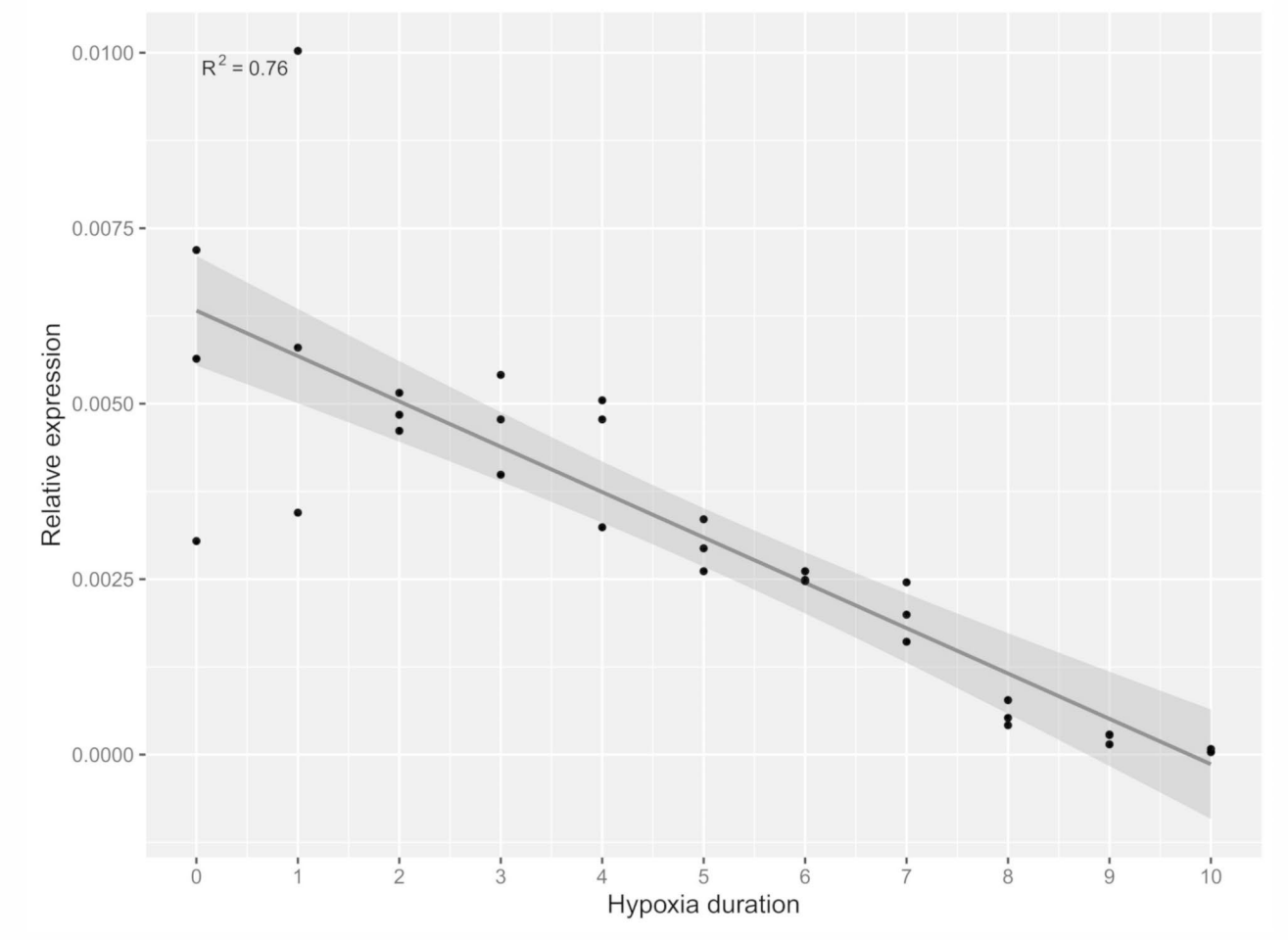
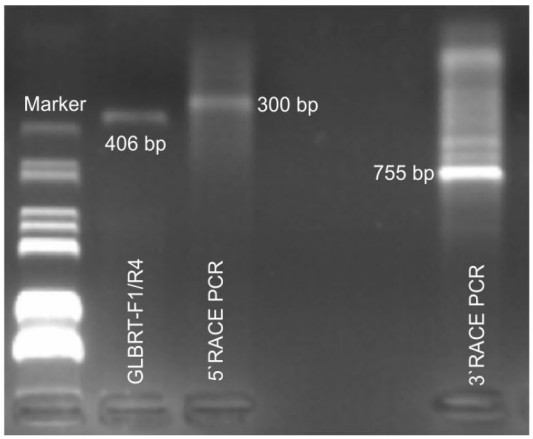


Fig. 4 *Globin* gene expression profiles in different hypoxic environments

Within the scope of this study, it was determined that the possible *globin* gene product obtained from *B. terrestris* was 19.27 kDa weight protein and the isoelectric point (pI) of the protein was found to be pH = 7.60. GRAVY value was calculated as -0.241, indicating that this protein mostly contains hydrophilic amino acids. Potential protein structure was hypothesized to consist of eight α-helixes. No signal peptide sequence was found as a result of Signal-BLAST and Signal-3 L localization analyses, and this protein was predicted to localize in cytoplasm based on data from PROTTER and DeepLoc 2.0 Accordingly, this potential globin product localized in cytoplasm might be functional in the storage of oxygen similar to myoglobin protein. Tertiary structure of the protein was given in Fig. 6 and pointed that the positions of functional significance are histidine E7, histidine F8 amino acids, which are important for O_2 binding potential of heme group with oxygenation.

It is determined that 1- Met, 2- Gly, 3- Thr, 4- Phe, 5- Leu, 7- Phe, 15- Asp, 42- Asp, 57- Thr, 115- Glu, 118- Lys and 120- Arg amino acids may be interaction and/ or binding sites of protein. According to the D- value estimates, the most of the amino acid changes are conserved (25 of 29). Others were not located near the active site of protein, so were unlikely to affect the character of the protein. This increases the conservation of protein at the amino acid level, indicating the potential functionality of this protein sequence in *Bombus*, rather than existing as a junk or residual protein sequence.

It is determined that with the Iss ≥ Iss.c indication, the *globin* gene was saturated in total Insecta (0.8343 ≥ 0.7776); since the gene still evolves in Hymenoptera (0.1921 ≥ 0.8205). Transitional mutations were determined to be more frequent than transversional mutations in *globin* gene, and the most encountered variation was found to be G to A (17,7906), and the least one as A to T (1,5037) change (Online Resource 3).



ACCGCGTGGTATTAGTAGTACATAGCGACGATCAAGCGGAGGGGTGGATTTAAAATAGCATTAGATACATAGGCAGTGCC
GTCAGTCGGACTCGAAGAACGAGTAGTCGTCGTCCAACCGTGTCAGCGATTGAGCGAATTGAAACGACAGTACGTGGCAC
GTGACATATCATACAGAATTACGGGTTTCGTTCTTTTTCGACGATATCTGGCCACCGAACATGGGAACGTTCCTTCGTTT
TTTCGGATTCTCGTCATCGGACGATAACAGAATCGATGAAGCTACAGGGTTGACGGAGAAACAGAAGAAATTAGTGCAGA
ACACGTGGGCCGTTATCAGAAAGGACGAGGTGGCTTCCGGGATTGCCGTGATGACCACATTCTTTAAAACATATCCAGAA
TATCAGCGATATTTTAGCGCTTTCGCGGACGTTCCGTTCGACGAATTGCCGGCTAATAAACGGTTTCAAGCTCATTGTGT
CAGCGTAATTACGGCCTTAAATAGTGTTATCGATTCCTTGCACGACCCAGGATTAATGGAGGCGAGCCTAATCAGCTTGG
GCGAGAGGCATAAAAGACGTGGCCAAACGAAAGAAGAATTTGAGAATTTGAAAGGAGTAGTGTTGAAAGTGCTTTCTCAA
GCACTAGGGAAACAATACACACCGGAAGTGGCCGAGGCGTGGAGCAAAACCTTGGACGGGGTATTTGCGAAGATATATCA
GGTTTTCTCTTCTTGAAGGACGAAACGATTGAACGCACTAGAAAAAGATGAATCTCGTTGATATAATTAACTCTAAATAA
CGAATCTCCATAATAGTTGTTAAGAATTAAATGTTGTAAATAAAATGCACTGTTTTACATGTTTGCACAGTTACGATCAC
TAGAGCTACATTGGACGAAAATGGAAGAAAATGTTCCACAAATTATCCGTGAATATCGAACAGAAATATATACGTTTATA
TCTTCAACACTTAATAGATTTTTGCGCACACTAATGTTAACGTATTGTATAATGCATGTATATATGGACATGGTTTTATA
TCATTTTATTATTT

5`upstream region Start Codon Open Reading Frame Stop Codon 3`upstream region

Fig. 5 Results of RACE system analysis

Table 1 Relative synonymous codon usage (RSCU) table of putative protein

Relative synonymous codon usage (RSCU)											
Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	6	0.92	UCU(S)	3	1.29	UAU(Y)	4	1.6	UGU(C)	1	2
UUC(F)	7	1.08	UCC(S)	2	0.86	UAC(Y)	1	0.4	UGC(C)	0	0
UUA(L)	3	1.29	UCA(S)	1	0.43	UAA(*)	0	0	UGA(*)	1	3
UUG(L)	7	3	UCG(S)	2	0.86	UAG(*)	0	0	UGG(W)	2	1
CUU(L)	2	0.86	CCU(P)	0	0	CAU(H)	2	1.33	CGU(R)	2	1.5
CUC(L)	0	0	CCC(P)	0	0	CAC(H)	1	0.67	CGC(R)	0	0
CUA(L)	2	0.86	CCA(P)	2	1.6	CAA(Q)	4	1	CGA(R)	1	0.75
CUG(L)	0	0	CCG(P)	3	2.4	CAG(Q)	4	1	CGG(R)	1	0.75
AUU(I)	2	0.86	ACU(T)	0	0	AAU(N)	3	1.2	AGU(S)	1	0.43
AUC(I)	4	1.71	ACC(T)	2	0.73	AAC(N)	2	0.8	AGC(S)	5	2.14
AUA(I)	1	0.43	ACA(T)	4	1.45	AAA(K)	10	1.54	AGA(R)	3	2.25
AUG(M)	3	1	ACG(T)	5	1.82	AAG(K)	3	0.46	AGG(R)	1	0.75
GUU(V)	4	1.14	GCU(A)	5	1.43	GAU(D)	3	0.67	GGU(G)	0	0
GUC(V)	1	0.29	GCC(A)	4	1.14	GAC(D)	6	1.33	GGC(G)	2	0.8
GUA(V)	3	0.86	GCA(A)	1	0.29	GAA(E)	6	1	GGA(G)	4	1.6
GUG(V)	6	1.71	GCG(A)	4	1.14	GAG(E)	6	1	GGG(G)	4	1.6

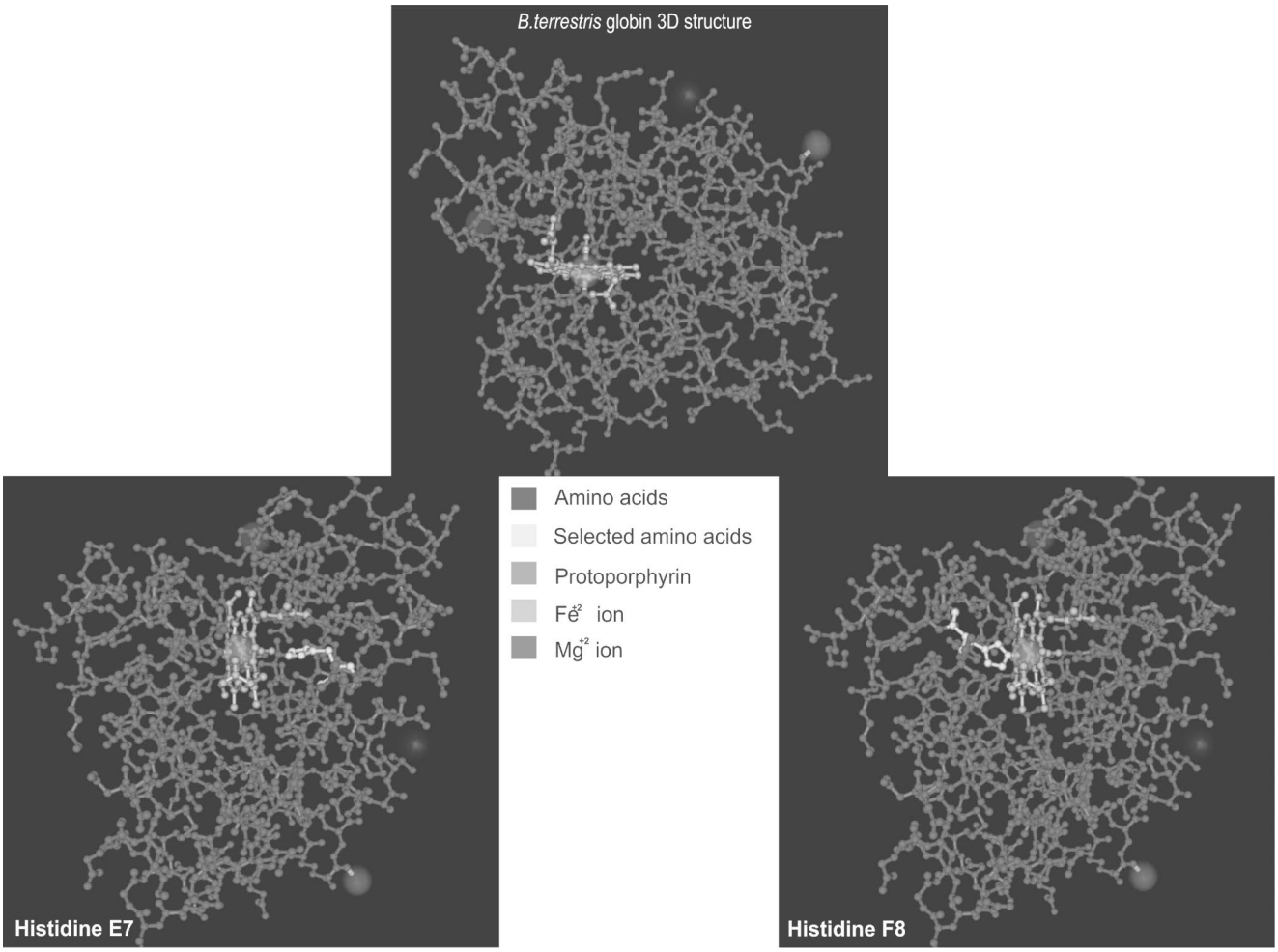


Fig. 6 Tertiary structure of potential globin protein in *B. terrestris.* Positions of conserved and functionally remarkable amino acids were pointed

Conclusion

In this study, we showed that the *globin* gene in *B. terrestris* was expressed based on the real- time PCR data, it was determined that the gene can be transcribed to mRNA. The presence of globin gene expression was determined based on real time qPCR and RACE system, and it was confirmed by SRA data. The potential globin ORF sequence was 516 bp nucleotide and 171 bp amino acid long in *B. terrestris*. This data was deposited in the GenBank database under accession number OP312897. On the other hand, expression profiles have been investigated in different hypoxic conditions at different developmental stages (larva, pupae and adult) and in different tissue types (abdomen, thorax, head and legs). It was determined that the *globin* gene is transcribed in all developmental stages, but only in abdomen. In expression studies performed in different hypoxia environments, it was determined that globin expression decreases with decreasing amounts of oxygen.

As can be understood from other studies on globin characterization and globin evolution in the literature, globin, an ancestral protein, is also represent in insects. Insects with different life strategies have members and/or members of the *globin* gene family according to their oxygen needs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11756-023-01389-4.

Acknowledgements We thank Sivas Cumhuriyet University (CÜBAP F-403) for providing financial support. We are grateful to Dr. Merve Nur AYDEMİR (Department of Molecular Biology and Genetics, Tokat Gaziosmanpaşa University) and Dr. Mahir BUDAK (Sivas Cumhuriyet University) for their valuable contributions.

Author contributions All authors contributed to the conceptualization and methodology. Material preparation, data collection and analysis were performed by Habeş Bilal Aydemir. The first draft of the manuscript was written by Habeş Bilal Aydemir and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by Sivas Cumhuriyet University (Grant number F-403).

Data Availability Not applicable.

Code Availability Not applicable.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

Conflict of interest All authors declare that they have no conflicts of interest.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Buchan DWA, Minneci F, Nugent TCO et al (2013) Scalable web services for the PSIPRED protein analysis workbench. Nucleic Acids Res 41:W349–W357. https://doi.org/10.1093/nar/gkt381

Burmester T (2015) Evolution of respiratory proteins across the pancrustacea. Integr Comp Biol 55:792–801. https://doi.org/10.1093/icb/icv079

Burmester T, Hankeln T (1999) A globin gene of *Drosophila melanogaster*. Mol Biol Evol 16:1809–1811. https://doi.org/10.1093/oxfordjournals.molbev.a026093

Burmester T, Hankeln T (2007) The respiratory proteins of insects. J Insect Physiol 53:285–294. https://doi.org/10.1016/j.jinsphys.2006.12.006

Burmester T, Hankeln T (2014) Function and evolution of vertebrate globins. Acta Physiol 211:501–514. https://doi.org/10.1111/apha.12312

Burmester T, Storf J, Hasenjager A et al (2006) The *hemoglobin* genes of *Drosophila*. FEBS J 273:468–480. https://doi.org/10.1111/j.1742-4658.2005.05073.x

Chen K, Kurgan L (2007) PFRES: protein fold classification by using evolutionary information and predicted secondary structure. Bioinformatics 23:2843–2850. https://doi.org/10.1093/bioinformatics/btm475

Flogel U, Merx MW, Godecke A et al (2001) Myoglobin: a scavenger of bioactive NO. Proc Natl Acad Sci USA 98:735. https://doi.org/10.1073/pnas.98.2.735

Flogel U, Godecke A, Klotz L-O, Schrader AJ (2004) Role of myoglobin in the antioxidant defense of the heart. FASEB J 18:1156–1158. https://doi.org/10.1096/fj.03-1382fje

Gleixner E, Abriss D, Adryan B et al (2008) Oxygen-induced changes in hemoglobin expression in *Drosophila*. FEBS J 275:5108–5116. https://doi.org/10.1111/j.1742-4658.2008.06642.x

Gleixner E, Ripp F, Gorr TA et al (2016) Knockdown of *Drosophila hemoglobin* suggests a role in O_2 homeostasis. Insect Biochem Mol Biol 72:20–30. https://doi.org/10.1016/j.ibmb.2016.03.004

Grantham R (1974) Amino acid difference formula to help explain protein evolution. Science 185:862–864. https://doi.org/10.1126/science.185.4154.862

Hamdane D, Kiger L, Dewilde S et al (2003) The redox state of the cell regulates the ligand binding affinity of human neuroglobin and cytoglobin. J Biol Chem 278:51713–51721. https://doi.org/10.1074/jbc.M309396200

Hankeln T, Jaenicke V, Kiger L et al (2002) Characterization of *Drosophila* Hemoglobin. J Biol Chem 277:29012–29017. https://doi.org/10.1074/jbc.M204009200

Hankeln T, Klawitter S, Krämer M, Burmester T (2006) Molecular characterization of hemoglobin from the honeybee *Apis mellifera*. J Insect Physiol 52:701–710. https://doi.org/10.1016/j.jinsphys.2006.03.010

Hardison RC (1996) A brief history of hemoglobins: plant, animal, protist, and bacteria. Proc Natl Acad Sci USA 93:5675–5679. https://doi.org/10.1073/pnas.93.12.5675

De Henau S, Tilleman L, Vangheel M et al (2015) A redox signalling globin is essential for reproduction in *Caenorhabditis elegans*. Nat Commun 6:8782. https://doi.org/10.1038/ncomms9782

Hendgen-Cotta UB, Merx MW, Shiva S et al (2008) Nitrite reductase activity of myoglobin regulates respiration and cellular viability in myocardial ischemia-reperfusion injury. Proc Natl Acad Sci USA 105:10256–10261. https://doi.org/10.1073/pnas.0801336105

Herhold HW, Davis SR, Grimaldi DA (2020) Transcriptomes reveal expression of hemoglobins throughout insects and other Hexapoda. PLoS ONE 15:e0234272. https://doi.org/10.1371/journal.pone.0234272

Kawaoka S, Katsuma S, Meng Y et al (2009) Identification and characterization of *globin* genes from two lepidopteran insects, *Bombyx mori* and *Samia cynthia ricini*. Gene 431:33–38. https://doi.org/10.1016/j.gene.2008.11.004

Kearse M, Moir R, Wilson A et al (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Koch J, Burmester T (2016) Membrane-bound globin X protects the cell from reactive oxygen species. Biochem Biophys Res Commun 469:275–280. https://doi.org/10.1016/j.bbrc.2015.11.105

Lee SM, Lee SB, Park CH, Choi J (2006) Expression of heat shock protein and *hemoglobin* genes in *Chironomus tentans* (Diptera, Chironomidae) larvae exposed to various environmental pollutants: a potential biomarker of freshwater monitoring. Chemosphere 65:1074–1081. https://doi.org/10.1016/j.chemosphere.2006.02.042

Li LC, Dahiya R (2002) MethPrimer: Designing primers for methylation PCRs. Bioinformatics 18:1427–1431. https://doi.org/10.1093/bioinformatics/18.11.1427

Nath BB (2018) Extracellular hemoglobin and environmental stress tolerance in chironomus larvae. J Limnol 77:104–112. https://doi.org/10.4081/jlimnol.2018.1805

Ofran Y, Rost B (2007) Protein–protein interaction hotspots carved into sequences. PLoS Comput Biol 3:e119. https://doi.org/10.1371/journal.pcbi.0030119

Omasits U, Ahrens CH, Müller S, Wollscheid B (2014) Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics 30:884–886. https://doi.org/10.1093/bioinformatics/btt607

Picotti P, Dewilde S, Fago A et al (2009) Unusual stability of human neuroglobin at low pH - molecular mechanisms and biological significance. FEBS J 276:7027–7039. https://doi.org/10.1111/j.1742-4658.2009.07416.x

Stothard P (2000) The sequence Manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28:1102–1104. https://doi.org/10.2144/00286ir01

Thumuluri V, Armenteros JJA, Johansen AR et al (2022) DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. Nucleic Acids Res 50:1–7. https://doi.org/10.1093/nar/gkac278

Trent JT, Hargrove MS (2002) A ubiquitously expressed human hexacoordinate hemoglobin. J Biol Chem 277:19538–19545. https://doi.org/10.1074/jbc.M201934200

Wang Y, Geer LY, Chappey C et al (2000) Cn3D: sequence and structure views for Entrez. Trends Biochem Sci 25:300–302. https://doi.org/10.1016/S0968-0004(00)01561-9

Weatherall DJ (1983) Hemoglobin for beginners Hemoglobin: structure, function, evolution and pathology. Cell 33:644–645

Weber RE, Vinogradov SN (2001) Nonvertebrate hemoglobins: functions and molecular adaptations. Physiol Rev 81:569–628. https://doi.org/10.1152/physrev.2001.81.2.569

Xia X (2001) DAMBE: Software package for data analysis in molecular biology and evolution. J Hered 92:371–373. https://doi.org/10.1093/jhered/92.4.371

Yadav R, Kundu S, Sarkar S (2015) *Drosophila* glob1 expresses dynamically and is required for development and oxidative stress response. Genesis 53:719–737. https://doi.org/10.1002/dvg.22902