

RESEARCH ARTICLE

Open Access



# Diversity, evolution, and function of myriapod hemocyanins

Samantha Scherbaum<sup>1</sup>, Nadja Hellmann<sup>2</sup>, Rosa Fernández<sup>3,4</sup>, Christian Pick<sup>1</sup> and Thorsten Burmester<sup>1\*</sup> 

## Abstract

**Background:** Hemocyanin transports O<sub>2</sub> in the hemolymph of many arthropod species. Such respiratory proteins have long been considered unnecessary in Myriapoda. As a result, the presence of hemocyanin in Myriapoda has long been overlooked. We analyzed transcriptome and genome sequences from all major myriapod taxa – Chilopoda, Diplopoda, Symphyla, and Pauropoda – with the aim of identifying hemocyanin-like proteins.

**Results:** We investigated the genomes and transcriptomes of 56 myriapod species and identified 46 novel full-length hemocyanin subunit sequences in 20 species of Chilopoda, Diplopoda, and Symphyla, but not Pauropoda. We found in *Cleidogona* sp. (Diplopoda, Chordeumatida) a hemocyanin-like sequence with mutated copper-binding centers, which cannot bind O<sub>2</sub>. An RNA-seq approach showed markedly different hemocyanin mRNA levels from ~ 6 to 25,000 reads per kilobase per million reads. To evaluate the contribution of hemocyanin to O<sub>2</sub> transport, we specifically studied the hemocyanin of the centipede *Scolopendra dehaani*. This species harbors two distinct hemocyanin subunits with low expression levels. We showed cooperative O<sub>2</sub> binding in the *S. dehaani* hemolymph, indicating that hemocyanin supports O<sub>2</sub> transport even at low concentration. Further, we demonstrated that hemocyanin is > 1500-fold more highly expressed in the fertilized egg than in the adult.

**Conclusion:** Hemocyanin was most likely the respiratory protein in the myriapod stem-lineage, but multiple taxa may have independently lost hemocyanin and thus the ability of efficient O<sub>2</sub> transport. In myriapods, hemocyanin is much more widespread than initially appreciated. Some myriapods express hemocyanin only at low levels, which are, nevertheless, sufficient for O<sub>2</sub> supply. Notably, also in myriapods, a non-respiratory protein similar to insect storage hexamerins evolved from the hemocyanin.

**Keywords:** Arthropoda, Evolution, Hemocyanin, Myriapoda, Phenoloxidase, Subunit diversity

## Background

Oxygen (O<sub>2</sub>) is required for efficient generation of metabolic energy. In many animals, O<sub>2</sub> is transported in the body fluid (blood or hemolymph) by specific binding-proteins. Such respiratory proteins have long been considered unnecessary in Myriapoda, which appeared to be well equipped for diffusive O<sub>2</sub> transport by the tracheal system [1]. In the past years, it has become evident that at least some myriapods possess hemocyanin (Hc) for O<sub>2</sub> transport [2, 3].

Hc is the respiratory copper-protein typically found in many arthropod and molluscan species [4–10]. Molluscan and arthropod Hcs do not share significant sequence similarities and are most likely the result of convergent

evolution from different types of tyrosinases [7, 11]. Arthropod Hcs form hexamers or oligohexamers (up to 8 x 6mers), which are composed of similar or identical subunits of ~ 75 kDa [4–7, 9]. In each subunit, O<sub>2</sub> binds to a pair of Cu<sup>+</sup>-ions, which are coordinated by six histidine residues of the protein chain (type III copper-center). In addition to its prominent role in O<sub>2</sub>-transport and -storage, arthropod Hc may carry out other functions, such as a source for antimicrobial peptides [12], as phenoloxidase [13] or as storage protein that provides amino acids and metabolic energy [14].

Arthropod Hcs are members of a protein superfamily that additionally includes arthropod phenoloxidases (POs), which are copper-containing enzymes involved in various immunological functions and cuticle building [10, 15], and various proteins that are involved in energy and amino acid storage, such as pseudo-hemocyanins

\* Correspondence: [thorsten.burmester@uni-hamburg.de](mailto:thorsten.burmester@uni-hamburg.de)

<sup>1</sup>Institute of Zoology, University of Hamburg, D-20146 Hamburg, Germany  
Full list of author information is available at the end of the article



(cryptocyanins) of some decapod crustaceans, hexapod hexamerins, and dipteran hexamerin receptors [6, 9, 10, 15–18]. Sequences that share significant similarities with the arthropod proteins have been found in tunicates, hemichordates, sponges, and fungi [19–21]. Although functional analyses are missing, these proteins probably act as phenoloxidases.

Hc structure, function, and subunit evolution have been thoroughly studied in Chelicerata and Crustacea [5–7, 10, 22–26]. In the past about 15 years, Hc has also been identified in Hexapoda [27, 28] as well as in Onychophora [29], which are the sister taxon of the (eu-)arthropods. The first evidence for the presence of Hc in Myriapoda was found in the centipede *Scutigera longicornis* (Scutigeraomorpha, Chilopoda) [30]. Mangum and colleagues [3] demonstrated that the Hc from *Scutigera coleoptrata* is a 6×6-mer, which is composed of four distinct subunit types [31, 32]. The presence of Hc in *Scutigera* was attributed to the high activity of this species and its blind-ending tracheal system [3]. However, a structurally similar Hc, which occurs as a mixture of 6×6-mers and 3×6-mers, was identified in the hemolymph of *Spirostreptus* sp. (Spirostreptida, Diplopoda) [2] and *Archispirostreptus gigas* [33]. In *Polydesmus angustus* (Polydesmida, Diplopoda) Hc occurs as a 3×6-mer [34].

Here, we present a thorough survey of Hc in 56 myriapod species, which has been made possible by the recent availability of transcriptome and genome data [35–39]. Employing an RNA-seq approach, we found large differences in mRNA levels across species. By specifically analyzing the Hc of the centipede *Scolopendra dehaani*, we demonstrated that Hc can contribute to O<sub>2</sub> transport in the hemolymph even at low expression levels and that Hc may have a specific role in early myriapod development.

## Methods

### Databases and hemocyanin sequences

Short read data (Illumina or 454) of transcriptomes from 54 myriapod species were obtained from the SRA database at NCBI (<https://www.ncbi.nlm.nih.gov/sra/>) (Table 1). The transcriptomes were assembled with Trinity [40] using the standard settings. Contigs with  $\geq 500$  bp were kept and searched for Hc and PO cDNA sequences using a locally installed BLAST tool [41]. The derived sequences were verified by back-mapping of the reads from the corresponding species employing the CLC Genomics Workbench 11.0.0 (<https://www.qiagen-bioinformatics.com/>). This approach was also used to extend partial sequences. Few incomplete or ambiguous Hc sequences that could not be resolved were discarded. The genome of *Strigamia maritima* Leach, 1817 [38] was searched and analyzed for Hc and PPO sequences at <http://www.ensembl.org/>. Genomic sequences of

*Trigoniulus corallinus* were obtained from [39]. Scaffolds and contigs that cover the single Hc gene were assembled by hand. The Hc coding sequences were predicted by AUGUSTUS (<http://bioinf.uni-greifswald.de/webaugustus/>) [42] and GENSCAN (<http://genes.mit.edu/GENSCAN.html>) [43], and verified by the aid of a multiple sequence alignment with other myriapod Hc sequences. The cDNA sequences were translated into proteins with the translate tool at the ExPASy Server of the Swiss Institute of Bioinformatics (<http://web.expasy.org/translate/>).

### Cloning and sequencing of *S. dehaani* hemocyanins and phenoloxidase cDNAs

An adult female centipede *S. dehaani* Brandt, 1840 (Scolopendromorpha, Chilopoda) was obtained from a local pet shop. Selected tissues of *S. dehaani* were dissected, shock-frozen in liquid N<sub>2</sub> and kept at  $-80$  °C until use. Total RNA was extracted either according to the method by Holmes and Bonner [44] or with the RNeasy Kit (Qiagen, Hilden, Germany) according to the manufacturers' instructions. The quality of the RNA was checked by measuring the OD 260/280 ratio and by gel electrophoresis. cDNA was obtained using the SuperScript™ III RNase H-Reverse Transcriptase Kit (Invitrogen, Karlsruhe, Germany), employing an oligo(dT)-primer. Partial *S. dehaani* Hc and PO sequences were obtained with PCR using sets of primers that had been generated on the basis of partial cDNAs [35] (Additional file 1: Table S1). Missing 5' and 3' ends were completed by the RACE technique using the GeneRacer™ kit (Invitrogen) according to the manufacturer's instructions. The PCR products were cloned into the pGEM-T vector (Promega, Mannheim, Germany) and sequenced by a commercial service (GATC, Konstanz, Germany).

### Sequence analyses and phylogenetic studies

N-terminal signal sequences required for export into the extracellular space were predicted with SignalP 4.1 [45]. An alignment of the amino acid sequences of myriapod Hcs and prophenoloxidases (PPOs) (Additional file 2: Table S2) was constructed with MAFFT 7 [46] with the G-INS-i method and the BLOSUM 62 matrix. The final alignment covered 57 Hc and 16 PPO sequences and 868 characters (Additional file 3: Figure S1). An additional alignment was generated by eliminating poorly aligned sections by Gblocks v0.91b [47], employing the options for a less stringent selection (smaller final blocks, gap positions within the final blocks, less strict flanking positions). This approach resulted in an alignment of 531 amino acids. The best-fitting models of amino acid sequence evolution, LG [48] and WAG [49], were selected with ProtTest [50] under the Akaike Information Criterion. MrBayes 3.2.6 [51] was used for Bayesian phylogenetic analysis. The LG model was

**Table 1** List of short read sequences used in this study

Species	Order	Class	SRA	Tissue	Sequencing	Hemocyanin	RPKM
<i>Hanseniella</i> sp.		Symphyla	SRX1734405, SRX3326582	whole body, body tissue	Illumina	1 Hc	328.1
<i>Scutigere</i> sp.		Symphyla	SRX1734406	whole body	Illumina	1 Hc	997.7
<i>Symphylella vulgaris</i>		Symphyla	SRX246919	whole body	454	n.d.	0
<i>Symphylella</i> sp.		Symphyla	SRX3256428	trunk	Illumina	n.d.	0
<i>Pauropus huxleyi</i>		Pauropoda	SRX3257465	whole body	Illumina	n.d.	0
<i>Eudigraphis taiwaniensis</i>	Polyxenida	Diplopoda	SRX1734393	body tissue	Illumina	n.d.	0
<i>Polyxenus lagurus</i>	Polyxenida	Diplopoda	SRX390592	whole body	454	n.d.	0
<i>Cyliosoma</i> sp.	Sphaerotheriida	Diplopoda	SRX1734403	body tissue	Illumina	n.d.	0
<i>Glomeris marginata</i>	Glomerida	Diplopoda	SRX1638914	body tissue	Illumina	n.d.	0
<i>Glomeris pustulata</i>	Glomerida	Diplopoda	SRX246920	whole body	454	n.d.	0
<i>Glomeridesmus</i> sp.	Glomeridesmida	Diplopoda	SRX326775	n.a.	Illumina	n.d.	0
<i>Brachycybe lecontii</i>	Playtodesmida	Diplopoda	SRX326776	n.a.	Illumina	n.d.	0
<i>Petaserpes</i> sp.	Polyzoniida	Diplopoda	SRX326777	n.a.	Illumina	n.d.	0
<i>Cleidogona</i> sp.	Chordeumatida	Diplopoda	SRX326780	n.a.	Illumina	5 Hc	549.3
<i>Abacion magnum</i>	Callipodida	Diplopoda	SRX326781	n.a.	Illumina	4 Hc	232.2
<i>Chamberlinius hualienensis</i>	Polydesmida	Diplopoda	DRX028808	animals without gut	454	3 Hc	2145.4
<i>Polydesmus angustus</i>	Polydesmida	Diplopoda	SRX390267	whole body	454	3 Hc	195.8
<i>Pseudopolydesmus</i> sp.	Polydesmida	Diplopoda	SRX326779	n.a.	Illumina	3 Hc	460.6
<i>Prostemmiulus</i> sp.	Stemmiulida	Diplopoda	SRX326782	n.a.	Illumina	4 Hc	25,587.7
<i>Narceus americanus</i>	Spirobolida	Diplopoda	SRX1638916	body tissue	Illumina	n.d.	0
<i>Trigoniulus corallinus</i>	Spirobolida	Diplopoda	SRX700727	whole organism	Illumina (genome)	1 Hc	n.a.
<i>Cambala annulata</i>	Spirostreptida	Diplopoda	SRX326783	n.a.	Illumina	4 Hc	1183.5
<i>Cylindroiulus punctatus</i>	Julida	Diplopoda	SRX1734404	body tissue	Illumina	n.d.	0
<i>Scutigera coleoptrata</i>	Scutigermorpha	Chilopoda	SRX462011	n.a.	Illumina	5 Hc	22,719.8
<i>Scutigera weberi</i>	Scutigermorpha	Chilopoda	SRX1637773	body tissue	Illumina	4 Hc	1904.3
<i>Sphendononema guildingii</i>	Scutigermorpha	Chilopoda	SRX1637754	body tissue	Illumina	4 Hc	3238.3
<i>Thereuopoda</i> cf. <i>longicornis</i>	Scutigermorpha	Chilopoda	SRX450808	venom gland	454	n.d.	0
<i>Craterostigma tasmanianus</i>	Craterostigmomorpha	Chilopoda	SRX461877	n.a.	Illumina	n.d.	0
<i>Craterostigma crabilli</i>	Craterostigmomorpha	Chilopoda	SRX1638413	body tissue	Illumina	n.d.	0
<i>Anopsobius giribeti</i>	Lithobiomorpha	Chilopoda	SRX1638351	body tissue	Illumina	n.d.	0
<i>Eupolybothrus cavernicolus</i>	Lithobiomorpha	Chilopoda	ERX311347	whole organism	Illumina	n.d.	0
<i>Lithobius forficatus</i>	Lithobiomorpha	Chilopoda	SRX462145, SRX270896	n.a., whole organism	Illumina + 454	n.d.	0
<i>Lithobius</i> sp.	Lithobiomorpha	Chilopoda	SRX326784	n.a.	Illumina	n.d.	0
	Lithobiomorpha	Chilopoda	SRX1638281	body tissue	Illumina	n.d.	0

**Table 1** List of short read sequences used in this study (Continued)

Species	Order	Class	SRA	Tissue	Sequencing	Hemocyanin	RPKM
<i>Paralamyctes validus</i>							
<i>Notiphilides grandis</i>	Geophilomorpha	Chilopoda	SRX1734365	body tissue	Illumina	n.d.	0
<i>Henia brevis</i>	Geophilomorpha	Chilopoda	SRX1734368	body tissue	Illumina	n.d.	0
<i>Himantarium gabrielis</i>	Geophilomorpha	Chilopoda	SRX461787	n.a.	Illumina	n.d.	0
<i>Hydroschendyla submarina</i>	Geophilomorpha	Chilopoda	SRX1638908	body tissue	Illumina	n.d.	0
<i>Mecistocephalus guildingii</i>	Geophilomorpha	Chilopoda	SRX1638910	body tissue	Illumina	n.d.	0
<i>Stenotaenia linearis</i>	Geophilomorpha	Chilopoda	SRX1638912	body tissue	Illumina	n.d.	0
<i>Strigamia maritima</i>	Geophilomorpha	Chilopoda	SRX530370-SRX530372	mixed	Illumina (genome)	n.d.	0
<i>Tygarrup javanicus</i>	Geophilomorpha	Chilopoda	SRX1638906	body tissue	Illumina	n.d.	0
<i>Akymnopellis chilensis</i>	Scolopendromorpha	Chilopoda	SRX1638823	body tissue	Illumina	n.d.	0
<i>Alipes grandidieri</i>	Scolopendromorpha	Chilopoda	SRX205685	n.a.	Illumina	2 Hc	19.6
<i>Cormocephalus westwoodi</i>	Scolopendromorpha	Chilopoda	SRX273033	venom gland	454	n.d.	0
<i>Cryptops hortensis</i>	Scolopendromorpha	Chilopoda	SRX457664	n.a.	Illumina	1 Hc	338.0
<i>Ethmostigmus rubripes</i>	Scolopendromorpha	Chilopoda	SRX275322, SRX423979, SRX272980, SRX275340	venom gland + epidermis	454	n.d.	0
<i>Newportia adisi</i>	Scolopendromorpha	Chilopoda	SRX1638658	body tissue	Illumina	3 Hc	573.4
<i>Rhysida longipes</i>	Scolopendromorpha	Chilopoda	SRX1638865	body tissue	Illumina	n.d.	0
<i>Scolopendra alternans</i>	Scolopendromorpha	Chilopoda	SRX275341	venom gland	454	n.d.	0
<i>Scolopendra morsitans</i>	Scolopendromorpha	Chilopoda	SRX273031	venom gland	454	n.d.	0
<i>Scolopendra mutilans</i>	Scolopendromorpha	Chilopoda	SRX286707, SRX286708	whole organism	Illumina	3 Hc	589.7
<i>Scolopendropsis bahiensis</i>	Scolopendromorpha	Chilopoda	SRX1734364	body tissue	body tissue	2 Hc	273.2
<i>Scolopocryptops sexspinosus</i>	Scolopendromorpha	Chilopoda	SRX1638758	body tissue	Illumina	fragments	n.a.
<i>Scolopendra dehaani</i>	Scolopendromorpha	Chilopoda	SRX390596	body tissue (trunk)	454	2 Hc	74.1
<i>Theatops spinicaudus</i>	Scolopendromorpha	Chilopoda	SRX1734363	body tissue	Illumina	1 Hc	6

n.a. not available, n.d. not detected, RPKM cumulative RPKM

coded as GTR model with fixed priors using the prset command of MrBayes by specifying the parameters aarevmatpr and statefreqpr. Metropolis-coupled Markov chain Monte Carlo sampling was performed with one cold and three heated chains in two independent runs for 5 million generations on the CIPRES web portal [52]. Prior probabilities were equal, starting trees were random, and tree sampling was performed every 1000th generation. The final average standard deviation of split frequencies was < 0.005, suggesting convergence of the chains. Posterior probability densities of the nodes were estimated after discarding the initial 25% of the trees as burnin.

#### Quantitative real-time reverse transcription-PCR

Reverse transcription was performed with 775 ng total RNA from eggs and fat body of *S. dehaani* employing the SuperScript™ III RNase H-Reverse Transcriptase Kit (Invitrogen) and oligo(dT)<sub>20</sub> primer according to the manufacturer's instructions. Quantitative real-time reverse transcription PCR (qRT-PCR) was performed with primer sets specific for Hc subunits, PO,  $\beta$ -actin and RPLP0 (Additional file 1: Table S1) using the Power SYBR Green PCR Master Mix and the 7500 Real-Time PCR System (Applied Biosystems, Darmstadt). qRT-PCR reactions were performed in technical triplicates.

Amplification was carried out using a standard PCR protocol (95 °C for 15 s, 58 °C for 15 s, and 72 °C for 30 s; 40 cycles) and fluorescence was measured at the last step of each cycle. The specificity of the amplification reactions was validated by analysis of the dissociation curve. The mRNA copy numbers were calculated with the standard curve approach, which employs a dilution series with plasmids carrying the respective cDNA sequences [53]. Calculations were performed with the 7500 Software 2.0.6 (Applied Biosystems).

#### RNA-Seq

RNA-Seq analyses were performed with the CLC Genomics Workbench. The reads of the individual transcriptomes were mapped to the Hc cDNA sequences. The following parameters were applied: Masking mode = No masking, Update contigs = No, Match score = 1, Mismatch cost = 2, Cost of insertions and deletions = Linear gap cost, Insertion cost = 3, Deletion cost = 3, Length fraction = 0.5, Similarity fraction = 0.95, Global alignment = No, Auto-detect paired distances = Yes, Non-specific match handling = Map randomly. The mRNA levels of the Hc subunit were calculated as RPKM (Reads Per Kilobase exon model per Million reads).

#### SDS-PAGE and western blotting

Proteins concentrations from total hemolymph, egg and fat body extracts were determined photometrically. The proteins were separated by SDS-PAGE on a 10% gel with standard conditions [54]. The gels were stained with 0.1% Coomassie Brilliant Blue dissolved in 10% acetic acid/ 25% isopropanol. For Western blotting, proteins were transferred onto nitrocellulose. Non-specific binding sites were blocked with 4% non-fat dry milk in TBS (20 mM Tris-HCl, pH 7.5, 150 mM NaCl). An antiserum raised in rabbits against *S. coleoprata* Hc [32] was diluted 1:10,000 in 5% non-fat dry milk in TBS and used for detection overnight at 4 °C. After four successive washing steps with 0.1% Tween-20 in TBS, the secondary antibody (goat  $\alpha$ -rabbit Fab; Dianova, Hamburg) was applied in a 1:10,000 dilution for 1 h at room temperature. After four additional washing steps, the Hc bands were detected with nitroblue tetrazolium and 5-Bromo-4-chloro-3-indolyl phosphate in 100 mM Tris-HCl, pH 9.5, 100 mM NaCl, in the dark. The intensity of the Hc bands was estimated with the ImageJ program (<https://imagej.nih.gov/ij/>).

#### Oxygen binding curves

Oxygen-binding curves were determined by the polarographic-fluorometric method [55], which bases on the fluorescence of deoxygenated Hc upon excitation with light with a wavelength of 290 nm. The intensity of

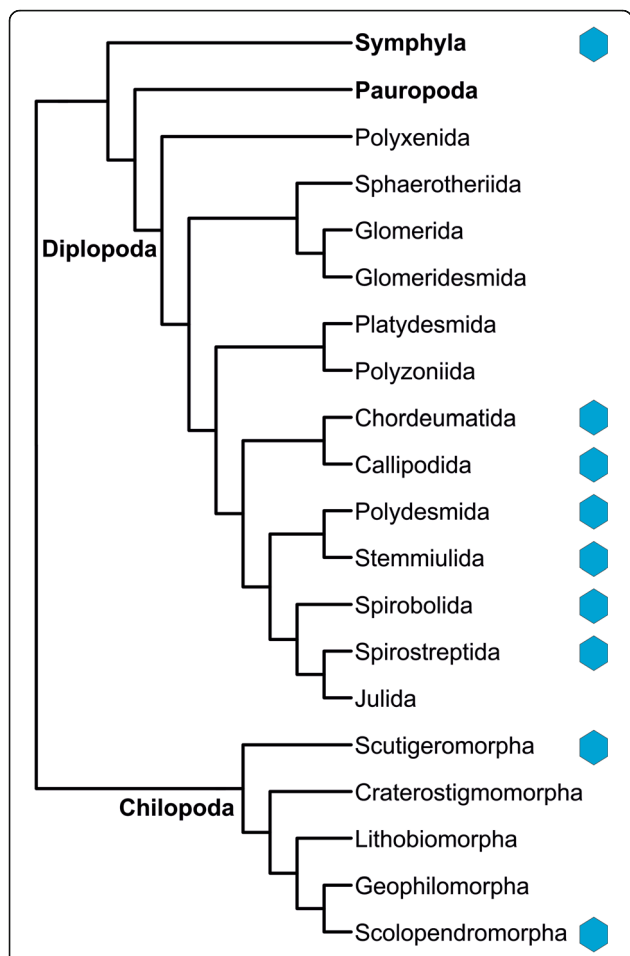
this fluorescence linearly decreases with increasing O<sub>2</sub> saturation of Hc. The fluorescence was measured with Hitachi F4500 (Binniger Analytic, Germany) at 338 nm, while the oxygen concentration of the Hc solution was determined simultaneously with an oxygen electrode (Microelectrodes. Inc., Bedford, USA) equipped with a home-built amplifier. Hemolymph was diluted twofold with 100 mM Tris-HCl, 20 mM MgCl<sub>2</sub>, 20 mM CaCl<sub>2</sub> directly after sampling. The diluted hemolymph was centrifuged (10 min, 12,000 x g) and the supernatant was further diluted into Ringer solution (500 mM NaCl, 12 mM KCl, 12 mM CaCl<sub>2</sub>, 20 mM MgCl<sub>2</sub>, 10 mM Tris/HCl) at the indicated pH. The resulting O<sub>2</sub> binding curves were used to determine the half-saturation pressure P<sub>50</sub> and, after conversion by a Hill plot, cooperativity.

## Results

### Occurrence of hemocyanin sequences in Myriapoda

We screened the available transcriptomes of 54 myriapod species [35–37] for the presence of Hc or PPO (Table 1). We found putative Hc sequences in the transcriptomes of 20 species belonging to Chilopoda (Scutigermorpha and Scolopendromorpha), Diplopoda (Chordeumatida, Callipodida, Polydesmida, Stemmiulida, Spirobolida, and Spirostreptida) and Symphyla (Fig. 1). We included in the survey the transcriptomes of the venom glands of several scolopendromorphs and scutigermorphs, which were found devoid of Hc mRNA. A Hc gene could also be assembled from the genome of *T. corallinus* (Spirobolida) [39]. By contrast, no Hc gene was found in the genome of the geophilomorph *S. maritima* [38]. Therefore, a total of 21 of the 56 (54 transcriptomes + two genomes) investigated myriapod species harbor at least one putative Hc genes.

The contigs that included Hc were extracted and the Hc cDNA sequences were deduced (Additional file 4: Data S1). In some cases, contigs were re-assembled or completed by back-mapping of the original Illumina reads. By this approach, we obtained 50 full-length Hc subunits sequences from the genomes and transcriptomes of 20 myriapod species. In the transcriptome of *Scolopocryptops sexspinosus* (Scolopendromorpha), only few Hc fragments were found, which could not be assembled. This species was ignored in our analyses. The sequences of the Hc subunits A – D of *S. coleoprata* obtained before [32] were identical to that in the transcriptomes, except three isofunctional replacements in subunit B (ScoHcB). For ScoHcX, a highly divergent subunit that is not included in the native Hc [32], only a cDNA fragment was found, which displays multiple substitutions. In summary, this study revealed 46 novel Hc cDNA-sequences. The transcriptome of *S. dehaani* showed the presence of Hc-containing contigs, which



**Fig. 1** The occurrence of hemocyanin in myriapods. The phylogenetic interrelationships of the myriapod classes and orders were taken from [37]; the position of the Pauropoda was added according to a recent study [36]

could be assigned to two distinct subunits. The full-length sequences were obtained by 5' and 3' RACE methods.

The cDNAs were translated *in silico* into proteins, which were included in an alignment that contained all other available myriapod Hc sequences [32–34, 56], resulting in a dataset of 57 putative myriapod Hc subunit sequences. These were assigned to specific subunit types on the basis of phylogenetic analyses (see below). We also added 16 PPO sequences identified in this or a previous study [34]. In all Hc subunits but HcBI2 (CleHcBI2) of *Cleidogona* sp. (Chordeumatida) the six histidines required for copper-binding and thus O<sub>2</sub> transport are conserved (Additional file 3: Figure S1; Fig. 2). In both copper-binding sites (CuA and CuB) of CleHcBI2, one or two copper-coordinating histidines are replaced by another amino acid. Except for a low content of histidines (22 vs. 47 in CleHcBI1), no peculiarities in amino acid composition could be found in CleHcBI2.

**Molecular phylogeny of myriapod hemocyanin subunits**

The multiple sequence alignment of the myriapod Hc and PPO sequences were subject to Bayesian phylogenetic reconstructions. As previous studies demonstrated the monophyly of the myriapod Hc subunits in respect to the Hc subunits of other arthropod subphyla [32, 34], we restricted the present phylogenetic analyses to the sequences of the Myriapoda. We employed a full alignment and, also, an alignment in which poorly aligned sections have been removed by Gblocks [47], and two different models of amino acid evolution (WAG and LG). The PPOs were considered as the outgroup.

The four resulting trees were largely similar (Fig. 3). In all trees, we found five different types of subunits, which we named according to the subunit types of the

**CuA**

	*	*		*
ClePPO1	YFREDISLNSH <b>HW</b> HLAYPYVWLPSQ	-	GGIKDRKGE <b>LF</b> YMHQ <b>Q</b> VLAR	
CleHcD	YFTEDVGLNAH <b>HY</b> HWHLVLPALWK	-	-	GDDLDRKGE <b>HF</b> YWH <b>Q</b> HKQLLAR
CleHcC	YYTEDVGLNAH <b>HY</b> HWHTMNSPLWKDSF	-	-	GGHKDRKGE <b>FF</b> YMHQ <b>Q</b> MLAR
CleHcBII	YFTEDLGLNAN <b>HY</b> QWHA>VNPATWKGAYMGVKE <b>RR</b> GEAFYWH <b>R</b> QMLAR			
CleHcBI1	YFTEDLGMNSH <b>HY</b> HWHLHPGIWNEKF	-	-	G-HKDRNGE <b>LF</b> YMH <b>R</b> Q <b>M</b> VAR
CleHcBI2	YFTEDHGMNS <b>Y</b> YR <b>W</b> HLVLPRIWNEKF	-	-	G-HKDRKGE <b>HF</b> YMH <b>N</b> R <b>Q</b> MVAR

**CuB**

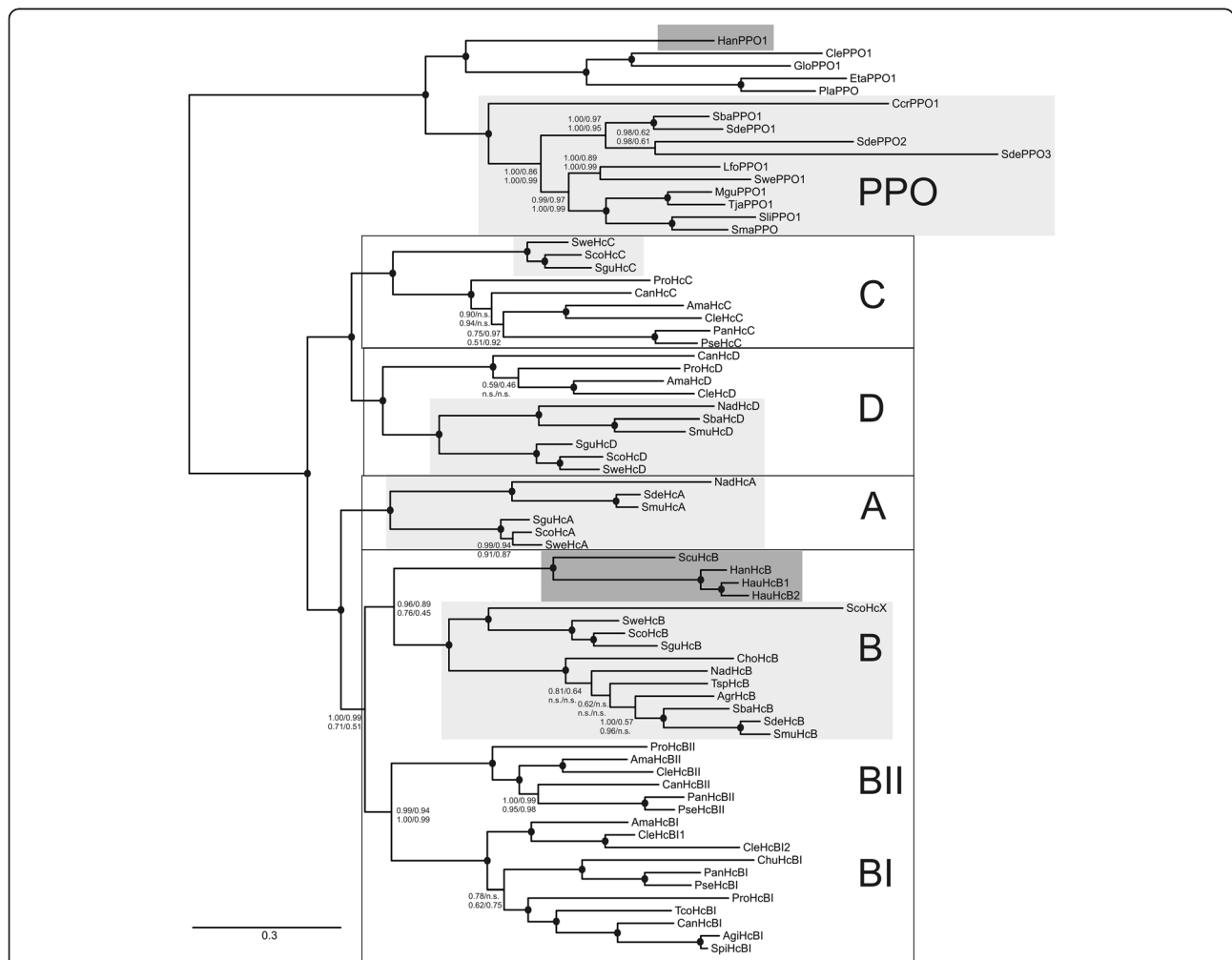
ClePPO1	SVNRTFYGD <b>F</b> EN <b>F</b> GHVIGQIHDPDGKYGTDSGVMGFSSTAMR <b>D</b> PIFYR <b>W</b> HKAIDN <b>I</b> F
CleHcD	SKNKELYGNIENYAHVMAAKVTD <b>P</b> YDRFGGNYGVMYDVATSARDPLFYS <b>W</b> H <b>K</b> FDN <b>I</b> F
CleHcC	SVNCHLYGNIENYAHVIAAKITD <b>P</b> LEAFGETAGAMYDVATSARDPLFYR <b>W</b> H <b>K</b> YIGN <b>F</b>
CleHcBII	SVNFNFGNLENYAHVLSRIVD <b>P</b> DGKYEQDNGAMYDIATSARDPLFY <b>K</b> W <b>H</b> KFDN <b>I</b> F
CleHcBI1	SVNHEFYGNLENYAHVIAARIAD <b>P</b> DGKFKEDSGVMADATTAARDPLFYS <b>W</b> H <b>K</b> FDN <b>I</b> F
CleHcBI2	SINRQFYGNLLSAAHAI <b>T</b> ARI <b>T</b> DPDGKYQ <b>Q</b> ESSVMADAATAARDPLFYS <b>W</b> H <b>K</b> FDN <b>L</b> F

**Fig. 2** Copper-binding sites A and B of hemocyanin subunits and phenoloxidase of *Cleidogona* sp. Conserved residues are shaded in grey. The six copper-binding histidines are shaded in black and denoted by asterisks. Note the replacements of three histidines in CleHcBI2, which cannot bind copper

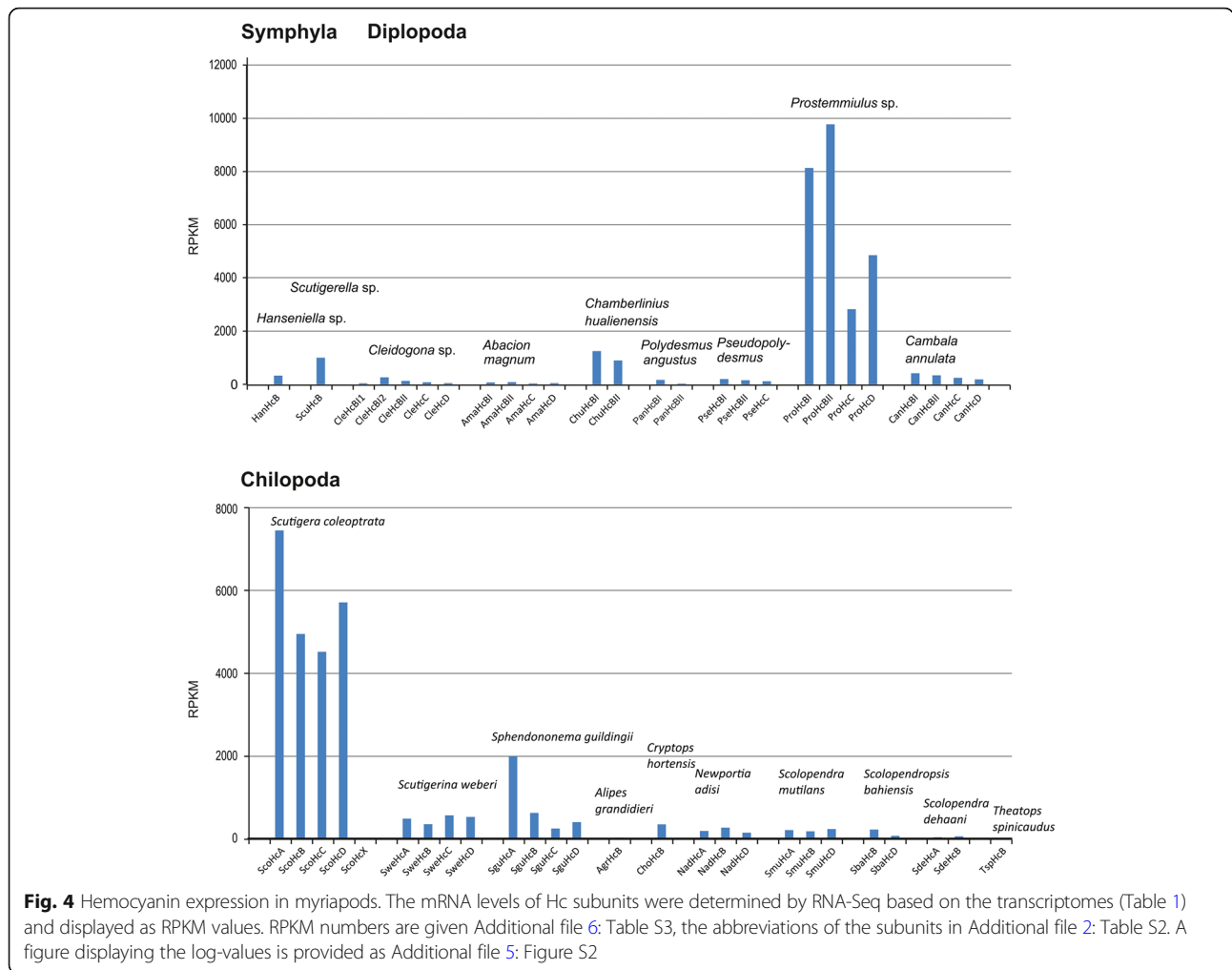
centipede *S. coleoptrata* A-D [31, 32]. In phylogenetic analyses, A and B-type subunits on the one hand and C and D-type subunits on the other hand, form two distinct clades (Fig. 3). Within the C and D-type subunits, chilopod and diplopod Hc subunits are strictly separated. Symphylans only have B-type subunits, which associated with the chilopod B-type subunits, albeit with poor support (0.45 to 0.96 Bayesian posterior probability; BPP). In the Diplopoda, we did not find subunit type A, but two well-supported clades of distinct B-type subunits, which were thus named BI and BII. It should be noted that B-type subunits were found in all investigated myriapod species that have Hc. The copper-free Hc subunit HcBI2 of *Cleidogona* sp. is associated with the “typical” subunit HcBI1 of the same species.

### Estimating Hc subunit mRNA levels by RNA-seq

We studied the expression of Hc subunit mRNA by RNA-Seq, employing the publically available transcriptomes of 19 species (Table 1). The origin of the RNA used for RNA-Seq varied, although in most cases whole organisms or mixed body tissue have been used. We found highly divergent expression levels, which varied several orders of magnitude (Fig. 4; Additional file 5: Figure S2; Additional file 6: Table S3). The lowest Hc mRNA level was observed in the scolopendromorph *Theatops spinicaudus*, which has a single Hc subunit that displays 6 RPKM. *Prostemmiulus* sp. (Stemmiulida) had a cumulative RPKM of to 25,587.7. High Hc expression levels were also found in the Scutigermorpha, with *S. coleoptrata* having the highest Hc expression levels



**Fig. 3** Phylogenetic analysis of the myriapod hemocyanin subunits. A Bayesian phylogenetic tree was inferred from an alignment of myriapod Hc and PPO amino acid sequences (Additional file 5: Figure S2). The displayed tree derived from the analysis of the full 868 amino acid (aa) alignment assuming the WAG model. The numbers at the nodes indicate the Bayesian support values according to the 868 aa alignment/WAG model, 531 aa alignment/WAG model, 868 aa alignment/LG model, 531 aa alignment/LG model. The black dots indicate support values > 0.98 in all four analyses. The bar represents 0.3 PAM distance. Chilopod Hc subunits and PPO are shaded in light grey; the symphylan proteins are shaded in dark grey; the different Hc subunit types are boxed and indicated on the right-hand side. See Additional file 2: Table S2 for the abbreviations



(cumulative RPKM 22719.8). In most species, the RPKM of the different Hc subunits were in the same range, except the HcX-subunit of *S. coleoptrata*, which has a more than 1000-fold lower mRNA level than the other Hc subunits of this species. However, HcX is not a component of the native Hc and has an unknown function [32].

**O<sub>2</sub> binding properties of *S. dehaani* hemocyanin**

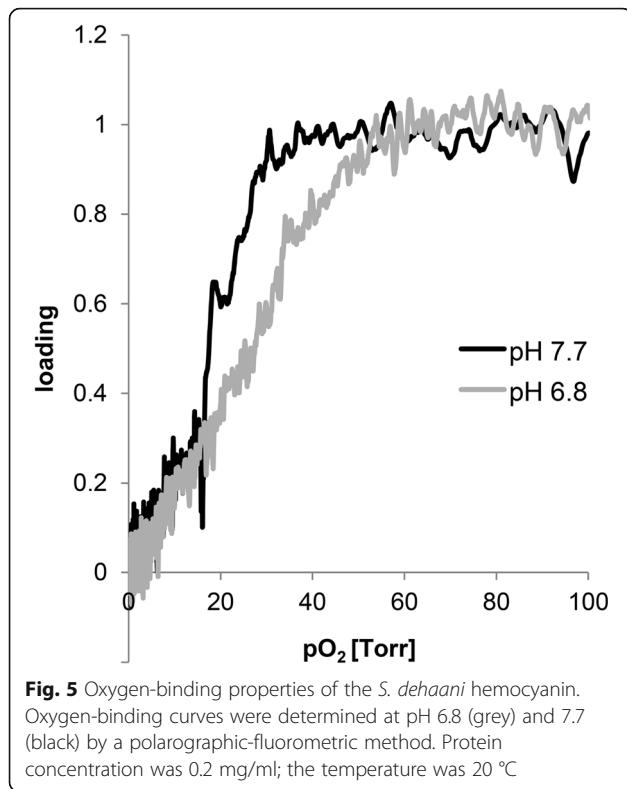
Equilibrium O<sub>2</sub> binding curves of *S. dahaani* Hc were determined by measuring the O<sub>2</sub>-dependent fluorescence of the hemolymph of this species. Because there is no other respiratory protein in the hemolymph of this species and oxygen-dependent changes of the tryptophan fluorescence at this low oxygen concentration is specific for Hc, this approach is valid. We employed two different pH values in our measurements: At pH 7.7, the P<sub>50</sub> was ~19 Torr (2.6 kPa), at pH 6.8 ~27 Torr (3.6 kPa) (Fig. 5). The reduced O<sub>2</sub> affinity at low pH indicates that *S. dahaani* Hc displays the typical Bohr

effect observed in hemocyanins. The sigmoidal O<sub>2</sub>-binding curve indicates cooperative O<sub>2</sub>-binding and a Hill coefficient (*h*) of 1.9 (pH 7.7) and 1.6 (pH 6.8) was calculated.

**mRNA and protein expression of *S. dehaani* hemocyanin**

Here, we investigated a female specimen of *S. dehaani* that carried fertilized eggs. Relative expression of Hc and PPO mRNA levels was determined using qRT-PCR in the hepatopancreas and the eggs. Notably, we found 3400-fold (SdeHcA) and 1500-fold (SdeHcB) higher mRNA levels in the oocyte compared to the hepatopancreas (Fig. 6a). The differences in the levels of PPOs were less pronounced, with factors ranging from 0.4 to 237. In addition, proteins were isolated from hemolymph, eggs and hepatopancreas, and analyzed by Western blotting. A single band with a mass of 80 kDa was detected (Fig. 6b). The relative band intensities were quantified with the software ImageJ, showing six-fold higher Hc protein levels in the egg than in the hepatopancreas.



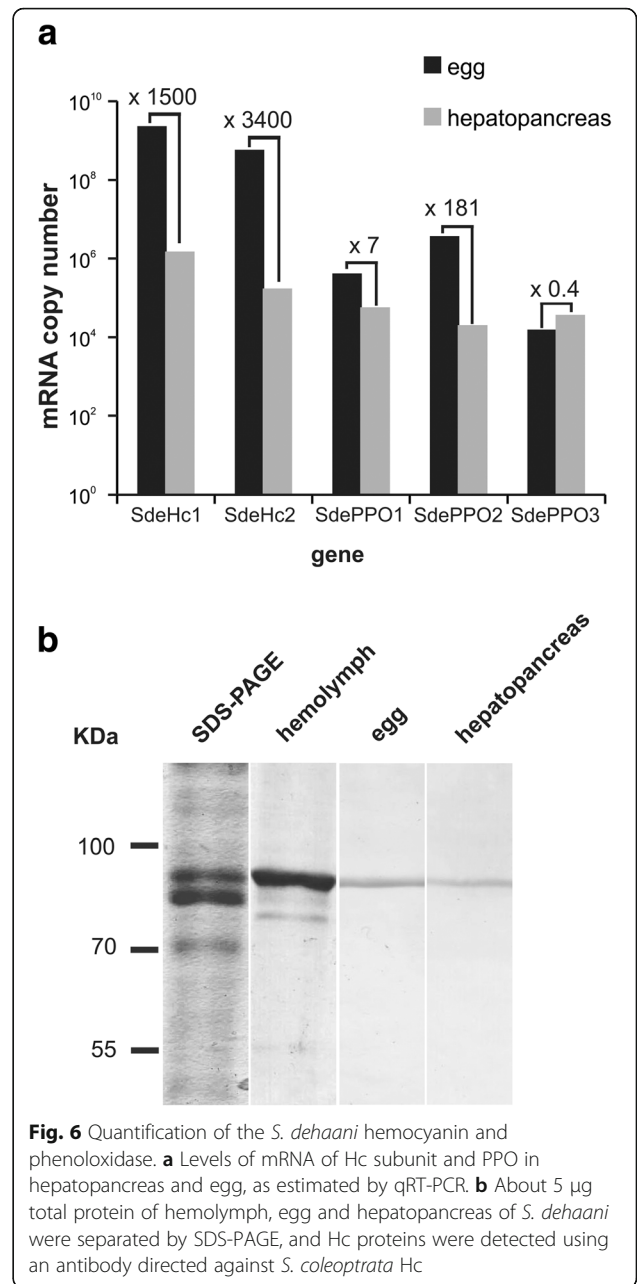


**Discussion**

For a long time, it has been assumed that – regardless of species and/or developmental stage – myriapods do not have Hc or any other respiratory protein. Only the Scutigermorpha were considered as exceptions [3, 30]. However, more recent data have suggested that Hc is also present in other myriapods [2, 33, 34, 56]. A thorough survey on the occurrence, subunit diversity, and expression pattern was missing. Also, it was unknown why Hc was lost in some myriapod taxa, and whether Hc can significantly contribute to O<sub>2</sub> supply.

**Widespread occurrence, losses and divergent expression levels of hemocyanins in Myriapoda**

The examination of the available transcriptomes and genomes of the myriapods [35–39] showed that *Hc* genes are present in three of the four myriapod classes, i.e., Chilopoda, Diplopoda, and Symphyla (Table 1; Fig. 1). No evidence for Hc was found in the transcriptome of Pauropoda. These findings confirm previous notions that Hc was present in the myriapod stem-lineage. Mapping of the occurrence of Hc onto the myriapod phylogenetic tree (Fig. 1) did not reveal a clear pattern. However, we must consider that a *Hc* gene may only be expressed in some developmental stages or certain tissues (see below), and might have been missed in this survey because the corresponding samples were not available. Thus, conclusive evidence for the absence of a *Hc* gene



could only be derived from the genome of *S. maritima*. In those other species, in which no Hc mRNA was identified in the transcriptomes, at least a role of Hc in O<sub>2</sub> supply in the adult can be excluded.

The phylogenetic pattern of the presence/absence suggests that multiple independent losses of the Hc gene occurred during myriapod evolution. Losses of *Hc* genes were also observed in certain taxa of Chelicerata, Crustacea and Hexapoda [5, 22, 28, 57]. Within Diplopoda, Hc appears to be restricted to the Eugnatha (except Julida), while Hc is absent in the early diverging diplopod orders. In the chilopods, Hc is present in

Scutigermorpha and Scolopendromorpha. While Scutigermorpha split from the other Chilopoda early in evolution, Scolopendromorpha occupy a terminal position as the sister group to Geophilomorpha. Thus, there is no true correlation between the presence of Hc and myriapod phylogeny (Fig. 1). With the exception of Scutigermorpha (see below), there is also no apparent morphological, physiological or ecological explanation for the presence of Hc, as the species with Hc have quite different body plans, behavior, and habitats. Thus, there must be others – yet unknown – features that have made a respiratory protein redundant in some species and explain its retention in others.

We also observed divergent expression rates across the myriapod taxa. Also in this context, we must consider that some variations in the RPKM may be due to differences in tissues or developmental stages used for transcriptome generations. There is no clear pattern across the taxa (Fig. 4), although in chilopods there is a tendency of high Hc mRNA levels in Scutigermorpha and low Hc mRNA levels in Scolopendromorpha. The high mRNA Hc levels in Scutigermorpha agrees with previous observations [3, 30]. A requirement of an effective O<sub>2</sub> transporting system in the Scutigermorpha can be explained by the high activity of this species, which is also reflected by the presence of a peculiar, highly branched circulatory system and tracheal lungs [3, 58]. The highest Hc mRNA levels were found in the diplopod *Prostemmiulus* sp. (Stemmiulida). To the best of our knowledge, there is at present no specific characteristics that could explain the high expression of Hc mRNA in this species. Notably, *S. coleoptrata* and *Prostemmiulus* sp. display approximately tenfold higher Hc mRNA levels (> 20,000 RPKM) than any other myriapod species. The very low Hc mRNA in several Scolopendromorpha (*S. dehaani*, *Alipes grandidieri* and *Theatops spinicaudus*) with RPKM < 100 raised the question whether Hc contributes to O<sub>2</sub> transport, or whether this protein may have another function in those species (see below).

#### Myriapod hemocyanin subunit diversity and evolution

Previous studies have demonstrated that distinct Hc subunit types occur in chelicerates, crustaceans, and hexapods [5, 10, 22, 28, 59]. Up to eight distinct subunit-types may occur in a single Hc molecule [25]. The presence of distinct subunit types is a common feature of arthropod Hc; they emerged independently in the different subphyla, and thus the subunit-types may have a long and independent evolutionary history of several hundred million years [10, 22, 28, 60, 61]. The presence of multiple distinct subunits has been associated with the controlled assembly to distinct quaternary structure, containing up to 48 subunits, and thus allowing establishment and regulation of

high cooperativity in Hc or – if differentially expressed – may reflect specific needs during development and in response to environmental changes.

Most chelicerates have conserved 4 × 6-mer or 8 × 6-mer Hcs with seven distinct subunit types that emerged ~ 540 million years ago (MYA) [22, 25, 62]. The pattern in crustaceans is more complex and further complicated by the paraphyletic nature of crustacean and hexapod Hc subunits [60, 61, 63].

Myriapod species have between one and five distinct Hc subunits. Because the subunit HcX of *S. coleoptrata* is not present in the native Hc protein [32] and the copper-free subunit HcBI2 of *Cleidogona* sp. is also probably not part of the respiratory protein, the maximum number of distinct subunits in the native (oligo-)hexameric Hc is probably four. Notably, a B-type subunit was found in all myriapod species investigated here, suggesting that this polypeptide may be the central building block of the native Hc. The phylogenetic studies showed that orthologous C- and D-type subunits occur in both Diplopoda and Chilopoda. A-type subunits appear to be restricted to Chilopoda; the HcX subunit of *S. coleoptrata* is a B-type variant. In Diplopoda, two paralogous HcB subunits are present (HcBI and HcBII). Most likely, the duplication of the *HcB* gene was the response to the loss of the HcA subunit in this taxon. In a previous study [34], we identified two clades of Hc subunits in the myriapods, one which are built by the subunit types A and B and one which contains only C and D subunits. We could confirm this result in the current study. Further, the tree topology implies that the four distinct subunit types diverged before the Diplopoda, Symphyla and Chilopoda separated more than 500 MYA ago [35, 37, 64]. In case that multiple subunits occur within a single species, mRNA expression analyses showed that their RPKM values are similar within an order of magnitude. However, from the expression data, it is difficult to estimate the exact subunit composition of the (oligo-) hexameric Hc protein.

#### A myriapod hemocyanin-like protein that does not bind O<sub>2</sub>

*Cleidogona* sp. (Diplopoda:Chordeumatida) harbors a phenoloxidase and five sequences that share significant similarities with the typical diplopod Hc subunits. Notably, we found that the HcBI subunit is duplicated. Closer investigation showed that in one of these subunits (which we called CleHcBI2) three of the six copper-binding histidines are replaced by another amino acid (CuA: H→Y, H→N; CuB: H→L; Fig. 2). Therefore, CleHcBI2 cannot bind copper and thus O<sub>2</sub>, and is unlikely to be integrated into Hc, or has a PO-like function. Phylogenetic analyses (Fig. 3) showed an about 4.5-fold faster evolutionary rate of

CleHcBI2 compared to CleHcBI1, which is most likely due to the relief of the constraints imposed by the function of O<sub>2</sub> transport. We speculate that CleHcBI2 carries out a similar function as the insect hexamerins or decapod pseudo-hemocyanins (cryptocyanins), Hc-related proteins which are used mainly for the storage of energy and amino acids [10, 16–18, 65]. Many hexamerins accumulate specific amino acids [18, 66], a feature that was not observed in CleHcBI2. The only notable difference in amino acid composition between CleHcBI2 and CleHcBI1 was the reduced relative amount of histidines (3.38% vs. 7.14%) beyond the replacement of the copper-binding sites. The phylogenetic tree (Fig. 3) suggests that this copper-free protein is not a common feature of myriapods or diplopods, but rather evolved specifically within the Chordeumatida. Nevertheless, it is remarkable that Hc has lost three times independently in different arthropod subphyla its respiratory function and probably evolved into a storage protein. This suggests that Hc has particular structural advantages, which may be for example their high stability or the ability to accumulate many amino acids with low osmotic impact due to their large size [17].

#### Hemocyanin supports O<sub>2</sub> supply even at low concentration

The low expression of Hc in many myriapod species raised the question whether this protein can in principle support O<sub>2</sub> transport. We, therefore, analyzed the Hc of the centipede *S. dehaani*. This protein consists of two distinct subunits (SdeHcA and SdeHcB), and expression analyses showed RPKM values of 25.6 and 48.5, respectively (Fig. 4; Additional file 6: Table S3). We showed that the hemolymph and thus the Hc of *S. dehaani* does reversibly bind O<sub>2</sub>. Moreover, it displays a Bohr effect and cooperative O<sub>2</sub>-binding. These features strongly suggest a respiratory function of Hc in *S. dehaani*.

*S. dehaani* Hc displays a low O<sub>2</sub> affinity (P<sub>50</sub> = 19 Torr [2.6 kPa] to 27 Torr [3.6 kPa]; Fig. 5), which is higher than that of *S. coleoptrata* (P<sub>50</sub> = 55 Torr [7.32 kPa] at pH 7.5) [3] but lower than the affinities measured for the diplopod Hcs (*Spirostreptus* Hc P<sub>50</sub> = 4.7 Torr [0.63 kPa] at pH 7.5; *A. gigas* Hc (P<sub>50</sub> = 3.45 Torr [0.46 kPa] at pH 8.1) [2, 33]. A similar pattern was found for the cooperativity, with high Hill coefficient for of *S. coleoptrata* ( $h = 8.9$  at pH 7.5), an intermediate value for *S. dehaani* ( $h = 1.6$  to 1.9) and low cooperativity for ( $h = 1.3 \pm 0.2$ ) for *Spirostreptus* Hc. These values show the flexibility of the O<sub>2</sub> binding behavior of myriapod Hc, which most likely reflect differences in habitat and life-style. For example, a low O<sub>2</sub> affinity and high cooperativity of *S. coleoptrata* Hc may be adaptive in efficient O<sub>2</sub> release in this highly active species, whereas the high O<sub>2</sub>

affinity of *Spirostreptus* Hc may be interpreted as an adaptation for O<sub>2</sub> storage function or efficient extraction of O<sub>2</sub> in the sub-terrestrial environment. *S. dehaani* Hc appears to be better adapted to O<sub>2</sub> release.

#### The embryo may require high levels of Hc

We studied a female *S. dehaani* specimen with fertilized eggs. The exact developmental stage of the egg is unknown. Notably, we found much more Hc mRNA and also slightly enhanced Hc protein levels in the eggs than in the hepatopancreas, which is the principal site of Hc synthesis in adult crustaceans, hexapods and myriapods [27, 56, 67]. The relative Hc protein amount in the egg is probably underestimated due to the high concentration of yolk proteins. Our findings suggest that high levels of Hc mRNA are required for the early development of the *S. dehaani*. An embryo-specific role of Hc has also been demonstrated in the hexapods [68–70]. The accumulation of Hc mRNA in the fertilized egg may be required to meet the O<sub>2</sub> requirements of the early development of the embryo and the restriction of O<sub>2</sub> diffusion across the eggshell. In adult *S. dehaani*, the tracheal system additionally supports O<sub>2</sub> supply, thus a lower concentration of Hc is required.

#### Conclusions

Our results demonstrated that Hc was most likely the standard respiratory protein in the myriapod stem-lineage. Although Hc has been lost in certain taxa, it is still much more widespread in myriapods than initially appreciated. In previous studies, the presence of Hc in some myriapod species may have been overlooked due to the restriction to early developmental stages. Furthermore, some myriapods express Hc only at very low levels. Nevertheless, it is likely that Hc's function is to support O<sub>2</sub> supply by the hemolymph since the cooperative binding behavior is retained. Notably, in myriapods a non-respiratory protein evolved from Hc, which may act as a storage protein similar to insect hexamerins.

#### Additional files

**Additional file 1 Table S1.** List of primer sequences used in this study. (PDF 64 kb)

**Additional file 2 Table S2.** List of sequences used in this study. (DOCX 24 kb)

**Additional file 3: Figure S1.** Multiple sequence alignment of myriapod hemocyanins and phenoloxidasases in FASTA format. (PDF 101 kb)

**Additional file 4: Data S1.** Nucleotide sequences of the Hc and PPO cDNAs and gene identified in this study. (ZIP 81 kb)

**Additional file 5: Figure S2.** The mRNA levels of Hc subunits were determined by RNA-Seq based on the transcriptomes (Table 1) and displayed as log RPKM values. (TIF 3 mb)

**Additional file 6: Table S3.** Expression of hemocyanin subunit mRNA in myriapods. (PDF 72 kb)

## Abbreviations

Hc: Hemocyanin; MYA: Million years ago;  $P_{50}$ : Half-saturation pressure; PPO: Prophenoloxidase; RPKM: Reads Per Kilobase exon model per Million reads; see Additional file 2: Table S2 for abbreviations of the hemocyanin subunits

## Acknowledgments

We thank Andrej Fabrizio (Hamburg University, Germany) for his help with the bioinformatic software and Danilo Harms (Hamburg University, Germany) for his advice on centipede reproduction. We acknowledge the support by Gonzalo Giribet (Harvard University, USA).

## Funding

This work is supported by the Deutsche Forschungsgemeinschaft (Bu956/9). The funding body had no role in the design of the study and collection, analysis, interpretation of data, or in writing the manuscript.

## Availability of data and materials

All supporting data are available in the Additional files 1, 2, 3, 4, 5 and 6. The short read sequences listed in Table 1 are available from the SRA database at NCBI: <https://www.ncbi.nlm.nih.gov/sra>

## Authors' contributions

SS and CP obtained the sequences and biochemical data from *S. dehaani*. NH collected the O<sub>2</sub>-binding data. RF obtained the transcriptome sequences. TB analyzed the sequence and expression data, and wrote the manuscript. All authors read and approved the final manuscript.

## Ethics approval and consent to participate

We followed the regulations of the German Animal Welfare Act (TierSchG).

## Consent for publication

Does not apply.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Institute of Zoology, University of Hamburg, D-20146 Hamburg, Germany. <sup>2</sup>Institute for Biophysics, Johannes Gutenberg University of Mainz, D-55099 Mainz, Germany. <sup>3</sup>Museum of Comparative Zoology & Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA. <sup>4</sup>Bioinformatics & Genomics Unit, Center for Genomic Regulation, 08004 Barcelona, Spain.

Received: 6 January 2018 Accepted: 24 June 2018

Published online: 05 July 2018

## References

- Brusca RC, Brusca GJ. Invertebrates. Sinauer: Sunderland; 2003.
- Jaenicke E, Decker H, Gebauer W, Markl J, Burmester T. Identification, structure, and properties of hemocyanins from diplopod myriapoda. *J Biol Chem*. 1999;274(41):29071–4.
- Mangum CP, Scott JL, Black RE, Miller KI, Van Holde KE. Centipede hemocyanin: its structure and its implications for arthropod phylogeny. *Proc Natl Acad Sci U S A*. 1985;82(11):3721–5.
- Mangum CP. Oxygen transport in invertebrates. *Am J Phys*. 1985;248(5 Pt 2):R505–14.
- Markl J. Evolution and function of structurally diverse subunits in the respiratory protein hemocyanin from arthropods. *Biol Bull*. 1986;171:90–115.
- Markl J, Burmester T, Decker H, Savel-Niemann A, Harris JR, Süling M, Naumann U, Scheller K. Quaternary and subunit structure of *Calliphora* arylphorin as deduced from electron microscopy, electrophoresis, and sequence similarities with arthropod hemocyanin. *J Comp Physiol B*. 1992; 162(8):665–80.
- van Holde KE, Miller KI. Hemocyanins. *Adv Protein Chem*. 1995;47:1–81.
- Markl J. Evolution of molluscan hemocyanin structures. *Biochim Biophys Acta*. 2013;1834(9):1840–52.
- Burmester T. Evolution of respiratory proteins across the Pancrustacea. *Integr Comp Biol*. 2015;55(5):792–801.
- Burmester T. Origin and evolution of arthropod hemocyanins and related proteins. *J Comp Physiol B*. 2002;172(2):95–107.
- Burmester T. Molecular evolution of the arthropod hemocyanin superfamily. *Mol Biol Evol*. 2001;18(2):184–95.
- Destoumieux-Garçon D, Saulnier D, Garnier J, Jouffrey C, Bulet P, Bachère E. Crustacean immunity. Antifungal peptides are generated from the C terminus of shrimp hemocyanin in response to microbial challenge. *J Biol Chem*. 2001;276(50):47070–7.
- Nagai T, Osaki T, Kawabata S. Functional conversion of hemocyanin to phenoloxidase by horseshoe crab antimicrobial peptides. *J Biol Chem*. 2001; 276(29):27166–70.
- Depledge MH, Bjerregaard P. Haemolymph protein composition and copper levels in decapod crustaceans. *Helgoländer Meeresun*. 1989;43:207–23.
- Burmester T, Scheller K. Common origin of arthropod tyrosinase, arthropod hemocyanin, insect hexamerin, and dipteran arylphorin receptor. *J Mol Evol*. 1996;42(6):713–28.
- Terwilliger NB, Dangott L, Ryan M. Cryptocyanin, a crustacean molting protein: evolutionary link with arthropod hemocyanins and insect hexamerins. *Proc Natl Acad Sci U S A*. 1999;96(5):2013–8.
- Burmester T. Identification, molecular cloning, and phylogenetic analysis of a non-respiratory pseudo-hemocyanin of *Homarus americanus*. *J Biol Chem*. 1999;274(19):13217–22.
- Burmester T. Evolution and function of the insect hexamerins. *Eur J Entomol*. 1999;96:213–25.
- Immesberger A, Burmester T. Putative phenoloxidases in the tunicate *Ciona intestinalis* and the origin of the arthropod hemocyanin superfamily. *J Comp Physiol B*. 2004;174(2):169–80.
- Martín-Durán JM, de Mendoza A, Sebé-Pedrós A, Ruiz-Trillo I, Hejnol A. A broad genomic survey reveals multiple origins and frequent losses in the evolution of respiratory hemerythrins and hemocyanins. *Genome Biol Evol*. 2013;5(7):1435–42.
- Aguilera F, McDougall C, Degnan BM. Origin, evolution and classification of type-3 copper proteins: lineage-specific gene expansions and losses across the Metazoa. *BMC Evol Biol*. 2013;13:96.
- Rehm P, Pick C, Borner J, Markl J, Burmester T. The diversity and evolution of chelicerate hemocyanins. *BMC Evol Biol*. 2012;12:19.
- Marxen JC, Pick C, Kwiatkowski M, Burmester T. Molecular characterization and evolution of haemocyanin from the two freshwater shrimps *Caridina multidentata* (Stimpson, 1860) and *Atyopsis moluccensis* (De Haan, 1849). *J Comp Physiol B*. 2013;183:613–24.
- Terwilliger NB, Ryan M, Phillips MR. Crustacean hemocyanin gene family and microarray studies of expression change during eco-physiological stress. *Integr Comp Biol*. 2006;46(6):991–9.
- Martin AG, Depoix F, Stohr M, Meissner U, Hagner-Holler S, Hammouti K, Burmester T, Heyd J, Wriggers W, Markl J. *Limulus polyphemus* hemocyanin: 10 Å cryo-EM structure, sequence analysis, molecular modelling and rigid-body fitting reveal the interfaces between the eight hexamers. *J Mol Biol*. 2007;366(4):1332–50.
- Cong Y, Zhang Q, Woolford D, Schweikardt T, Khant H, Dougherty M, Ludtke SJ, Chiu W, Decker H. Structural mechanism of SDS-induced enzyme activity of scorpion hemocyanin revealed by electron cryomicroscopy. *Structure*. 2009;17(5):749–58.
- Hagner-Holler S, Schoen A, Erker W, Marden JH, Rupprecht R, Decker H, Burmester T. A respiratory hemocyanin from an insect. *Proc Natl Acad Sci U S A*. 2004;101(3):871–4.
- Pick C, Schneuer M, Burmester T. The occurrence of hemocyanin in Hexapoda. *FEBS J*. 2009;276:1930–41.
- Kusche K, Ruhberg H, Burmester T. A hemocyanin from the Onychophora and the emergence of respiratory proteins. *Proc Natl Acad Sci U S A*. 2002; 99(16):10545–8.
- Sundara Rajulu G. Presence of haemocyanin in the blood of a centipede *Scutigera longicornis* (Chilopoda: Myriapoda). *Curr Sci*. 1969;7:168–9.
- Gebauer W, Markl J. Identification of four distinct subunit types in the unique 6 x 6 hemocyanin of the centipede *Scutigera coleoptrata*. *Naturwissenschaften*. 1999;86(9):445–7.
- Kusche K, Hembach A, Hagner-Holler S, Gebauer W, Burmester T. Complete subunit sequences, structure and evolution of the 6 x 6-mer hemocyanin

- from the common house centipede, *Scutigera coleoptrata*. *Eur J Biochem*. 2003;270(13):2860–8.
33. Damsgaard C, Fago A, Hagner-Holler S, Malte H, Burmester T, Weber RE. Molecular and functional characterization of hemocyanin of the giant African millipede, *Archispirostreptus gigas*. *J Exp Biol*. 2013;216:1616–23.
  34. Pick C, Scherbaum S, Hegedus E, Meyer A, Saur M, Neumann R, Markl J, Burmester T. Structure, diversity and evolution of myriapod hemocyanins. *FEBS J*. 2014;281(7):1818–33.
  35. Rehm P, Meusemann K, Borner J, Misof B, Burmester T. Phylogenetic position of Myriapoda revealed by 454 transcriptome sequencing. *Mol Phylogenet Evol*. 2014;77:25–33.
  36. Fernández R, Edgecombe GD, Giribet G. Phylogenomics illuminates the backbone of the Myriapoda Tree of Life and reconciles morphological and molecular phylogenies. *Sci Rep*. 2018;8(1):83.
  37. Fernández R, Edgecombe GD, Giribet G. Exploring phylogenetic relationships within myriapoda and the effects of matrix composition and occupancy on phylogenomic reconstruction. *Syst Biol*. 2016;65(5):871–89.
  38. Chipman AD, Ferrier DE, Brena C, Qu J, Hughes DS, Schroder R, Torres-Oliva M, Znassi N, Jiang H, Almeida FC, et al. The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*. *PLoS Biol*. 2014;12(11):e1002005.
  39. Kenny NJ, Shen X, Chan TT, Wong NW, Chan TF, Chu KH, Lam HM, Hui JH. Genome of the rusty millipede, *Trigoniulus corallinus*, illuminates diplopod, myriapod, and arthropod evolution. *Genome Biol Evol*. 2015;7(5):1280–95.
  40. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29(7):644–52.
  41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10.
  42. Hoff KJ, Stanke M. WebAUGUSTUS—a web service for training AUGUSTUS and predicting genes in eukaryotes. *Nucleic Acids Res*. 2013;41(Web Server issue):W123–8.
  43. Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA. *J Mol Biol*. 1997;268(1):78–94.
  44. Holmes DS, Bonner J. Preparation, molecular weight, base composition, and secondary structure of giant nuclear ribonucleic acid. *Biochemistry*. 1973;12(12):2330–8.
  45. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;8(10):785–6.
  46. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80.
  47. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*. 2000;17(4):540–52.
  48. Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol*. 2008;25(7):1307–20.
  49. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*. 2001;18(5):691–9.
  50. Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*. 2005;21(9):2104–5.
  51. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*. 2001;17(8):754–5.
  52. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: Proceedings Of The Gateway Computing Environments Workshop (GCE): 14 Nov. 2010; New Orleans, LA; 2010. p. 1–8.
  53. Pfaffl MW, Hageleit M. Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. *Biotechnol Lett*. 2001;23:275–82.
  54. Lämmler UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970;227(5259):680–5.
  55. Loewe R. Hemocyanins in spiders. 5. Fluorimetric recording of oxygen binding curves, and its application to analysis of allosteric interactions in *Eurytelma californicum* hemocyanin. *J Comp Physiol*. 1978;128:161–8.
  56. Kusche K, Burmester T. Diplopod hemocyanin sequence and the phylogenetic position of the Myriapoda. *Mol Biol Evol*. 2001;18(8):1566–73.
  57. Burmester T, Hankeln T. The respiratory proteins of insects. *J Insect Physiol*. 2007;53(4):285–94.
  58. Wirkner CS, Pass G. The circulatory system in Chilopoda: functional morphology and phylogenetic aspects. *Acta Zool*. 2002;83:193–202.
  59. Markl J, Stöcker W, Runzler R, Precht E. Immunological correspondences between the hemocyanin subunits of 86 arthropods: evolution of a multigene protein family. In: Linzen B, editor. *Invertebrate oxygen carriers*. Heidelberg: Springer; 1986. p. 281–92.
  60. Marxen JC, Pick C, Oakley TH, Burmester T. Occurrence of hemocyanin in ostracod crustaceans. *J Mol Evol*. 2014;79(1–2):3–11.
  61. Pinnow P, Fabrizius A, Pick C, Burmester T. Identification and characterisation of hemocyanin of the fish louse *Argulus* (Crustacea: Branchiura). *J Comp Physiol B*. 2016;186(2):161–8.
  62. Ballweber P, Markl J, Burmester T. Complete hemocyanin subunit sequences of the hunting spider *Cupiennius salei*: recent hemocyanin remodeling in entelegyne spiders. *J Biol Chem*. 2002;277(17):14451–7.
  63. Ertas B, von Reumont BM, Wagele JW, Misof B, Burmester T. Hemocyanin suggests a close relationship of remipedia and hexapoda. *Mol Biol Evol*. 2009;26(12):2711–8.
  64. Rehm P, Borner J, Meusemann K, von Reumont BM, Simon S, Hadrys H, Misof B, Burmester T. Dating the arthropod tree based on large-scale transcriptome data. *Mol Phylogenet Evol*. 2011;61(3):880–7.
  65. Burmester T. Expression and evolution of hexamerins from the tobacco hornworm, *Manduca sexta*, and other Lepidoptera. *Insect Biochem Mol Biol*. 2015;62:226–34.
  66. Burmester T, Hankeln T. Function and evolution of vertebrate globins. *Acta Physiol (Oxford)*. 2014;211(3):501–14.
  67. Markl J, Decker H. Molecular structure of the arthropod hemocyanins. *Adv Comp Environm Physiol*. 1992;13:325–76.
  68. Sánchez D, Ganfornina MD, Gutierrez G, Bastiani MJ. Molecular characterization and phylogenetic relationships of a protein with potential oxygen-binding capabilities in the grasshopper embryo. A hemocyanin in insects? *Mol Biol Evol*. 1998;15(4):415–26.
  69. Pick C, Schneuer M, Burmester T. Ontogeny of hemocyanin in the ovoviparous cockroach *Blattella germanica* suggests an embryo-specific role in oxygen supply. *J Insect Physiol*. 2010;56(5):455–60.
  70. Chen B, Ma R, Ma G, Guo X, Tong X, Tang G, Kang L. Haemocyanin is essential for embryonic development and survival in the migratory locust. *Insect Mol Biol*. 2015;24(5):517–27.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

