Die Evolution der Hämocyaningene in Apogastropoda

Einflussfaktor oder Bedingung für die enorme Diversität der Heterobranchia und Caenogastropoda?

Dissertation

zur Erlangung des Grades
Doktor der Naturwissenschaften

Am Fachbereich Biologie
der Johannes Gutenberg-Universität Mainz

Gabriela Giannina Schäfer
geb. am 27.01.1992 in Wiesbaden

Mainz, April 2021

Dekan: [redacted]

1.Berichterstatter: [redacted]

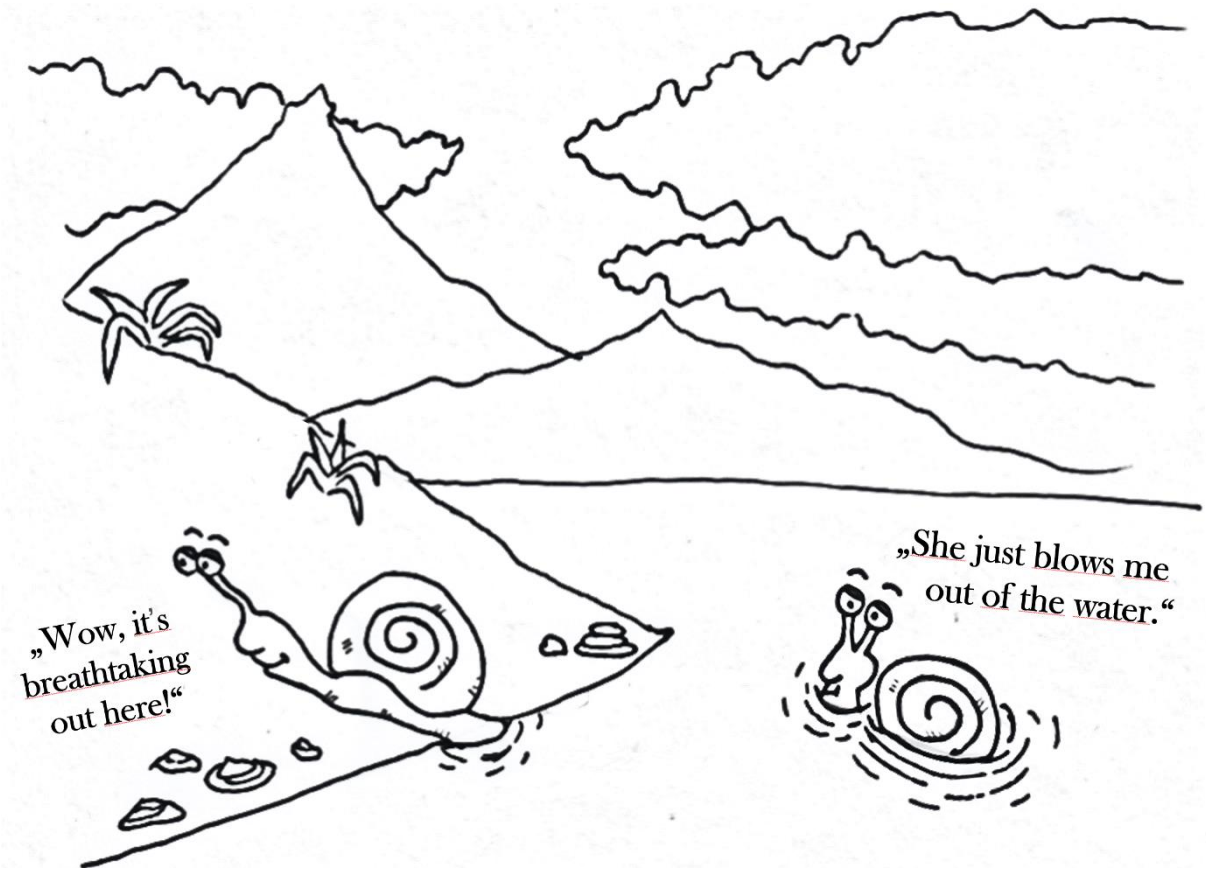
2.Berichterstatter: [redacted]

Tag der mündlichen Prüfung: 08.07.2021

für meine Mutter

„Im Licht der Evolution betrachtet ist die Biologie vielleicht die intellektuell befriedigendste und inspirierendste Wissenschaft.“

Theodosius Dobzhansky, 1973



Inhaltsverzeichnis

Zusammenfassung 1

Abstract 2

1 **Einleitung** 3

1.1 Apogastropoda 4

1.1.1 Heterobranchia 5

1.1.2 Caenogastropoda 7

1.1.3 Aufklärung der Phylogenie 9

1.2 Mollusken-Hämocyanin 10

1.2.1 Grundaufbau der Mollusken-Hämocyanine 10

1.2.2 Die Gene der Mollusken-Hämocyanine 12

1.2.3 Die Evolution der Mollusken-Hämocyaningene 14

1.3 Zielsetzung der Arbeit 17

2 **Publikationen** 21

2.1 Hemocyanin genes as indicators of habitat shifts in Panpulmonata 25

2.2 Hemocyanins of Muricidae: New 'Insights' Unravel an Additional Highly Hydrophilic 800 kDa Mass Within the Molecule 33

2.3 The evolution of hemocyanin genes in Tectipleura: a multitude of conserved introns in highly diverse gastropods 47

2.4 The evolution of hemocyanin genes in Caenogastropoda: Gene duplications and intron accumulation in highly diverse gastropods 67

2.5 Responsiveness of metallothionein and hemocyanin genes to cadmium and copper exposure in the garden snail *Cornu aspersum* 87

3 **Fazit & Ausblick** 101

4 **Literaturverzeichnis** 103

5 **Quellenverzeichnis** 109

6 **Danksagung** 111

7 **Lebenslauf** 113

8 **Erklärung** 115

Zusammenfassung

Mollusken bilden nach den Arthropoden den zweitgrößten Tierstamm und sind weltweit in nahezu allen Lebensräumen verbreitet. Sie umfassen acht Klassen, von denen die Gastropoda (Schnecken) die mit Abstand größte darstellt. Gastropoden sind sowohl in Salz- und Süßwasser als auch an Land und in diversen Übergangshabitaten zu finden und kommen sogar in extremen Lebensräumen wie der Tiefsee, heißen Quellen, Wüsten und der Arktis und Antarktis vor. Die enorme Diversität dieser Klasse lässt sich insbesondere auf viele Habitatswechsel und damit einhergehende, immens vielfältige Anpassungen der Heterobranchia und Caenogastropoda zurückführen. Diese bilden gemeinsam die Apogastropoda und umfassen über 60.000 der insgesamt etwa 68.000 bisher beschriebenen Gastropoda-Arten.

In diesem Zusammenhang habe ich in meiner Doktorarbeit die Evolution der Mollusken-Hämocyaningene innerhalb der Apogastropoda untersucht. Hämocyanine sind die Sauerstofftransportproteine der meisten Vertreter dieses Tierstamms und bestehen charakteristischer Weise aus ca. 400 kDa großen Untereinheiten, die sich zu Dekameren oder Di- und Multidekameren zusammenlagern. Als Sauerstofftransporter stellen diese enorm großen Proteine eine wichtige Verbindung zwischen der Umwelt der Tiere und ihrem Metabolismus dar und müssen wegen der Abhängigkeit ihrer Sauerstoffaffinität von der Temperatur und dem Sauerstoffpartialdruck an die Lebensbedingungen der Tiere angepasst sein. Aus diesem Grund war es das Ziel meiner Arbeit, die Hämocyaningene der Apogastropoda genauer zu analysieren und auf mögliche Adaptationen, die während der Evolution dieser Schnecken-Gruppe aufgetreten sind, zu untersuchen.

Meine Untersuchungen ergaben, dass sowohl innerhalb der Heterobranchia als auch innerhalb der Caenogastropoda mehrfach unabhängig voneinander Duplikationen der Hämocyaningene aufgetreten sind und zu verschiedenen paralogen Hämocyaningenen geführt haben. Weiterhin zeigen meine Ergebnisse eine strukturelle Besonderheit eines Hämocyanins der Muricidae (Caenogastropoda), die eine Abweichung von allen bisher beschriebenen Mollusken-Hämocyaninen darstellt und einen Einblick in die Evolution spezieller Hämocyanin-Strukturen bieten kann. Meine Analysen der Exon-Intron-Strukturen verschiedener Hämocyaningene der Tectipleura (Heterobranchia) und Caenogastropoda ergaben, dass sich während der Evolution der einzelnen Apogastropoda-Gruppen viele Introns innerhalb der Hämocyaningene gebildet haben und diese sich hierin stark von den Genarchitekturen anderer bisher analysierter Mollusken-Hämocyaningene (10 – 15 Introns) unterscheiden. Während die Hämocyanin-Genstrukturen der Caenogastropoda zwischen verschiedenen Gruppen, sowie zwischen paralogen Genen innerhalb derselben Art stark variieren (28 – 61 Introns), sind die der Tectipleura hoch konserviert (52/53 Introns). Meine Arbeit zeigt somit viele evolutive Erneuerungen der Hämocyaningene innerhalb der Apogastropoda, die mehrfach unabhängig voneinander aufgetreten sind und demnach mögliche Anpassungen an neue Lebensräume darstellen.

Abstract

Molluscs form the second largest animal phylum after the arthropods and are distributed in almost all habitats worldwide. They comprise eight classes, of which the Gastropoda (including snails and slugs) is by far the largest. Gastropods inhabit marine, limnic and terrestrial biotopes as well as various transitional habitats and even occur in extreme habitats such as the deep sea, hot springs, deserts and the Arctic and Antarctic. The enormous diversity of this class can be attributed in particular to many habitat shifts and the diverse adaptations of Heterobranchia and Caenogastropoda. Together, these two groups of gastropods form the Apogastropoda and comprise over 60,000 of the approximately 68,000 gastropod species described so far.

In this context, I investigated the evolution of the molluscan hemocyanin genes within the Apogastropoda during my doctoral thesis. Hemocyanins are the oxygen transport proteins of most molluscs and characteristically consist of 400 kDa subunits which form decamers or di- and multi-decamers. As oxygen transporters, these extremely large proteins represent an important link between the animals' environment and their metabolism. Due to the dependence of their oxygen affinity on temperature and oxygen partial pressure, they must be well adapted to the living conditions of the animals. For this reason, the aim of my work was to analyze hemocyanin genes of Apogastropoda in more detail and to examine possible adaptations that occurred during the evolution of this group of gastropods.

My investigations showed that within Heterobranchia as well as within Caenogastropoda duplications of hemocyanin genes occurred several times independently and led to different paralogous hemocyanin genes. Furthermore, my results show a structural peculiarity of a hemocyanin of Muricidae (Caenogastropoda) which represents a deviation from all previously described molluscan hemocyanins and can offer an insight into the evolution of special hemocyanin structures. My analyses of the exon-intron structures of various hemocyanin genes of Tectipleura (Heterobranchia) and Caenogastropoda showed that during the evolution of the individual Apogastropoda groups many introns were gained within these genes. Thus, hemocyanin gene architectures of Apogastropoda differ remarkably from those of other molluscan hemocyanin genes analyzed so far (10 – 15 introns). While the hemocyanin gene structures of Caenogastropoda vary widely between different groups as well as between paralogous genes within the same species (28-61 introns), those of Tectipleura are highly conserved (52/53 introns). Thus, my thesis shows many evolutionary renewals of the hemocyanin genes within the Apogastropoda which occurred several times independently and therefore represent possible adaptations to new habitats.

1

Einleitung

Weichtiere oder Mollusca (lateinisch *molluscus* „weich“) sind eine äußerst artenvielfältige, formenreiche und in nahezu allen Lebensräumen vorkommende Tiergruppe. Sie bilden nach den Gliederfüßern (Arthropoda) den zweitgrößten Stamm im Tierreich und umfassen insgesamt acht Klassen (Abb. 1A; Kocot et al., 2011; Smith et al., 2011) – darunter Kopffüßer (Cephalopoda), Muscheln (Bivalvia) und Käferschnecken (Polyplacophora). Die bei weitem größte Klasse der Mollusken stellen die Schnecken (Gastropoda) dar (Abb. 1A). Ihre enorme Artenvielfalt lässt sich unter anderem damit erklären, dass sie als einzige Klasse innerhalb der Weichtiere – zusätzlich zu Salz- und Süßwasser – das Land als Lebensraum erobert haben (Abb. 1B; vgl. Ponder et al., 2019). Das Vorkommen in so vielen verschiedenen Habitaten, die sich in ihren Extremen von der Tiefsee über Wüsten bis in die Polarmeere erstrecken, lässt erkennen, dass während der Evolution der verschiedenen Gastropoda-Gruppen vielfache physiologische Veränderungen sowie Verhaltens-anpassungen aufgetreten sein müssen (Mordan und Wade, 2008). Je nach Lebensraum unterscheiden sich beispielsweise die Anforderungen an Osmoregulation und Temperaturanpassung der Tiere ganz erheblich (vgl. Vermeij und Dudley, 2000; Symanowski und Hildebrandt, 2010; Chao et al., 2020). Gleichzeitig erfordern die verschiedenen Habitate unterschiedliche Formen der Bewegung und der Atmung (Vermeij und Wesselingh F. P., 2002). Hierbei sind neben strukturellen Veränderungen der Organe (z.B. Evolution einer Lunge) auch molekularphysiologische Aspekte wichtig, wie die Adaptation einzelner Proteine an veränderte Anforderungen (Hsia et al., 2013). Ziel meiner Doktorarbeit war deshalb die Untersuchung eines für die Atmung enorm wichtigen Proteins – dem Sauerstofftransporter Hämocyanin – welches ich im Hinblick auf seine Evolution in der besonders diversen Schneckengruppe der Apogastropoda genauer analysieren wollte. Die Hintergründe hierzu werde ich im Folgenden eingehender erläutern und präzisieren.

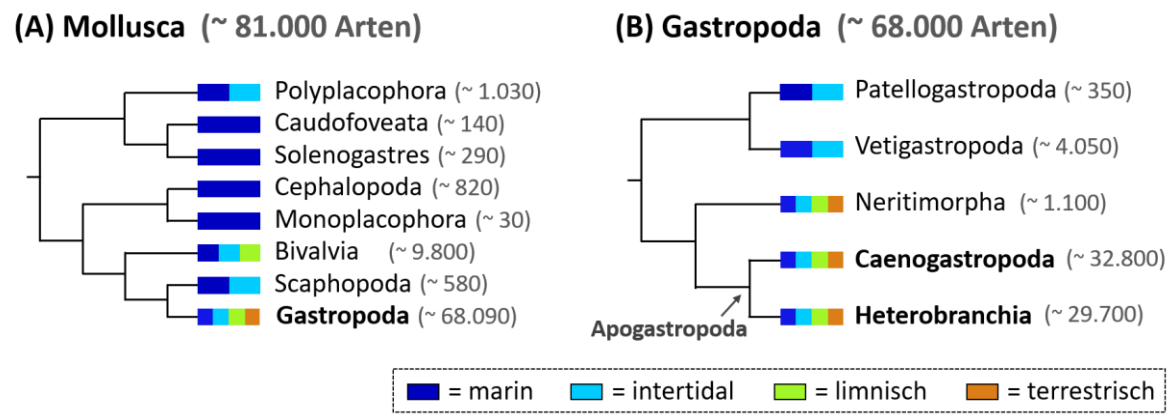


Abb. 1: Systematik der Weichtiere (Mollusca). Neben den phylogenetischen Zusammenhängen zeigen die Stammbäume die ungefähre Anzahl der beschriebenen rezenten Arten sowie die von der Gruppe besiedelten Lebensräume unterteilt in marin (dunkelblau), intertidal (hellblau), limnisch (grün) und terrestrisch (braun). Die Anzahl der Arten sind in Klammern hinter dem jeweiligen Taxon angegeben und basieren auf den Angaben der online Datenbank WoRMS – *World Register of Marine Species* (2020). (**A**) Der Stamm der Mollusken umfasst insgesamt acht Klassen (Kocot et al., 2011; Smith et al., 2011) und knapp 81.000 rezente Arten. Die mit Abstand größte Molluskengruppe stellen die Gastropoda dar (**B**), die wiederum fünf Großgruppen umfasst (Zapata et al., 2014). In dieser Arbeit habe ich Hämocyanine der Heterobranchia und Caenogastropoda untersucht. Diese bilden gemeinsam die Gruppe der Apogastropoda. Die aktuell anerkannte Systematik beider Gruppen sind in Abb. 3 angegeben.

1.1 Apogastropoda

Phylogenetische Analysen innerhalb der Gastropoda werden durch deren rasante Radiation, die zu der enormen Diversität dieser Gruppe geführt hat, und das Aussterben diverser Schneckenlinien erschwert, weshalb die Systematik innerhalb dieser Klasse der Mollusken bis heute kontrovers diskutiert wird (Colgan et al., 2007; Ponder et al., 2008; 2019). Unstrittig ist jedoch, dass ihre traditionelle Einteilung in die drei Gruppen Prosobranchia (Vorderkiemer), Opisthobranchia (Hinterkiemer) und Pulmonata (Lungenschnecken) phylogenetisch nicht korrekt ist, da diese zwar Organisationsebenen angeben, jedoch keine Monophyla darstellen (Haszprunar, 1988; Ponder und Lindberg, 1997).

Die Gruppe der Apogastropoda wurde in der 1997 von Ponder und Lindberg veröffentlichten und auf morphologischen Merkmalen basierenden Systematik als ein monophyletisches Taxon eingeführt, das die zwei Monophyla Caenogastropoda und Heterobranchia umfasst. Molekular-phylogenetische Analysen bestätigen diese systematischen Zusammenhänge (Aktipis et al., 2008; Zapata et al., 2014; Cunha und Giribet, 2019). Heterobranchia und Caenogastropoda stellen die zwei größten der insgesamt fünf Schneckengruppen dar (siehe Artenanzahl in Abb. 1B), sodass über 60.000 der insgesamt über 68.000 beschriebenen rezenten Schneckenarten den Apogastropoda angehören (WoRMS Editorial Board, 2020). Während die drei anderen Schneckengruppen Patellogastropoda, Vetigastropoda und Neritimorpha eine geringere Artenanzahl sowie eine kleinere Formen- und Habitatvielfalt aufweisen (mit Ausnahmen einiger terrestrischer Neritimorpha ausschließlich marin

oder intertidal lebend), sind Heterobranchia und Caenogastropoda sehr divers und haben fast jeden Lebensraum von der Tiefsee bis in hohe Gebirgslagen besiedelt (Überblick in Ponder et al. 2019).

1.1.1 Heterobranchia

Der Name Heterobranchia setzt sich aus dem altgriechischen ἕτερος (hetero) = verschieden und dem griechischen βράγχια (bránchia) = Kieme zusammen und kann deshalb mit Verschiedenkiemer übersetzt werden (Gosselck et al., 2009). Bereits 1840 von Gray erwähnt, erhielt dieser Begriff Mitte der 1980er Jahre Einzug in die moderne Systematik der Gastropoda. Haszprunar (1985, 1988) wählte ihn dabei zur Beschreibung seines neuen phylogenetischen Konzepts, welches viele der traditionell als Opisthobranchia und Pulmonata bezeichneten Schnecken sowie einige Familien der Prosobranchia gemeinsam als eine monophyletische Gruppe beschreibt[1]. Diese Bezeichnung spiegelt die vielfältigen Anpassungen an unterschiedliche Habitate wider, die diese Gruppe zu der wohl vielfältigsten Gruppe der Schnecken machen. Die Heterobranchia umfassen zwar eine geringere Artenzahl als die Caenogastropoda (Abb. 1B), sind morphologisch allerdings noch diverser und weisen eine enorme Vielfalt unterschiedlichster Farben, Formen und besonderer Merkmale auf (Beispiele in Abb. 2). Ihre Gehäuse beispielsweise variieren nicht nur in Größe, Form und Farbe, sondern sind auf vielfältige Weise in unterschiedlichen Gruppen der Heterobranchia unabhängig voneinander verloren gegangen, stark zurückgebildet oder durch Überwachsung des Mantels nach innen verlagert worden. Während ihrer Evolution haben Heterobranchia vielfach unabhängig voneinander die verschiedensten Lebensräume erobert und umfassen mit Abstand die meisten landlebenden Gastropoda.

Obwohl viele phylogenetische Zusammenhänge innerhalb dieser sehr diversen Schneckengruppe noch ungeklärt sind, werden die einzelnen Ordnungen der Heterobranchia mittlerweile in durch mehrere Analysen gut unterstützte Großgruppen eingeteilt (Jörger et al., 2010; Kano et al., 2016). Danach gliedern sie sich in sehr basale („lower“) Heterobranchia und die Euthyneura, die wiederum in Acteonacea, Ringipleura und Tectipleura unterteilt werden (siehe Abb. 3A). Die enorme Arten- und Lebensraumvielfalt geht dabei insbesondere auf die Gruppe der Tectipleura (über 26.000 Arten nach den Angaben der online Datenbank WoRMS – World Register of Marine Species (2020); Abb. 3A) zurück, die sich in Euopisthobranchia und Panpulmonata gliedern (Jörger et al., 2010; Kocot et al., 2013; Wägele et al., 2014; Kano et al., 2016). Die von mir durchgeführten Untersuchungen von Heterobranchia Hämocyaninen beschränken sich deshalb auf diese äußerst vielfältige Schneckengruppe.

[1] Zuvor wurden bereits mehrfach große Ähnlichkeiten zwischen Opisthobranchia und Pulmonata beschrieben und einige Linien bereits unter dem Namen „Euthyneura“ zusammengefasst (z.B. in Boettger (1954) und Hyman L. H. (1967))

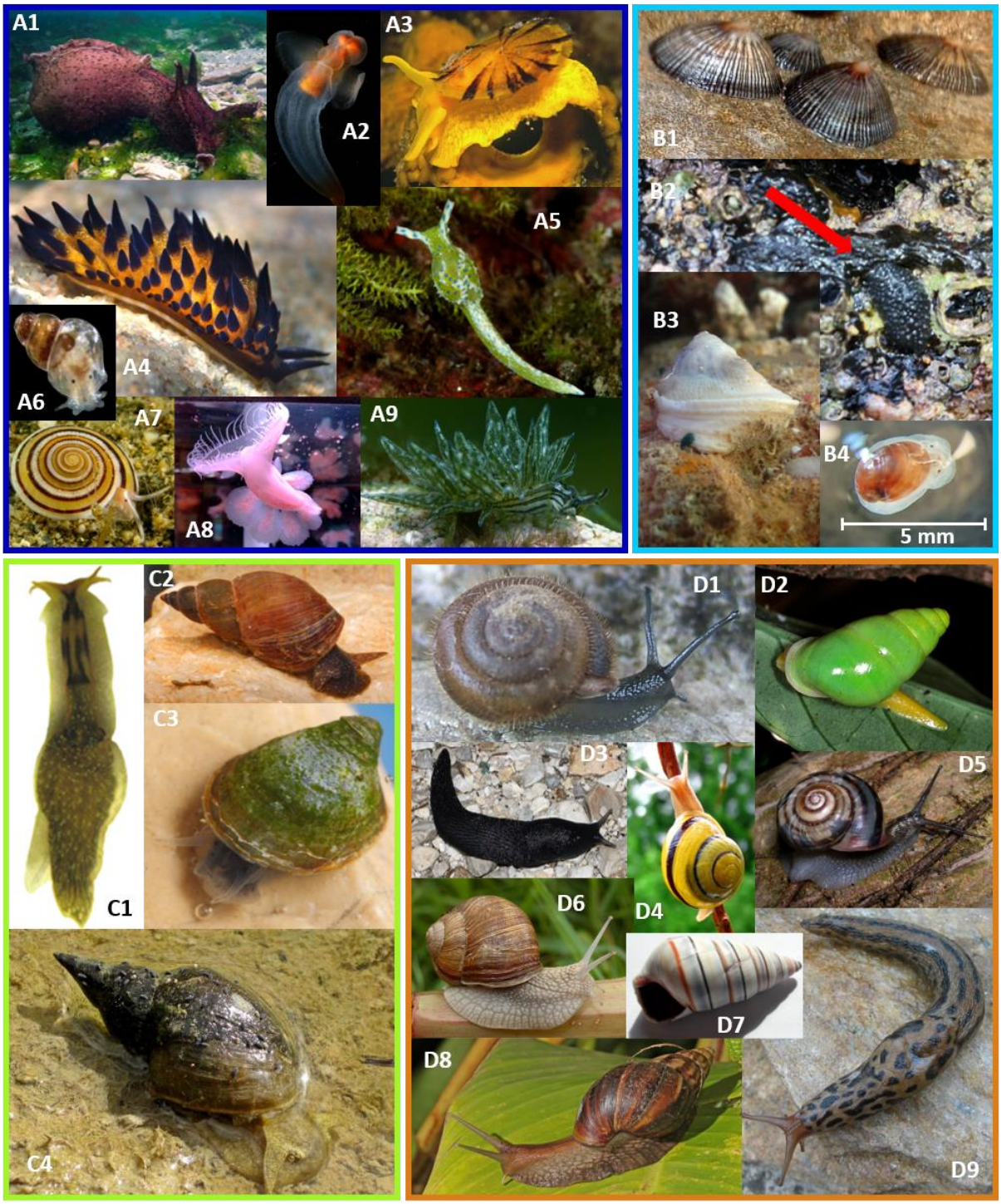


Abb. 2: Beispiele einiger Heterobranchia. Die abgebildeten Schnecken stellen eine kleine Auswahl unterschiedlicher Heterobranchia-Arten dar, die die Vielfältigkeit dieser Gruppe verdeutlichen soll, jedoch keinerlei Vollständigkeit beansprucht. Die abgebildeten Schnecken stammen aus verschiedenen Gruppen der Heterobranchia und sind unabhängig von phylogenetischen Zusammenhängen nach ihren Habitaten angeordnet und farblich eingerahmt: Meer (dunkelblau), Gezeitenzone (hellblau), Süßwasser (grün), Land (braun). Für diese Lebensräume sind folgende Schnecken exemplarisch abgebildet: marine Schnecken: *Aplysia californica* (**A1**), *Clione limacina* (**A2**), *Tylodina perversa* (**A3**), *Placida cremoniana* (**A4**), *Oxynoe viridis* (**A5**), *Rissoella opalina* (**A6**), *Architectonica perspectiva* (**A7**), *Melibe leonina* (**A8**), *Aplysiopsis elegans* (**A9**); intertidal lebende Schnecken: *Siphonaria spec.* (**B1**), *Onchidella celtica* (**B2**), *Trimusculus mauritianus* (**B3**), *Otina ovata* (**B4**); limnische Schnecken: *Acochlidium fijiiensis* (**C1**), *Stagnicola palustris* (**C2**), *Ancylus fluviatilis* (**C3**), *Lymnaea stagnalis* (**C4**); terrestrische Schnecken: *Isognomostoma isognomostomos* (**D1**), *Edentulina spec.* (**D2**), *Limax cinereoniger* (**D3**), *Cepaea nemoralis* (**D4**), *Euhadra peliomphala* (**D5**), *Helix pomatia* (**D6**), *Liguus virigineus* (**D7**), *Achatina fulica* (**D8**), *Limax maximus* (**D9**) (alle Bildnachweise im Quellenverzeichnis).

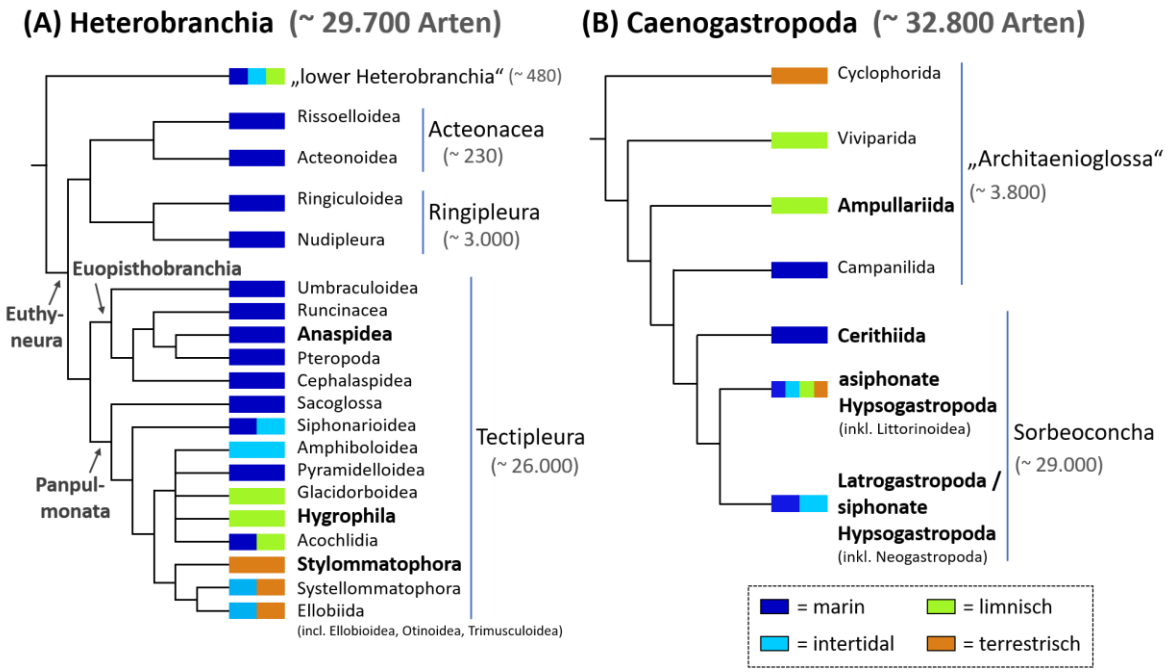


Abb. 3: Systematik der Apogastropoda-Gruppen Heterobranchia und Caenogastropoda. Neben den phylogenetischen Zusammenhängen zeigen die Stammbäume die ungefähre Anzahl rezenter Arten sowie die von der jeweiligen Gruppe besiedelten Lebensräume unterteilt in marin (dunkelblau), intertidal (hellblau), limnisch (grün) und terrestrisch (braun). Die Artenanzahl ist in Klammern hinter dem jeweiligen Taxon angegeben und basiert auf den Angaben der online Datenbank WoRMS (2020). (**A**) Die Gruppe der Heterobranchia umfasst ca. 29.700 beschriebene rezente Arten, die sich neben den „lower Heterobranchia“ in Acteonacea, Ringipleura und Tectipleura gliedern lassen (Jörger et al., 2010; Kocot et al., 2013; Kano et al., 2016). In meiner Arbeit habe ich einige Hämocyanine aus der Gruppe der Tectipleura untersucht. Diese umfasst die beiden Monophyla Panpulmonata und Euopisthobranchia. (**B**) Die Gruppe der Caenogastropoda ist mit ca. 32.800 rezenten beschriebenen Arten die größte der fünf Großgruppen der Schnecken. Sie lässt sich in die monophyletische Gruppe der Sorbeoconcha und die paraphyletischen „Architaenioglossa“ unterteilen (vgl. Ponder und Lindberg, 1997, Colgan et al. , 2007, Bouchet et al. 2017 und Ponder et al. 2019). Die exakten phylogenetischen Zusammenhänge der einzelnen Großgruppen zueinander konnten trotz vielfacher Analysen bis heute nicht eindeutig aufgelöst werden. Die hier abgebildete Phylogenie basiert auf dem Stammbaum nach Ponder et al. (2019), der aus den Ergebnissen aktueller morphologischer und molekularer Analysen abgeleitet wurde. Untersucht wurden Hämocyanine aus den fett gedruckten Heterobranchia- und Caenogastropoda-Gruppen.

1.1.2 Caenogastropoda

Die Bezeichnung Caenogastropoda wurde bereits 1960 von Cox zur Beschreibung der Systematik der Gastropoda verwendet. Allerdings fasste er darunter einige zueinander paraphyletische Prosobanchia-Gruppen zusammen. In seiner heutigen Bedeutung als eine der fünf großen monophyletischen Schneckengruppen (Zapata et al., 2014) wurde der Begriff 1988 von Haszprunar und 1997 von Ponder und Lindberg eingeführt. Caenogastropoda stellen die artenvielfältigste Gruppe der Gastropoda dar (~ 32.800 beschriebene rezente Arten nach WoRMS Editorial Board, 2020; Abb. 1B). Ihre Vertreter haben zumeist gut erkennbare Schneckenhäuser unterschiedlichster Formen und Farben, die in Größen zwischen 1 und 900 mm vorkommen (Abb. 4). Während ein Großteil der Caenogastropoda marin lebt, gibt es einige Gruppen, die unabhängig voneinander Süßwasser und/oder Land besiedelt und sich dort verbreitet haben (Abb. 3B; vgl. Übersicht in Ponder et al., 2019).

Abb. 4: Beispiele einiger Caenogastropoda. Die abgebildeten Schnecken stellen eine kleine Auswahl unterschiedlicher Caenogastropoda-Arten dar, die die Vielfältigkeit dieser Gruppe verdeutlichen soll, jedoch keinerlei Vollständigkeit beansprucht. Die abgebildeten Schnecken stammen aus verschiedenen Gruppen der Caenogastropoda und sind unabhängig von phylogenetischen Zusammenhängen nach ihren Habitaten angeordnet und farblich eingerahmt: Meer (dunkelblau), Gezeitenzone (hellblau), Süßwasser (grün), Land (braun). Für diese Lebensräume sind folgende Schnecken exemplarisch abgebildet: marine Schnecken: *Lentigo lentiginosus* (**A1**), *Cyphoma gibbosum* (**A2**), *Phenacovolva rosea* (**A3**), *Chicoreus palmarosae* (**A4**), *Lambis scorpius* (**A5**), *Dentiovula dorsuosa* (**A6**), *Arestorides argus* (**A7**), *Rapana venosa* (**A8**); intertidal lebende Schnecken: *Telescopium telescopium* (**B1**), *Batillaria estuarina* (**B2**), *Nodilittorina pyramidalis* (**B3**), *Crepidula fornicata* (**B4**), *Nucella lapillus* (**B5**); limnische Schnecken: *Pomacea bridgesii* (**C1**), *Sadleriana fluminensis* (**C2**), *Oncomelania hupensis* (**C3**), *Melanoides tuberculata* (**C4**), *Anentome helena* (**C5**), *Melanopsis praemorsa* (**C6**), *Emmericia patula* (**C7**), *Elimia clara* (**C8**); terrestrische Schnecken: *Pupinella rufa* (**D1**), *Rhiostoma huberi* (**D2**), *Cochlostoma auritum* (**D3**), *Terebralia palustris* (**D4**), *Pomatias elegans* (**D5**), *Cochlostoma septemspirale* (**D6**), *Cyclophorus perdix* (**D7**) (alle Bildnachweise im Quellenverzeichnis).

Die Unterteilung der Caenogastropoda in einige Großgruppen wie z.B. Ampullariida und Neogastropoda ist aufgrund morphologischer (Strong, 2003; Simone, 2011) und molekularer (Harasewych et al., 1998; Colgan et al., 2007) Stammbäume mehrfach bestätigt. Die Phylogenie der einzelnen Gruppen untereinander ist hingegen oftmals nicht eindeutig aufgelöst. Insbesondere innerhalb der Großgruppe der Hypsogastropoda, die beispielsweise die Ordnungen Neogastropoda und Littorinida einschließt, ist außer der Unterteilung in zwei Monophyla (solche mit und solche ohne Siphon, siehe Abb. 3B) bis heute keine genaue Auflösung der Systematik möglich (Colgan et al., 2007; Ponder et al., 2008; 2019). Der in Abb. 3B dargestellte Stammbaum entspricht der Phylogenie nach Ponder et al. (2019), die auf der gemeinsamen Betrachtung verschiedener Ergebnisse aktueller morphologischer und molekularen Analysen basiert.

1.1.3 Aufklärung der Phylogenie

Die rasante Radiation der Heterobranchia und Caenogastropoda und die vielfach unabhängig voneinander aufgetretenen evolutiven Veränderungen, die zu der enormen Diversität der Apogastropoda geführt haben, erschweren die Rekonstruktion der phylogenetischen Zusammenhänge (Wägele et al., 2008; Ponder et al., 2008; Mordan und Wade, 2008; Zapata et al., 2014; Frýda, 2021). Zur Erforschung der äußerst komplexen Systematik und der weiteren Aufklärung der Evolution dieser vielfältigen Schneckengruppe wurden bereits vielfach phylogenetische Analysen mit sowohl morphologischen als auch molekularen Markern durchgeführt. Diese haben bisher jedoch zu keinen einheitlichen Ergebnissen bezüglich der genauen systematischen Beziehungen innerhalb der jeweiligen Gruppen geführt (vgl. Haszprunar, 1988; Bang et al., 2000; Simone, 2001; Strong, 2003; Collin, 2005; Colgan et al., 2007; Wägele et al., 2008; Jörger et al., 2010; Dinapoli und Klussmann-Kolb, 2010; Zapata et al., 2014; Wägele et al., 2014; Kano et al., 2016; Romero et al., 2016; Bouchet et al., 2017; Cunha und Giribet, 2019). Um Unstimmigkeiten zwischen den einzelnen Ergebnissen unterschiedlicher Analysen aufklären zu können, müssen nach Medina und Collins (2003), Wägele et al. (2008) und Sigwart und Lindberg (2015) immer mehr solcher unterschiedlicher Methoden kombiniert sowie verschiedene Erkenntnisse sukzessive zusammengetragen und miteinander ausgewertet werden. Neben ribosomalen und mitochondrialen Markern müssen hierbei auch nukleäre Gene phylogenetisch untersucht werden, um die Informationsdichte bezüglich evolutiver Veränderungen der Mollusken zu erhöhen (Kocot et al., 2013). In meiner Arbeit habe ich deshalb die Analysen der Mollusken-Hämocyanine, die als molekulare Marker genutzt werden können (Streit et al., 2006; Lieb und Todt, 2008; Wägele et al., 2008; Warnke et al., 2011; Kocot et al., 2011), auf Apogastropoda ausgeweitet, um sie auf mögliche neue Informationen zur Evolution dieser Schneckengruppe zu untersuchen.

1.2 Mollusken-Hämocyanin

Hämocyanin ist das Sauerstofftransport-Protein der meisten Mollusken und Arthropoden. In beiden Tierstämmen liegt es extrazellulär in der Hämolymphe vor und färbt diese blau (van Holde und Miller, 1995). Grund für diese Blaufärbung ist die Ausbildung von Histidin-Kupfer(II)-Komplexen in oxygenierten Hämocyanin-Molekülen, deren Grundstruktur bei den Hämocyaninen der beiden Tierstämme gleich ist. Dabei liegen jeweils zwei Kupfer-Ionen vor, die über sechs Histidin-Reste komplexiert sind und je ein Sauerstoff-Molekül reversibel binden können (van Holde und Miller, 1995). Abgesehen von diesen Kupferzentren unterscheiden sich die Hämocyanine der Mollusken und der Arthropoden strukturell sehr stark (sowohl in der Primär-, als auch der Tertiär- und Quartärstruktur) und werden als zwei getrennte Proteinfamilien angesehen (Burmester, 2001). In meiner Arbeit habe ich ausschließlich Mollusken-Hämocyanine untersucht.

1.2.1 Grundaufbau der Mollusken-Hämocyanine

Neben einigen zum Teil gruppenspezifischen Unterschieden der Mollusken-Hämocyanine ist der im Folgenden beschriebene prinzipielle Aufbau dieses Sauerstofftransporters hoch konserviert (vgl. Übersichtsartikel zur Struktur der Mollusken-Hämocyanine: van Holde und Miller, 1995; Markl, 2013; Kato et al., 2018). Die Grundstruktur bildet dabei ein partiell gefüllter Hohlzylinder, der aus zehn gleichen Untereinheiten besteht (Abb. 5A). Mit einer Höhe von ca. 18 nm und einem Durchmesser von etwa 35 nm weisen diese 4 MDa Hämocyanin-Dekamere Dimensionen eines mittelgroßen Viruspartikels auf und können mittels Transmissionselektronenmikroskopie sichtbar gemacht werden. Die elektronenmikroskopische Aufnahme in Abbildung 5B zeigt, dass Hämocyanine in der Hämolymphe von Mollusken sowohl als einzelne Dekamere als auch als Di-, Tri- oder Multidekamere vorliegen. Hämocyanin-Dekamere weisen dabei immer eine fünffache Symmetrie auf, die darauf zurückzuführen ist, dass Hämocyanin-Untereinheiten zunächst zu Dimeren aggregieren, die die eigentliche repetitive Einheit der Dekamere darstellen (siehe gelb/goldene Hervorhebung in Abb. 5A, Siezen und van Bruggen, 1974; Markl, 2013; Gai et al., 2015).

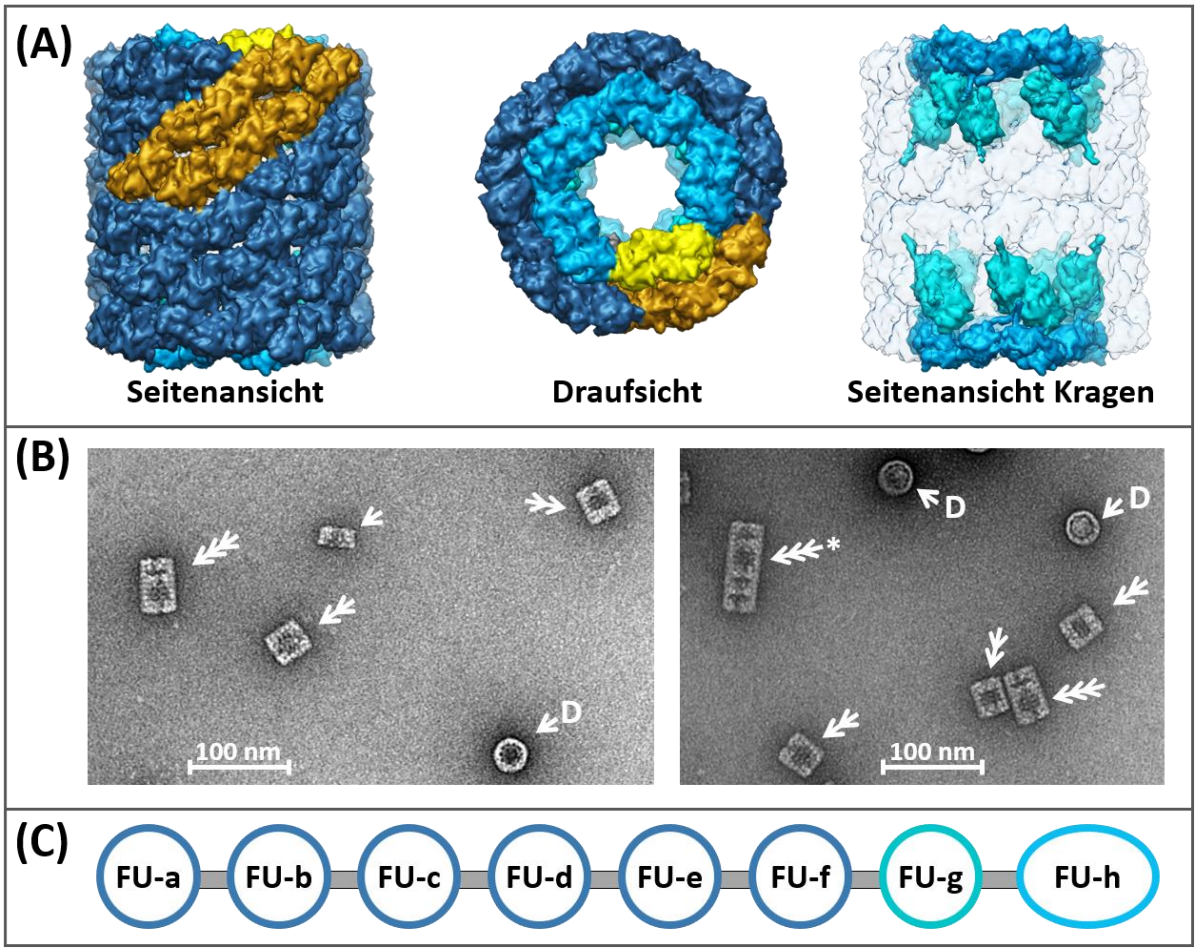


Abb. 5: Grundaufbau der Gastropoda-Hämocyanine. Die Grundstruktur der Mollusken-Hämocyanine stellen Dekamere dar, die partiell gefüllte Hohlzylinder ausbilden und sich häufig zu Didekameren, aber auch zu Tri- oder Multidekameren zusammenlagern. (**A**) zeigt den Aufbau eines Didekamers. Dieser besteht aus einer äußeren Wand (dunkelblau) und einem in die Mitte des Hohlzylinders ragenden Kragen (hellblau und cyanfarben). Die repetitive Grundeinheit stellen hierbei Dimere dar, die zuerst aggregieren, bevor sich fünf von ihnen zu Dekameren zusammenlagern. Ein solches Dimer ist in der Seitenansicht und der Draufsicht jeweils golden (Wand) und gelb (Kragen) hervorgehoben. Die Abbildung des Didekamers basiert auf der Elektronendichtekarte (9 Å) des Hämocyanins KLH1 der großen Kalifornischen Schlüssellochschnecke *Megathura crenulata* und ist verändert nach Schäfer et al. (2021). Die transmissionselektronen-mikroskopischen Aufnahmen in (**B**) zeigen verschiedene Hämocyanin-Moleküle der Schneckenart *Nucella lapillus*, die für Gastropoden-Hämocyanine charakteristisch sind. Die Anzahl der nebenstehenden Pfeile beschreibt die Art der Moleküle: Einzelpfeil = Dekamer; Doppelpfeil = Didekamer; Trippelpfeil = Tridekamer; Trippelpfeil mit Sternchen = Multidekamer; Pfeil mit „D“ = Draufsicht. In (**C**) ist der charakteristische Aufbau einer Hämocyanin-Untereinheit (engl. *subunit*) mit acht funktionellen Einheiten (FU-a bis FU-h; FU aus dem Englischen *functional unit*) dargestellt. Diese einzelnen Proteindomänen bilden globuläre Substrukturen, die über kurze Linkerpeptide (grau dargestellt) miteinander verbunden sind. Die Farbgebung entspricht der in (**A**), da FU-a bis FU-f der zehn Untereinheiten eines Dekamers die Wand des Hohlzylinders formen und FU-g und FU-h den Kragen ausbilden.

Die einzelnen 400 kDa Untereinheiten der Mollusken-Hämocyanine umfassen im Grundaufbau acht paraloge Domänen, die als *funktionelle Einheiten* FU-a bis FU-h (FU aus dem Englischen *functional unit*) bezeichnet werden. Sie bilden jeweils eine globuläre Substruktur der Polypeptidkette und sind durch kurze sogenannte Linkerpeptide miteinander verbunden (Abb. 5C; Lang, 1988; van Holde und Miller, 1995). Jede dieser FUs beinhaltet dabei eine wie oben beschriebene zwei Kupfer-Ionen umfassende Sauerstoff-Bindungsstelle (van Holde und Miller, 1995). Die Anzahl der funktionellen Einheiten weicht aufgrund von Verlusten und/oder Duplikationen einiger dieser FUs bei

verschiedenen Mollusken-Hämocyaninen vom Grundaufbau ab (Übersicht in Markl, 2013). Cephalopoda-Hämocyanine enthalten beispielsweise keine FU-h (Lang, 1988; van Holde und Miller, 1995; Miller et al., 1998; Bergmann et al., 2006; Thonig et al., 2014). Der Grundaufbau der Gastropoda-Hämocyanine, auf die in meiner Arbeit der Fokus gerichtet ist, umfasst die oben beschrieben regulären acht FUs (van Holde und Miller, 1995).

Die einzelnen funktionellen Einheiten bestehen aus etwa 420 Aminosäuren, die zu einer molekularen Masse von ca. 45 bis 50 kDa pro FU führen. Eine Ausnahme stellt dabei die ca. 60 kDa schwere FU-h dar, die C-terminal etwa 100 zusätzliche Aminosäuren umfasst (Markl, 2013). Der Aufbau verschiedener funktioneller Einheiten ist sehr ähnlich, da viele Bereiche der Primärstruktur, insbesondere solche, die am Aufbau der Kupferbindungsstellen beteiligt sind, hoch konserviert sind. Mittels multipler Sequenzvergleiche (Sequenzalignments) zeigten Miller et al. (1998), dass diese stark konservierten Sequenzmotive in allen paralogen FUs eines Hämocyanins von *Enteroctopus dofleini* vorhanden sind. Dies unterstützt die bereits 1995 von van Holde und Miller aufgestellte Hypothese, dass die acht Domänen umfassende Hämocyanin-Untereinheit durch Genduplikationen eines monomeren Sauerstofftransport-Proteins, welches einer funktionellen Einheit entspricht, entstanden ist. Drei aufeinanderfolgende Genduplikationen mit anschließenden Genfusionen können dabei zu der Mollusken-Hämocyanin-Untereinheit mit insgesamt acht FUs geführt haben (van Holde und Miller, 1995). Weitere Analysen ergaben, dass die konservierten Sequenzbereiche auch in Hämocyaninen weiterer Cephalopoda-Arten und Vertretern anderer Molluskenklassen wie z.B. den Gastropoda, Bivalvia und Caudofoveata, enthalten sind (Lieb et al., 1999; 2000; 2004; Altenhein et al., 2002; Bergmann et al., 2006; 2007; Lieb und Todt, 2008). Darüber hinaus zeigten multiple Sequenzalignments, dass die einzelnen funktionellen Einheiten, die sich bezüglich ihrer Position in den verschiedener Mollusken-Hämocyanine entsprechen (z.B. FU-c der Hämocyanine HtH und OdH), höhere Sequenzidentitäten aufzeigen, als paraloge FUs des gleichen Hämocyanins (z.B. FU-a, FU-b, … FU-h von HtH). Sie sind daher als ortholog zueinander anzusehen (Lieb et al., 1999; 2000; Lieb und Markl, 2004). Demnach wird angenommen, dass ein Hämocyanin mit acht funktionellen Einheiten bereits vor der Aufspaltung der Mollusken in unterschiedliche Klassen in einem gemeinsamen Vorfahren entstanden ist (Lieb et al., 1999; 2000).

1.2.2 Die Gene der Mollusken-Hämocyanine

Die Untergliederung der Polypeptidkette in globuläre funktionelle Einheiten spiegelt sich bereits in der Struktur der Gene wider. Der Grundaufbau eines für ein 400 kDa Hämocyanin kodierenden Gens besteht dabei aus acht Exons, die jeweils für eine funktionelle Einheit kodieren und durch sogenannte *Linker-Introns* voneinander getrennt sind (Lieb et al., 2001). Die Position dieser stets in Phase 1 vorliegenden Linker-Introns ist über alle bisher untersuchten Mollusken-Hämocyanine hinweg konserviert und liegt innerhalb der für die Linkerpeptide kodierenden Genabschnitte

(Abb. 5C; Lieb et al., 2001). Innerhalb der für die einzelnen funktionellen Einheiten kodierenden Sequenzabschnitte können dabei weitere Introns liegen, die diese Abschnitte in mehrere Exons unterteilen. Diese Introns werden als *interne Introns* bezeichnet und unterscheiden sich zwischen verschiedenen Mollusken-Gruppen sowohl in Anzahl, Position und Intronphase voneinander (Abb. 6; Lieb et al., 2001; Altenhein et al., 2002; Bergmann et al., 2006; Yao et al., 2019).

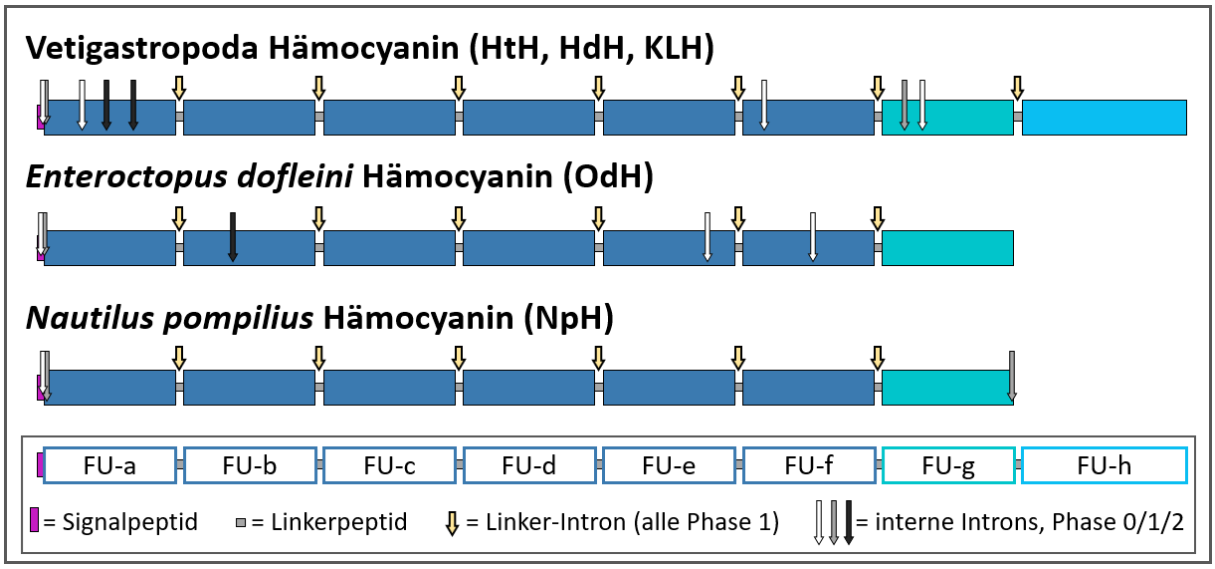


Abb. 6: Exon-Intron-Strukturen der Mollusken-Hämocyanine. Dargestellt ist die Position der Introns in Bezug auf die kodierende Sequenz der Hämocyaningene. Die einzelnen FU-kodierenden Bereiche der Mollusken-Hämocyanine sind klassenübergreifend durch sogenannte *Linker-Introns* voneinander getrennt, sodass sie im Grundaufbau einzelne Exons darstellen. Diese Exons sind zum Teil in mehrere Exons unterteilt. Introns, die innerhalb der einzelnen FU-kodierenden Sequenzen liegen, werden als *interne Introns* bezeichnet. Linker-Introns liegen sowohl in Gastropoda- als auch in Cephalopoda-Hämocyaningenen in Phase 1 vor und befinden sich in allen bisher untersuchten Genstrukturen an gleicher Position bezüglich der cDNA. Im Gegensatz hierzu unterscheiden sich die verschiedenen Hämocyaningene bezüglich der Anzahl, Position und Phase der internen Introns. Bisher wurden Hämocyanin-Genstrukturen folgender Arten veröffentlicht: *Haliotis tuberculata* (HtH1+2; Lieb et al., 2001; Altenhein et al., 2002), *Haliotis diversicolor* (HdH1-3; Yao et al., 2019), *Megathura crenulata* (KLH1+2; Altenhein et al., 2002), *Enteroctopus dofleini* (OdH; Lieb et al., 2001), *Nautilus pompilius* (NpH; Bergmann et al., 2006).

Abbildung 6 zeigt die bisher veröffentlichten Hämocyanin-Genstrukturen der Vetigastropoda *Haliotis tuberculata*, *Haliotis diversicolor* und *Megathura crenulata* sowie der Cephalopoda *Enteroctopus dofleini* und *Nautilus pompilius*. Die Exon-Intron-Strukturen der Vetigastropoda-Hämocyanine beinhalten acht interne Introns (inklusive der im Signalpeptid) und sind hinsichtlich der Lage der Introns in Bezug zur kodierenden Sequenz identisch zueinander (Lieb et al., 2001; Altenhein et al., 2002; Yao et al., 2019). Die Gene der beiden Cephalopoda-Hämocyanine weisen eine jeweils andere Genstruktur mit drei (*N. pompilius*) bzw. fünf (*E. dofleini*) internen Introns auf. Bisher unveröffentlichte Ergebnisse aus zwei meiner Doktorarbeit vorausgegangenen Untersuchungen unserer Arbeitsgruppe zeigen, dass die Hämocyanin-Genstrukturen des Kalifornischen Seehasen (*Aplysia californica*; Streit, 2008) und der Spitzschlammschnecke (*Lymnaea stagnalis*; Schäfer, 2017) eine deutlich größere Anzahl an internen Introns (46-47) beinhalten. Für *A. californica* konnte nur ein Hämocyanin identifiziert werden (Streit, 2008), während für *L. stagnalis* zwei Hämocyanine

entdeckt wurden. Die Exon-Intron-Strukturen beider Gene von *L. stagnalis* unterschieden sich voneinander in nur einem Intron, wobei die Genstruktur mit 46 internen Introns der des Hämocyanins von *A. californica* entspricht (Schäfer, 2017). *A. californica* ist eine marin lebende Schnecke, die der Gruppe der Euopisthobranchia angehört. *L. stagnalis* hingegen ist eine limnisch lebende Lungenschnecke aus der Gruppe der Panpulmonata. Damit gehören beide Schneckenarten zu den Tectipleura und stellen die einzigen bisher untersuchten Hämocyanin-Genstrukturen innerhalb der Heterobranchia dar (Abb. 3A). Auch bei *Pomacea canaliculata*, einem Vertreter der Caenogastropoda, wurde eine Vielzahl interner Introns in Hämocyaningenen nachgewiesen (Chiumiento et al., 2020). Auf die entsprechenden Ergebnisse von Chiumiento et al., die eine Anzahl interner Introns zwischen 20 und 25 beschreiben, gehe ich an dieser Stelle nicht genauer ein, da sie Unstimmigkeiten enthalten, die ich in meiner Arbeit genauer untersucht und in Kapitel 2.2 korrigiert habe.

1.2.3 Die Evolution der Mollusken-Hämocyaningene

Als Sauerstofftransporter stellen Hämocyanine eine wichtige Verbindung zwischen der Außenwelt und dem Metabolismus der Mollusken dar. Durch physiologische Faktoren wie dem pH-Wert und Umwelteinflüsse wie der Temperatur und dem Sauerstoffpartialdruck werden diese respiratorischen Proteine beispielsweise in ihrer Sauerstoffaffinität stark beeinflusst (Miller und van Holde, 1974; Miller, 1985), sodass eine molekulare Anpassung an solche Umweltfaktoren grundlegend für die Physiologie der Tiere ist (Oellermann et al., 2015). Veränderungen der Umwelt in Bezug auf Temperatur und Sauerstoffpartialdruck, wie sie oft mit der Besiedelung neuer Lebensräume einhergehen, müssen deshalb unter anderem zu Adaptationen der Hämocyanine führen.

Eine Anpassung an verschiedene Umweltbedingungen stellt beispielsweise eine differentielle Expression unterschiedlicher Hämocyaningene dar. Mehrere Hämocyanin-Isoformen, die durch unterschiedliche Gene exprimiert werden, wurden in Gastropoda (Senozan et al., 1981; Markl et al., 1991; Altenhein et al., 2002; Gebauer et al., 1999; Velkova et al., 2010), in Cephalopoda (Miller et al., 1998; Melzner et al., 2007), in Bivalvia (Bergmann et al., 2007) und in Caudofoveata (Lieb und Todt, 2008) gefunden. Diese paralogen Hämocyanine in den einzelnen Molluskenklassen sind durch Genduplikationen entstanden, die in den verschiedenen Taxa mehrfach unabhängig voneinander stattgefunden haben (Lieb und Markl, 2004). Strobel et al. (2012) zeigten, dass die Sauerstoffversorgung bei S*epia officinalis* während der Ontogenese durch die unterschiedliche Expression verschiedener Hämocyanine optimierte wurde. Ein hauptsächlich embryonal exprimiertes Hämocyanin gewährleistet dabei eine gute Sauerstoffversorgung trotz niedrigem Sauerstoffpartialdruck im Ei. Unterschiedliche, Isoform-spezifische Eigenschaften, wie sie z.B. bei den zwei Hämocyaninen von *Megathura crenulata* beobachtet wurden (verschiedene P_{50}-Werte, Swerdlow et al., 1996), beruhen dabei auf unterschiedlichen Aminosäuresequenzen, die auf

verschieden evolvierte Gene zurückzuführen sind. Unterschiede innerhalb der Primärstrukturen, die unter anderem zu unterschiedlichen Sauerstoffbindungsaffinitäten führen, wurden auch bei Mollusken, die in verschiedenen Lebensräumen vorkommen, beobachtet (Melzner et al., 2007): Hämocyanine von Cephalopoden, die in sauerstoffarmen Gebieten oder in Habitaten mit schwankendem Sauerstoffgehalt leben, haben häufig höhere Sauerstoffaffinitäten, um genügend Sauerstoff binden und in die Zielgewebe transportieren zu können, als solche aus sauerstoffreicheren Gewässern (Finke et al., 1996; Seibel et al., 1999), die meist eine höhere Metabolismusrate aufrecht erhalten und deshalb möglichst viel Sauerstoff an die verschiedenen Gewebe abgeben müssen (Pörtner et al., 1991). Diese Beispiele verdeutlichen die Notwendigkeit der Anpassung der Sauerstoffaffinitäten verschiedener Hämocyanine an die entsprechenden Umweltbedingungen.

Neben der Evolution mehrerer Hämocyanin-Paraloge, die zum Teil differentiell exprimiert werden oder stark unterschiedliche physiologische Eigenschaften aufweisen, gibt es auch Hämocyanine mit strukturellen Abweichungen ihrer Untereinheiten vom oben beschriebenen charakteristischen Grundaufbau (z.B. sechs oder zwölf statt acht funktionelle Einheiten; Lieb et al., 2006; 2010). Solche strukturellen Unterschiede haben ebenfalls starken Einfluss auf die Funktion der Proteine und damit auch auf die Physiologie der Mollusken. In der Caenogastropoda-Gruppe Cerithioidea hat sich beispielsweise das sogenannte Mega-Hämocyanin entwickelt (Lieb et al., 2010; Gatsogiannis et al., 2015). Hierbei handelt es sich um ein Tridekamer, dessen äußere Dekamere aus typischen 400 kDa Hämocyanin-Untereinheiten mit den funktionellen Einheiten FU-a bis FU-h bestehen. Zwischen diesen Dekameren befindet sich ein Dekamer, das aus 550 kDa Polypeptidketten aufgebaut ist (Abb. 7). In dem für diese Untereinheit kodierenden Gen sind die Sequenzabschnitte für FU-g und FU-h verloren gegangen. Durch anschließende Domänenduplikationen entwickelten sich sechs zusätzliche FU-f, sodass eine für das Mega-Hämocyanin charakteristische Untereinheit entstanden ist, die insgesamt zwölf funktionelle Einheiten umfasst (Abb 7; Gatsogiannis et al., 2015). Das Mega-Hämocyanin der Cerithioidea weist demnach bei gleicher Größe wie ein Tridekamer, das aus drei charakteristischen 4 MDa Hämocyanindekameren aufgebaut ist, 40 zusätzliche Sauerstoff-bindungsstellen auf. Bei gleicher Viskosität und gleichem kolloid-osmotischem Druck der Hämolymphe kann hierdurch mehr Sauerstoff transportiert werden (Gatsogiannis et al., 2015). Durch unterschiedliche Expression der 400 und der 550 kDa Untereinheiten bilden sich in der Hämolymphe der Cerithioidea dabei unterschiedliche Verhältnisse aus Mega-Tridekameren und typischen Didekameren, was die Sauerstofftransport-Kapazität beeinflusst und so eine Adaptation an verschiedene Umweltbedingungen sein kann. Lieb et al. (2010) stellten die Hypothese auf, dass dieses Phänomen die adaptive Radiation innerhalb dieser Schneckengruppe beschleunigte.

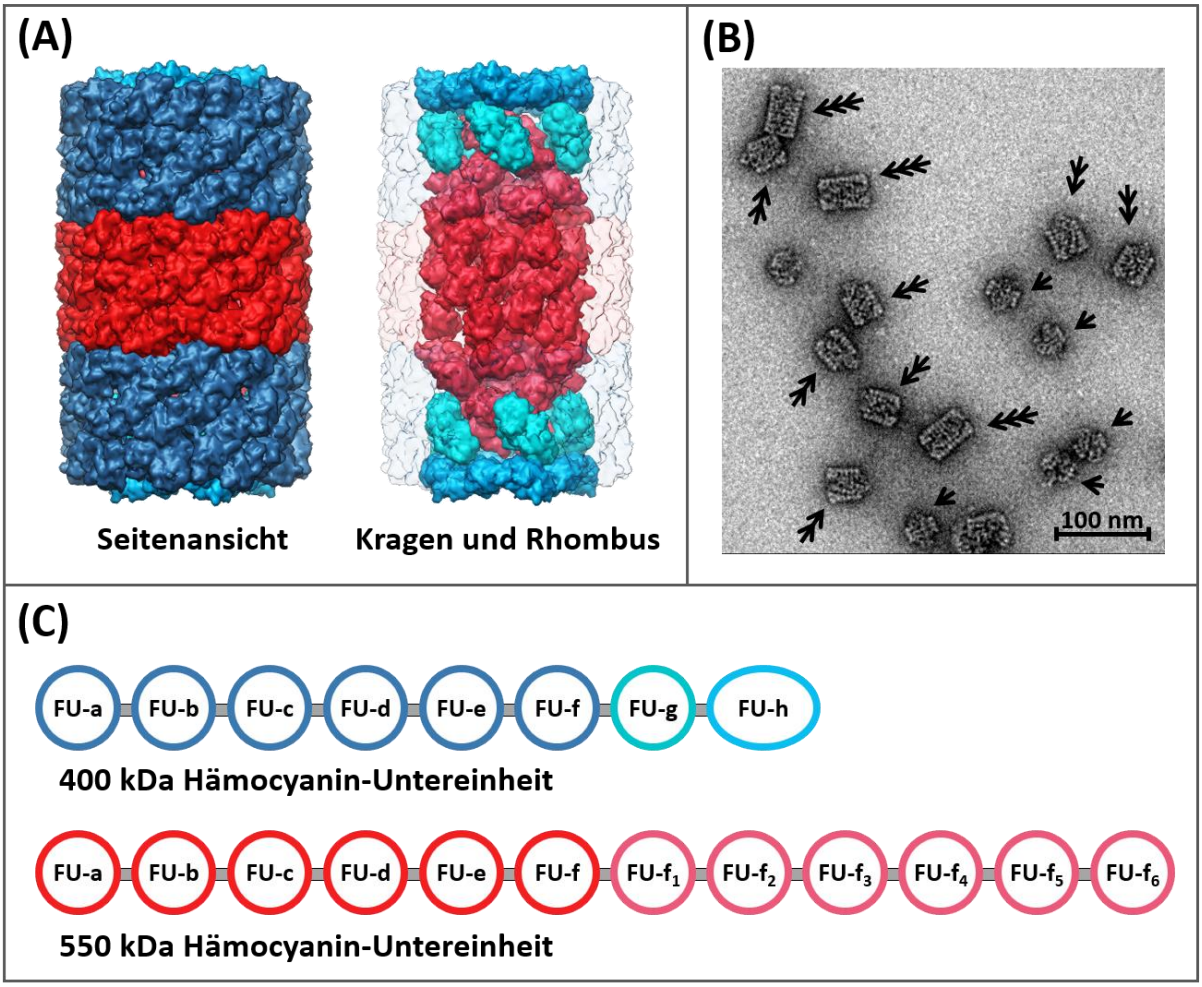


Abb. 7: Aufbau des Mega-Hämocyanins von *Melanoides tuberculata*. (**A**) Das Mega-Hämocyanin von *Melanoides tuberculata* stellt ein Tridekamer dar, welches aus zwei typischen 4 MDa Hämocyanin-Dekameren (außen, blau dargestellt) und einem 5,5 MDa Mega-Hämocyanin-Dekamer (mittig, rot dargestellt) besteht. Die äußere Wand wird durch die funktionellen Einheiten FU-a bis FU-f aller drei Dekamere aufgebaut. Der äußere Kragen (hellblau und cyanfarben) setzt sich aus den funktionellen Einheiten FU-g und FU-h der 400 kDa Untereinheiten der außen liegenden Dekamere zusammen. Der innere Rhombus, der die Mitte des Zylinders ausfüllt (rot), besteht aus FU-f_1 bis FU-f_6 der 550 kDa Untereinheiten des mittleren Mega-Hämocyanin-Dekamers. Die Abbildung des Mega-Hämocyanins basiert auf einem von Jürgen Markl mit CHIMERA simulierten 3D-Volumen bei einer 7 Å-Auflösung auf Basis des pseudoatomaren Modells nach Gatsogiannis et al. (2015) und ist nach Schäfer et al. (2021) verändert. Die transmissionselektronenmikroskopische Aufnahme in (**B**) zeigt verschiedene Hämocyanin-Moleküle der Schneckenart *M. tuberculata*. Es handelt sich hierbei um einzelne 5,5 MDa Dekamere (Einzelpfeil), Didekamere, aus je einem 4 MDa und einem 5,5 MDa Dekamer bestehen (Doppelpfeile), oder wie oben beschriebene Tridekamere mit zwei 4 MDa Dekameren und einem 5,5 MDa Dekamer (Trippelpfeil). In (**C**) ist sowohl der charakteristische Aufbau einer 400 kDa Hämocyanin-Untereinheit mit acht funktionellen Einheiten (blau) als auch der Aufbau einer 550 kDa Mega-Hämocyanin-Untereinheit mit zwölf funktionellen Einheiten (rot) dargestellt. Diese einzelnen Proteindomänen bilden globuläre Substrukturen, die über kurze Linkerpeptide (grau dargestellt) miteinander verbunden sind. Die Farbgebung der funktionellen Einheiten entspricht der für (**A**) beschriebenen Darstellung.

Die Abhängigkeit der Hämocyanine von den gegebenen Umweltbedingungen zeigt, dass die Evolution der Mollusken mit der Entstehung neuer Arten, sich verändernden Umweltbedingungen und diversen Habitatswechseln eng mit der Adaptation der Hämocyanine verknüpft sein muss. Aufgrund ihrer enormen Größe mit über 10.000 Nukleotiden, die für über 3.400 Aminosäuren kodieren, enthalten diese Sauerstofftransport-Proteine außerdem viele phylogenetische Informationen und können für Sequenzanalysen zur Aufklärung der Systematik der Mollusken

herangezogen werden. Wie oben bereits beschrieben sind zur Auflösung genauer phylogenetischer Zusammenhänge Kombinationen verschiedener phylogenetischer Marker notwendig, sodass Hämocyanine wie bereits bisher (z.B. Kocot et al., 2011) auch in Zukunft weiter in solchen Analysen miteinbezogen werden sollten (Wägele et al., 2008).

1.3 Zielsetzung der Arbeit

Duplikationen und Verluste einzelner funktioneller Einheiten, die Entstehung neuer Gene und die Veränderungen der physiologischen Eigenschaften durch die Evolution der Aminosäuresequenzen stellen eine Vielzahl von Veränderungen der Hämocyanine dar, die während der Evolution der Mollusken aufgetreten sind. Die vielfältigen Adaptationen des Sauerstofftransporters sind vermutlich eng mit der großen Diversität der Lebensräume, die von Mollusken besiedelt wurden, verknüpft (vgl. Kapitel 1.2.3). Dies gilt insbesondere für die vielfältige Gruppe der Gastropoda. Die Besiedelung von sehr diversen terrestrischen Biotopen sowie von Flüssen und Seen, aber auch von unterschiedlichen Bereichen des Meeres (Tiefsee, Gezeitenzonen, Brackwasser usw.) gehen unter anderem mit starken Temperaturunterschieden und -veränderungen einher. Anpassungen sind dementsprechend nicht nur an neue Temperaturen, sondern auch an stärkere Temperaturschwankungen nötig, wie sie beispielsweise an Land oder in der Gezeitenzone vorkommen (Steele, 1985). Wie in Kapitel 1.2.3 beschrieben beeinflusst die Temperatur die Sauerstoffaffinität der Mollusken-Hämocyanine sehr stark (Miller, 1985), sodass innerhalb der Gastropoda eine große Anzahl weiterer Adaptationen des Hämocyanins zu erwarten sind, die die Eroberung der unterschiedlichen Lebensräume ermöglicht haben. Auf Grundlage dieser Annahme war es Ziel meiner Doktorarbeit, die Hämocyanine verschiedener Schneckengruppen, die sich bezüglich ihrer Lebensbedingungen und Habitate voneinander unterscheiden, zu untersuchen. Da Apogastropoda die mit Abstand arten- und formenreichsten sowie die weitverbreitetsten Schneckengruppen Heterobranchia und Caenogastropoda beinhaltet, richtete ich den Fokus meiner Untersuchungen auf diese Gastropoda-Gruppe. Die Analysen ihrer Hämocyanine umfassten dabei folgende Zielsetzungen:

(i) In den untersuchten Schneckenarten sollte die Anzahl an Hämocyaningenen bestimmt und die Nukleotidsequenzen der kodierenden Genabschnitte erarbeitet werden, um daraus abgeleitete Primärstrukturen für phylogenetische Untersuchungen zu nutzen. Diese sollten zeigen, wie die einzelnen Gene evolviert sind.

(ii) Die durchgeführten phylogenetischen Untersuchungen der Hämocyanine sollten außerdem genutzt werden, um zu analysieren, ob Genduplikationen, wie sie bei anderen Mollusken gefunden wurden (vgl. Kapitel 1.2.3), auch in der Gruppe der Heterobranchia und Caenogastropoda aufgetreten sind. Durch Sequenzvergleiche und Stammbaumanalysen sollte hierbei genau untersucht werden, wie weit solche Genduplikationen von Hämocyaninen

innerhalb der Apogastropoda verbreitet sind und ob sie mehrfach unabhängig voneinander stattgefunden haben oder auf einige wenige Duplikationen zurückzuführen sind.

(iii) Weiterhin sollte mit diesen Analysen untersucht werden, ob es innerhalb der Apogastropoda weitere Hämocyanine mit strukturellen Abweichungen vom Grundaufbau, wie sie zum Beispiel beim Mega-Hämocyanin der Cerithioidea vorkommen (vgl. Kapitel 1.2.3), gibt.

(iv) Zusätzlich zu den kodierenden Sequenzen war es Ziel meiner Doktorarbeit auch die Exon-Intron-Strukturen der Hämocyaningene zu analysieren. Die große Anzahl an Introns, wie sie in den Hämocyaningenen von *Aplysia californica* und *Lymnaea stagnalis* gefunden wurde, (Streit, 2008; Schäfer, 2017), führt zu der Frage, wann diese Introns entstanden sind und ob sie Ausnahmen oder ein möglicherweise weit verbreitetes Phänomen der Gene innerhalb der Apogastropoda darstellen. Um diese Fragen zu klären, sollten außerdem die Genstrukturen weiterer Hämocyanine aus unterschiedlichen Heterobranchia- und Caenogastropoda-Gruppen untersucht werden (vgl. Kapitel 1.2.3, Abb. 3).

Wie in Kapitel 1.1 erläutert, stellt die Gruppe der Apogastropoda eine sehr artenvielfältige und diverse Schneckengruppe dar, sodass ich für diese Untersuchungen gezielt einige Arten und Gruppen auswählen musste. Da meine Arbeit auf der Hypothese basiert, dass insbesondere Habitatswechsel eine starke Adaptation des Hämocyanins erfordern, fokussieren sich meine Untersuchungen innerhalb der Heterobranchia auf verschiedene Arten der Tectipleura. Diese repräsentieren die artenreichste Gruppe der Heterobranchia und besiedeln mit Abstand die vielfältigsten Lebensräume (vgl. Abb. 3A und Kapitel 1.1.1). Dabei habe ich aus unterschiedlichen Untergruppen der Tectipleura, die sowohl Vertreter der Euopisthobranchia als auch der Panpulmonata umfassen und aus unterschiedlichen Lebensräumen stammen (Abb. 3A), Hämocyanin-Sequenzen assembliert, phylogenetisch miteinander verglichen und Genduplikationen untersucht (Kapitel 2.1). Zudem habe ich auch die Exon-Intron-Strukturen dieser Hämocyaningene erarbeitet und analysiert (Kapitel 2.3).

Des Weiteren habe ich die Hämocyanine von Vertretern aus den im Folgenden beschriebenen vier großen Caenogastropoda-Gruppen auf Sequenz- und Strukturebene analysiert. Durch biochemische und molekularbiologische Untersuchungen der Hämocyanine der Muricidae (siphonate Hypsogastropoda) konnte ich eine bisher unbekannte Hämocyanin-Struktur beschreiben (Kapitel 2.2). Diese Hämocyanine habe ich ebenso wie einige Hämocyanine der Caenogastropoda-Gruppen Ampullariida, Cerithiida und der asiphonaten Hypsogastropoda in phylogenetischen Analysen miteinander verglichen, Genduplikationen innerhalb dieser Gruppen untersucht und Exon-Intron-Strukturen der Gene analysiert. Der Vergleich der hieraus gewonnenen Ergebnisse mit den Erkenntnissen zu den Hämocyaningenen der Tectipleura ermöglichte schließlich erste allgemeine Rückschlüsse auf die Evolution der Apogastropoda-Hämocyaningene (Kapitel 2.4).

In Kooperation mit der Arbeitsgruppe um Prof. Reinhard Dallinger aus Innsbruck haben wir in einem weiteren Projekt außerdem Hämocyanine und Metallothioneine gemeinsam auf kupferabhängige Expressionsveränderungen untersucht (Kapitel 2.5). Hintergrund für diese Untersuchung ist die Annahme, dass Kupfer-Metallothioneine an der Hämocyaninsynthese beteiligt sein könnten (Dallinger et al., 2005), da beide Moleküle Kupfer als Liganden enthalten und ihr Syntheseort die Rhogozyten sind (Chabicovsky, 2003; Albrecht et al., 2001; Martin et al., 2011).

2

Publikationen

Die Ergebnisse meiner Dissertation haben schließlich zu fünf Publikationen geführt, in denen wir die Evolution der Hämocyaningene in Apogastropoda analysiert und mögliche Zusammenhänge mit der Diversität dieser Schneckengruppe diskutiert haben. Die Veröffentlichungen beziehen sich auf die Hämocyanine der Tectipleura als Vertreter der Heterobranchia (Kapitel 2.1, 2.3, 2.5) sowie verschiedene Gruppen der Caenogastropoda (Kapitel 2.2, 2.4) und sind im Folgenden kurz zusammengefasst:

Die erste Publikation (Kapitel 2.1: HEMOCYANIN GENES AS INDICATORS OF HABITAT SHIFTS IN PANPULMONATA?) stellt die erste überhaupt veröffentlichte phylogenetische Analyse von Hämocyaninen aus der Gruppe der Apogastropoda dar. Dabei untersuchen wir elf verschiedene Hämocyaningene von insgesamt fünf Tectipleura-Arten und zeigen, dass unabhängig voneinander in mindestens drei verschiedenen phylogenetischen Gruppen der Tectipleura ein bis zwei Genduplikationen stattgefunden haben. Diese mehrfach aufgetretenen Genduplikationen stellen einen wesentlichen Unterschied zur Evolution der Hämocyaningene der Lepetellida (Vetigastropoda) dar, da diese höchstwahrscheinlich bereits viel früher in einem gemeinsamen Vorfahren entstanden und bis heute nicht wieder verloren gegangen sind (Lieb und Markl, 2004). Die Duplikationen der Hämocyaningene innerhalb der Tectipleura hingegen sind als später aufgetretene evolutive Erneuerungen anzusehen, die möglicherweise verschiedene Adaptationen der einzelnen Tectipleura-Gruppen an ihre unterschiedlichen Habitate oder Lebensbedingungen darstellen.

In der zweiten Publikation (Kapitel 2.2: HEMOCYANINS OF MURICIDAE: NEW 'INSIGHTS' UNRAVEL AN ADDITIONAL HIGHLY HYDROPHILIC 800 KDA MASS WITHIN THE MOLECULE) charakterisieren wir die Hämocyanine der beiden Muricidae-Arten *Nucella lapillus* und *Rapana venosa*, wobei der Fokus der Veröffentlichung auf den strukturellen Besonderheiten der Hämocyanine NlH2 und RtH2 (benannt nach dem alternativ verwendeten Namen *Rapana thomasiana*) liegt. Diese beiden orthologen Hämocyanin-Untereinheiten besitzen innerhalb der FU-g 340 (NlH2) bzw. 118 (RtH2) hauptsächlich hydrophile Aminosäuren, die deutlich von dem in der Regel hoch konservierten Grundaufbau der Mollusken-Hämocyanine (vgl. Markl, 2013) abweichen. Diese haben wir sowohl biochemisch als auch mittels Sequenzierung identifiziert und untersucht. Des Weiteren deuten unsere elektronenmikroskopischen Aufnahmen auf atypische Massen innerhalb einiger Didekamere dieser Muricidae-

Hämocyanine hin, die bisher für kein anderes Mollusken-Hämocyanin beschrieben wurden und höchstwahrscheinlich auf eben diese zusätzlichen Aminosäuren zurückzuführen sind. Diese bislang unbekannte Struktur für ein Mollusken-Hämocyanin wird im Hinblick auf die Evolution spezieller Hämocyanin-Architekturen, insbesondere dem Mega-Hämocyanin der Cerithioidea, diskutiert.

In der dritten Publikation (Kapitel 2.3: THE EVOLUTION OF HEMOCYANIN GENES IN TECTIPLEURA: A MULTITUDE OF CONSERVED INTRONS IN HIGHLY DIVERSE GASTROPODS) zeigen wir, dass Genstrukturen der Tectipleura-Hämocyanine mit 53 Introns pro Untereinheit eine deutlich größere Anzahl an Introns umfassen als bisher analysierte Mollusken-Hämocyanine (9 – 33 Introns; Lieb et al., 2001; Altenhein et al., 2002; Bergmann et al., 2006; Yao et al., 2019; Chiumiento et al., 2020). Durch die Analyse von Hämocyaninen der Vertreter beider Tectipleura-Gruppen Euopisthobranchia und Panpulmonata weisen wir nach, dass diese Exon-Intron-Struktur bereits in einem gemeinsamen Vorfahren der Tectipleura vor über 200 Millionen Jahren entstanden und seitdem in den untersuchten Gruppen gleichgeblieben ist. Als mögliche Erklärung für diese starke Konservierung der Genarchitektur diskutieren wir mögliche evolutive Vorteile und Zusammenhänge mit der enormen Diversität der Tectipleura.

In der vierten Publikation (Kapitel 2.4: THE EVOLUTION OF HEMOCYANIN GENES IN CAENOGASTROPODA: GENE DUPLICATIONS AND INTRON ACCUMULATION IN HIGHLY DIVERSE GASTROPODS) zeigen wir mittels phylogenetischer Analysen, dass auch innerhalb der Caenogastropoda mehrfach unabhängig voneinander Hämocyanin-Genduplikationen aufgetreten sind und in allen untersuchten Gruppen (Ampullariida, Cerithiida und Hypsogastropoda) zu mindestens zwei Hämocyaningenen pro Art geführt haben. Des Weiteren charakterisieren wir Hämocyanin-Genstrukturen innerhalb dieser Caenogastropoda-Gruppen und zeigen, dass diese sich im Vergleich zu Tectipleura-Hämocyaninen deutlich in Anzahl und Lage der Introns voneinander unterscheiden (21 – 54 Introns pro charakteristischem Hämocyaningen). Diese variieren sowohl zwischen Hämocyaninen verschiedener Caenogastropoda-Gruppen als auch zwischen paralogen Hämocyaninen innerhalb derselben Art. Auf Grundlage der Maximum-Parsimony-Theorie und phylogenetischer Zusammenhänge leiten wir ein mögliches Szenario für die Evolution der Hämocyanin-Genstrukturen der untersuchten Apogastropoda ab. Dieses zeigt, dass die Anhäufung der Introns vermutlich kontinuierlich und graduell während der Evolution verschiedener Linien der Caenogastropoda stattgefunden und womöglich bereits in einem gemeinsamen Vorfahren aller Apogastropoda (Heterobranchia und Caenogastropoda) begonnen hat. Diese Ergebnisse diskutieren wir schließlich hinsichtlich der in den vorherigen Publikationen aufgestellten Hypothesen (vgl. Kapitel 2.1, 2.2 und 2.3) zu Zusammenhängen zwischen der Evolution der Hämocyaningene und der enormen Diversität dieser Schneckengruppe.

In der fünften Publikation (Kapitel 2.5: RESPONSIVENESS OF METALLOTHIONEIN AND HEMOCYANIN GENES TO CADMIUM AND COPPER EXPOSURE IN THE GARDEN SNAIL *CORNU ASPERSUM*) zeigen wir, dass kupferexponierte Tiere der Tectipleura-Art *Cornu aspersum* erhöhte Transkriptionsraten des Kupfer-Metallothioneingens sowie eines Hämocyaningens (CaH αD) aufweisen. Diese Ergebnisse werden bezüglich einer möglichen Funktionsänderung dieses Hämocyanins zur Kupfer-Detoxifikation sowie hinsichtlich der Hypothese, dass Kupfer-Metallothionein als Donator von Kupfer-Ionen für die Hämocyanin-Synthese dienen, diskutiert.

2.1

Hemocyanin genes as indicators of habitat shifts in Panpulmonata?

Gabriela Giannina Schäfer
Veronika Pedrini-Martha
Raimund Schnegg
Reinhard Dallinger
Daniel John Jackson
Bernhard Lieb

Molecular Phylogenetics and Evolution
Januar 2019; Volume 130, Pages 99-103

Online: Oktober 2018

https://doi.org/10.1016/j.ympev.2018.10.014

Mein Beitrag:

Ich habe sowohl die cDNA-Sequenzen assembliert und analysiert als auch die phylogenetischen Analysen (multiple Sequenzvergleiche und Stammbaum-Berechnung) durchgeführt. Außerdem verfasste ich die Erstversion des Manuskripts und arbeitete die Kommentare der Mitautoren ein.

Molecular Phylogenetics and Evolution 130 (2019) 99–103

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Short Communication

Hemocyanin genes as indicators of habitat shifts in Panpulmonata?

Gabriela Giannina Schäfer[a], Veronika Pedrini-Martha[b], Raimund Schnegg[b], Reinhard Dallinger[b], Daniel John Jackson[c], Bernhard Lieb[a,*]

[a] Johannes Gutenberg-University of Mainz, Institute of Molecular Physiology, Johann-Joachim-Becher-Weg 7, D-55128 Mainz, Germany
[b] Institute of Zoology and Center of Molecular Biosciences, University of Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria
[c] Georg-August University of Göttingen, Department of Geobiology, Goldschmidtstr. 3, 37077 Göttingen, Germany

ARTICLE INFO

Keywords:
Hemocyanin
Panpulmonata
Tectipleura
Evolution
Mollusc
Gene duplication

ABSTRACT

Hemocyanin is the primary respiratory protein for the majority of the Mollusca and therefore directly interfaces with the physiological requirements of each species and the environments to which they are adapted. Hemocyanin is therefore likely to have been evolutionarily imprinted by significant habitat shifts. In the gastropod clade Panpulmonata (> 30,000 species) major realm transitions have occurred multiple times independently and may have contributed to the diversification of this group. Yet, little is known about the adaptive changes linked to these habitat shifts. In order to gain deeper insight into the evolution of panpulmonate hemocyanins and to infer possible impacts associated with those scenarios, we have assembled and analysed hemocyanin isoforms from 4 panpulmonate species: (i) *Helix pomatia, (ii) Cantareus aspersus* (both Helicidae, Stylommatophora), (iii) *Arion vulgaris* (Arionidae, Stylommatophora) and (iv) *Lymnaea stagnalis* (Lymnaeidae, Hygrophila). Additionally, we describe a new hemocyanin isoform within the genome of the euopisthobranch *Aplysia californica.* Using these newly acquired hemocyanin data, we performed a phylogenetic analysis that reveals independent duplication events of hemocyanin within lineages that correlate with significant habitat shifts.

1. Introduction

The Gastropoda is the most diverse class within the phylum of Mollusca with an enormous number of species and representatives occupying a wide variety of habitats and displaying vastly different lifestyles (Pechenik, 2015). One significant factor that supported the diversification of the Gastropoda was the evolutionary innovation of air breathing (Mordan and Wade, 2008). This ability enabled the invasion of terrestrial and limnic habitats and occurred several times independently (Vermeij and Dudley, 2000). Within the Panpulmonata, a transition from marine to freshwater habitats occurred at least four times, once in Glacidorbidae and Acochlidia and twice in Hygrophila (Dinapoli and Klussmann-Kolb, 2010; Jörger et al., 2010). Terrestrial habitats were also colonised multiple times independently within this clade, e.g. twice in Ellobiidae (Jörger et al., 2010; Kano et al., 2015; Romero et al., 2016a). Accordingly, Panpulmonata represent a highly diverse but monophyletic clade as defined in the reclassification of the Heterobranchia by Jörger et al. in 2010, and further corroborated by Kocot et al. (2013), Wägele et al. (2014) and Kano et al. (2016).

The diverse environmental adaptations required with each habitat transition have been suggested to contain phylogenetic information about internal panpulmonate relationships. However, many relationships within Panpulmonata are poorly supported and remain unresolved due to a lack of knowledge about the evolution of morphological adaptations (Barker, 2001; Mordan and Wade, 2008). For example, lungs are generally accepted to be convergent adaptations due to habitat shifts, while it is debated whether pallial cavities and pneumostomes are plesiomorphic or homoplastic (Dayrat and Tillier, 2002; Jörger et al., 2010; Kocot et al., 2013; Schrödl, 2014). Jörger et al. (2010) summarised that it is hard to unambiguously identify morphological adaptations as synapomorphies or homoplasies for Panpulmonata. In contrast, on the genetic level Romero et al. (2016b) discovered adaptive changes in the mitochondrial complexes I and III of panpulmonates suggesting convergent evolution caused by positive selection. These adaptive changes are assumed to increase metabolic efficiency, e.g. by increased energy production and reduced reactive oxygen species (ROS) excess, allowing the host to manage abiotic stress associated with new habitats.

The invasion of land and freshwater habitats by early panpulmonates likely necessitated many such adaptations, especially concerning

* Corresponding author.
E-mail address: lieb@uni-mainz.de (B. Lieb).

https://doi.org/10.1016/j.ympev.2018.10.014
Received 10 August 2018; Received in revised form 10 October 2018; Accepted 10 October 2018
Available online 13 October 2018

respiratory and metabolic systems. As a further step towards characterising these adaptations on the molecular level, we have started to analyse hemocyanin, the respiratory protein of panpulmonates.

Hemocyanin is the blue, copper-containing oxygen transport molecule for most molluscs. Therefore, it represents an important connection between the environment and the metabolism of these animals, and thus has a great impact on their physiologies and lifestyles. In most gastropods, the 400 kDa polypeptide subunits form decameric complexes, which float freely in the hemolymph and possess molecular masses of up to 4 megadaltons or more. In the freshwater snail *Lymnaea stagnalis* mono-, di- and tridecamers have been found (Markl, 2013). Phylogenetic analyses revealed that the first hemocyanin subunit evolved during the Precambrian prior to the split of the phylum of Mollusca into the different classes (Lieb et al., 2000). Even though the structure and ancestry of hemocyanin has been studied for decades (cf. Markl, 2013), little is known about the evolution of these proteins in Panpulmonata. In marine gastropods one or two isoforms are known (Swerdlow et al., 1996; Keller et al., 1999; de Ioannes et al., 2004; Lieb et al., 2004). However, Lontie (1983) characterised three distinct hemocyanin subunits from the terrestrial gastropod *Helix pomatia* and Velkova et al. (2010) isolated three distinct hemocyanin subunits from the species *Helix lucorum*. Full length cDNA sequences of two hemocyanin isoforms, namely α_D-HlH and β-HlH, have been resolved, and a third isoform, α_N-HlH, has been partially determined (de Smet et al., 2011). Hitherto, these are the only known hemocyanin sequences of Panpulmonata. The closest relative to *H. lucorum* with a determined full length hemocyanin sequence is *Aplysia californica.* This species belongs to the Euopisthobranchia, which forms the sister group to Panpulmonata within the Tectipleura (Kano et al., 2016). To further trace the evolution of hemocyanins within the latter group, we characterised 11 novel tectipleura hemocyanin sequences, calculated a molecular tree and identified different duplication events.

2. Material and methods

2.1. Hemocyanin sequences of Stylommatophora

2.1.1. RNA isolation and next generation sequencing

Adult snails of the species *Helix pomatia* and *Cantareus aspersus* were obtained from the Wiener Schneckenmanufaktur e.U., Austria. Adult snails of *Arion vulgaris* were collected in the surroundings of Innsbruck (Tyrol, Austria; n = 4 for *H. pomatia* and *C. aspersus*; n = 1 for *A. vulgaris*), and the hepatopancreatic tissues (approx. 1 mg fresh wt.) were dissected out and stored in RNAlater™ (Invitrogen by Thermo Fisher Scientific, Waltham, MA, USA) at −80 °C. Additionally, two adult and four pooled juvenile individuals of *H. pomatia* and *C. aspersus* were sampled for long-distance (LD) PCR and stored in RNAlater™.

Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) with on-column DNase I digestion (Qiagen) according to manufacturer's instructions. RNA concentrations were estimated by Nanodrop (Thermo Fisher Scientific). 600–1000 ng total RNA was used for cDNA synthesis using the SMARTer™ Race cDNA Amplification Kit (Takara Bio, Otsu, Shiga, Japan). For transcriptome generation, 1245–1477 ng total RNA from each individual of *H. pomatia* and *C. aspersus* were sent to StarSeq (Mainz, Germany) for RNA seq (Illumina Next Seq500). Total RNA of *Arion vulgaris* was sent to the Duke University (California, USA) for mRNA sequencing (see more details in Dvorak et al., 2018).

2.1.2. In silico assembly and confirmation of assembled hemocyanin subunits of Helix pomatia and Cantareus aspersus via LD PCR

Bioinformatics were performed using Geneious 9.1.8 (Kearse et al., 2012). Firstly, raw reads were quality trimmed and sequencing adapters were removed. Trimmed reads were then mapped to the three known hemocyanin sequences from *Helix lucorum* (Acc. numbers: JF752343.1, JF752344.1, JF752345.1). Mapped reads were then assembled resulting in several contigs which represented various fragments of different hemocyanin isoforms. To fill gaps between fragments, reads were mapped iteratively against contig ends and the contigs were extended via local assemblies of the mapped reads. This procedure was reiterated until all gaps were closed. Primers targetting these assembled hemocyanin sequences were designed using CLC main workbench (Version 6.9). The complete coding region of all three hemocyanin subunits of *Helix pomatia* and *Cantareus aspersus* were PCR-amplified using the Platinum™ SuperFi™ Green PCR Master Mix (Invitrogen) (primers and parameters see supplementary Table 1). PCR products were visualised on a 0.8% agarose gel (Biozym; Hessisch Oldendorf, Germany). Gene-specific bands were cut out, purified (PCR purification kit - Invitrogen) and cloned using the TOPO™ XL-2 Complete PCR Cloning Kit (Invitrogen). Two to three plasmids of two samples (adult or pooled juveniles) were purified using the QIAprep Spin Miniprep Kit (Qiagen) and sequenced via Sanger sequencing by Microsynth (Balgach, Switzerland) (sequencing primers see supplementary Table 1).

2.2. Hemocyanin sequences of the Hygrophila Lymnaea stagnalis

2.2.1. Compiling the cDNA sequences of hemocyanin isoforms of Lymnaea stagnalis from gDNA data

The previously reported draft genome of *L. stagnalis* (Davison et al., 2016) was searched using BLAST. For these searches two partial contigs, kindly provided by Hisayo Sadamoto (Tokushima Bunri University, Shido, Sanuki-City, Kagawa, Japan), were used as queries. In this way we detected genomic scaffolds containing two different hemocyanin isoforms. The full length sequences of these isoforms were confirmed with additional transcriptomic data reported in Herlitze et al., 2018.

2.3. Hemocyanin sequences of the Euopisthobranchia Aplysia californica

2.3.1. Analysis of hemocyanin sequences in Aplysia californica

We identified and assembled one *A. californica* hemocyanin subunit by analysing publicly available transcriptome data (Acc. PRJNA77701) by iteratively mapping as described above for the hemocyanins of Stylommatophora. As a reference we used the hemocyanin sequences of *L. stagnalis* determined in this study. With the help of genomic data (Acc. PRJNA13635) we confirmed this sequence and also assembled a second hemocyanin sequence. This second sequence was partly identified within the draft genome assembly "Sept. 2008 *Aplysia californica*" (Broad Institute version Aplcal2.0) (Knudsen et al., 2006). Gaps within this genome assembly have also been filled up by iteratively mapping of the genomic NGS raw reads. Geneious 9.1.8 (Kearse et al., 2012) was used for all analyses of *Aplysia californica* data.

2.4. Sequence data analyses and phylogenetic analyses

Putative signal peptides of all sequences were predicted using SignalP 4.1 Server (Petersen et al., 2011). For phylogenetic analyses, two trees were calculated based on the hemocyanin subunits listed in supplementary Table 2. Amino acid sequences were aligned by the algorithm implemented in Geneious 9.1.8 (Kearse et al., 2012) using the default parameters. We used MEGA version 7 (Kumar et al., 2016) to determine $WAG + G + I$ to be the best model of evolution for both alignments. Maximum Likelihood analyses were conducted by PhyML 3.0 (Guindon et al., 2010) implemented in Geneious (Kearse et al., 2012) using the $WAG + G + I$ model. Branch support was evaluated using 1000 bootstrap replicates. The first tree (Fig. 1) is based on the functional units (FUs) a-g corresponding to 2908 amino acid positions from all taxa with the exception of *Arion vulgaris* (we were only able to obtain partial sequence data from *A. vulgaris* encompassing hemocyanin FUs f to h). Therefore, the second tree (supplementary Fig. 1) was calculated using only FUs f and g from all taxa (supplementary Table 2). The hemocyanins of *A. vulgaris* are therefore indicated with dashed lines in Fig. 1. Bayesian inference was conducted using MrBayes 3.2.6

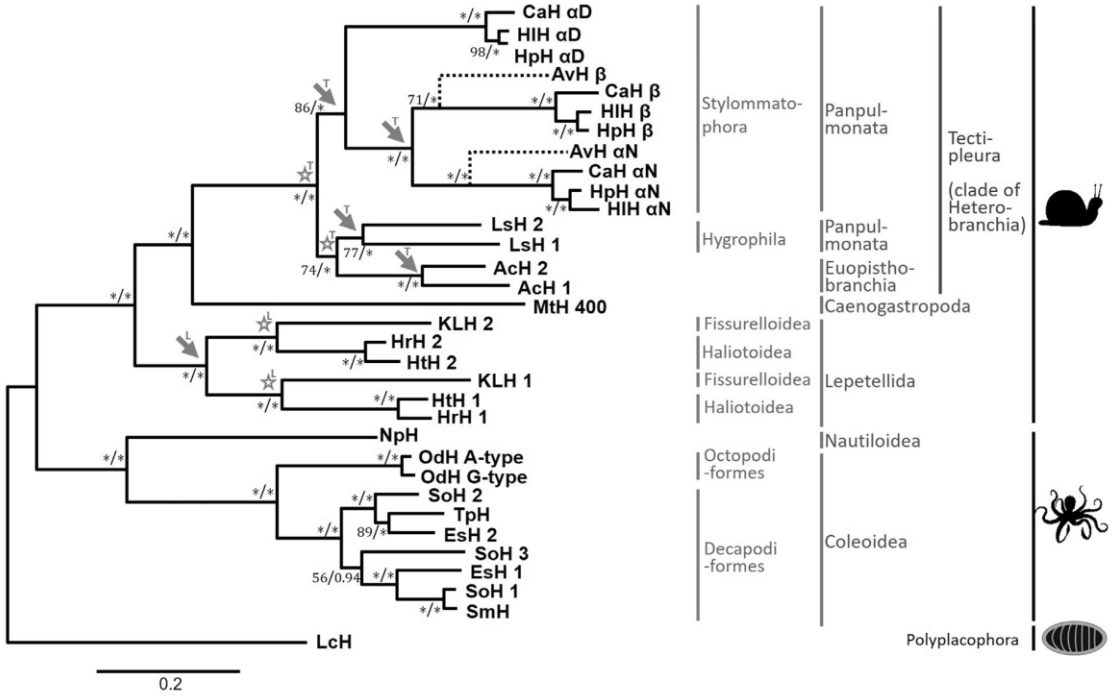


Fig. 1. Maximum likelihood tree of molluscan hemocyanins. The tree is based on an amino acid sequence alignment and calculated by PhyML (Guindon et al.) implemented in Geneious (Kearse et al., 2012) using the WAG + I + G model. It includes hemocyanin sequences (FU-a – FU-g) of 17 molluscan species: *Cantareus aspersus* (CaH αD, CaH αN, CaH β), *Helix lucorum* (HlH αD, HlH αN, HlH β), *Helix pomatia* (HpH αD, HpH αN, HpH β), *Lymnaea stagnalis* (LsH 1, LsH 2), *Aplysia californica* (AcH 1, AcH 2), *Melanoides tuberculata* (MtH 400), *Megathura crenulata* (KLH 1, KLH 2), *Haliotis rubra* (HrH 1, HrH 2), *Haliotis tuberculata* (HtH 1, HtH 2), *Nautilus pompilius* (NpH), *Enteroctopus dofleini* (OdH A-type, OdH G-type), *Sepia officinalis* (SoH 1, SoH 2, SoH 3), *Todarodes pacificus* (TpH), *Euprymna scolopes* (EsH 1, EsH 2) and *Sepiella inermis* (SmH). The accession numbers of these sequences are listed in supplementary Table 2. The tree is rooted with the hemocyanin sequence (FU-a – FU-g) of *Lepidochitona cinerea* (LcH). The two hemocyanin sequences of *Arion vulgaris* (AvH αN, AvH β) are included (dotted lines) from another PhyML analysis (WAG + I + G) that only analysed FU-f/g (see supplementary Fig. 1). Nodes congruent with those computed by Bayesian analyses and are labelled with bootstrap percentages (based on 1000 replicates) from PhyML analyses and posterior probabilities computed by MrBayes (BS/PP). Asterisks indicate fully supported nodes (BP = 100, PP = 1.0). Arrows symbolise *hemocyanin* gene duplication events of Tectipleura (arrow^{T}) and Lepetellida (arrow^{L}). The stars symbolise speciation events before the hemocyanin gene duplications in Tectipleura (star^{T}) and after the hemocyanin gene duplications in Lepetellida (star^{L}), respectively.

(Huelsenbeck and Ronquist, 2001) implemented in Geneious (Kearse et al., 2012). This analysis was performed with two parallel runs using four Monte Carlo Markov Chains (MCMC) of one million generations each and with a subsampling frequency of 500. Burn-in was set to 250,000. Bayesian nodes congruent with those computed using PhyML and posterior probabilities are shown in both trees (Fig. 1 and supplementary Fig. 1).

3. Results

In this study, we determined the complete coding sequences of three hemocyanin subunits derived from two Helicidae (*Helix pomatia* and *Cantareus aspersus*) and verified them via PCR. The identified sequences are orthologous to those found in *Helix lucorum* by de Smet et al. (2011): HpH αD (MH485358), HpH αN (MH485359), HpH β (MH485360) and CaH αD (MH485355), CaH αN (MH485356), CaH β (MH485357). We also identified two partial hemocyanin isoform sequences in *Arion vulgaris.* One sequence includes the coding sequence of FU-e to FU-h (AvH β, MH485362) whereas the second one comprises the coding sequence for FU-f to FU-h (AvH α_N, MH485361). Screening the draft genome of the Hygrophila *L. stagnalis*, we identified two hemocyanin sequences (FU-a to FU-h) (LsH1: MH485363; LsH2: MH485364).

By using publicly available genome and transcriptome data, we were also able to update the coding sequence of the single *Aplysia californica* hemocyanin AcH1 published in 2004 by Lieb et al. The newly resolved sequence (AcH1: BK010575) confirms most of the previous sequence (AcH: AJ556169.1): the identity between both sequences is 99.4% and 98.6% on nucleotide and amino acid levels, respectively. In addition to this already known hemocyanin sequence, we identified a second *hemocyanin* gene coding for another subunit within the genome of *A. californica*. The deduced cDNA codes for the eight functional units a–h and comprises characteristic hemocyanin regions. We therefore named the gene *AcH2* (BK010576).

The full length sequences identified in this study comprise about 10,300 nucleotides each coding for ca. 3430 amino acids. The identified signal peptides include 19 amino acids on average. For the exact data of each sequence see supplementary Table 2.

Trees generated using maximum likelihood and Bayesian analyses are well supported overall and are largely congruent with each other (see Fig. 1 and supplementary Fig. 1).

4. Discussion

The internal phylogenetic relationships and historical habitat shifts within Panpulmonata have been much debated (Jörger et al., 2010; Kocot et al., 2013; Kano et al., 2016). Transitions to land and freshwater habitats occurred multiple times independently within this diverse group of heterobranchs (e.g. Dinapoli and Klussmann-Kolb, 2010; Romero et al., 2016a). Studies analysing these scenarios have traditionally focused on the origin and significance of distinct morphological adaptations to the different habitats, but it has been difficult to clearly

identify morphological changes as homologous or analogous (Jörger et al., 2010). Using molecular data, Romero et al. (2016b) identified mitochondrial gene adaptations as clearly convergent and thus indicative of multiple independent invasions of the land or freshwater, hinting at a more complicated history of habitat invasion than previously appreciated.

The genetic findings of Romero et al. (2016b) motivated us to also employ molecular data to clarify scenarios of hemocyanin evolution within the Panpulmonata. Interestingly, we identified duplications of hemocyanin genes in each panpulmonate species we investigated. To analyse whether these isoforms of hemocyanin reflect habitat shifts in Panpulmonata, we calculated a molecular phylogenetic tree of all Tectipleura (Panpulmonata + Euopisthobranchia) hemocyanin isoforms (Fig. 1). Arrows in Fig. 1 represent hemocyanin gene duplication events and stars represent speciation events.

The two distinct isoforms of hemocyanin present in all surveyed Lepetellida (previously reported in Lieb and Markl, 2004 and now including the hemocyanin isoforms we identified in *Haliotis rufescens*) apparently originate from a single gene duplication. This duplication event (Fig. 1, arrow^{L}) must have taken place in a common ancestor of Haliotoidea and Fissurelloidea before the speciation event of the Lepetellida into those clades (Fig. 1, stars^{L}). The distinct clustering of the Lepetellida hemocyanins 1 and 2 (KLH1|HtH1|HrH1 and KLH2|HtH2|HrH2) in our phylogenetic tree (supported by 100% bootstrap and posterior probabilities) fully corroborates these findings.

The pattern of hemocyanin duplication events in Panpulmonata seems to be more complex. In contrast to the single hemocyanin gene duplication within the order Lepetellida, we unravelled multiple duplication events within Tectipleura (Fig. 1, arrows^{T}). The two hemocyanin isoforms of *Lymnaea stagnalis* (Hygrophila) group together in one distinct branch, whereas the hemocyanin isoforms of Stylommatophora form separate branches. This reflects that those duplications ocurred independently from each other after the diversification of the order Panpulmonata into different clades (Fig. 1, stars^{T}).

In *A. vulgaris* we found two hemocyanin isoforms (AvH) which group together with the αN - and β-isoforms of Helicidae (Fig. 1, dotted lines). According to high bootstrap value and posterior probability, one *A. vulgaris* hemocyanin isoform appears to be an ortholog of αN. The second isoform clusters together with the helicid β-isoforms with a posterior probability of 1.0 but a bootstrap value of only 71%. The low bootstrap support, most probably due to many differences in the helicid and arionid β-isoform sequences, provides some evidence for a continuing evolutionary process. We could not identify a third subunit in *A. vulgaris* (that would be expected to be an ortholog of αD). However, the existence of such a third isoform cannot be excluded, considering that our transcriptomic data were derived from a single adult specimen. Thus, expression at another developmental stage cannot be eliminated. Regardless whether a third isoform is present or lacking, such a change in hemocyanin expression patterns unravels a still ongoing evolution within the Stylommatophora.

Independent duplications of the hemocyanin gene within Panpulmonata potentially reflect convergent evolutionary processes and a significant selective pressure for multiple isoforms of this protein in air-breathing panpulmonates. From the invasion of new habitats to the evolution of new respiratory organs, multiple hemocyanin isoforms could be expected to support adaptation to new ways of life. This is perhaps most strikingly illustrated by the independent duplications of hemocyanin in Hygrophila (air-breathing freshwater snails such as *L. stagnalis*) and Stylommatophora (air-breathing terrestrial snails). It will be informative to further investigate the differences within and between Hygrophila and Stylommatophora hemocyanin isoforms in order to clarify whether the same functional requirements have been reached by convergent evolution, or whether completely different roles are fulfilled by each of these molecules.

In contrast to the panpulmonate subunits, in *A. californica* just one single hemocyanin subunit has been detected to be expressed (Lieb et al., 2004). Here we report a second hemocyanin isoform sequence within in the genome of *A. californica*. However, since this isoform has not been detected as being expressed (Lieb et al., 2004), it remains unclear whether it is a pseudogene, or is active in an as yet unsurveyed tissue or developmental stage. Nevertheless, we have included the deduced primary structure of this hemocyanin in our phylogenetic analysis, where it groups together with the other *A. californica* isoform. This duplication therefore occurred independently, potentially reflecting a common tendency in Tectipleura.

Our results suggest that multiple hemocyanin isoforms might have assisted the panpulmonates to adapt to new habitats and novel ways of life. While the precise functions of discrete isoforms of hemocyanin remain unknown, differential ontogenetic activities of these genes have been previously studied in several species. Streit et al. (2003) for example identified distinct tissue-specific spatial expression patterns for the two hemocyanin isoforms of *Haliotis asinina.* Those results hint at functional differences in the two subunits in Lepetellida. Different expression patterns and different biochemical properties have also been reported in the three isoforms of the cuttlefish *Sepia officinalis*. Melzner et al. (2007) identified different theoretical isoelectric points for the two isoforms SoH 1 and SoH 2 which might therefore have different oxygen affinities and sensitivities for pH and temperature. Additionally, Strobel et al. (2012) found development-specific expression patterns of all isoforms in *S. officinalis* and a decrease in SoH 3 levels during ontogeny from embryos through hatchlings to adults. By analysing different embryonic stages Thonig et al. (2014) further validated the findings of Strobel et al. and identified the absence of significant amounts of SoH 1 and SoH 2 within early developmental stages and an increase in expression of both isoforms later on. Since sepia eggs possess relatively hard shells the egg-mass is hypoxic in comparison to its environment (Gutowska and Melzner, 2009). SoH 3 may therefore be expected to have a higher affinity for oxygen than SoH1 and SoH2 to compensate this hypoxic conditions within the eggs. Embryonic respiratory proteins are known to exist in other metazoan groups, e.g. crustaceans and mammals (Fantoni et al., 1981; Brown and Terwilliger, 1998). They have different oxygen affinities compared to adult isoforms and therefore help to adapt to the different conditions during embryogenesis. While validating the sequences of *H. pomatia* and *C. aspersus* via PCR we also observed potentially higher expression levels of the α_N-subunits in juveniles.

Overall, this study highlights the extensive lineage specific duplications of hemocyanin that have occurred within Panpulmonata. These independent duplications may have enabled different clades to invade new habitats. There is clearly now a need to functionally characterise the properties of paralogous hemocyanin isoforms. Such studies will allow us to more clearly understand the evolutionary forces that supported the independent duplications of this important gene. Furthermore hemocyanin genes from additional panpulmonate lineages should be included in phylogenetic analyses to see whether they provide more information about the evolution of those genes and their hosts.

Acknowledgements

This study was supported by a cooperative research project granted by the Austrian Science Foundation (FWF), project number I 13032-B21 and the German Research Foundation (DFG), project number LI998/15-1. Additional support came from the Wagner and the Feldbausch Foundation (both to B. Lieb, University of Mainz). We thank Hisayo Sadamoto (Tokushima Bunri University) for kindly sharing unpublished cDNA sequences with us.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2018.10.014.

References

Barker, G.M. (Ed.), 2001. The Biology of Terrestrial Molluscs. CABI, Wallingford.

Brown, A., Terwilliger, N.B., 1998. Ontogeny of hemocyanin function in the dungeness crab cancer magister: hemolymph modulation of hemocyanin oxygen-binding. J. Exp. Biol. 201 (Pt 6), 819–826.

Davison, A., McDowell, G.S., Holden, J.M., Johnson, H.F., Koutsovoulos, G.D., Liu, M.M., Hulpiau, P., van Roy, F., Wade, C.M., Banerjee, R., Yang, F., Chiba, S., Davey, J.W., Jackson, D.J., Levin, M., Blaxter, M.L., 2016. Formin is associated with left-right asymmetry in the pond snail and the frog. Curr. Biol.: CB 26 (5), 654–660. https://doi.org/10.1016/j.cub.2015.12.071.

Dayrat, B., Tillier, S., 2002. Evolutionary relationships of euthyneuran gastropods (Mollusca): a cladistic re-evaluation of morphological characters. Zool. J. Linn. Soc. 135 (4), 403–470. https://doi.org/10.1046/j.1096-3642.2002.00018.x.

Dinapoli, A., Klussmann-Kolb, A., 2010. The long way to diversity–phylogeny and evolution of the Heterobranchia (Mollusca: Gastropoda). Mol. Phylogenet. Evolut. 55 (1), 60–76. https://doi.org/10.1016/j.ympev.2009.09.019.

Dvorak, M., Lackner, R., Niederwanger, M., Rotondo, C., Schnegg, R., Ladurner, P., Pedrini-Martha, V., Salvenmoser, W., Kremser, L., Lindner, H., García-Risco, M., Calatayud, S., Albalat, R., Palacios, Ò., Capdevila, M., Dallinger, R., 2018. Metal binding functions of metallothioneins in the slug Arion vulgaris differ from metal-specific isoforms of terrestrial snails. Metallomics. https://doi.org/10.1039/c8mt00215k.

Fantoni, A., Farace, M.G., Gambari, R., 1981. Embryonic hemoglobins in man and other mammals. Blood 57 (4), 623–633.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59 (3), 307–321.

Gutowska, M.A., Melzner, F., 2009. Abiotic conditions in cephalopod (Sepia officinalis) eggs: embryonic development at low pH and high pCO2. Mar. Biol. 156 (3), 515–519. https://doi.org/10.1007/s00227-008-1096-7.

Herlitze, I., Marie, B., Marin, F., Jackson, D.J., 2018. Molecular modularity and asymmetry of the molluscan mantle revealed by a gene expression atlas. GigaScience 7 (6). https://doi.org/10.1093/gigascience/giy056.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinform. (Oxford, England) 17 (8), 754–755.

de Ioannes, P., Moltedo, B., Oliva, H., Pacheco, R., Faunes, F., de Ioannes, A.E., Becker, M.I., 2004. Hemocyanin of the molluscan Concholepas concholepas exhibits an unusual heterodecameric array of subunits. J. Biol. Chem. 279 (25), 26134–26142. https://doi.org/10.1074/jbc.M400903200.

Jörger, K.M., Stöger, I., Kano, Y., Fukuda, H., Knebelsberger, T., Schrödl, M., 2010. On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. BMC Evol. Biol. 10, 323. https://doi.org/10.1186/1471-2148-10-323.

Kano, Y., Brenzinger, B., Nützel, A., Wilson, N.G., Schrödl, M., 2016. Ringiculid bubble snails recovered as the sister group to sea slugs (Nudipleura). Sci. Rep. 6, 30908. https://doi.org/10.1038/srep30908.

Kano, Y., Neusser, T.P., Fukumori, H., Jörger, K.M., Schrödl, M., 2015. Sea-slug invasion of the land. Biol. J. Linn. Soc. 116 (2), 253–259. https://doi.org/10.1111/bij.12578.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics (Oxford, England) 28 (12), 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

Keller, H., Lieb, B., Altenhein, B., Gebauer, D., Richter, S., Stricker, S., Markl, J., 1999. Abalone (Haliotis tuberculata) hemocyanin type 1 (HtH1). Organization of the approximately 400 kDa subunit, and amino acid sequence of its functional units f, g and h. Eur. J. Biochem. 264 (1), 27–38.

Knudsen, B., Kohn, A.B., Nahir, B., McFadden, C.S., Moroz, L.L., 2006. Complete DNA sequence of the mitochondrial genome of the sea-slug, Aplysia californica: conservation of the gene order in Euthyneura. Mol. Phylogenet. Evolut. 38 (2), 459–469. https://doi.org/10.1016/j.ympev.2005.08.017.

Kocot, K.M., Halanych, K.M., Krug, P.J., 2013. Phylogenomics supports Panpulmonata: opisthobranch paraphyly and key evolutionary steps in a major radiation of gastropod molluscs. Mol. Phylogenet. Evolut. 69 (3), 764–771. https://doi.org/10.1016/j.ympev.2013.07.001.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870–1874. https://doi.org/10.1093/molbev/msw054.

Lieb, B., Altenhein, B., Markl, J., 2000. The sequence of a gastropod hemocyanin (HtH1 from Haliotis tuberculata). J. Biol. Chem. 275 (8), 5675–5681. https://doi.org/10.1074/jbc.275.8.5675.

Lieb, B., Boisguerin, V., Gebauer, W., Markl, J., 2004. cDNA sequence, protein structure, and evolution of the single hemocyanin from Aplysia californica, an opisthobranch gastropod. J. Mol. Evol. 59 (4), 536–545. https://doi.org/10.1007/s00239-004-2646-3.

Lieb, B., Markl, J., 2004. Evolution of molluscan hemocyanins as deduced from DNA sequencing. Micron 35 (1–2), 117–119. https://doi.org/10.1016/j.micron.2003.10.035.

Lontie, R., 1983. Components, functional units, and active sites of Helix pomatia. Life Chem. Rep. (suppl) 1 (109), 120.

Markl, J., 2013. Evolution of molluscan hemocyanin structures. Biochim. Biophys. Acta 1834 (9), 1840–1852. https://doi.org/10.1016/j.bbapap.2013.02.020.

Melzner, F., Mark, F.C., Pörtner, H.-O., 2007. Role of blood-oxygen transport in thermal tolerance of the cuttlefish, Sepia officinalis. Integrat. Comparat. Biol. 47 (4), 645–655. https://doi.org/10.1093/icb/icm074.

Mordan, P., Wade, C., 2008. Heterobranchia II: the Pulmonata. In: Ponder, W. (Ed.), Phylogeny and Evolution of the Mollusca. University of California Press, pp. 409–426.

Pechenik, J.A., 2015. Biology of the Invertebrates. McGraw-Hill Education, New York, NY, pp. 606.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8 (10), 785–786. https://doi.org/10.1038/nmeth.1701.

Romero, P.E., Pfenninger, M., Kano, Y., Klussmann-Kolb, A., 2016a. Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. Mol. Phylogenet. Evolut. 97, 43–54. https://doi.org/10.1016/j.ympev.2015.12.014.

Romero, P.E., Weigand, A.M., Pfenninger, M., 2016b. Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life. BMC Evol. Biol. 16 (1), 164. https://doi.org/10.1186/s12862-016-0735-8.

Schrödl, M., 2014. Opinion: time to say “Bye-bye Pulmonata”? SPIXIANA 37 (2), 161–164.

de Smet, L., Dimitrov, I., Debyser, G., Dolashka-Angelova, P., Dolashki, A., van Beeumen, J., Devreese, B., 2011. The cDNA sequence of three hemocyanin subunits from the garden snail Helix lucorum. Gene 487 (2), 118–128. https://doi.org/10.1016/j.gene.2011.07.030.

Streit, K., Lieb, B., Markl, J., 2003. Keyhole limpet hemocyanin: genes, cDNA and evolution. 95th Annual Meeting of the DZG, 2003, Berlin, Germany.

Strobel, A., Hu, M.Y.A., Gutowska, M.A., Lieb, B., Lucassen, M., Melzner, F., Portner, H.O., Mark, F.C., 2012. Influence of temperature, hypercapnia, and development on the relative expression of different hemocyanin isoforms in the common cuttlefish Sepia officinalis. J. Exp. Zool.. Part A Ecol. Genet. Physiol. 317 (8), 511–523. https://doi.org/10.1002/jez.1743.

Swerdlow, R.D., Ebert, R.F., Lee, P., Bonaventura, C., Miller, K.I., 1996. Keyhole limpet hemocyanin: structural and functional characterization of two different subunits and multimers. Comp. Biochem. Physiol. B: Biochem. Mol. Biol. 113 (3), 537–548. https://doi.org/10.1016/0305-0491(95)02091-8.

Thonig, A., Oellermann, M., Lieb, B., Mark, F.C., 2014. A new haemocyanin in cuttlefish (Sepia officinalis) eggs: sequence analysis and relevance during ontogeny. EvoDevo 5 (1), 6. https://doi.org/10.1186/2041-9139-5-6.

Velkova, L., Dimitrov, I., Schwarz, H., Stevanovic, S., Voelter, W., Salvato, B., Dolashka-Angelova, P., 2010. Structure of hemocyanin from garden snail Helix lucorum. Comparat. Biochem. Physiol. Part B Biochem. Mol. Biol. 157 (1), 16–25. https://doi.org/10.1016/j.cbpb.2010.04.012.

Vermeij, G.J., Dudley, R., 2000. Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems. Biol. J. Linn. Soc. 70 (4), 541–554. https://doi.org/10.1111/j.1095-8312.2000.tb00216.x.

Wägele, H., Klussmann-Kolb, A., Verbeek, E., Schrödl, M., 2014. Flashback and foreshadowing—a review of the taxon Opisthobranchia. Org. Divers Evol. 14 (1), 133–149. https://doi.org/10.1007/s13127-013-0151-5.

2.2

Hemocyanins of Muricidae: New ‘Insights’ Unravel an Additional Highly Hydrophilic 800 kDa Mass Within the Molecule

Gabriela Giannina Schäfer
Lukas Jörg Grebe
Frank Depoix
Bernhard Lieb

Journal of Molecular Evolution
Februar 2021; Volume 89, Pages 62-72

Online: Januar 2021

https://doi.org/10.1007/s00239-020-09986-6

Mein Beitrag:

Ich habe die Untersuchung von *Nucella lapillus* als intertidal lebender Vertreter der Caenogastropoda konzipiert und die Probensammlung geplant und mit Kollegen durchgeführt. Außerdem isolierte ich die DNA und RNA für die Sequenzierung, assemblierte die NGS-Daten, analysierte die Nukleotidsequenzen und führte die biochemischen Analysen durch. Ich habe die Erstversion des Manuskripts geschrieben und anschließend die Kommentare von Mitautoren und Reviewern eingearbeitet.

Journal of Molecular Evolution
https://doi.org/10.1007/s00239-020-09986-6

ORIGINAL ARTICLE

Hemocyanins of Muricidae: New 'Insights' Unravel an Additional Highly Hydrophilic 800 kDa Mass Within the Molecule

Gabriela Giannina Schäfer[1] · Lukas Jörg Grebe[1] · Frank Depoix[1] · Bernhard Lieb[1]

Received: 17 October 2020 / Accepted: 17 December 2020

Abstract
Hemocyanins are giant oxygen transport proteins that freely float within the hemolymph of most molluscs. The basic quaternary structure of molluscan hemocyanins is a cylindrical decamer with a diameter of 35 nm which is built of 400 kDa subunits. Previously published results, however, showed that one out of two hemocyanin subunits of *Rapana venosa* encompasses two polypeptides, one 300 kDa and one 100 kDa polypeptide which aggregate to typical 4 MDa and 8 MDa hemocyanin (di-) decamer molecules. It was shown that the polypeptides are bound most probably by one or more cysteine disulfide bridges but it remained open if these polypeptides were coded by one or two genes. Our here presented results clearly showed that both polypeptides are coded by one gene only and that this phenomenon can also be found in the gastropod *Nucella lapillus*. Thus, it can be defined as clade-specific for Muricidae, a group of the very diverse Caenogastropoda. In addition, we discovered a further deviation of this hemocyanin subunit within both species, namely a region of 340 mainly hydrophilic amino acids (especially histidines and aspartic acids) which have not been identified in any other molluscan hemocyanin, yet. Our results indicate that, within the quaternary structure, these additional amino acids most probably protrude within the inner part of didecamer cylinders, forming a large extra mass of up to 800 kDa. They presumably influence the structure of the protein and may affect the functionality. Thus, these findings reveal further insights into the evolution and structures of gastropod hemocyanins.

Keywords Hemocyanins · Muricidae · Gastropoda · Histidines · Protein evolution · Disulfide bridge

Introduction

Hemocyanins are blue, copper-containing oxygen transporters that freely float in the hemolymph of many arthropods and molluscs and, thus, are central proteins of the physiology of most species of the two largest animal phyla. Arthropod and molluscan hemocyanins are both very large multimeric proteins that share the same binuclear copper active site (van Holde and Miller 1995). Their primary, tertiary, and quaternary structures, on the other hand, are completely different. Burmester (2001) showed that they represent two distinct protein superfamilies. Since our work is focused on molluscan hemocyanins solely, we concentrate on the characteristics of the latter one. Molluscan hemocyanins form partly hollow cylinders consisting of decamers that can aggregate to didecamers or multidecamers (van Holde and Miller 1995; Markl 2013). The cylinders consist of a wall and an inner collar which is located typically at one end of the cylinder (Fig. 1a). With a diameter of 35 nm and a height of ≥ 18 nm, these large proteins can easily be seen in the transmission electron microscope. Hemocyanins also exhibit enormously large primary structures of up to 550 kDa (Lieb et al. 2010; Gatsogiannis et al. 2015). Each subunit covers several paralogous domains called functional units (FUs) which contain one oxygen-binding site each. The basic gastropod hemocyanin subunit consists of eight of these domains termed FU-a to FU-h, forming a polypeptide subunit of ca. 400 kDa (Fig. 1c). Each FU comprises about 420 amino acids with exception of the C-terminal FU-h that

Handling editor: **Willie Swanson**.

Supplementary Information The online version of this article (https://doi.org/10.1007/s00239-020-09986-6) contains supplementary material, which is available to authorized users.

✉ Gabriela Giannina Schäfer
schaefga@uni-mainz.de

[1] Institute of Molecular Physiology, Johannes Gutenberg-University of Mainz, Johann-Joachim-Becher-Weg 7, 55128 Mainz, Germany

Published online: 13 January 2021

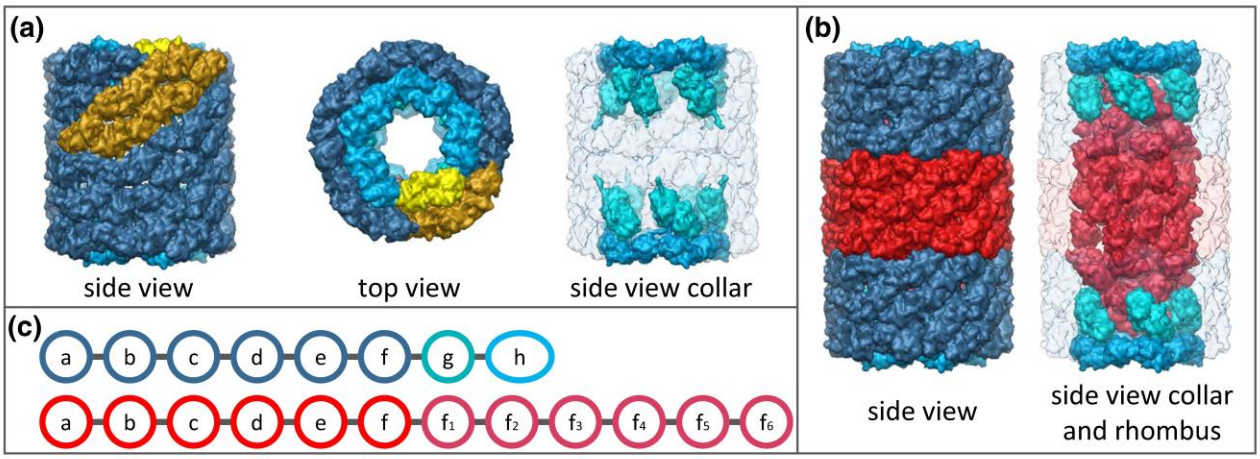


Fig. 1 Gastropod hemocyanin didecamer and mega-hemocyanin tridecamer. **a** A typical gastropod hemocyanin didecamer is shown with a dark blue wall (FU-a–FU-f) and a cyan (FU-g) and light blue (FU-h) collar. It is based on the 9 Å model of KLH_1 (Gatsogiannis and Markl 2009, PDB: 4BED). In the side and top view, one hemocyanin subunit dimer, which represents the repetitive unit within a molluscan hemocyanin decamer, is highlighted in golden with a yellow collar. In the side view of the collar, the wall is transparent to reveal the positions of FU-g and FU-h. **b** Mega-hemocyanin based on 3D volumes simulated by J. Markl in CHIMERA at 7 Å resolution on the basis of the pseudoatomic model presented in Gatsogiannis et al. (2015). The mega-hemocyanin comprises two typical hemocyanin decamers colored as described for (**a**). Between them, the 550 kDa subunit decamer is shown in red. In the right model, the wall is transparent to reveal the rhombus and the peripheral collar complex. **c** Scheme of a typical 400 kDa gastropod hemocyanin subunit (blue) and a 550 kDa mega-hemocyanin subunit build from multiple FUs (FU-h is about 100 amino acids larger). Color coding as described in **a** and **b**

contains an additional tail of about 100 amino acids (van Holde and Miller 1995; Markl 2013).

In 2010, Lieb et al. discovered a gastropod hemocyanin that largely varies from the basic structure described above. It could be found exclusively within freshwater cerithioid snails and comprises two decamers formed by canonical 400 kDa hemocyanin subunits and additionally a central decamer consisting of 550 kDa subunits (Fig. 1c). In contrast to the typical partly hollow cylinder, this so-called mega-hemocyanin forms a cylinder which is almost completely filled (Fig. 1b). As the gastropod hemocyanin archetype, the 400 kDa subunit of this mega-hemocyanin tridecamer comprises the eight functional units FU-a to FU-h (Lieb et al. 2010). Thereby, FU-a to FU-f form the canonical wall of the cylinder while FU-g and FU-h build an inner collar. The 550 kDa subunit, however, is lacking FU-g and FU-h but covers, additionally to the characteristic FU-a to FU-f, six further functional units which are paralogous to FU-f and thus are termed FU-f_1 to FU-f_6 (Fig. 1c). They increase the inner collar which is normally built from FU-g and FU-h from ~ 100 to ~ 300 kDa (Lieb et al. 2010; Gatsogiannis et al. 2015). The resulting rhombus rises into the hollow parts of the flanking 400 kDa subunits and fills the tridecamer cylinder.

While these mega-hemocyanins do not exceed the outer dimension of other hemocyanin tridecamers consisting of three 400 kDa subunits, they carry 40 additional oxygen-binding sites. This increases the efficiency of oxygen transport without raising the viscosity or the colloid-osmotic pressure of the hemolymph (Gatsogiannis et al. 2015). Lieb et al. (2010) hypothesized that this may boost adaptive radiation of cerithioid snails, because mega-hemocyanins can evolve into high affinity as well as medium or low affinity forms which help to adapt to different habitats that could otherwise not be invaded (e.g., *Terebralia palustris* lives in hypoxic mangrove mud). In addition to these mega-hemocyanin tridecamers, typical decamers as well as di- and multidecamers that are comprised of exclusively 400 kDa subunits can be found in the hemolymph of Cerithioidea. Lieb et al. (2010) hypothesized that differential expression of the 400 kDa and 550 kDa subunits may help to respiratory acclimatize to different conditions.

Cerithioidea belong to Caenogastropoda, a group of very diverse gastropods living in all kinds of habitats. They have conquered land and freshwater several times independently and therefore have undergone a multitude of drastic evolutionary adaptations (Ponder et al. 2008). Another caenogastropod hemocyanin that has been intensely analyzed, so far, comes from *Rapana venosa* (e.g., Idakieva et al. 1993, 2012; Dolashka et al. 1996; Gebauer et al. 1999; Cheng et al. 2006). This species belongs to Muricidae, a clade of Neogastropoda which represent a sister group of Cerithioidea.

Gebauer et al. (1999) isolated two types of hemocyanin from the hemolymph of *R. venosa*, precisely RtH1 and RtH2 (named after the synonym *Rapana thomasiana*). Both paralogous hemocyanins represent archetypical 400 kDa

subunits encompassing the functional units FU-a to FU-h (Gebauer et al. 1999) and constitute didecamers that form characteristic hemocyanin cylinders (Georgieva et al. 2005; Cheng et al. 2006). Gebauer et al. (1999) showed that RtH1 results in one 400 kDa band by SDS-PAGE as other typical gastropod hemocyanins. Biochemical analyses of RtH2, on the other hand, showed that RtH2 splits into two fragments under reducing conditions. They appear as two protein bands on SDS gels, one representing a 280 kDa fragment that covers FU-a–FU-f (RtH2.1), the other one a 120 kDa fragment constituting FU-g and FU-h (RtH2.2). In the absence of reducing agents, however, only one 400 kDa protein band can be seen for both hemocyanins. These results suggest the presence of one or more disulfide bridges linking both the wall-constituting FU-a to FU-f polypeptide and the collar-building FU-g/h fragment. Whether these two polypeptides arise due to a post-translational cleavage of the peptide or due to two different genes coding for two separate short subunits remained unclear until now (Gebauer et al. 1999). The evolution of a gene without FU-g and FU-h, however, could indicate a process similar to that which took place during the evolution of the mega-hemocyanin within the sister group Cerithioidea. This gene probably first lost FU-g and FU-h before it gained the additional domains through duplications of FU-f (Gatsogiannis et al. 2015).

To figure out the reason for the two hemocyanin fragments of RtH2, the intent of this study was to analyze the mRNA sequences of both hemocyanins of *Rapana venosa* and to check whether RtH2 is encoded by one or two genes. In addition, this study should reveal whether the separation of RtH2 into two fragments is a species- or genus-specific exception or a widespread phenomenon within the diverse family of Muricidae. Therefore, we additionally investigated the hemocyanins of *Nucella lapillus* by transmission electron microscopy (TEM), 3D reconstruction, SDS-PAGEs, and sequence analyses. In contrast to the marine *R. venosa,* this species of Muricidae lives intertidal and belongs to a different subfamily of Muricidae (Ocenebrinae instead of Rapaninae).

Methods

Animal Sampling and Hemocyanin Purification

Three individuals of *Nucella lapillus* were caught near to Morgat at the western Atlantic coast of Brittany, France, and stored in sea water. One individual of *Lymnaea stagnalis* was collected in a pond in Mainz, Germany. The hemolymph of one individual each was collected by cutting the foot muscle after the animals were kept on ice for 30 min. Pefabloc® SC was used as protease inhibitor (1 mM). Blood cells were removed by centrifugation at 800×*g* and 4 °C for 30 min. Hemocyanin was pelleted by ultracentrifugation at 4 °C for 2 h. The pellet was resuspended in a stabilizing buffer (50 mM Tris–HCl, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.4) and stored at 4 °C.

RNA and DNA Extraction and Next-Generation Sequencing

One individual of *N. lapillus* was sacrificed to isolate DNA from tissue of the foot using E.Z.N.A.® Mollusc DNA Kit (Omega Bio-Tek, Norcross, GA, USA). For RNA extraction mixed tissues including foot, hepatopancreas and mantle were prepared. RNA was isolated applying E.Z.N.A.® Total RNA Kit I (Omega Bio-Tek). Both DNA and RNA were purity checked and quantified via Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA) and sent to StarSeq (Mainz, Germany) for next-generation sequencing (NGS, Illumina Next Seq500) and library preparation.

In Silico Assembly of Hemocyanin Subunits

For hemocyanins of *N. lapillus*, transcriptomic NGS data which were sequenced as described above were used to assemble hemocyanin subunits. For those of *R. venosa,* we have used publicly available transcriptomic data (Acc. SRR2086477). Bioinformatic sequence analyses were performed using Geneious 9.1.8 (Kearse et al. 2012). Paired-end reads were set, raw reads were quality trimmed, and sequencing adapters were removed. Transcriptomic reads were then mapped to the previously published 400 kDa hemocyanin cDNA sequence of *Melanoides tuberculata* (KC405575) with a nucleotide identity of 70%. Mapped reads were isolated and used as references for iteratively mapping of further transcriptomic reads to prolongate the various hemocyanin fragments. To assure an accurate assembly, highly sensitive mapping settings were used (minimum overlap: 60 nucleotides; minimum overlap identity: 99%; maximum mismatches: 1%). This procedure was reiterated until all gaps between these fragments were closed and hemocyanin fragments of both species could be assembled to full-length hemocyanin coding sequences. For *N. lapillus*, also genomic NGS data were used to fill sequence gaps and to confirm assemblies.

Since both species include two hemocyanins and additionally each hemocyanin cDNA includes eight paralogous functional unit domains with some highly conserved sequence sections, this method may incorrectly merge different hemocyanin cDNA sequences. This could lead to sequence hybrids. Therefore, it is of general importance for sequence assemblies of molluscan hemocyanins to check the resulting sequences carefully and to manually examine highly identical sequence sections to preclude wrong assemblies. In this study, we verified the resulted sequences

by (i) additionally mapping the total dataset of reads to the resulted hemocyanin sequences with low sensitive mapping settings which allow misassembly detection (small number of required overlapping nucleotides, here 25 nucleotides, and high number of nucleotide mismatches allowed beyond the overlapping region, here 60% of the total read length), (ii) aligning the resulted sequences to check for highly identical sequence sections (for results see Supplement 1), and (iii) using paired-end reads for sequence assemblies which enabled us to check that repetitive sequences are spanned by paired-mates.

In addition to transcriptomic data, genomic NGS data (for *N. lapillus sequenced* by StarSeq, for *R. venosa* publicly available data: SRR5371534) were used to check the genome for further hemocyanin subunits by mapping them to the compiled hemocyanin sequences with low mapping identity.

Sequence Alignment and Phylogenetic Tree

Amino acid sequences were aligned by the MUSCLE algorithm implemented in MEGA version 7 (Kumar et al. 2016). Geneious 9.1.8 (Kearse et al. 2012) was used to display the alignment with the annotations depicted in Fig. 2. We used MEGA version 7 (Kumar et al. 2016) to determine WAG + G + F to be the best model of evolution for the alignment and to conduct the maximum likelihood analyses based on this model. Branch support was evaluated using 100 bootstrap replicates. Additionally to the hemocyanins of *N. lapillus* and *R. venosa*, these analyses include the deduced amino acid sequences of the following hemocyanin cDNAs to enable phylogenetic classification and the interrelation of NlH1 + NlH2 and RtH1 + RtH2: KLH1 and KLH2 (CAG28309.2, CAG28310.1, *Megathura crenulata* / keyhole limpet), LsH1 and LsH2 (AYO86691.1, AYO86692.1, *Lymnaea stagnalis*), MtH_{400} (AGX25261.1, *Melanoides tuberculata*), and NpH (CAF03590.1, *Nautilus pompilius*; used to root the phylogenetic tree).

Biochemical Analysis

To dissociate hemocyanin into its subunits, freshly purified samples of NlH and LsH were dialyzed overnight against a 130 mM glycine/NaOH buffer at pH 9.6. Hemocyanin of *Rapana venosa* was used from the studies of Gebauer et al. (1999). SDS-PAGE was performed in a 10% polyacrylamide gel according to Laemmli (1970) using β-mercaptoethanol as reducing agent. One sample of NlH was prepared without β-mercaptoethanol. Protein bands were stained using colloidal Coomassie® brilliant blue R250 (Serva).

Transmission Electron Microscopy (TEM) and Image Processing

Negative staining TEM was done as described previously by Harris (2007) using continuous carbon films and 1% uranyl acetate with a protein concentration of 0.1 mg/ml. A FEI Tecnai 12 (bioTwin) transmission electron microscope was used at an accelerating voltage of 120 kV. Images were taken with a 1392 × 1040 SIS Megaview camera. For the 3D reconstruction, images were automatically collected using the LEGINON system (Suloway et al. 2005) and processed using cisTEM (Grant et al. 2018; Zivanov et al. 2018). 9,800 particles were used from 271 micrographs.

Results

In order to analyze the hemocyanins of *Rapana venosa* and *Nucella lapillus* more deeply, we sequenced and assembled transcriptomic NGS data to obtain their coding sequences and additionally analyzed hemocyanins we extracted from the hemolymph of *N. lapillus* by SDS-PAGEs and transmission electron microscopy (TEM).

NGS Data Reveal an Atypical Hydrophilic Region Within Hemocyanin Subunits RtH2 and NlH2

Assembling the NGS data, we were able to construct two complete hemocyanin coding sequences of *Rapana venosa* (RtH1: BK014286; RtH2: BK014286) and two of *Nucella lapillus* (NlH1: MT939254; NlH2: MT939255). Details for each cDNA and the obtained primary structure are shown in Table 1. Assembled sequences were verified as described under *methods* section (detailed results on sequence comparison and confirmation in Supplement 1). The analysis revealed that the two hemocyanin polypeptides of both species are encoded by one gene each. Further paralogous hemocyanins were neither detected within transcriptomic nor in genomic data.

The alignment shown in Fig. 2a reveals that RtH2 and NlH2 encompass amino acids within the N-terminal region of FU-g which are additional to the highly conserved polypeptide structure of typical FU-gs or FUs in general (green bar in Fig. 2a). Compared to the sequences of other gastropod hemocyanin FUs, FU-g is extended by a sequence of 118 amino acids in RtH2 (13.6 kDa) and 340 amino acids in NlH2 (41.4 kDa), respectively. These sequences, which have not been observed in any other molluscan hemocyanin yet, are remarkably histidine and aspartic acid rich. Those two amino acids represent over 60% of the amino acids within that sequence section of RtH2 and over 70% within that of NlH2. Thus, in both polypeptides, this part of the sequence is extremely hydrophilic (color coding in Fig. 2a).

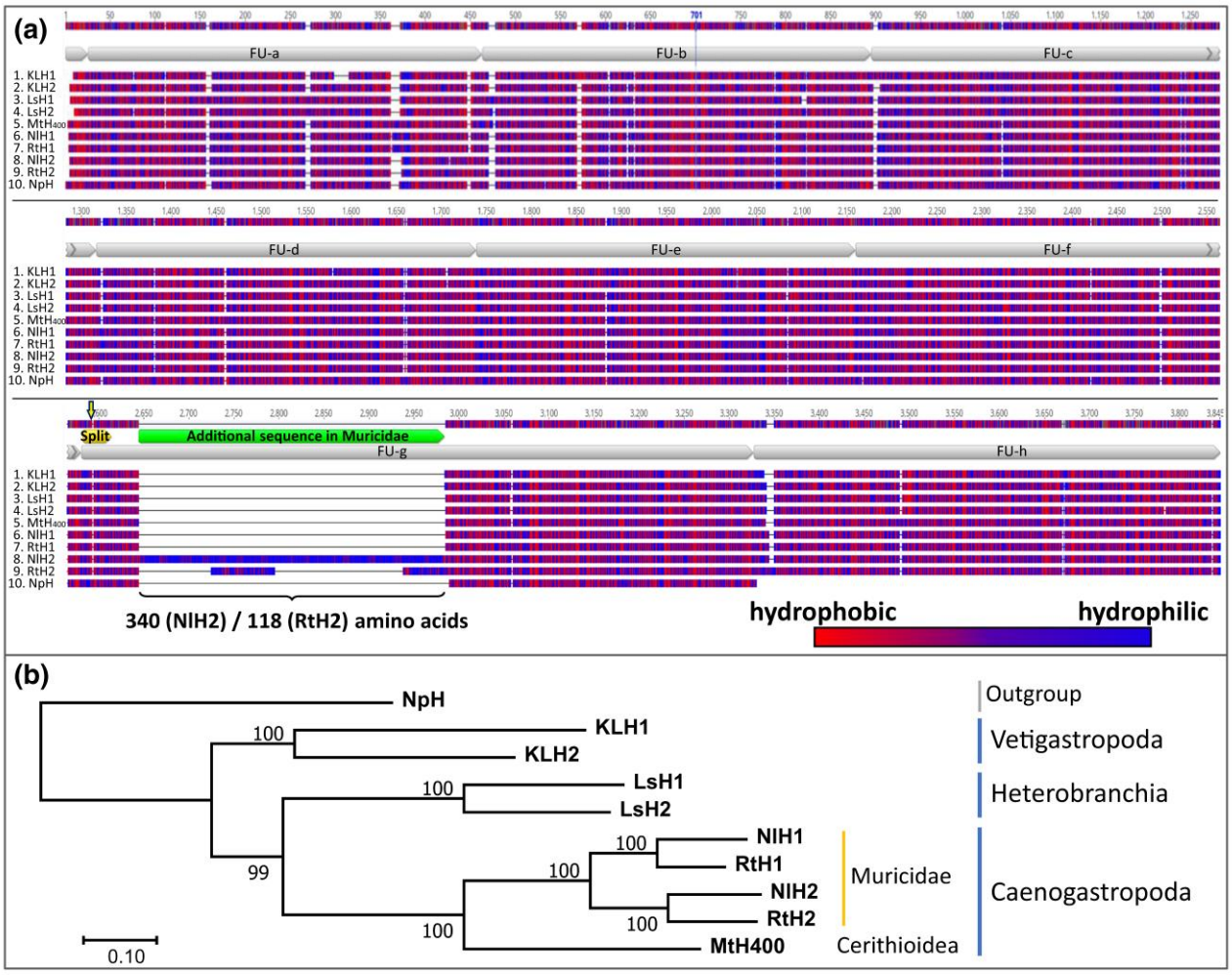


Fig. 2 Phylogenetic sequence analysis. The sequence alignment (**a**) and the maximum likelihood-based phylogenetic tree (**b**) include nine gastropod hemocyanins and one hemocyanin of *Nautilus pompilius* (NpH) as a representative of a non-gastropod mollusc. The hemocyanin sequences included are from the vetigastropod *Megathura crenulata* (KLH, from keyhole limpet), the heterobranch *Lymnaea stagnalis* (LsH) and the caenogastropods *Melanoides tuberculata* (MtH_{400}), *Nucella lapillus* (NlH), and *Rapana venosa* (RtH). NlH and RtH are derived from our study. **a** The alignment includes the full-length amino acid sequences of hemocyanins. Amino acids are color coded according to their hydrophobicity (red most hydrophobic; blue most hydrophilic). Their functional units are shown with gray bars below the consensus sequence (top line). The arrow above the yellow bar in FU-g marks the position in which RtH2 is split into two fragments according to Gebauer et al. (1999). The green bar highlights the additional highly hydrophilic sequence sections found in RtH2 and NlH2. The detailed alignment is presented in supplement 2. **b** The maximum likelihood tree is not intended to represent a full phylogeny but only reflecting the interrelation of the analyzed hemocyanins of Muricidae within the selected hemocyanins. It is based on the alignment in **a**, calculated with the WAG+G+F model and rooted with NpH (*Nautilus pompilius hemocyanin*)

Table 1 Hemocyanins of *Rapana venosa* and *Nucella lapillus*

	RtH1	RtH2	NlH1	NlH2
Accession number	BK014286	BK014287	MT939254	MT939255
Length of coding sequence (bp)	10,320	10,659	10,323	11,307
Length of primary structure (aa)	3440	3553	3441	3769
Molecular weight (kDa)	394	409	396	437

Shown are lengths of coding sequences in base pairs (bp); number of amino acids (aa) for the deduced primary structure of the polypeptides and the calculated molecular weight in kDa

The phylogenetic analysis (Fig. 2b) reveals that RtH1 and NlH1 as well as RtH2 and NlH2 are orthologous genes. Thus, the duplication event which led to these two genes most probably took place in a precursor of both Muricidae species.

Polypeptide Split of Hemocyanin Subunits is Muricidae-Specific

Total NlH was purified from the hemolymph of *N. lapillus* and dissociated into its subunits by dialysis. By SDS-PAGE analysis, we were able to identify three distinct protein bands in the Coomassie-stained gel (Fig. 3a, lane 2). For calibration, the known hemocyanins of *Lymnaea stagnalis* which represents typical 400 kDa hemocyanins were used as protein marker (lane 1). Furthermore, we included the

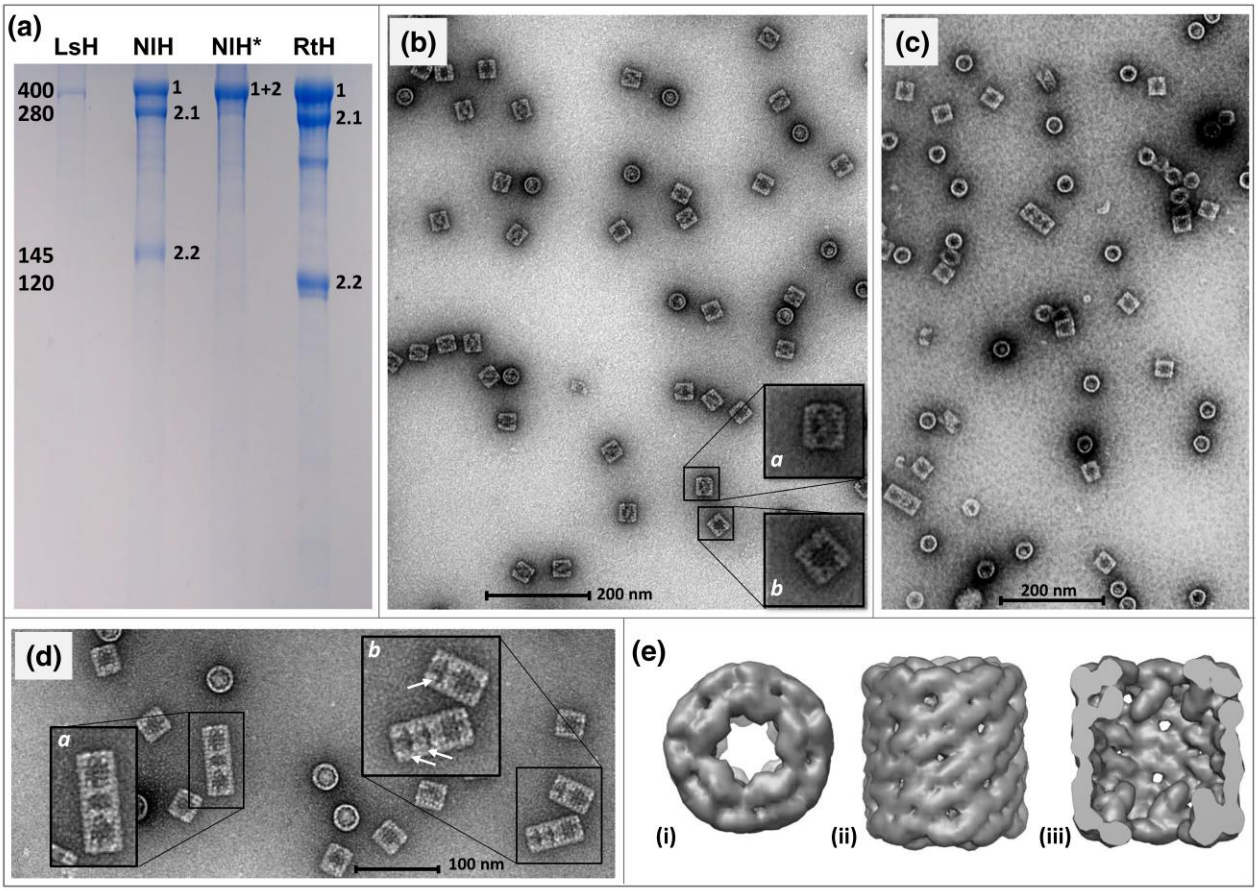


Fig. 3 Analysis of SDS-PAGE, transmission electron microscopy (TEM), and 3D reconstruction of the hemocyanin of *Nucella lapillus*. **a** Coomassie-stained gel obtained by SDS-PAGE with and without (asterisk) β-mercaptoethanol. LsH: *Lymnaea stagnalis* hemocyanin (marker for 400 kDa subunit). NlH: *Nucella lapillus* hemocyanin. NlH*: NlH under non-reducing conditions. RtH: *Rapana venosa* hemocyanin. **b–d** TEM of negatively stained hemocyanin from *Nucella lapillus* (**b, d**) and *Megathura crenulata* (**c**). **b** The electron micrograph shows hemocyanin molecules of *N. lapillus* with two appearances of didecamers which may represent two different types of hemocyanin: One side view of each potential type is shown in subwindows on the right side in a higher magnification. Both types build cylinders of about 35 nm in diameter. While one appears partly hollow as typical for molluscan hemocyanins (*b*), the other regularly appearing type seems to be semi-filled (*a*) as it has not been observed before for molluscan hemocyanin didecamers. For comparison, a micrograph of hemocyanins of *M. crenulata* (KLH) is shown in **c**. For KLH, no didecamer side view appears to be semi-filled (compare also overview of negatively stained hemocyanins from different species in Markl (2013)). The electron micrograph in **d** depicts hemocyanin multidecamers of *N. lapillus* (higher magnification in the subwindows). The pentadecamer in **d**(*a*) comprises two typical didecamers (wall and inner collar on both ends) and additionally a monodecamer between them (*). This monodecamer (marked with the asterisk) appears to lack a gastropod hemocyanin characteristic inner collar. In comparison to that the tridecamer and tetradecamer in **d**(*b*) include—besides to one didecamer each—typical monodecamers which comprise an inner collar (marked with arrows). **e** 13 Å structure of NlH in top view (*i*), side view (*ii*), and cut-open side view (*iii*)

hemocyanins of *R. venosa* in the analysis (lane 4) to enable comparisons with the results of Gebauer et al. (1999). Figure 3a shows that under reducing conditions, the sample of NlH includes one typical 400 kDa hemocyanin (NlH1), as well as one hemocyanin fragment of about 280–290 kDa as *R. venosa* does (NlH2.1). The third protein fragment runs similar but clearly above the 120 kDa RtH2.2 fragment of *R. venosa* (NlH2.2 = 145 kDa). As Gebauer et al. (1999) have already shown for RtH, the NlH sample without β-mercaptoethanol (NlH* in Fig. 3a) results in only one protein band of about 400 kDa. The gel additionally shows a fourth protein band for the hemocyanins of *R. venosa* which appears when *RtH* was stored for several months (Gebauer et al. 1999) and most probably contains degradation products (e.g., fragments of three FUs).

Transmission Electron Microscopy Indicates an Extra Mass Within one Type of NlH Didecamers

The electron micrograph of the negatively stained total native NlH in Fig. 3b shows hemocyanin didecamer cylinders, as they are typical for molluscan hemocyanins. While some of the particles appear as canonical 400 kDa hemocyanins (partly hollow cylinders; Fig. 3b(*b*)), others seem to be semi-filled with an additional mass (semi-filled cylinders; Fig. 3b(*a*)). The comparison of the observed hemocyanin cylinders of *N. lapillus* with those of KLH (Fig. 3c) and transmission electron micrographs of other hemocyanins (e.g., overview in Markl (2013)) shows that the observed appearances most probably do not result from a typically hollow 400 kDa hemocyanin. Such semi-filled didecamers have not been detected for any other gastropod hemocyanin before.

In addition to didecamers, the hemocyanin of *N. lapillus* forms multidecamers. Examples are shown in the electron micrograph in Fig. 3d. The tri- and tetradecamer in Fig. 3d(*b*) include solely (di)decamers with an inner collar (marked with arrows) as typical for hemocyanins. In contrast to that, the pentadecamer in Fig. 3d(*a*) comprises a decamer which appears to lack an inner collar (marked with an asterisk).

To receive more information on the hypothetical additional mass inside of the hemocyanin didecamer cylinders, we rendered a 3D reconstruction of NlH. Using approximately 9,800 didecamers from negative stain TEM, we obtained one 13 Å model with a structure similar to that of KLH and without an additional mass in the center (Fig. 3e). Thereby, it was impossible to distinguish between the two potentially different types of molecules (filled and unfilled). In order to ensure that the proposed additional masses are not only misinterpreted views of the reconstructed unfilled molecule, we compared the images of TEM with back projections from the 3D model. The back projections clearly did not reveal any indications of the proposed extra mass within the center of the didecamer. Thus, we conclude that the reconstruction failed to resolve an inner mass. This may imply that if there is an additional mass, it will most likely be highly flexible within the center of the protein. It is a well-known phenomenon of single particle analyses not to resolve highly flexible parts of molecules because this method is founded on image averages of the same views of molecules. Structural flexibilities will be averaged out when merged to a 3D model (Orlova and Saibil 2010; Durand et al. 2013).

Discussion

As many other molluscs, the marine gastropod *Rapana venosa* has two paralogous 400 kDa hemocyanins (Gebauer et al. 1999), both including the functional units FU-a to FU-h. In contrast to all other known molluscan hemocyanins, however, one of these hemocyanin subunits, precisely RtH2, consists of two polypeptides (RtH2.1 + RtH2.2) which most probably are bound by one or more disulfide bridges (Gebauer et al. 1999). Thus, our particular focus of attention was directed towards the second hemocyanin isoform of *Rapana venosa*. To find out if this peculiarity is species-specific, we additionally investigated hemocyanins of *Nucella lapillus*, another species of Muricidae which in contrast to *R. venosa* represents an intertidal living gastropod.

We were able to assemble the full-length coding sequences of two hemocyanins within both species and our analyses revealed that *N. lapillus* possesses two hemocyanins (NlH1 and NlH2) that are orthologous to RtH1 and RtH2. Biochemical investigations by SDS-PAGEs showed that NlH2 consists of two polypeptides, as well (Fig. 3a). NlH2.1 is similar in size as RtH2.1 (~ 280–290 kDa) and most probably comprises FU-a to FU-f. In contrast to that, NlH2.2 (~ 145 kDa) is about ~ 25 kDa larger than RtH2.2 (~ 120 kDa; FU-gh). The difference in size of NlH2.2 and RtH2.2 can be explained by the lengths of the amino acid sequences deduced from the cDNA. NlH2 is about 220 amino acids longer than RtH2 and has a calculated molecular weight which is about 28 kDa higher than that of RtH2 (Table 1). This corresponds to the difference between the protein bands of the gel (~ 25 kDa; Fig. 3a). The alignment, shown in Fig. 2a, reveals that the varying number of amino acids are located within FU-g and thus are part of RtH2.2 and NlH2.2, respectively. Additionally, our results proved that RtH2 and NlH2 are encoded by one gene each and that the amino acid chain must post-translationally be split into two polypeptides.

In comparison with other molluscan hemocyanins, both species we analyzed in this study include additional amino

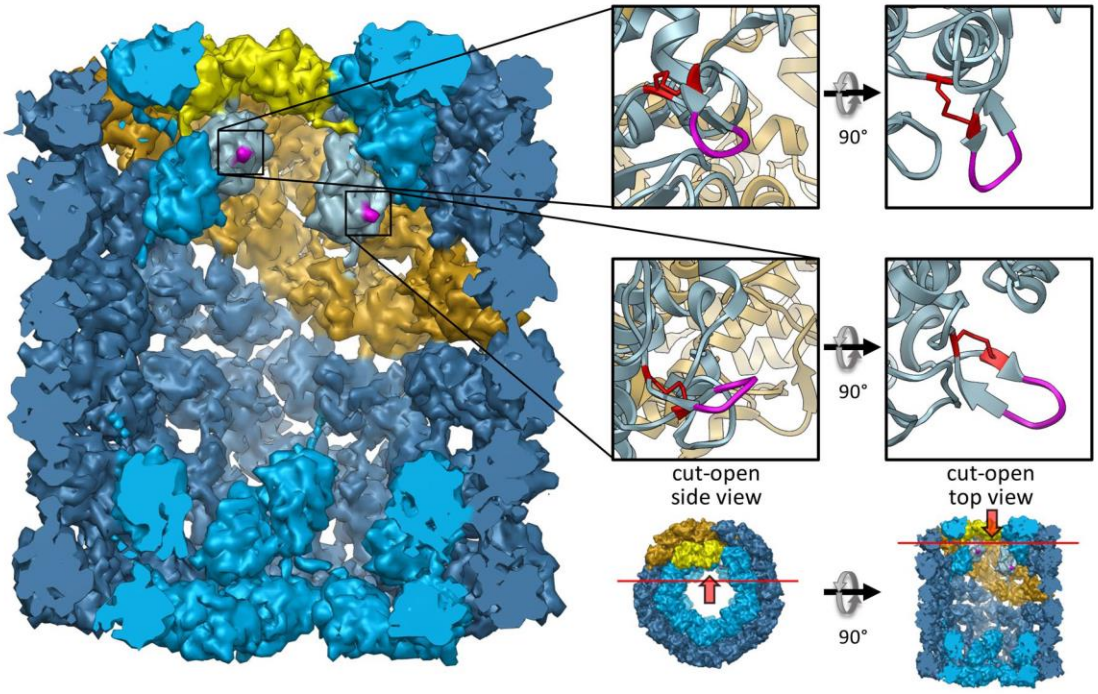


Fig. 4 Location of the additional amino acids of RtH2 and NlH2 marked in a typical molluscan hemocyanin didecamer molecule. Shown is the model of a KLH1 didecamer based on a 9 Å cryoEM structure (Gatsogiannis and Markl 2009, PDB: 4BED). On the left side, the inner surface of the cylinder is shown by a cut-open side view (the cutting surface is represented by the red line in the top view model, next to it at the bottom). Color coding: dark blue: wall (FU-a–FU-f); light blue: inner collar (Fu-g and FU-h). One subunit dimer is highlighted in golden (wall) and yellow (collar) whereas FU-g of this dimer is distinctly marked in silver. The pink part of FU-g represents the position of the primary structure which is interrupted by the additional amino acids in RtH2 and NlH2. On the right, magnifications of the ribbon of this section are depicted. Left: cut-open side view. Right: cut-open top view. The smaller density models below illustrate the lines of sight (red arrows) and the cutting surfaces (red lines). The loop marked in pink is conserved in hemocyanins over all molluscan classes that have been analyzed so far and typically covers four to seven amino acids (five in KLH1). It is stabilized by a disulfide bridge marked in red. Within one hemocyanin isoform of both Muricidae species, however, it is interrupted by 118 (RtH2) and 340 (NlH2) amino acids, respectively. The folding of these additional amino acids and the resulting structure of the extra mass within the hemocyanin cylinder is unknown and, thus, not included in this figure

acids in their primary structure which are located within the N-terminal region of FU-g (Fig. 2a). These additional sequences encompass 118 amino acids in RtH2 and 340 amino acids in NlH2 which correspond to 13.6 kDa and 41.4 kDa, respectively. The hemocyanin didecamer model in Fig. 4 shows the location of the amino acids in FU-g within the KLH molecule where the primary structure of NlH2 and RtH2 indicates the insertion of additional amino acids (marked in pink). The depicted model shows the hemocyanin of the gastropod *Megathura crenulata* (KLH1). Nevertheless, the structures are well comparable because our electron microscopical results on NlH as well as the studies of Cheng et al. (2006) on RtH show that their structures of didecamers correspond to the basic composition of hemocyanin cylinders. This is highly conserved over all molluscan hemocyanins (cf. e.g., Orlova et al. 1997; Meissner et al. 2007; Gatsogiannis et al. 2007; Gatsogiannis and Markl 2009; Markl 2013; Gai et al. 2015). Since the amino acids which frame the extra polypeptide section in RtH2 and NlH2 are located at the inner surface of the didecamer cylinder, we hypothesize that the additional amino acids reach into the inner part of the molecule and form an extra mass within the center of the hemocyanin didecamer.

The strong polarity (Fig. 2a) of the additional sequence section of the polypeptide chain further indicates that the extra mass reaches into the cavity of the molecule because this facilitates an interaction with the solvent (Zhou and Pang 2018). Furthermore, we would expect that the highly hydrophilic sequence sections would most probably change the folding of the proteins significantly, if they reached into the cylinder wall or collar structure, because additional amino acids would influence hydrophobic interactions of nonpolar amino acid residues. These, however, are typically located within proteins and are eminently important to stabilize their folding (Schreiber et al. 2009; Nick Pace et al. 2014). Since the cylindrical structures observed for NlH and RtH are rather typical for hemocyanins (Fig. 3b, d; Cheng

et al. 2006), this also supports the hypothesis that the additional amino acids reach within the center of the molecule.

By TEM, we identified two appearances of NlH didecamers. While one represents a typical didecamer, the other one potentially comprises an additional mass lying within its center. Negative staining may cause artifacts and thus these results have to be considered carefully. Comparing the results with other negative staining TEM results, however, such an inner mass has not been detected yet for any other hemocyanin and may also indicate that the additional amino acids reach within the center of the hemocyanin molecules. It has to be regarded that such an additional mass was not detected for the hemocyanins of *R. venosa* (Georgieva et al. 2005; Cheng et al. 2006). This could be traced back to the size of the additional mass, which is reduced for RtH2. The additional hydrophilic sequence section of NlH2 is 2.5 to 3 times larger than that of RtH2 and would aggregate to an extra mass encompassing 828 kDa within a total didecamer molecule of NlH2. This even corresponds to the size of two complete hemocyanin subunits. In a didecamer composed of RtH2, in contrast, the potential emerging additional mass would correspond to only 272 kDa which could easily be undetected by analyzing electron micrographs.

Although we were able to achieve a reconstruction with a resolution of 13 Å from negative stained electron micrographs of NlH, the additional mass (~828 kDa) could not be resolved in a 3D model (Fig. 3e). This suggests a highly variable structure, since it is a well-known aspect of single particle analysis that very flexible parts will be averaged out during image processing (Orlova and Saibil 2010; Durand et al. 2013). Such a high flexibility of the additional sequence section could result from its composition of repetitive amino acid motifs that encompass about 70% histidines and aspartic acids in NlH2. These amino acids can interact not only with the solvent (polarized bonds and ionization) but also with each other due to van der Waals interactions and π–π stacking of the aromatic systems (Trevino et al. 2007; Liao et al. 2013). Thereby, the large number of identical amino acids can facilitate a multitude of interactions between varying residues within one and between several subunits. This could result in different formations and a highly variable structure of the additional mass within the center of a hemocyanin didecamer cylinder and might therefore be an explanation why the additional mass cannot be resolved within a 3D reconstruction.

Despite these presumably flexible extra masses, the electron micrographs clearly show an inner collar in both types of molecules which are typical for gastropod hemocyanins (Fig. 3b, d; cf. Markl 2013). Also, the final model obtained by a 3D reconstruction includes a typical inner collar. Since it was not possible to distinguish between the two types of hemocyanins, the reconstruction indicates that—apart from the presumably highly flexible structures which are averaged out—the two potentially different molecules are very similar to each other. This substantiates the assumption that NlH1 and NlH2 both form an inner collar within the hemocyanin didecamers and therefore indicates that the two polypeptide fragments of NlH2 (FU-abcdef and FU-gh) are most probably congregated within the didecameric molecule.

Besides didecamers, we observed multidecamers of hemocyanins (Fig. 3d). Among them, we found multidecameric structures which incorporate decamers without an inner collar (asterisk in Fig. 3d). This suggests that these decamers are built by subunits consisting solely of FU-a to FU-f. Since we did not find any further hemocyanin genes, such decamers lacking an inner collar must descend from the genes of NlH1 or NlH2. One possible explanation for that phenomenon could be an alternative splicing. We, however, suggest that NlH2 could result in such decamers without an inner collar due to a post-translational loss of the protein fragment NlH2.2 which was detected by SDS-PAGEs (Fig. 3a) and presumably comprises FU-g and FU-h. Our results show that under non-reducing conditions, NlH2 does not split into two fragments as it does during SDS-PAGEs including β-mercaptoethanol (Fig. 3a). This suggests the presence of one or more inter-FU disulfide bridges between the two hemocyanin polypeptide chains NlH2.1 and NlH2.2 and confirms the results of Gebauer et al. (1999) which revealed the same phenomenon for RtH2. The non-appearance or loss of the hypothesized disulfide bridges in some cases could eventually cause decamers without inner collars. Subunits without FU-g and FU-h can form stable hemocyanin decamers as shown in *Biomphalaria glabrata*, a gastropod which uses multimeric hemoglobin as oxygen carrier but additionally still expresses "rudimentary" hemocyanins (Lieb et al. 2006).

All molluscan hemocyanins that have been analyzed so far have disulfide bridges within their functional units (Gielens et al. 1997; Cuff et al. 1998; Georgieva et al. 2004; Bergmann et al. 2007). An inter-FU disulfide bond as hypothesized by Gebauer et al. (1999) and additionally indicated by our here presented results, however, has not been proven so far. We have neither obtained an atomic structure of NlH2 or RtH2 nor is there any hemocyanin model with a comparable structure inside of the cylinder published so far which could be used to fit the sequence into a model. Thus, it is impossible to assert a disulfide bridge between different FUs. However, our sequence data show that hemocyanins of both analyzed species comprise a cysteine additionally to the highly conserved cysteines which typically build intra-FU disulfide bonds. This additional cysteine is located in the N-terminal region of FU-d (see position 1348 in alignment, Supplement 2) and has not been detected in any other molluscan hemocyanin, yet. The figure in Supplement 3 shows that the position where this cysteine evolved is located on the inner surface of the wall of the KLH1 didecamer

cylinder. Thus, it might be accessible from the center of the molecule. One highly speculative scenario we propose is that this cysteine may potentially interact in a disulfide bond with one of the cysteines in FU-g which typically form the disulfide bridge that stabilizes the loop shown in Fig. 4 (marked in pink). In RtH2 and NlH2, these two cysteines are spatially separated within the primary structure from each other due to the incorporation of the additional amino acids (in the conserved hemocyanin amino acid sequence, only eight amino acids are located between them, while it would be 118 and 340 in RtH2 and NlH2, respectively; Supplement 2). This may, for example, impede the primordial disulfide bridge in quaternary structure and may enable another disulfide bond, e.g., with the described cysteines in FU-d. Mass spectroscopic analysis could help to reveal the structure of RtH2 and NlH2 and to provide further information on the role of the specific disulfide bridge marked in pink in Fig. 4, which most probably stabilizes the core domain including the oxygen-binding site of all functional units (Georgieva et al. 2004).

All the results of this study indicate a diversification of FU-g within one hemocyanin isoform of both Muricidae species, *Nucella lapillus* and *Rapana venosa*. Although we were not able to resolve the three-dimensional structure of the extra mass, the additional amino acids included in the primary structure inevitably must change the folding and structure of this functional unit. This may also be associated with a functional change or loss. Since the two cysteines which form one of the three highly conserved disulfide bonds (Cuff et al. 1998) are spatially separated from each other due to the additional amino acids, it is uncertain if they are still able to build a disulfide bridge within NlH2 and RtH2. The disulfide bridges, however, are necessary for the oxygen-binding capacity of hemocyanins which has been shown to get lost completely after disulfide bond reduction (Topham et al. 1999). If FU-g in this hemocyanin isoform of Muricidae, for example, has lost its oxygen-binding site, this would most probably have reduced the evolutionary constraint on it which furthermore could open the way for the loss of this functional unit within the gene. Such losses of single FUs are facilitated by the modular structure of the hemocyanin genes and occurred several times within the evolution of molluscs (e.g., Cephalopoda (Miller et al. 1990) and Planorbidae (Lieb et al. 2006)). Losses of evolutionary constraint as described in the hypothetical scenario above could be the driving force within the evolution of hemocyanins with varying FU composition. For example, within Cerithioidea, another group of Caenogastropoda, the loss of FU-g and FU-h most probably facilitated the evolution of their mega-hemocyanin. This, however, presumably enabled adaptation to new habitats that could otherwise not be colonized (Lieb et al. 2010; Gatsogiannis et al. 2015) and thus may have had a great impact on the evolution and the large diversity of this group of gastropods.

Acknowledgements The authors would like to thank Arne Möller for supporting TEM and for performing single-particle analyses and image processing. Additional thanks to Julia and Christoph Markl for generating back projections.

Author Contributions GGS prepared genomic DNA and RNA for NGS, conducted biochemical investigations, and analyzed as well as interpreted nucleotide sequences. She was a major contributor to the conception of the analyses. LJG conducted parts of the NGS analyses. FD collected *N. lapillus* and prepared, performed, as well as interpreted TEM. BL conceived the project and supervised this study as principle investigator. GGS wrote the first version of the manuscript that was commented, revised, and approved by all authors.

Funding Open Access funding enabled and organized by Projekt DEAL. This study was supported by a grant from the German Research Foundation (DFG) to B. Lieb, project number LI998/15-1. Additional support came from the Wagner and the Feldbausch Foundation (both to B. Lieb, University of Mainz) as well as from the Johannes Gutenberg-University as internal research funding to B. Lieb and F. Depoix.

Data Availability The cDNA sequences obtained during the current study are available in NCBI under the following accession numbers: NlH1: MT939254; NlH2: MT939255; RtH1: BK014286; RtH2: BK014287.

Compliance with Ethical Standards

Competing interests The authors declare that they have no competing interests.

References

Bergmann S, Markl J, Lieb B (2007) The first complete cDNA sequence of the hemocyanin from a bivalve, the protobranch Nucula nucleus. J Mol Evol 64:500–510

Burmester T (2001) Molecular evolution of the arthropod hemocyanin superfamily. Mol Biol Evol 18:184–195

Cheng K, Koeck PJB, Elmlund H, Idakieva K, Parvanova K, Schwarz H et al (2006) Rapana thomasiana hemocyanin (RtH): comparison of the two isoforms, RtH1 and RtH2, at 19A and 16A resolution. Micron 37:566–576

Cuff ME, Miller KI, van Holde KE, Hendrickson WA (1998) Crystal structure of a functional unit from Octopus hemocyanin. J Mol Biol 278:855–870

Dolashka P, Genov N, Parvanova K, Voelter W, Geiger M, Stoeva S (1996) Rapana thomasiana grosse (gastropoda) haemocyanin: spectroscopic studies of the structure in solution and the conformational stability of the native protein and its structural subunits. Biochem J 315(Pt 1):139–144

Durand A, Papai G, Schultz P (2013) Structure, assembly and dynamics of macromolecular complexes by single particle cryo-electron microscopy. J Nanobiotechnol 11(Suppl 1):S4

Gai Z, Matsuno A, Kato K, Kato S, Khan MRI, Shimizu T et al (2015) Crystal structure of the 3.8-MDa respiratory supermolecule hemocyanin at 3.0 A resolution. Structure (London, England: 1993) 23:2204–2212

Gatsogiannis C, Markl J (2009) Keyhole limpet hemocyanin: 9-A CryoEM structure and molecular model of the KLH1 didecamer reveal the interfaces and intricate topology of the 160 functional units. J Mol Biol 385:963–983

Gatsogiannis C, Moeller A, Depoix F, Meissner U, Markl J (2007) Nautilus pompilius hemocyanin: 9 A cryo-EM structure and molecular model reveal the subunit pathway and the interfaces between the 70 functional units. J Mol Biol 374:465–486

Gatsogiannis C, Hofnagel O, Markl J, Raunser S (2015) Structure of mega-hemocyanin reveals protein origami in snails. Structure (London, England: 1993) 23:93–103

Gebauer W, Stoeva S, Voelter W, Dainese E, Salvato B, Beltramini M et al (1999) Hemocyanin subunit organization of the gastropod Rapana thomasiana. Arch Biochem Biophys 372:128–134

Georgieva DN, Genov N, Perbandt M, Voelter W, Betzel C (2004) Contribution of disulfide bonds and calcium to Molluscan hemocyanin stability. Zeitschrift fur Naturforschung C, J Biosci 59:281–287

Georgieva D, Schwark D, Nikolov P, Idakieva K, Parvanova K, Dierks K et al (2005) Conformational states of the Rapana thomasiana hemocyanin and its substructures studied by dynamic light scattering and time-resolved fluorescence spectroscopy. Biophys J 88:1276–1282

Gielens C, de Geest N, Xin XQ, Devreese B, van Beeumen J, Préaux G (1997) Evidence for a cysteine-histidine thioether bridge in functional units of molluscan haemocyanins and location of the disulfide bridges in functional units d and g of the betaC-haemocyanin of Helix pomatia. Eur J Biochem 248:879–888

Grant T, Rohou A, Grigorieff N (2018) cisTEM, user-friendly software for single-particle image processing. eLife 7:e35383

Harris JR (2007) Negative staining of thinly spread biological samples. Methods Mol Biol (Clifton, NJ) 369:107–142

Idakieva K, Severov S, Svendsen I, Genov N, Stoeva S, Beltramini M et al (1993) Structural properties of Rapana thomasiana grosse hemocyanin: isolation, characterization and N-terminal amino acid sequence of two different dissociation products. Compar Biochem Physiol B 106:53–59

Idakieva K, Meersman F, Gielens C (2012) Reversible heat inactivation of copper sites precedes thermal unfolding of molluscan (Rapana thomasiana) hemocyanin. Biochem Biophys Acta 1824:731–738

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S et al (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics (Oxford, England) 28:1647–1649

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Liao S-M, Du Q-S, Meng J-Z, Pang Z-W, Huang R-B (2013) The multiple roles of histidine in protein interactions. Chem Cent J 7:285

Lieb B, Dimitrova K, Kang H-S, Braun S, Gebauer W, Martin A et al (2006) Red blood with blue-blood ancestry: intriguing structure of a snail hemoglobin. Proc Natl Acad Sci 103:12011–12016

Lieb B, Gebauer W, Gatsogiannis C, Depoix F, Hellmann N, Harasewych MG et al (2010) Molluscan mega-hemocyanin: an ancient oxygen carrier tuned by a ~550 kDa polypeptide. Front Zool 7:14

Markl J (2013) Evolution of molluscan hemocyanin structures. Biochem Biophys Acta 1834:1840–1852

Meissner U, Gatsogiannis C, Moeller A, Depoix F, Harris JR, Markl J (2007) Comparative 11A structure of two molluscan hemocyanins from 3D cryo-electron microscopy. Micron 38:754–765

Miller KI, Schabtach E, van Holde KE (1990) Arrangement of subunits and domains within the Octopus dofleini hemocyanin molecule. Proc Natl Acad Sci 87:1496–1500

Nick Pace C, Scholtz JM, Grimsley GR (2014) Forces stabilizing proteins. FEBS Lett 588:2177–2184

Orlova EV, Saibil HR (2010) Methods for three-dimensional reconstruction of heterogeneous assemblies. In: Jensen JG (ed) Cryo-EM, Part B: 3-D reconstruction. Elsevier, North-Holland, pp 321–341

Orlova EV, Dube P, Harris JR, Beckman E, Zemlin F, Markl J et al (1997) Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 A resolution by electron cryomicroscopy and angular reconstitution. J Mol Biol 271:417–437

Ponder WF, Colgan DJ, Healy JM, Alexander N, Simone LRL, Mielke EE (2008) Caenogastropoda. In: Ponder W (ed) Phylogeny and evolution of the mollusca. University of California Press, Los Angeles, pp 331–383

Schreiber G, Haran G, Zhou H-X (2009) Fundamental aspects of protein-protein association kinetics. Chem Rev 109:839–860

Suloway C, Pulokas J, Fellmann D, Cheng A, Guerra F, Quispe J et al (2005) Automated molecular microscopy: the new Leginon system. J Struct Biol 151:41–60

Topham R, Tesh S, Westcott A, Cole G, Mercatante D, Kaufman G et al (1999) Disulfide bond reduction: a powerful, chemical probe for the study of structure-function relationships in the hemocyanins. Arch Biochem Biophys 369:261–266

Trevino SR, Scholtz JM, Pace CN (2007) Amino acid contribution to protein solubility: Asp, Glu, and Ser contribute more favorably than the other hydrophilic amino acids in RNase Sa. J Mol Biol 366:449–460

van Holde KE, Miller KI (1995) Hemocyanins. Adv Protein Chem 47:1–81

Zhou H-X, Pang X (2018) Electrostatic interactions in protein structure, folding, binding, and condensation. Chem Rev 118:1691–1741

Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E et al (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife. https://doi.org/10.7554/eLife.42166

2.3

The evolution of hemocyanin genes in Tectipleura: a multitude of conserved introns in highly diverse gastropods

Gabriela Giannina Schäfer
Veronika Pedrini-Martha
Daniel John Jackson
Reinhard Dallinger
Bernhard Lieb

BMC Ecology and Evolution
März 2021; Volume 21, Article number: 36

https://doi.org/10.1186/s12862-021-01763-3

Mein Beitrag:

Ich habe sowohl die Genstrukturen analysiert und interpretiert als auch einen wesentlichen Teil der Studie konzipiert. Des Weiteren verfasste ich die Erstversion des Manuskripts und arbeitete die Kommentare von Mitautoren und Reviewern ein.

Schäfer *et al. BMC Ecol Evo* (2021) 21:36
https://doi.org/10.1186/s12862-021-01763-3

BMC Ecology and Evolution

RESEARCH ARTICLE Open Access

The evolution of hemocyanin genes in Tectipleura: a multitude of conserved introns in highly diverse gastropods

Gabriela Giannina Schäfer[1*], Veronika Pedrini-Martha[2], Daniel John Jackson[3], Reinhard Dallinger[2] and Bernhard Lieb[1]

Abstract

Background: Hemocyanin is the oxygen transporter of most molluscs. Since the oxygen affinity of hemocyanin is strongly temperature-dependent, this essential protein needs to be well-adapted to the environment. In Tectipleura, a very diverse group of gastropods with > 27,000 species living in all kinds of habitats, several hemocyanin genes have already been analyzed. Multiple independent duplications of this gene have been identified and may represent potential adaptations to different environments and lifestyles. The aim of this study is to further explore the evolution of these genes by analyzing their exon–intron architectures.

Results: We have reconstructed the gene architectures of ten hemocyanin genes from four Tectipleura species: *Aplysia californica, Lymnaea stagnalis, Cornu aspersum* and *Helix pomatia*. Their hemocyanin genes each contain 53 introns, significantly more than in the hemocyanin genes of Cephalopoda (9–11), Vetigastropoda (15) and Caenogastropoda (28–33). The gene structures of Tectipleura hemocyanins are identical in terms of intron number and location, with the exception of one out of two hemocyanin genes of *L. stagnalis* that comprises one additional intron. We found that gene structures that differ between molluscan lineages most probably evolved more recently through independent intron gains.

Conclusions: The strict conservation of the large number of introns in Tectipleura hemocyanin genes over 200 million years suggests the influence of a selective pressure on this gene structure. While we could not identify conserved sequence motifs within these introns, it may be simply the great number of introns that offers increased possibilities of gene regulation relative to hemocyanin genes with less introns and thus may have facilitated habitat shifts and speciation events. This hypothesis is supported by the relatively high number of introns within the hemocyanin genes of *Pomacea canaliculata* that has evolved independently of the Tectipleura. *Pomacea canaliculata* belongs to the Caenogastropoda, the sister group of Heterobranchia (that encompass Tectipleura) which is also very diverse and comprises species living in different habitats. Our findings provide a hint to some of the molecular mechanisms that may have supported the spectacular radiation of one of Metazoa's most species rich groups.

Keywords: Hemocyanin, Gene structure, Intron, Intron gain, Evolution, Habitat shift, Adaptation, Tectipleura, Heterobranchia, Mollusca

Background

The Mollusca is the second largest animal phylum and comprises diverse species with an enormous array of different body forms, physiologies, habitats and behaviors. They are divided into eight major classes (e.g.

*Correspondence: schaefga@uni-mainz.de
[1] Institute of Molecular Physiology, Johannes Gutenberg-University of Mainz, Johann-Joachim-Becher-Weg 7, 55128 Mainz, Germany
Full list of author information is available at the end of the article

Cephalopoda, Bivalvia and Scaphopoda) of which Gastropoda represents by far the largest. Gastropoda encompass the five major groups Patellogastropoda, Vetigastropoda, Neritimorpha, Caenogastropoda and Heterobranchia of which especially the latter two are highly diverse. Together they form the clade Apogastropoda [1] and include about 68,800 different species [2]. Species of both groups can be found in most kinds of environments and exhibit a wide range of specializations. Adaptations of their respiratory systems, for example, were fundamental for colonization of different habitats. Besides respiratory organs, also the oxygen transporter of molluscs, namely hemocyanin, must have been adapted to new environments because the oxygen affinity of this protein is known to be temperature-dependent [3–6]. Therefore, changes to molluscan hemocyanin genes must reflect an essential part of the process of adaptation to new habitats to sustain a sufficient oxygen supply [3, 4]. Accordingly, the evolution of this protein is strongly linked to the evolution of molluscs. Analyzing hemocyanin genes and reconstructing their evolution will therefore deepen our understanding of the evolution of the Mollusca.

In this study we focus on hemocyanins of Tectipleura, a group of Heterobranchia with over 27,000 species. This clade was recovered by molecular phylogenetic analyses [7–9] and comprises Euopisthobranchia (e.g. sea hares (Anaspidea) and pteropod sea butterflies) and Panpulmonata (e.g. traditional pulmonates as Hygrophila and Stylommatophora; Fig. 1a). They have undergone multiple independent habitat shifts from sea to land or freshwater within different clades, and have therefore developed independent adaptations. Within Panpulmonata, for example, the evolution of lungs took place several times independently [7, 10]. In a previous analysis, we have reported about hemocyanin gene duplications that occurred multiple times independently within different Panpulmonata lineages [11]. Most probably, those duplication events represent convergent adaptations to their multiple habitat shifts and may have supported the extensive radiation of Tectipleura. Thus, the aim of this study was to analyze the genes of hemocyanins within Tectipleura in more detail.

Comprising approximately 3400 amino acids, the 400 kDa subunits of molluscan hemocyanins belong to the largest naturally occurring polypeptides (Fig. 2). Mono-, di- and multidecamers of these proteins are freely floating in the hemolymph and cause its blue color due to their binuclear copper active site. In several molluscan species, two or more hemocyanin genes have been found [12–14]. Previous studies identified the basic composition of hemocyanin subunits to be highly conserved [15]: Usually, each subunit comprises eight homologous protein domains termed functional units FU-a to FU-h (Fig. 2) which are connected by linker peptides of 10–20 amino acids. This eight-FU protein most probably evolved from a single-FU gene precursor through three domain duplications in a common ancestor of all molluscs before the radiation into different molluscan classes [16]. This repetitive structure in combination with the multitude of hemocyanin genes within many molluscan species impedes de novo assemblies of NGS (Next Generation Sequencing) data. Thus, automatically assembled hemocyanin gene or cDNA sequences are often error-prone. To analyze these genes which partially reflect the evolution of molluscs (see above), however, additional studies like this one are required.

The highly conserved and complex structure of hemocyanins is also reflected by the structure of their genes: In all molluscan hemocyanin genes that have been analyzed so far the encoded individual functional units are separated from each other by phase 1 introns (called linker introns) which lie at an almost equivalent point just upstream from the linker peptide coding regions (Fig. 2) [16–18]. Thereby, intron phases 0, 1 and 2 are defined as being located before the first, after the first and after the second nucleotide of a codon, respectively. Consequently, every FU is represented by at least one separate exon. The identical positions across all molluscan classes suggest that those introns originated through domain duplications which resulted in the first hemocyanin consisting of eight functional units [16]. Thus, these introns are most probably as old as the hemocyanin molecule itself.

Beyond these linker introns between the functional units, Lieb et al. [16] identified introns lying within these FUs, subdividing them into multiple exons. The number of these introns as well as their position within the sequence and their intron phases vary between hemocyanin genes of different molluscan lineages and between different functional units. To distinguish between those different kinds of introns within hemocyanin genes, they have been classified in two types called (i) linker introns (phase 1 introns located between single functional units) and (ii) internal introns (non-conserved introns of any phase within any functional unit, see Fig. 2 and [16]).

Full hemocyanin exon–intron architectures are known from the cephalopods *Enteroctopus dofleini* [16] and *Nautilus pompilius* [18] as well as of three species of Vetigastropoda (*Megathura crenulata, Haliotis tuberculata and H. diversicolor* [16, 17, 19, 20]). Figure 1b shows that hemocyanin gene structures differ between hemocyanins of these different groups. Hemocyanin genes within these groups and within the same species, on the other hand, are identical in terms of number and positions of introns (see also Additional file 1). The internal introns of hemocyanin genes most probably originated independently after the ancestors of these groups separated from each

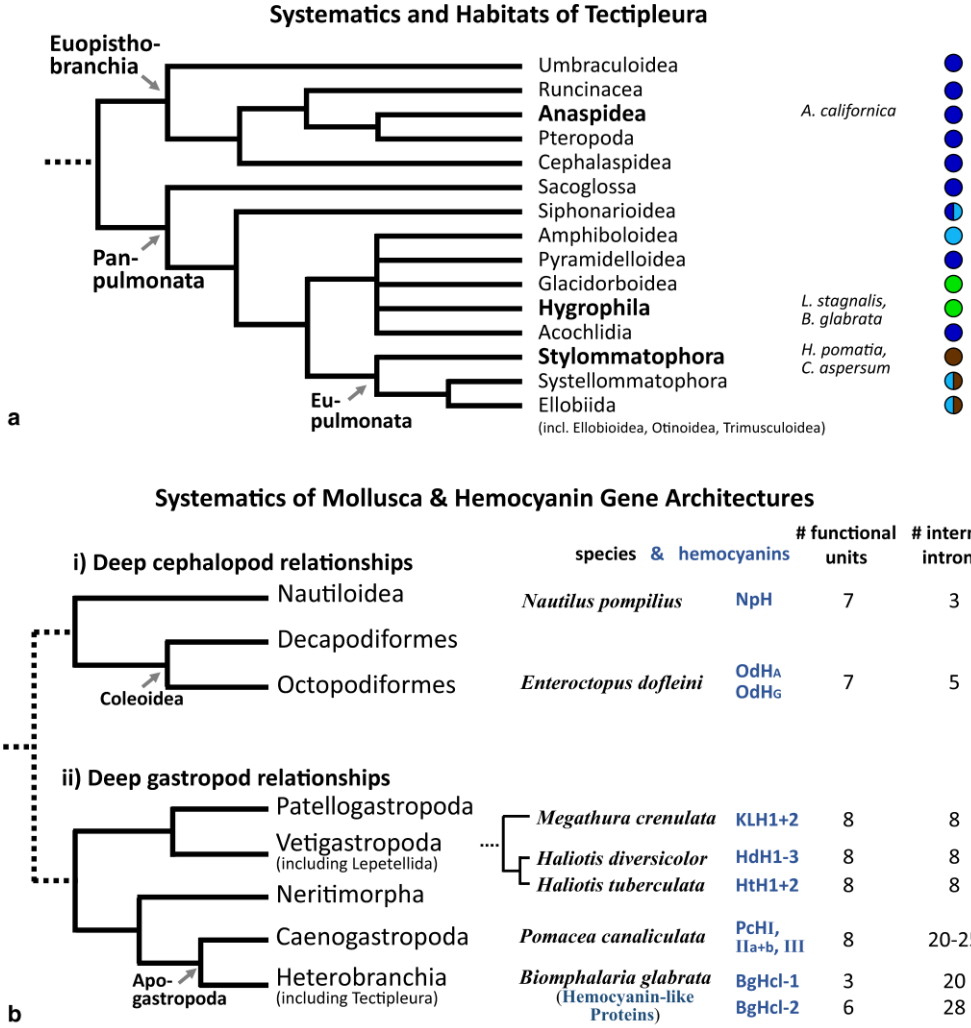


Fig. 1 Systematics of Tectipleura, their habitats and hemocyanin gene structures. **a** Tectipleura comprise the two large groups Euopisthobranchia and Panpulmonata. The systematics of both groups are shown as derived in molecular analysis by Jörger et al. [7], Schrödel [8] and Kano et al. [9]. Tectipleura included in this study are shown on the right site of the lineage they belong to (highlighted in bold). Habitats are represented by colored circles on the right. While all Euopisthobranchia are marine (dark blue), Panpulmonata also include intertidal (light blue), limnic (green) and terrestrial (brown) gastropods. **b** The number of internal introns and of functional units (right) of yet analyzed cephalopod and gastropod hemocyanin genes. For the distribution within functional units see Additional file 1. Species and the rough systematics of the major groups they belong to are shown on the left side. Hemocyanin-like proteins of *Biomphalaria glabrata* are included in this figure, as well

other (e.g. for the listed Vetigastropoda in a common ancestor of Lepetellida, the vetigastropod group to which *Megathura* and *Haliotis* belong to).

Only recently, Chiumiento et al. [21] showed that exon–intron structures of four hemocyanin genes of the caenogastropod *Pomacea canaliculata* differ not only from those of Lepetellida but also vary from each other. Furthermore, they contain a much larger number of introns than those known from Vetigastropoda or Cephalopoda (Fig. 1b). Since no other hemocyanins of Caenogastropoda are known yet, it is unclear whether the multitude of introns in hemocyanin genes of *P. canaliculata* is species-specific or widespread within that clade.

Apart from Lepetellida and *P. canaliculata*, no gastropod exon–intron architectures of full-length hemocyanin genes have yet been described. Peña and Adema [22], however, published the exon–intron architecture of two hemocyanin-like genes of *Biomphalaria glabrata*. This species belongs to Planorbidae, hygrophilid panpulmonates within Tectipleura which represents the only known gastropod family that uses hemoglobin (evolved from a gastropod myoglobin) instead of hemocyanin

Schäfer *et al. BMC Ecol Evo* (2021) 21:36

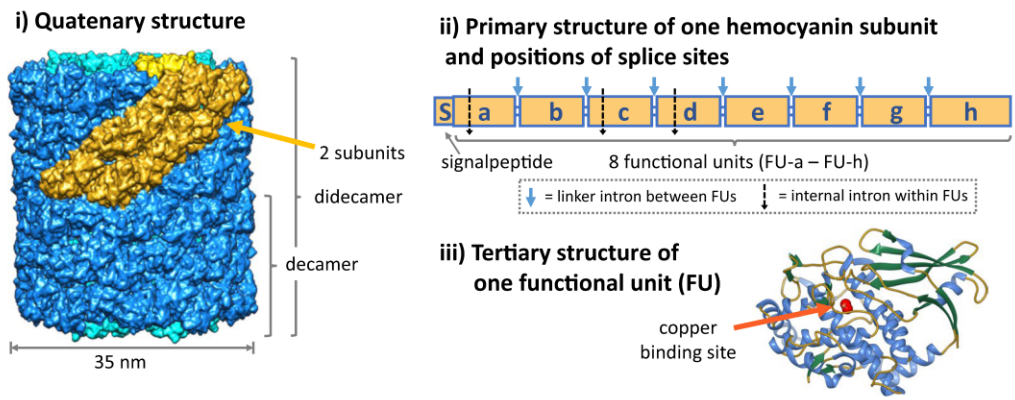


Fig. 2 Basic structure of molluscan hemocyanin. On protein level, hemocyanin represents a partly hollow cylinder (outer wall: dark blue; inner collar: light blue, PDB-ID: 4BED) with a diameter of 35 nm (**i**). This quaternary structure is built out of two decamers lying on top of each other. The repetitive unit is a subunit dimer colored in golden (collar: yellow). The primary structure of one subunit (**ii**) shows the composition of one signal peptide (S) and eight paralogous domains called functional units (FU-a – FU-h). Each functional unit forms a tertiary structure as shown in (**iii**) and comprises a copper binding site with two copper ions that can bind a dioxygen molecule reversibly (PDB-ID: 1JS8). Within the primary structure (**ii**) also splice sites are illustrated: blue arrows symbolize splice sites of linker introns that lie between the FUs (conserved at the same position and in phase 1 for all analyzed hemocyanins); dotted black arrows symbolize splice sites of internal introns. Their number and positions vary among the different gene structures

as their primary oxygen transporter [23]. Despite this, hemocyanin relics of *B. glabrata* have been detected in electron microscope analyses and by SDS-PAGEs by Lieb et al. [23] (hemocyanin without the inner collar). Furthermore, two genes that encode partial, incomplete hemocyanin-like proteins have been identified [22]: One (*BgHcl-1*) consisting of three FUs (FU-a, FU-b and FU-h), the other one (*BgHcl-2*) comprising six FUs (FU-a to FU-f). The gene architectures of both genes comprise the typical phase 1 linker introns between distinct functional units, leading to the conclusion that these genes may indeed be evolutionary remnants of previous full-length hemocyanin genes. Similar to the hemocyanin genes of *P. canaliculata*, both *BgHcl* genes contain a conspicuously larger number of internal introns per FU than those of Lepetellida and Cephalopoda (Fig. 1b). Since the exact number and the positions of internal introns in *BgHcl* are different to those in hemocyanins of *P. canaliculata*, the gene structures must have evolved independently within these different species of Apogastropoda.

The increased numbers of introns which are present in hemocyanin genes of both distinctly-related species of Apogastropoda may be one of a multitude of molecular adaptations that enabled the enormous radiation as well as the great biodiversity of Caenogastropoda and Heterobranchia. Thus, our intent was to reconstruct exon–intron architectures of hemocyanins of additional Tectipleura lineages to investigate if increased numbers of introns, as Peña & Adema [22] have discovered in *BgHcl* genes, are also present in full-length hemocyanins of Tectipleura. This should provide insight into any correlation between adaptation to altered living environments and different gene structures of hemocyanins within this clade of gastropods.

Results/discussion

We accomplished the full reconstruction of exon–intron architectures of ten hemocyanin coding sequences of four different Tectipleura species (*Aplysia californica, Lymnaea stagnalis, Cornu aspersum* and *Helix pomatia*) as well as three hemocyanin coding sequences of two Octopodoidea species (*Octopus bimaculoides*, *Octopus vulgaris*). The publicly available coding sequences of hemocyanins have been updated with annotations of exons under their accession numbers (see Methods). Genomic sequences used to generate gene architectures of hemocyanins of *C. aspersum* and *H. pomatia* were obtained and subsequently assembled, whereas for all other species we used publicly available genomic data (see Methods).

Gene structures of *Tectipleura hemocyanins*: a multitude of conserved introns

The gene structures of all analyzed Tectipleura hemocyanins are very similar (Fig. 3). With the exception of one hemocyanin gene from *Lymnaea stagnalis* (*LsH1*) that includes one additional intron, all full-length Tectipleura hemocyanin genes contain 53 introns within the

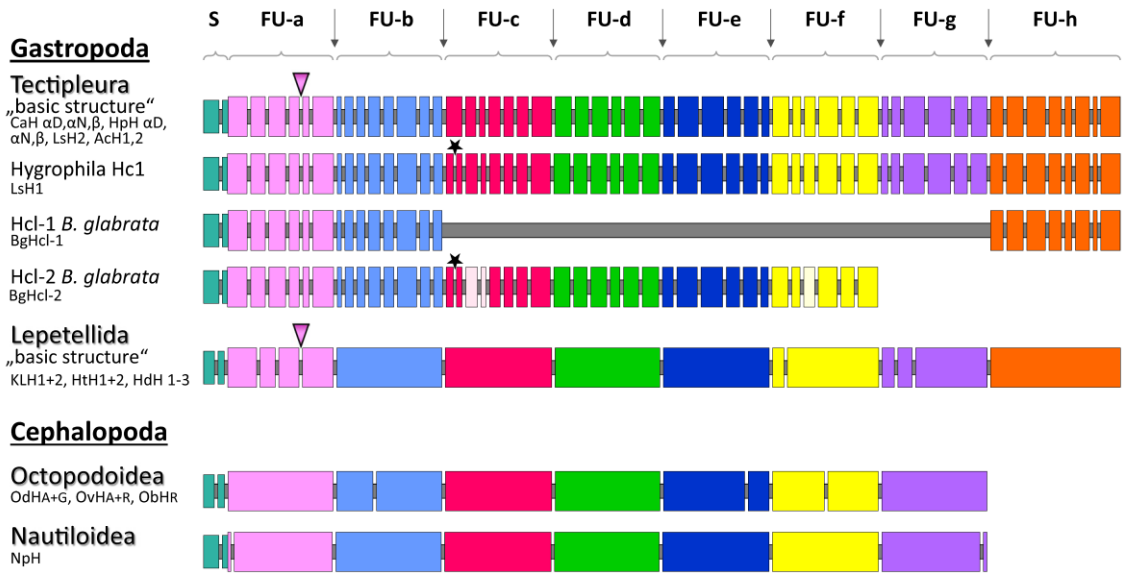


Fig. 3 Gene structures of molluscan hemocyanins. The comparison of exon–intron architectures illustrates the conspicuously larger number of introns in hemocyanin genes of Tectipleura, compared to Lepetellida and Cephalopoda. Big boxes represent exons, small grey ones between them, introns. Positions of plotted splice sites correlate with their position relative to the coding sequence of the FUs. Widths of small boxes are not relative to intron length. Exons of signal peptides (S) and of different FUs are colored differently. The additional Hygrophila-specific intron (only in one of their hemocyanin genes, Hc1) can be seen in FU-c (star) and represents the only difference to the basic structure of Tectipleura hemocyanins. It is also present in *BgHcl-2*. Linker introns are conserved in their position between all FUs of hemocyanins of all species and are marked with grey arrows on top. Pink arrowheads show the only internal introns which are located at the identical position within Tectipleura and Lepetellida. For *BgHcl-2* three internal introns are not fully proven by our analysis (brighter colored in FU-c and Fu-f). Gastropod gene structures are depicted for hemocyanins of Tectipleura (*CaH*: *Cornu aspersum*; *HpH*: *Helix pomatia*; *LsH*: *Lymnaea stagnalis*; *AcH*: *Aplysia californica*; *BgHcl*: *Biomphalaria glabrata*) and Lepetellida (*KLH*: *Megathura crenulata* (Keyhole limpet); *HtH*: *Haliotis tuberculata*; *HdH*: *Haliotis diversicolor*); for Cephalopoda hemocyanins of Octopodoidea (*OdH*: *Enteroctopus dofleini*; *OvH*: *Octopus vulgaris*; *ObH*: *Octopus bimaculoides*) and of Nautiloidea (*NpH*: *Nautilus pompilius*). Cephalopod hemocyanins differ from the basic molluscan hemocyanin structure by lacking FU-h

coding sequence. In addition to the number of introns, the intron positions with respect to the coding sequences and the phases of these introns are identical. We assume that this gene architecture is primordial for Tectipleura hemocyanins because we have found it in the two sister groups Euopisthobranchia and Panpulmonata which together form the clade of Tectipleura (listed species in Fig. 1a)). Corresponding to the molecular clock of Zapata et al. [1] and Kano et al. [9], this implies that this architecture of hemocyanin genes of Tectipleura arose at least 230 ± 50 million years ago (mya).

Figure 3 compares the distribution of introns among the different hemocyanin FUs of Tectipleura and of other molluscs. As in all genes of molluscan hemocyanins, linker introns are located between all functional units (grey arrows in Fig. 3). However, the number of internal introns within Tectipleura hemocyanin genes is strikingly higher than in any published gene architecture of a molluscan hemocyanin so far: For Tectipleura hemocyanins we identified 46 internal introns (in *LsH1* even 47) in addition to the seven conserved linker introns, whereas there are three (*Nautilus pompilius*) and five (Octopodoidea) in Cephalopoda and eight internal introns in Lepetellida (Vetigastropoda) [16–18, 20]. The number of internal introns of Tectipleura hemocyanins is greater than that found in hemocyanin genes of *Pomacea canaliculata*, which comprise 21 to 26 internal introns per gene according to Chiumiento et al. [21]. The *P. canaliculata* gene structures are intentionally omitted from Fig. 3 and the following analyses because we detected inconsistencies in their assembly and annotation (e.g. two out of four genes are represented in the wrong orientation and some splice sites were not or incorrectly identified). However, we discuss these results because they unambiguously report a higher number of introns in hemocyanin genes of *P. canaliculata* compared to those of Cephalopoda or Vetigastropoda.

The positions of internal introns are distributed along the complete coding sequences of Tectipleura hemocyanin genes and are identical between different Tectipleura species (except for one additional intron in the *LsH1* gene of *L. stagnalis*). Exon–intron architectures of the individual FUs of Tectipleura hemocyanins, on the other hand, differ from each other. Each FU-coding sequence

contains six to eight exons varying in size between 97 and 309 nucleotides with a median number of 189. As in other molluscan hemocyanins, internal introns of all three phases are present. The general phase distribution within Tectipleura hemocyanins is 52% of phase 0, 24% of phase 1 and 24% of phase 2. These results match the overall tendency that phase 0 is the most frequent intron phase among different genes of various species [24, 25].

The additional intron in the *LsH1* gene is Hygrophila-specific

LsH1 contains one additional intron in FU-c which has not been detected in the gene structures of any other Tectipleura other than Hygrophila (star in Fig. 3). Therefore, we confirmed the assembled sequence independently via PCR. We also uncovered the presence of this additional intron in one hemocyanin gene of *Radix balthica*, another species of Hygrophila. We searched for this specific nucleotide sequence section within genomic NGS data of *R. balthica* and found two corresponding hemocyanin sequences: one containing an additional intron as per *LsH1* and one without an intron at this position – like in *LsH2* and other Tectipleura hemocyanins. The hemocyanin genes to which those sequences belong to were named *RbH1* and *RbH2*, respectively. This additional intron is also present in *BgHcl-2*, as revealed by our analysis of this hemocyanin-like gene of *B. glabrata* (see below). Therefore, it can be considered as a Hygrophila-specific intron which evolved after the radiation event of this lineage from a Tectipleura ancestor.

Hemocyanin-like proteins of *Biomphalaria glabrata* and full-length hemocyanins of other Tectipleura share the same gene structure

Previous to this study, only the exon–intron architectures of two hemocyanin-like protein genes of *Biomphalaria glabrata* (*BgHcl*-1, *BgHcl*-2) have been published for species of the clade of Tectipleura [22]. *Biomphalaria glabrata* belongs to Planorbidae, a family of Hygrophila which represents an exception within the clade of Tectipleura. Instead of hemocyanin, *B. glabrata* uses a hemoglobin as oxygen transporter which most probably evolved exclusively within this gastropod family [23, 26]. Thus, the hemocyanin isoforms of these species are probably not essential any more for oxygen transport but may have been repurposed to novel functions (suggestions for further functions *cf.* [22]). Furthermore, the detected proteins contained only three and six functional units, respectively, instead of eight FUs which is typical for other gastropod hemocyanins [22, 23].

The comparison of exon–intron structures of Tectipleura hemocyanin genes with those of hemocyanin-like protein genes of *B. glabrata* (Fig. 3) shows that all splice sites of the *BgHcl*-1 gene are located at the same position in the corresponding FUs (a, b and h) of the analyzed Tectipleura hemocyanin genes [22]. Through our reanalysis of the draft *B. glabrata* genome assembly, which was used by Peña and Adema to determine the coding sequence of the *BgHcl*-2 gene [22], we identified sequences that were defined as introns but are characteristic of hemocyanin coding sequences. Therefore, we assume that these are genuine exons and thus we suggest an expanded coding sequence and a novel gene structure. The revised exon–intron architecture of *BgHcl*-2 is highly similar to the gene structure of other Tectipleura hemocyanins (Fig. 3).

Differences of this exon–intron structure to those of full-length hemocyanins of Tectipleura are only present in sequence sections that were poorly resolved within the published genome of *B. glabrata*: In FU-c and in FU-f we have found sequence motifs which are usually conserved in hemocyanins but could not be detected in the *BgHcl*-2 gene. Also, the number of bases was smaller than characteristic for those hemocyanin functional units. Within FU-c the undetected parts correspond to two Tectipleura exons (exon 3 + 4). The section of the genome assembly in which those exons would have been expected includes ambiguous nucleotides (NNN's). For FU-f the 3′ section of the third exon could not be determined due to stretches of NNN's. On the basis of this draft genome, it is currently impossible to determine whether those exons are missing in *BgHcl*-2 or if they are just not properly assembled. In Fig. 3 these exons are included but are highlighted with brighter colors. Additionally, two of the newly identified exons were found in a reverse orientation within the draft genome (FU-b exon 4 and FU-f exon 1). Since they correspond to nucleotide sequences and exon–intron structures of other Tectipleura hemocyanins, we propose that they constitute exons of the *BgHcl*-2 gene. Incorrectly oriented exons are common errors in automated computational draft genome assemblies. Including these hypothetical exons in a predicted cDNA sequence, the gene architecture of *BgHcl*-2 has the same exon–intron structure of FU-a to FU-f of the *LsH1* gene in terms of number, positions and phases of introns (including the additional intron in FU-c which has been found exclusively in *LsH1* and *RbH1* genes; stars in Fig. 3).

This strict conservation of gene structures is particularly interesting, since the amino acid sequences of both *BgHcl* proteins differ significantly in sequence motifs that are otherwise conserved through all analyzed molluscan hemocyanins including those of Tectipleura, Vetigastropoda and Cephalopoda (see differences highlighted in the alignment in Additional file 2). The exon–intron structures of *BgHcl* genes, which remained consistent despite

all these deviations from characteristic sequence features of other hemocyanin genes, highlights the strong conservation of introns within Tectipleura hemocyanin genes. Additionally, the identical splice site positions corroborate the hypothesis of Peña and Adema that those genes are descendants of the hemocyanin gene family. This suggests that the large number of introns should not be considered as a novel gain within hemocyanin-like genes and is, therefore, probably not linked to the evolution of novel functions of hemocyanin-like proteins in Planorbidae as suggested after their first discovery [22].

Accumulation of introns in Tectipleura is hemocyanin gene-specific

To exclude that intron accumulation is a general phenomenon in Tectipleura genomes, we compared average numbers of introns as well as exon sizes of over 15,000 genes of two Tectipleura species, specifically *Aplysia californica* (GCF_000002075.1) and *Radix auricularia* [27]), with those of two Octopoda species, namely *Octopus bimaculoides* (GCF_001194135.1) and *O. vulgaris* (GCF_006345805.1). We found no general increase in the number of introns within Tectipleura genomes (Additional file 3a). Additionally, we compared exon counts of orthologous genes of *A. californica* and *O. bimaculoides* as well as *A. californica* and *O. vulgaris* (results shown in Additional file 3b). These exon count comparisons must be considered carefully because the underlying data encompass all data published in NCBI and were thus produced by different studies using different strategies of sequencing and data processing. Hence, comparability of these data is limited (cf. [28]). However, the results exhibit trends of gene architecture evolution which do not comprise extensive intron gains or losses between orthologs of *A. californica* and the two *Octopus* species. While the differences in exon counts of hemocyanin genes between these species is 35 (only including FU-a – Fu-g of *AcH* for comparability), we have only identified 7 and 6 orthologous genes that vary in exon count in more than 25 introns between *A. californica* and the two *Octopus* species, respectively (shown in table in Additional file 3b). 50% of the numbers of differences in exon counts encompass only up to one intron change and 80% up to four (*O. bimaculoides*) and up to six intron changes (*O. vulgaris*), respectively. Although this is only an approximate trend based on a preliminary analysis, widely varying numbers of introns between orthologous genes of Octopoda and Tectipleura seem to be characteristic for only few genes and do not represent a general phenomenon within these genomes. Apparently, the extensive accumulation of introns we observe in hemocyanins of Tectipleura is gene-specific. The strong lineage-specific conservation of these introns across disparate molluscan clades, on the other hand, argues against a random variation of internal introns but rather suggests a significant selection pressure acting on this gene structures. We therefore performed more in-depth analyses on this aspect of these genes.

Highly conserved linker introns

Previous studies on molluscan hemocyanin gene structures revealed strong conservation of the so-called linker introns of hemocyanins (see above and [16]). As shown in Fig. 3, they are located within the linker peptide coding regions between all functional units of hemocyanins throughout different molluscan classes, including Tectipleura (grey arrows in Fig. 3). They are without exception phase 1 introns (located after the first nucleotide of a codon). Since they are shared by all modern molluscan clades, they must have evolved prior to the radiation of molluscs into different classes. Lieb et al. [16] assumed that this occurred during duplications of FUs through which the eight-FU hemocyanin molecule arose from an ancestral mono-FU. Despite the increasing amount of molluscan hemocyanin sequence data during the last decades, the chronological order of these duplication events and the exact evolutionary origin and function of linker introns remains unresolved.

Although linker introns of hemocyanin genes and their positions are conserved through all molluscan clades, we did not find any indications of concerted evolution, transposition or conserved regulatory elements by comparing their nucleotide sequences. Comparative studies of exon–intron structures of human genes identified high positional conservation as typical characteristics of introns carrying important functions [29]. Accordingly, these strictly conserved linker introns in molluscan hemocyanin genes might be functional, even though they vary in their sequences. Altenhein et al. [17] suggested, for example, that they could play a role for correct transcription of these giant polypeptides. However, since the nucleotide sequences of a multitude of these linker introns are not fully assembled yet, a final conclusion on their functional significance is still pending.

Internal introns

The similarity of hemocyanin gene architectures of Tectipleura suggests that their exon–intron structures arose in a common ancestor and stayed almost the same for $> 230 \pm 50$ mya (according to molecular clocks of [1] and [9]). Considering that linker introns of hemocyanin genes in Tectipleura match all other known hemocyanin genes of molluscs in terms of their positions and phases, we concentrated on the internal introns which lie within the functional units of hemocyanin genes. Splice sites of internal introns of molluscan hemocyanins have not been

found to be highly conserved in previous studies. Lieb et al. [16] showed that gene structures of hemocyanins in *Enteroctopus dofleini* (Cephalopoda, Octopodoidea) and in *Haliotis tuberculata* (Vetigastropoda) are totally different concerning numbers, positions, lengths and phases of their internal introns (Fig. 3). Bergman et al. [18] revealed that internal introns vary completely even between the cephalopod hemocyanins of *E. dofleini* and *Nautilus pompilius.* In contrast, the two different hemocyanin genes identified in *H. tuberculata* possess a highly similar genomic structure [17]. A comparable exon–intron architecture was also found for the three hemocyanin genes of *Haliotis diversicolor* [20] and the two of *Megathura crenulata* (*KLH1* and *KLH2*) [19]. Like Haliotoidea, *M. crenulata* belongs to Fissurelloidea, a group of Lepetellida (Vetigastropoda). Although the internal introns of hemocyanin genes through all these Lepetellida species vary in length and in their nucleotide sequences, they are located at the same position and in the same phase with respect to the coding sequence within all analyzed Lepetellida hemocyanins. This most probably reflects a common origin of evolution.

Lineage-specific conservation of hemocyanin gene structures contrasts with high variability between major molluscan clades

To investigate whether internal introns are also conserved within the hemocyanin genes of closely related cephalopod groups, we analyzed hemocyanin genes from the published genomes of *Octopus vulgaris* and *Octopus bimaculoides*. We found that the derived exon–intron architectures of their hemocyanin genes correspond to those found in *Enteroctopus dofleini* concerning intron number, location and phase (Fig. 3). Due to their identities with hemocyanin genes of *E. dofleini* (shown in Additional file 4), we named the hemocyanin genes of *O. vulgaris OvH* Type-A (XM_029780310.1) and Type-R (XM_029796515.1) and the hemocyanin gene of *O. bimaculoides ObH* Type-R (XM_014934350.1; XM_014934481.1). As shown above, Tectipleura also displays a very conserved exon–intron structure of their hemocyanin genes. We therefore conclude that lineage-specific positional conservation of internal introns might be a common feature of molluscan hemocyanins.

A first comparison of the conserved exon–intron structure of Tectipleura hemocyanin genes with those splice sites of hemocyanin genes of *P. canaliculata*, which we were able to verify, showed that their exon–intron structures differ substantially from each other (e.g. 80% of the splice sites of the gene *PcHI* are absent within Tectipleura hemocyanins). Hence, we assume that the accumulations of internal introns in hemocyanin genes of *P. canaliculata* and in those of Tectipleura evolved independently from each other and thus arose most probably after the radiation of Apogastropoda into Caenogastropoda and Heterobranchia.

High conservation of intron positions as found for hemocyanin genes of Tectipleura is a frequently occurring phenomenon in orthologous genes of animals [30–32] but it is particularly striking for hemocyanin genes of Tectipleura, since they feature 46 conserved internal introns instead of three (*Nautilus pompilius*), four (Octopodoidea) or seven (Lepetellida) and all of them stayed at the exact same position for more than 200 million years. These results are furthermore surprising as they do not correspond with the hemocyanin gene structures of Caenogastropoda elucidated by Chiumiento et al. [21]. On one hand, their results indicate that accumulation of introns as we have detected in Tectipleura is a common feature of hemocyanin genes of Apogastropoda, but on the other hand, the four hemocyanin genes of *P. canaliculata* vary in number and positions of introns and do not show such a strong conservation as we have identified for Tectipleura.

Against this background, the strict conservation of the gene structure of Tectipleura hemocyanins appears not to be random. Instead, it may rather have been caused by an evolutionary pressure towards conservation of this exon–intron structure due to its supposed functional roles. Chorev et al. [29] described a correlation between, on one hand, high positional conservation of introns as well as low intron loss rates and, on the other, intronic functions. This might also be the case for internal introns of Tectipleura hemocyanin genes.

So far, we were not able to assemble all introns and to analyze them on sequence levels, because of large and highly repetitive sequence sections. We did, however, analyze 107 intron sequences of the two genes of *LsH1* and *LsH2* found in a soon to be released genome of *Lymnaea stagnalis.* Our sequence similarity searches did not reveal conserved sequence motifs within these intron sequences, neither by using the NCBI tool 'BLAST' [33], nor by applying our self-created databases of the *LsH* intron sequences.

However, it seems that introns rarely exhibit elevated degrees of conservation detectable with conventional similarity searches [29]. Small conserved motifs like cis-regulatory elements which often contribute to transcriptional regulation [33, 34] are hard to detect within long intron sequences such as those deciphered in hemocyanin genes of *L. stagnalis* (min: 118 bp; max: 4,657 bp; mean value 531). To enable the identification of short functional sequence sections within long mostly non-functional intronic sequences, a significantly larger dataset would be needed. Besides these difficulties, many intron functions do not depend on sequence motifs.

Some functions rather depend on intron length [35] or on their positions along the mRNA [29, 36, 37].

In fact, the large number of internal introns with strictly conserved positions in hemocyanin genes of Tectipleura suggests that they may also exert functions that depend on their positions within the genes. The multitude of introns we have identified enables us to analyze the positions of internal introns in detail by drawing comparisons between internal introns of different lineages as well as of different FUs of hemocyanins of Tectipleura.

Recent evolution of internal introns

Generally, there are two possible scenarios of intron evolution: Introns within a nuclear gene can be very ancient or can be added later during evolution of the gene (see review [38]). The clear segmentation into functional units concatenated by linker introns in different molluscan lineages (Fig. 3) implies that the basic structure of molluscan hemocyanin genes evolved prior to or during the radiation of molluscs into different clades and then remained conserved. In fact, the origin of linker introns has already been described as ancient [16]. While they remained conserved during speciation events, internal introns changed independently from each other in different FUs and different taxa. Consequently, this led to clade-specific patterns of gene structures that became fixed more recently during evolution in a lineage-specific manner, as seen in Octopodoidea, Lepetellida and Tectipleura (Fig. 3).

The overlay of hemocyanin functional units in Fig. 4 shows the positions of internal introns through all molluscan hemocyanin genes analyzed in the present study. More precisely, introns from all functional units of *N. pompilius*, Octopodoidea, Lepetellida and Tectipleura are marked together within one model FU-coding sequence. Due to ambiguities of the respective report, exon–intron architectures of hemocyanin genes of *P. canaliculata* [21] were not included in this figure.

Overall, the distribution of internal introns illustrated in the model FU in Fig. 4 exhibits a number of splice sites from different molluscan groups or different FUs which are located at equal positions with respect to the coding sequence (multiple splice site positions). About 48% of the splice sites occur at identical positions in two or three different FUs (of the same or of different hemocyanin genes, see color coding). Thereby, only one of them is observed at the same position within the same functional unit of hemocyanin genes from different groups (FU-a of Lepetellida and Tectipleura hemocyanins; pink arrowheads in Fig. 3 + 4). It is the only splice site that is located at the identical position concerning the whole coding sequence of the hemocyanin polypeptide within two different lineages. We call it an "ortholog intron" of Lepetellida and Tectipleura. Since both groups are gastropods, the introns can be real orthologs which arose before the gastropod radiation into the different lineages and remained conserved until today.

All other multiple occurring splice sites, however, are located within different functional units of hemocyanin genes of the same or of different molluscan lineages. If these splice sites were homologous as well, they must be "paralogous" to each other because they would have derived from an ancient intron due to tandem duplications of FUs: Regarding the positions within different FUs as paralogs they must have been gained in the mono-FU hemocyanin ancestor and then duplicated during the evolution of the eight-FU protein (example represented in Fig. 5). Since this would implicate that these splice sites were lost within much more FUs than they would have stayed conserved, we do not support this hypothesis as origin. In contrast to that, independent intron gains within different FUs during more recent evolutionary times represent a much more parsimonious scenario (Fig. 5).

Additionally, 52% of the detected hemocyanin splice sites occur exclusively in one functional unit of one

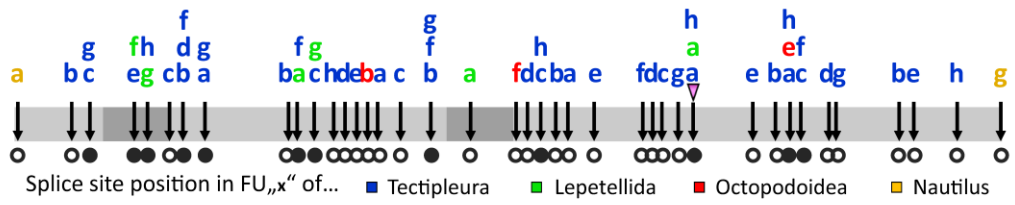


Fig. 4 Distribution of splice sites in hemocyanin FUs. Splice site positions (arrows) are shown within the overlay of FU-coding sequences of hemocyanin genes (grey bar; darker parts: copper binding sites). They are spread over the sequence in clusters. Letters above indicate in which FU and in which molluscan group (color) they are located. About 48% of all internal introns are located at identical positions as introns in one or two other FUs (multiple splice site positions; marked with black dots). Positions that are unique for one splice site are marked with an open dot. The pink arrowhead shows the only internal introns which are located at the identical position within Tectipleura and Lepetellida (same position in ortholog FU (FU-a))

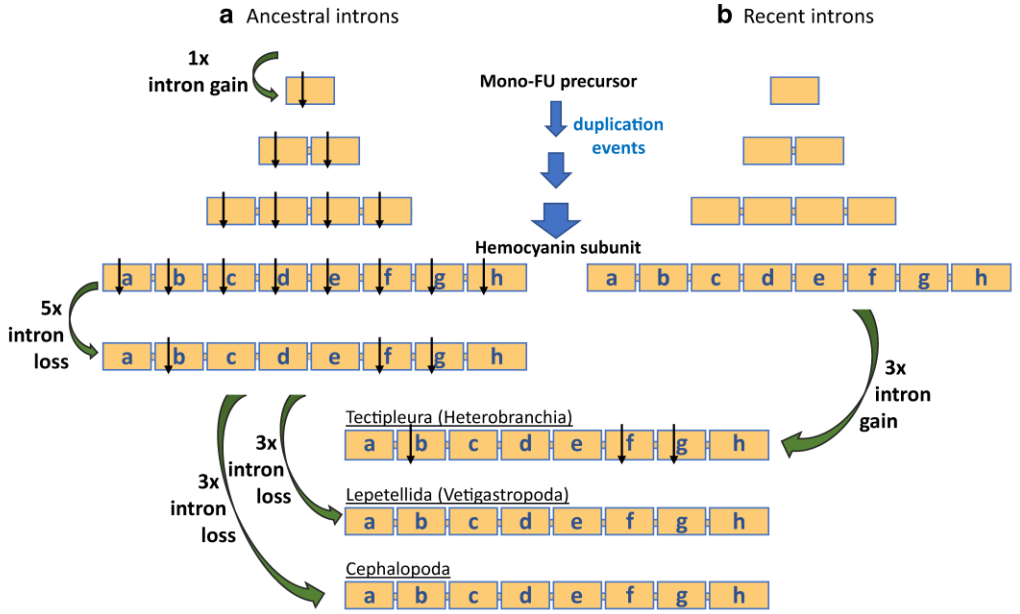


Fig. 5 Scenarios of intron evolution in hemocyanin genes. Intron evolution is exemplified by evolution of introns at identical splice sites of FU-b, FU-f, FU-g (*cf.* Fig. 4). Shown are two scenarios: **a** The introns at the identical position are derived from an ancestral intron which evolved within the mono domain precursor. In this scenario the intron maintained at the identical position during the evolution of the eight-FU hemocyanin. It then stayed conserved within FU-b, FU-f and FU-g of Tectipleura hemocyanins until today, whereas it was lost within the remaining Tectipleura FUs and all FUs of Vetigastropoda and Cephalopoda. In scenario (**b**), the introns evolved more recently after the radiation events of Gastropoda into Vetigastropoda and Heterobranchia. The intron gains took place independently from each other at the identical position of FU-b, FU-f and FU-g of Tectipleura. Scenario **a** includes at least 12 intron gain/loss events, scenario **b** only three intron gains. Thus, the more recent intron gains are the more parsimonious scenario

lineage-specific hemocyanin gene structure (unique splice sites). This can also be explained best by independent intron gains. Thus, we propose independent and recent gains for most internal introns of molluscan hemocyanin genes. This is further corroborated by a comparison of molluscan hemocyanin genes with splice sites of different Type-III copper protein genes (data not shown) that share the same origin with hemocyanins [39]. None of the 26 analyzed genes indicate ancestral intron evolution.

Identical intron positions due to conserved proto-splice sites?

Our analysis revealed highly conserved sequence motifs of exon–intron boundaries. The weblogos in Fig. 6 show the nucleotide distribution of exon and intron boundaries of internal splice sites (five nucleotides up- and downstream of each splice site). 99.3% of the intron sequences we analyzed start with the dinucleotide GT and end with AG. Therewith, they fulfill the Chambon's / GT-AG rule [40, 41]. This is in accordance with other studies that revealed GT-AG boundaries for over 98.5% of introns [42, 43]. As exceptions from this rule, the splice sites GC-AG were found twice in Hygrophila and Stylommatophora (Fig. 6). They represent 0.7% of all Tectipleura splice sites. This also fits to the results of Burset et al. [43] who described GC-AG as the second most frequent splice site covering 0.6% of all intron boundaries within their study.

In contrast to intronic splice sites, exonic splice sites vary greatly. However, an accumulation of proto-splice sites at exon boundaries can be detected within the consensus sequences represented by the weblogos (CAG at 3′ and GT at 5′ ends of the exons, Fig. 6). This tendency is stronger within Tectipleura hemocyanin genes than in those of Lepetellida, Octopodoidea and *N. pompilius* (taken together in Fig. 6). For Tectipleura as well as for the other three molluscan groups a distinction is made between splice site positions that occur just once (unique splice site positions) or in at least two different FU-coding sequences of the analyzed hemocyanin genes (multiple splice site positions; *cf.* open vs. filled dots in Fig. 4+6). Multiple splice site positions of both groups show an increased occurrence of AG|G proto-splice sites. In contrast, the presence of proto-splice sites at unique splice site positions is significantly less frequent in Tectipleura hemocyanins and not at all increased for non-Tectipleura hemocyanins.

Schäfer *et al. BMC Ecol Evo* (2021) 21:36

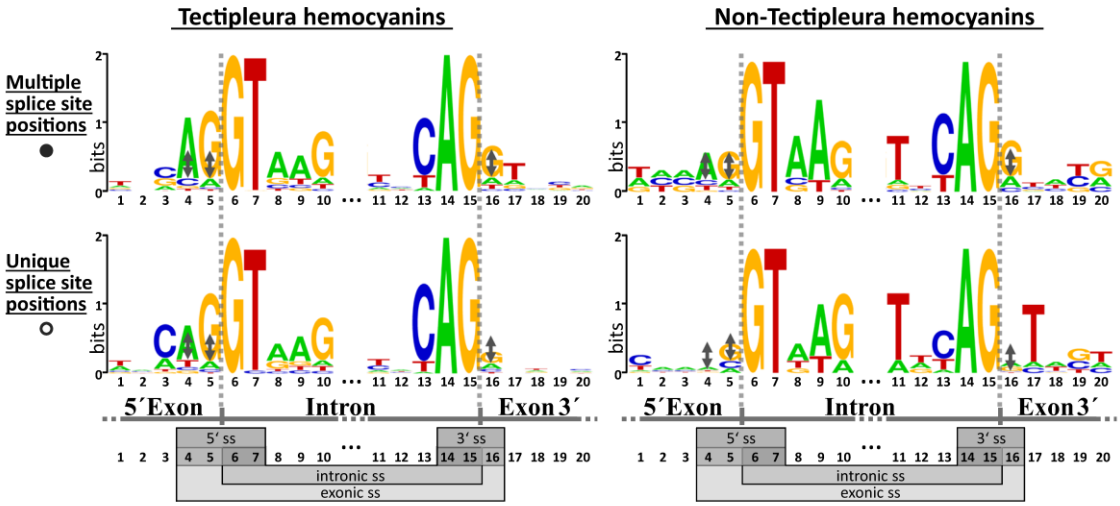


Fig. 6 Nucleotide distribution at splice sites (ss) of internal introns. Five nucleotides up- and downstream of each splice site are shown in weblogos. The height of the nucleotide letters indicates the frequency of each nucleic acid at the respective position. Sequence logos at the left depict splice sites of hemocyanins of Tectipleura; those at the right of Lepetellida, Octopodoidea and *Nautilus pompilius* (taken together as non-Tectipleura hemocyanins). Internal introns are further distinguished in introns at multiple (top, black dot, *cf.* Fig. 4) and introns at unique splice site positions (bottom, open dot, *cf.* Fig. 4). Multiple splice site positions are defined as positions along the cDNA where at least one different intron is located (*cf.* Fig. 4). This can be located in a different FU of the same gene structure type or in any FU of a different gene structure type. At unique splice site positions, no introns of other FUs/hemocyanins are located at the identical position with respect to the coding sequence (*cf.* Fig. 4). The intronic splice site GT-AG is highly conserved throughout all hemocyanins. Additionally, weblogos display the existence of the (exonic) proto-splice site AG|G. For better comparison of the frequency of AG|G in the four different weblogos, black double arrows of the same height are given. While all proto-splice site nucleotides of multiple splice site positions of Tectipleura and non-Tectipleura (upper ones) are higher than those reference arrows, some of the unique splice site positions are much smaller

These results display a larger frequency of proto-splice sites at positions that harbor introns more often. This is consistent with previous studies which showed that splice sites often include canonic nucleotide patterns like (C/A)AG|Gt [44–46]. Thereby, AG|G is significantly more frequent than any other exon boundary [47–50]. This is in accordance with our findings (Fig. 6). Although this splice site cannot be critical for correct splicing due to its relatively low conservation, a positive selection in its favor can be assumed. It has been hypothesized that canonical proto-splice sites might, for example, improve the efficiency of splicing while they simultaneously offer a wide flexibility for protein coding and the evolution of the coding sequence, if the respective introns are in phase 0 [50]. It is still unclear if the accumulation of specific splice site motifs arose due to more frequent intron gains or due to a higher positional conservation. Both reasons, however, might have led to the increased occurrence of multiple splice sites at such positions during the evolution of hemocyanin genes.

Thus, it can be assumed that internal introns which are located at identical positions in multiple FUs might have been gained convergently. This is in accordance with the maximum parsimony scenario of more recently and independent gains of internal introns described above. A correlation between highly conserved proto-splice sites and more frequent convergent intron gains has already been discovered by the studies on elongation factor genes of Hymenoptera by Klopfstein and Ronquist [51].

Splice site positions of internal introns do not correlate with module boundaries

The distribution of splice sites along the sequence of a model FU shows an accumulation of introns in clusters, whereas other sequence sections are intron-less (Fig. 4). Alignments of amino acid sequences showed that splice sites lie in highly conserved as well as in rarely conserved regions (alignments in Additional files 5 and 6). Additionally, our analyses revealed that they are located within sequences coding for alpha helices and beta sheets as well as for random coils (videos in Additional file 7). We could not identify specific positional conservation of internal introns that correspond to hypothetical module boundaries in ancient proteins as described by de Souza et al. [52] and, for example, as for linker introns which lie between the single FUs [16]. Accordingly, the clusters of introns cannot be explained by structural features nor by high or low conservation patterns of the coding sequence.

These clusters, however, could result from the existence of linker introns which lead to a regular fragmentation of hemocyanins in similar composed FUs. The coexistence of regular size distribution of exons which we found within these FUs of Tectipleura hemocyanins (97 – 309 nucleotides) must consequently have led to an incorporation of introns at similar positions with respect to the coding sequences of the eight different but relatively equal constructed FU fragments. Thereby, the roughly regular subdivision of FUs by internal introns leads to clusters in the overlay of several FUs.

In spite of our analyses, we are not able to deduce any importance for positions of internal intron splice sites which might be crucial for their possible functions. They rather seem to be distributed in a roughly regular manner within the FUs with the exception of introns at protosplice sites which most probably were gained several times convergently at those positions and remained there until today. Thus, it remains open why hemocyanin genes of Tectipleura comprise such a large number of introns in contrast to other molluscan hemocyanin genes.

Evolutionary pressure, adaptive radiation and habitat shifts

The conservation of such an extensive number of introns throughout long periods of evolution within different molluscan lineages is especially remarkable for Tectipleura, since they cover a distinctly larger number of introns. Although this increases the probability of intron loss or sliding, we did find none of them in the hemocyanin genes of Tectipleura species we analyzed in this study. As already described, Hygrophila comprise an additional internal intron within FU-c which is only conserved within this clade. However, these hemocyanin genes also comprise all primordial Tectipleura introns.

Furthermore, Tectipleura, whose radiation can be traced back to ~ 230 ± 50 mya (according to molecular clocks of [1] and [9]) represent a conspicuously larger clade (~ 27,100 species) compared to Lepetellida (~ 950 species) or Octopodoidea (~ 240 species) which also possess conserved exon–intron structures in their hemocyanin genes (for numbers see [2]). This indicates that Tectipleura were subjected to a stronger degree of radiation and an increased evolutionary rate. Such preconditions are normally expected to also increase the probability of changes on gene structures unless no evolutionary pressure would have acted upon these introns.

In addition to a multitude of speciation events, Tectipleura have undergone a variety of habitat shifts during their evolution which led to radical changes of their living conditions. In contrast to Vetigastropoda, which include very early branching and exclusively marine living gastropods, Tectipleura comprise very diverse snails and slugs living in marine (e.g. *A. californica*), freshwater (e.g. *L. stagnalis*), terrestrial (e.g. *H. pomatia*) as well as intermediate environments (Fig. 2). Consequently, they must have undergone several adaptations such as an increased capacity for osmoregulation, tolerance to fluctuating temperatures and varying degrees of water availability, as well as the capacity to breath air [10]. This resulted in a range of diverse and complex species. The respiratory systems of Tectipleura, for example, were modified in many ways through the evolution of lungs and of pneumostomes [7, 10], or through duplications of hemocyanin genes that took place [11]. Regarding all these changes during the evolution of different Tectipleura lineages, the exact conservation of so many splice sites within these genes appears extraordinary, indicating a high degree of evolutionary pressure for their conservation. In the absence of evolutionary constraints, however, a more continuous alteration of exon–intron structures of hemocyanins would have been expected including less sudden extreme changes between different molluscan lineages followed by a long period of high conservation of special gene structures. For Tectipleura, for example, the gene structure including the multitude of internal introns evolved after the radiation of gastropods into their main groups (including Vetigastropoda and Heterobranchia) but within a common ancestor of Tectipleura. According to the molecular clock of Zapata et al. [1], this implies that this gene structure which varies in 45 introns from that of Vetigastropoda evolved between ~ 480 ± 50 mya and ~ 230 ± 50 mya and thus within 250 ± 100 million years, whereas it then stayed strongly conserved for at least ~ 230 ± 50 million years. These strongly contrasting evolutionary rates may have occurred from changed evolutionary pressures.

In spite of the many conserved internal intron positions, however, we could neither identify any conserved sequence motifs within introns nor within splice site positions that correlate with module boundaries indicating specific functions. Consequently, no particular evolutionary pressure can be recognized acting on the hemocyanin gene architecture, neither with respect to intron sequences nor to their positions within the respective hemocyanin genes. However, a selection pressure that led to the strong intron conservation within Tectipleura hemocyanin genes, at least, cannot be excluded. This selective pressure could have acted, for example, on the number of introns which may bring advantages due to different functions or regulatory capacities.

One function of introns which has been proposed by Bonnet et al. [53] is that they help to protect genomes of eukaryotes from 'transcription-associated genetic instability'. The results of their study show that the recruitment of spliceosome onto the mRNA decrease R-loop

formation during transcription which represent stable hybrids between mRNA and DNA that lead to single-stranded DNA. The aggregation of many R-loops which can result from high expression rates of a gene, however, can cause DNA damage. Thus, high degrees of introns in genes, which lead to spliceosome recruitments, can help to protect DNA stability. R-loop prevention due to introns has been shown to be conserved for yeast as well as for human [53]. Therefore, it might also play a role in molluscs and could lead to an evolutionary pressure lying on introns as it might be the case for hemocyanins of Tectipleura.

Against the background that hemocyanin genes of *P. canaliculata* also possess an increased number of introns which most probably evolved independently from those of Tectipleura, our results also suggest a connection between a high rate of introns in hemocyanin genes and an enormous diversity of species and habitats. The multitude of introns in hemocyanin genes of Apogastropoda were gained most probably after the radiation into Heterobranchia and Caenogastropoda independently within these clades. Thereby, they potentially could have offered an exceptionally great chance of regulation and based on this also an increased rate of adaptive radiation and speciation.

Previous studies have shown that introns in general can provide a variety of regulation possibilities like incorporation of noncoding RNA genes [54–56], controlling of mRNA transport [57] and enhancement of gene expression [58]. Eghbalnia et al. [59], for example, analyzed exon:intron expression to further understand the role of exon and intron counts and concluded that changed intron expression may be involved in regulation of physiological processes. Overall, splicing can have regulatory effects on all levels of gene expression and may help to adapt to new living conditions [60, 61]. Gotic et al. [62] identified temperature-mediated splicing efficiency of pre-mRNA as an important control mechanism of gene expression of β-globins. It might be a widespread post-transcriptional mechanism to regulate mRNA accumulation and protein biosynthesis due to splicing factors [63]. For example, temperature-dependent splicing controls flowering at distinct temperatures in plants [64] or helps *Drosophila* to adapt to different temperatures [65]. Referring to hemocyanins, the large number of internal introns may allow a graduated temperature-dependent splicing which might help to regulate gene expression of hemocyanins at different temperatures. Such regulatory mechanisms offer great adaptability which is particularly important for limnic and especially terrestrial gastropods because they underlie a much stronger fluctuation of temperatures and other environmental conditions than marine molluscs. Temperature, on the other hand, deeply influences the oxygen affinity of hemocyanins [3–6]. This illustrates the significant effects that habitat shifts can have on the respiratory system and reflects the necessity of adaptations of these oxygen transporters to new living conditions.

Intron positions that possibly correlate with transitions to terrestrial habitats have also been described for plants [66]. Our study now provides some indication that such a correlation may also exist for Tectipleura, and maybe for animals in general. To adapt to decreased oxygen affinities (which result from increased temperatures), hemocyanin expression might, for example, be upregulated to ensure the transport of a sufficient amount of oxygen.

Previous analyses of hemocyanin genes [11] have disclosed a number of gene duplications which took place multiple times independently in different Tectipleura lineages. This suggests that multiple hemocyanin genes, which can be expressed differentially, might help to adapt to varying ecological conditions. This hypothesis was corroborated by the results of Chiumiento et al. (2020) who identified four hemocyanin genes in the *P. canaliculata* which represent further examples of multiple hemocyanin gene duplications associated with adaptations to air breathing lifestyles. Differential expression of isoform genes is an essential mechanism to adapt to different environmental conditions and has already been shown for hemocyanin genes of *Sepia officinalis* (Cephalopoda) [67]. In connection with different hemocyanin isoforms that possibly exhibit different oxygen affinities [11], a temperature-dependent splicing would increase hemocyanin variability which not only could enhance gastropod adaptability but also biodiversity and radiation. Due to the increased chance of adaptive radiation, the extensive number of conserved introns in hemocyanin genes may have supported diversification and speciation within Caenogastropoda and Heterobranchia, which represent by far the largest and most diverse groups of gastropods.

Conclusions

The hemocyanin genes of Tectipleura possess a significantly larger number of introns than any other analyzed molluscan hemocyanins. In this study we show that the majority of them most probably originated from intron gains which took place after the radiation into the major gastropod clades. Nevertheless, the exon–intron architectures of Tectipleura hemocyanin genes have been conserved for more than 200 million years concerning the number and positions of introns. The conservation of splice sites in these genes may be due to as yet unidentified functional roles. The increase in the number of introns in Tectipleura hemocyanin genes may have supported the frequent habitat shifts observed in this clade. For example, a multitude of introns could increase the

possibilities for alternative gene regulation and help a species to adapt to new living conditions, such as strong temperature fluctuations. This potential genetic plasticity may have supported the transition to new environments (e.g. land and freshwater) and the radiation of the Tectipleura. Our findings outline another dimension of the importance and the fundamental functions of introns and suggest the need for additional research. Thereby, molluscan hemocyanins represent a good tool to further investigate the role of introns during habitat shifts because the link between changed habitats and the need of adaptations of respiratory systems, including hemocyanins, is obvious.

Methods

Animal sampling and DNA isolation

Adult snails of *Cornu aspersum* and *Helix pomatia* were obtained from a commercial dealer (Wiener Schneckenmanufaktur e.U., Vienna, Austria). Three individuals of the species *Lymnaea stagnalis* were collected from a pond in Mainz. Three individuals of each species were anesthetized on ice for 20 min and subsequently sacrificed by quickly cutting off the head to minimize pain. Hepatopancreatic tissue was isolated on an ice cooled aluminum plate. Tissue aliquots were stored in RNAlater™ (Invitrogen by Thermo Fisher Scientific, Waltham, MA, USA) at − 80 °C. Samples were homogenized with a Precellys® homogenizer (Bertin Instruments, Montigny-le-Bretonneux, France). Subsequently, DNA was isolated applying the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). DNA integrity was checked on a 0.8% agarose gel (Biozym Scientific GmbH, Hessisch Oldendorf, Germany) and quantified via Nanodrop (ThermoFisher).

Compiling exon–intron structures of hemocyanin genes of *Cornu aspersum* and *Helix pomatia*

DNA samples from one adult snail of *C. aspersum* and *H. pomatia* were sent to StarSeq (Mainz, Germany) for NGS (Illumina Next Seq500) and library preparation. Bioinformatics were performed using Geneious 9.1.8 [68]. Sequencing adapters were removed and raw reads were quality trimmed. Processed genomic data were mapped to coding sequences of three known hemocyanins of *C. aspersum* (*CaH αD*: MH485355, *CaH αN*: MH485356, *CaH β*: MH485357) and *H. pomatia* (*HpH αD*: MH485358, *HpH αN*: MH485359, *HpH β*: MH485360). Sequence sections which were not covered by genomic NGS data or which were incongruous to cDNA sequences were used to separate nucleotide sequences in different parts which represent segments of different exons. Exon sequences were completed by repetitive mappings of genomic data to these sequence parts until non-cDNA sequences were assembled. At least ten base pairs of the 3′ and 5′ ends of each intron were assembled to assure that flanking sequences differ from those of neighboring cDNA sequences and therefore represent introns.

Compiling exon–intron structures of hemocyanin genes of *Lymnaea stagnalis*

Sequences of the previously published coding sequences of two hemocyanins found in *L. stagnalis* (*LsH1*: MH485363; *LsH2*: MH485364) were searched within the draft genome of *L. stagnalis* [69] using BLAST. Genomic scaffolds which included hemocyanin isoforms were aligned to amino acid sequences using GeneWise [70] to derive the exon–intron architectures of *LsH1* and *LsH2*. Those gene structures were verified by mapping genomic raw data (SRA (NCBI): ERR1083352 to ERR1083359) to the deduced exon sequences. The *LsH1*-specific intron which was additional to those found in other Tectipleura hemocyanins (Fig. 3, star) was additionally confirmed via PCR (see below).

Compiling exon–intron structures of hemocyanin genes of *Aplysia californica*

Large parts of the exon–intron architecture of *AcH1* have already been detected by Streit et al. [71] within the *Aplysia genome project* database Apl. Cal. 1.0. Undetected splice sites were determined via amplification and sanger sequencing. Now, we additionally verified this exon–intron structure by the version Apl. Cal. 2.0 of the *Aplysia genome project* database (released August 2006/ February 2009) and the UCSU Genome Browser [72]. Using the already published coding sequences (*AcH1*: BK010575 and *AcH2* BK010576), we furthermore determined the exon–intron structure of *AcH2*. This is a partial hemocyanin gene which has been identified in the genome, whereas no 3′ end (including a stop codon) was found. Also, no second hemocyanin could have been identified within the hemolymph of *Aplysia californica* [73]. Both exon–intron architectures have been fully verified with the help of genomic raw data (SRA (NCBI): SRX044044, SRX044049, SRX044050, SRX044058, SRX044081) as described for *LsH1* and *LsH2* (see above).

Compiling exon–intron structures of hemocyanin genes of two Octopus species

Sequences of previously published predicted nucleotide sequences of hemocyanins found in *Octopus vulgaris* (OvH_A: XM_029780310.1; OvH_R: XM_029796515.1) and *Octopus bimaculoides* (ObH_R: XM_014934350.1; XM_014934481.1) and corresponding scaffolds of genome assemblies (NC_043001.1; NC_043024.1; NW_014775317.1; NW_014779982.1) were used to

deduce exon–intron architectures via GeneWise [70] as described for *L. stagnalis* (see above).

Screening genomic NGS data of *Radix balthica* for Hygrophila-specific intron within FU-c

To verify whether the additional intron we found exclusively in *LsH1* is lineage-specific, we used genomic raw data of *Radix balthica* (SRA (NCBI): ERR2531849). We mapped the genomic NGS data to the additional exon as well as to the flanking exons of *LsH1* and to the corresponding sequence sections of *LsH2*. Afterwards, we analyzed whether bordering sequences represent hemocyanin-characteristic motifs which fit to cDNA sequences or if they constitute intronic parts.

Sequence analysis of introns of hemocyanin genes of *Lymnaea stagnalis*

We extracted 107 intron sequences of a soon to be released genome of *L. stagnalis* and searched sequence similarities to published nucleotide sequences using the NCBI tool 'BLAST' [74]. Additionally, we compared all those introns with each other to check them for conserved sequence motifs. Thereby, we created a database of all *LsH* intron sequences and investigated them using BLAST via Geneious 9.1.8 [68].

PCR confirmation of ambiguous hemocyanin sequence sections

For exon–intron borders which had a low assembly quality or which deviated from the Chambon's / GT-AG rule (Breathnach et al., 1978; Jacob & Gallinaro, 1989) in hemocyanin genes of *H. pomatia* and *C. aspersum*, gene-specific primers (Additional file 8, Table S3A)) were designed (CLC main workbench, Version 6.9) and respective gene regions were confirmed via Long Distance (LD) PCR for *H. pomatia*. For hemocyanin genes of *L. stagnalis* the additional intron of FU-c which appears exclusively in *LsH1* was PCR confirmed, too.

Long Fragments were PCR-amplified applying the Platinum™ SuperFi™ Green PCR Master Mix (Invitrogen by Thermo Fisher Scientific, Waltham, MA, USA) whereas shorter fragments were generated using the Advantage® 2 Polymerase Mix (Takara Bio Europe, Saint-Germain-en-Laye, France) (for PCR parameters see Additional file 8: Table S4). PCR products were visualized on a 0.8% agarose gel (Biozym, Hessisch Oldendorf, Germany). PCR products were cleaned up directly using the PCR clean-up kit (Qiagen, Hilden, Germany) or gene-specific bands were cut out and purified with the QIAquick Gel Extraction kit (Qiagen, Hilden, Germany). If possible, clean gene-specific products were sequenced directly by Microsynth (Balgach, Switzerland) using the same primers for sequencing as applied for LD PCR (Additional file 8: Table S3 (A)). Otherwise, they were cloned using the TOPO™ XL-2 Complete PCR Cloning Kit (Invitrogen by Thermo Fisher Scientific, Waltham, MA, USA) for long fragments or the TOPO® TA Cloning® Kit (Invitrogen by Thermo Fisher Scientific, Waltham, MA, USA) for shorter fragments. Three plasmids of one or two individuals were purified using the QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany) and sequenced via Sanger sequencing by Microsynth (Balgach, Switzerland) (sequencing primers in Additional file 8, Table S3 (B)).

Software used for sequence analysis

Alignments of amino acids (additional files 2, 5, 6) were compiled in MEGA 7 [75] using the muscle algorithm [76]. UCSF Chimera [77] was used to create illustrations of hemocyanins as well as the video that shows the overlay of splice site positions of hemocyanin FUs within the 3D reconstruction (additional file 7; PDB-ID: 1JS8). Sequence logos were generated by WebLogo Version 2.8 [78]. All sequences derived from LD PCR were analyzed with CLC main workbench (version 6.9).

Analyzing orthologs of *Aplysia californica* and two *Octopus* species

To enable rough comparison of intron evolution of other genes, we analyzed the number of introns in ortholog genes of *Aplysia californica* and (i) *Octopus bimaculoides* and (ii) *O. vulgaris*. We used OrthoVenn2 [79] for orthologous gene comparisons of all entries of the protein database in NCBI for the three species (as of 25.09.2020). We compared polypeptide lengths and exon counts of the results, filtered them to obtain all orthologous genes which possess entries on exon count and do not differ in polypeptide length in more than 5%. Additionally, we filtered paralogous genes if they encompass the same exon count. Comparing the differences in exon counts between orthologs of *A. californica* and the two *Octopus* species, we have only included the exon count of FU-a – FU-g of the gene coding for *AcH* to enable comparability for hemocyanins, because hemocyanin genes of octopuses do not encompass FU-h.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12862-021-01763-3.

Additional file 1: Figure S1. Hemocyanin gene structures.

Additional file 2: Figure S2. Sequence comparison of hemocyanin-like proteins of Biomphalaria glabrata with hemocyanins of gastropods and cephalopods.

Additional file 3. Table S1. Identities of cephalopod hemocyanins.

Additional file 4: Table S2. Identities of cephalopod hemocyanins.

Additional file 5: Figure S3. Splice site comparison of full hemocyanins.

Additional file 6: Figure S4. Splice site comparison of functional units.

Additional file 7. Movie files. Distribution of splice sites positions in a 3D model of a hemocyanin FU.

Additional file 8: Table S3. Primer description.

Abbreviations
FU: Functional unit; mya: Million years ago; NGS: Next Generation Sequencing; AcH: *Aplysia californica* Hemocyanin; BgHcl: *Biomphalaria glabrata* Hemocyanin-like; CaH: *Cornu aspersum* Hemocyanin; HdH: *Haliotis diversicolor* Hemocyanin; HpH: *Helix pomatia* Hemocyanin; HtH: *Haliotis tuberculata* Hemocyanin; KLH: Keyhole Limpet hemocyanin / *Megathura crenulata* hemocyanin; LsH: *Lymnaea stagnalis* Hemocyanin; NpH: *Nautilus pompilius* Hemocyanin; ObH: *Octopus bimaculoides* Hemocyanin; OdH: *Enteroctopus dofleini* Hemocyanin; OvH: *Octopus vulgaris* Hemocyanin; PcH: *Pomacea canaliculata* Hemocyanin; RbH: *Radix balthica* Hemocyanin.

Acknowledgements
Thanks to Nick Streit for preliminary analysis of hemocyanin genes of *Aplysia californica*. Furthermore, many thanks to Jennifer Aus dem Kahmen for screening genomic data of *Radix balthica* for the additional Hygrophila-specific intron and to Jonas Depoix for writing scripts for better data handling during orthologs analysis.

Authors' contributions
GGS analyzed and interpreted gene structures and was a major contributor to the conception of the analyses. VPM prepared genomic DNA for NGS, confirmed ambiguous sequence sections via PCR and has drafted parts of the manuscript (sections within Materials). DJJ provided data analyzed in this study and made substantial suggestions to the design of the bioinformatic analyses. RD and BL conceived the project and co-supervised this study as principle investigators. GSS wrote the first version of the manuscript that was commented, revised and approved by all authors. All authors read and approved the final manuscript.

Authors' information
GGS is a PhD student in BL's group. VPM is a postdoc working in the lab of RD. DJJ, RD and BL are group leaders.

Funding
Open Access funding enabled and organized by Projekt DEAL. This study was supported by a cooperative research project granted by the Austrian Science Foundation (FWF), project number I 130032-B21 (to RD) and the German Research Foundation (DFG), project number LI998/15-1 (to BL). Additional support came from the Wagner and the Feldbausch Foundation (both to BL, University of Mainz) and from the DFG to DJJ (JA 2108/6-1). The funders played no role in the study design, data collection and analysis, interpretation of data, or writing of the manuscript.

Availability of data and materials
The gene structures generated during the current study are available in NCBI under the accession numbers of their nucleotide sequences: *HpH αD*: MH485358; *HpH αN*: MH485359; HpH β: MH485360; *CaH αD*: MH485355; *CaH αN*: MH485356; *CaH β*: MH485357; *LsH1*: MH485363; *LsH2*: MH485364; *AcH1*: BK010575; *AcH2*: BK010576.
Furthermore, hemocyanin cDNA sequences and their annotations derived during this study as well as the generated genomic NGS data that cover exons and splice sites of hemocyanins of *Cornu aspersum* and *Helix pomatia* will be available in Dryad repository (https://doi.org/10.5061/dryad.b2rbnzsdj). For more information on the intronic sequences of *Lymnaea stagnalis* please contact the authors of this study as long as the data are not available online.

Ethics approval and consent to participate
No permissions were needed for animal collection because ***Helix pomatia*** and ***Cornu aspersum*** were obtained from a commercial dealer. Lymnaea stagnalis does not fall under the German Animal Protection Act §8 and is listed as "Least concern" under the International Union for Conservation of Nature (IUCN's) list of threatened species. Furthermore, we did not collect animals in a foreign country, so no further licenses were necessary. Since all experiments carried out in this study include non-cephalopod molluscs, this work is also exempt from regulations outlined by the University of Mainz Ethics Committee. We applied the 3R principles (replace, reduce, refine) in all of our animal work. To minimize pain of the snails, we first anesthetized them on ice for 20 min and subsequently sacrificed them by quickly cutting off the head.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Institute of Molecular Physiology, Johannes Gutenberg-University of Mainz, Johann-Joachim-Becher-Weg 7, 55128 Mainz, Germany. [2] Institute of Zoology and Center of Molecular Biosciences, University of Innsbruck, Technikerstraße 25, 6020 Innsbruck, Austria. [3] Department of Geobiology, Georg-August-University of Göttingen, Goldschmidtstr. 3, 37077 Göttingen, Germany.

Received: 19 April 2020 Accepted: 8 February 2021
Published online: 04 March 2021

References

1. Zapata F, Wilson NG, Howison M, Andrade SCS, Jörger KM, Schrödl M, et al. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proc Biol Sci. 2014;281:20141739. https://doi.org/10.1098/rspb.2014.1739.
2. Flanders Marine Institute. MolluscaBase. 2020–04–17. http://www.molluscabase.org/aphia.php?p=browser. Accessed 17 April 2020.
3. Brix O, Bårdgard A, Cau A, Colosimo A, Condò SG, Giardina B. Oxygen-binding properties of cephalopod blood with special reference to environmental temperatures and ecological distribution. J Exp Zool. 1989;252:34–42. https://doi.org/10.1002/jez.1402520106.
4. Brix O, Colosimo A, Giardina B. Temperature dependence of oxygen binding to cephalopod haemocyanins: Ecological implications. Marine and Freshwater Behaviour and Physiology. 1995;25:149–62. https://doi.org/10.1080/10236249409378914.
5. Mangum CP. Gas Transport in the Blood. In: Gilbert DL, Adelman WJ, Arnold JM, editors. Squid as Experimental Animals. Boston, MA: Springer US; 1990. p. 443–468. doi:https://doi.org/10.1007/978-1-4899-2489-6_2010.
6. Burnett LE, Scholnick DA, Mangum CP. Temperature Sensitivity of Molluscan and Arthropod Hemocyanins. The Biological Bulletin. 1988;174:153–62. https://doi.org/10.2307/1541782.
7. Jörger KM, Stöger I, Kano Y, Fukuda H, Knebelsberger T, Schrödl M. On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. BMC Evol Biol. 2010;10:323. https://doi.org/10.1186/1471-2148-10-323.
8. Schrödl M. Opinion: Time to say "Bye-bye Pulmonata"? SPIXIANA. 2014;37:161–4.
9. Kano Y, Brenzinger B, Nützel A, Wilson NG, Schrödl M. Ringiculid bubble snails recovered as the sister group to sea slugs (Nudipleura). Sci Rep. 2016;6:30908. https://doi.org/10.1038/srep30908.
10. Mordan P, Wade C. Heterobranchia II: The Pulmonata. In: Ponder W, editor. Phylogeny and Evolution of the Mollusca: University of California Press; 2008. p. 409–426. doi:https://doi.org/10.1525/california/9780520250925.003.0015.
11. Schäfer GG, Pedrini-Martha V, Schnegg R, Dallinger R, Jackson DJ, Lieb B. Hemocyanin genes as indicators of habitat shifts in Panpulmonata? Mol Phylogenet Evol. 2018;130:99–103. https://doi.org/10.1016/j.ympev.2018.10.014.
12. Miller KI, Cuff ME, Lang WF, Varga-Weisz P, Field KG, van Holde KE. Sequence of the Octopus dofleini hemocyanin subunit: structural and evolutionary implications. J Mol Biol. 1998;278:827–42. https://doi.org/10.1006/jmbi.1998.1648.
13. Boteva R, Severov S, Genov N, Beltramini M, Filipii B, Ricchelli F, et al. Biochemical and functional characterization of Rapana thomasiana

hemocyanin. Comp Biochem Physiol B. 1991;100:493–501. https://doi.org/10.1016/0305-0491(91)90210-5.
14. Senozan N, Landrum J, Bonaventura J, Bonaventura C. Hemocyanin of the giant keyhole limpet, Megathura crenulata. In: Lamy J, Lamy J, editors. Invertebrate oxygen-binding proteins: Structure, active site, and function : Proceedings of a workshop sponsored by the European Molecular Biology Organization, held in Tours, France, August 20–24 1979. New York: Dekker; 1981. p. 703–717.
15. Markl J. Evolution of molluscan hemocyanin structures. Biochim Biophys Acta. 2013;1834:1840–52. https://doi.org/10.1016/j.bbapap.2013.02.020.
16. Lieb B, Altenhein B, Markl J, Vincent A, van Olden E, van Holde KE, Miller KI. Structures of two molluscan hemocyanin genes: significance for gene evolution. Proc Natl Acad Sci U S A. 2001;98:4546–51. https://doi.org/10.1073/pnas.071049998.
17. Altenhein B, Markl J, Lieb B. Gene structure and hemocyanin isoform HtH2 from the mollusc Haliotis tuberculata indicate early and late intron hot spots. Gene. 2002;301:53–60. https://doi.org/10.1016/S0378-1119(02)01081-8.
18. Bergmann S, Lieb B, Ruth P, Markl J. The hemocyanin from a living fossil, the cephalopod *Nautilus pompilius*: protein structure, gene organization, and evolution. J Mol Evol. 2006;62:362–74. https://doi.org/10.1007/s00239-005-0160-x.
19. Altenhein B, Lieb B, Awenius C, Markl J. Gene Structure of Gastropod Hemocyanin. Zoology Suppl. III. Proceedings of the 93th Annual Meeting Bonn, Germany; 2000.
20. Yao T, Zhao M-M, He J, Han T, Peng W, Zhang H, et al. Gene expression and phenoloxidase activities of hemocyanin isoforms in response to pathogen infections in abalone Haliotis diversicolor. Int J Biol Macromol. 2019;129:538–51. https://doi.org/10.1016/j.ijbiomac.2019.02.013.
21. Chiumiento IR, Ituarte S, Sun J, Qiu JW, Heras H, Dreon MS. Hemocyanin of the caenogastropod Pomacea canaliculata exhibits evolutionary differences among gastropod clades. PLoS ONE. 2020;15:e0228325. https://doi.org/10.1371/journal.pone.0228325.
22. Peña JJ, Adema CM. The Planorbid Snail Biomphalaria glabrata Expresses a Hemocyanin-Like Sequence in the Albumen Gland. PLoS ONE. 2016;11:e0168665. https://doi.org/10.1371/journal.pone.0168665.
23. Lieb B, Dimitrova K, Kang H-S, Braun S, Gebauer W, Martin A, et al. Red blood with blue-blood ancestry: Intriguing structure of a snail hemoglobin. Proc Natl Acad Sci. 2006;103:12011–6. https://doi.org/10.1073/pnas.0601861103.
24. Long M, Deutsch M. Association of intron phases with conservation at splice site sequences and evolution of spliceosomal introns. Mol Biol Evol. 1999;16:1528–34. https://doi.org/10.1093/oxfordjournals.molbev.a026065.
25. Fedorov A, Suboch G, Bujakov M, Fedorova L. Analysis of nonuniformity in intron phase distribution. Nucleic Acids Res. 1992;20:2553–7. https://doi.org/10.1093/nar/20.10.2553.
26. Figuerdo EA, Gomez MV, Heneine IF, Santos IO, Hargreaves FB. Isolation and physicochemical properties of the hemoglobin of biomphalaria glabrata (Mollusca, Planorbidae). Comp Biochem Physiol B. 1973;44:481–91. https://doi.org/10.1016/0305-0491(73)90022-9.
27. Schell T, Feldmeyer B, Schmidt H, Greshake B, Tills O, Truebano M, et al. An annotated draft genome for Radix auricularia (Gastropoda, Mollusca). Genome Biol Evol. 2017. https://doi.org/10.1093/gbe/evx032.
28. Mardis ER. The challenges of big data. Dis Models Mech. 2016;9:483–5. https://doi.org/10.1242/dmm.025585.
29. Chorev M, Joseph Bekker A, Goldberger J, Carmel L. Identification of introns harboring functional sequence elements through positional conservation. Sci Rep. 2017. https://doi.org/10.1038/s41598-017-04476-0.
30. Shah DM, Hightower RC, Meagher RB. Genes encoding actin in higher plants: intron positions are highly conserved but the coding sequences are not. J Mol Appl Genet. 1983;2:111–26.
31. Rogozin IB, Wolf YI, Sorokin AV, Mirkin BG, Koonin EV. Remarkable Interkingdom Conservation of Intron Positions and Massive, Lineage-Specific Intron Loss and Gain in Eukaryotic Evolution. Curr Biol. 2003;13:1512–7. https://doi.org/10.1016/S0960-9822(03)00558-X.
32. Chorev M, Carmel L. Computational identification of functional introns: high positional conservation of introns that harbor RNA genes. Nucleic Acids Res. 2013;41:5604–13. https://doi.org/10.1093/nar/gkt244.
33. Das D, Clark TA, Schweitzer A, Yamamoto M, Marr H, Arribere J, et al. A correlation with exon expression approach to identify cis-regulatory elements for tissue-specific alternative splicing. Nucleic Acids Res. 2007;35:4845–57. https://doi.org/10.1093/nar/gkm485.
34. Narlikar L, Ovcharenko I. Identifying regulatory elements in eukaryotic genomes. Brief Funct Genomic Proteomic. 2009;8:215–30. https://doi.org/10.1093/bfgp/elp014.
35. Thummel CS, Burtis KC, Hogness DS. Spatial and temporal patterns of E74 transcription during Drosophila development. Cell. 1990;61:101–11. https://doi.org/10.1016/0092-8674(90)90218-4.
36. Cheng J, Belgrader P, Zhou X, Maquat LE. Introns are cis effectors of the nonsense-codon-mediated reduction in nuclear mRNA abundance. Mol Cell Biol. 1994;14:6317–25. https://doi.org/10.1128/mcb.14.9.6317.
37. Nagy E, Maquat LE. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem Sci. 1998;23:198–9. https://doi.org/10.1016/s0968-0004(98)01208-0.
38. Rogozin IB, Carmel L, Csuros M, Koonin EV. Origin and evolution of spliceosomal introns. Biol Direct. 2012;7:11. https://doi.org/10.1186/1745-6150-7-11.
39. Decker H, Terwilliger N. Cops and robbers: putative evolution of copper oxygen-binding proteins. J Exp Biol. 2000;203:1777–82.
40. Breathnach R, Benoist C, O'Hare K, Gannon F, Chambon P. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. Proc Natl Acad Sci USA. 1978;75:4853–7. https://doi.org/10.1073/pnas.75.10.4853.
41. Jacob M, Gallinaro H. The 5′ splice site: phylogenetic evolution and variable geometry of association with U1RNA. Nucleic Acids Res. 1989;17:2159–80. https://doi.org/10.1093/nar/17.6.2159.
42. Mount SM. A catalogue of splice junction sequences. Nucleic Acids Res. 1982;10:459–72. https://doi.org/10.1093/nar/10.2.459.
43. Burset M, Seledtsov IA, Solovyev VV. Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic Acids Res. 2000;28:4364–75. https://doi.org/10.1093/nar/28.21.4364.
44. Dibb NJ, Newman AJ. Evidence that introns arose at proto-splice sites. EMBO J. 1989;8:2015–21. https://doi.org/10.1002/j.1460-2075.1989.tb03609.x.
45. Rogers JH. How were introns inserted into nuclear genes? Trends Genet. 1989;5:213–6. https://doi.org/10.1016/0168-9525(89)90084-x.
46. Lehmann J, Eisenhardt C, Stadler PF, Krauss V. Some novel intron positions in conserved Drosophila genes are caused by intron sliding or tandem duplication. BMC Evol Biol. 2010;10:156. https://doi.org/10.1186/1471-2148-10-156.
47. Sverdlov AV, Rogozin IB, Babenko VN, Koonin EV. Reconstruction of ancestral protosplice sites. Curr Biol. 2004;14:1505–8. https://doi.org/10.1016/j.cub.2004.08.027.
48. Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV. Analysis of evolution of exon-intron structure of eukaryotic genes. Brief Bioinformatics. 2005;6:118–34. https://doi.org/10.1093/bib/6.2.118.
49. Sverdlov AV, Rogozin IB, Babenko VN, Koonin EV. Conservation versus parallel gains in intron evolution. Nucleic Acids Res. 2005;33:1741–8. https://doi.org/10.1093/nar/gki316.
50. Ruvinsky A, Ward W. Intron Framing Exonic Nucleotides: A Compromise Between Protein Coding and Splicing Constraints. TOEVOLJ. 2007;2:7–12. https://doi.org/10.2174/1874404400802010007.
51. Klopfstein S, Ronquist F. Convergent intron gains in hymenopteran elongation factor-1α. Mol Phylogenet Evol. 2013;67:266–76. https://doi.org/10.1016/j.ympev.2013.01.015.
52. de Souza SJ, Long M, Schoenbach L, Roy SW, Gilbert W. Intron positions correlate with module boundaries in ancient proteins. Proc Natl Acad Sci U S A. 1996;93:14632–6. https://doi.org/10.1073/pnas.93.25.14632.
53. Bonnet A, Grosso AR, Elkaoutari A, Coleno E, Presle A, Sridhara SC, et al. Introns protect eukaryotic genomes from transcription-associated genetic instability. Mol Cell. 2017;67(608–621):e6. https://doi.org/10.1016/j.molcel.2017.07.002.
54. Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA. 2005;11:241–7. https://doi.org/10.1261/rna.7240905.
55. Brown JWS, Marshall DF, Echeverria M. Intronic noncoding RNAs and splicing. Trends Plant Sci. 2008;13:335–42. https://doi.org/10.1016/j.tplants.2008.04.010.

56. Rearick D, Prakash A, McSweeny A, Shepard SS, Fedorova L, Fedorov A. Critical association of ncRNA with introns. Nucleic Acids Res. 2011;39:2357–66. https://doi.org/10.1093/nar/gkq1080.
57. Valencia P, Dias AP, Reed R. Splicing promotes rapid and efficient mRNA export in mammalian cells. Proc Natl Acad Sci U S A. 2008;105:3386–91. https://doi.org/10.1073/pnas.0800250105.
58. Callis J, Fromm M, Walbot V. Introns increase gene expression in cultured maize cells. Genes Dev. 1987;1:1183–200. https://doi.org/10.1101/gad.1.10.1183.
59. Eghbalnia HR, Wilfinger WW, Mackey K, Chomczynski P. Coordinated analysis of exon and intron data reveals novel differential gene expression changes. Sci Rep. 2020;10:15669. https://doi.org/10.1038/s41598-020-72482-w.
60. Parenteau J, Abou ES. Introns: Good Day Junk Is Bad Day Treasure. Trends Genet. 2019;35:923–34. https://doi.org/10.1016/j.tig.2019.09.010.
61. Jo B-S, Choi SS. Introns: The Functional Benefits of Introns in Genomes. Genomics Inform. 2015;13:112–8. https://doi.org/10.5808/GI.2015.13.4.112.
62. Gotic I, Omidi S, Fleury-Olela F, Molina N, Naef F, Schibler U. Temperature regulates splicing efficiency of the cold-inducible RNA-binding protein gene Cirbp. Genes Dev. 2016;30:2005–17. https://doi.org/10.1101/gad.287094.116.
63. James AB, Calixto CPG, Tzioutziou NA, Guo W, Zhang R, Simpson CG, et al. How does temperature affect splicing events? Isoform switching of splicing factors regulates splicing of LATE ELONGATED HYPOCOTYL (LHY). Plant Cell Environ. 2018;41:1539–50. https://doi.org/10.1111/pce.13193.
64. Airoldi CA, McKay M, Davies B. MAF2 Is Regulated by Temperature-Dependent Splicing and Represses Flowering at Low Temperatures in Parallel with FLM. PLoS ONE. 2015;10:e0126516. https://doi.org/10.1371/journal.pone.0126516.
65. Evantal N, Anduaga AM, Bartok O, Patop IL, Weiss R, Kadener S. Thermosensitive alternative splicing senses and mediates temperature adaptation in Drosophila. bioRxiv 2018. doi: https://doi.org/10.7554/eLife.44642.
66. Teich R, Grauvogel C, Petersen J. Intron distribution in Plantae: 500 million years of stasis during land plant evolution. Gene. 2007;394:96–104. https://doi.org/10.1016/j.gene.2007.02.011.
67. Thonig A, Oellermann M, Lieb B, Mark FC. A new haemocyanin in cuttlefish (Sepia officinalis) eggs: Sequence analysis and relevance during ontogeny. Evodevo. 2014;5:6. https://doi.org/10.1186/2041-9139-5-6.
68. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9. https://doi.org/10.1093/bioinformatics/bts199.
69. Davison A, McDowell GS, Holden JM, Johnson HF, Koutsovoulos GD, Liu MM, et al. Formin Is Associated with Left-Right Asymmetry in the Pond Snail and the Frog. Curr Biol. 2016;26:654–60. https://doi.org/10.1016/j.cub.2015.12.071.
70. Birney E, Clamp M, Durbin R. GeneWise and Genomewise. Genome Res. 2004;14:988–95. https://doi.org/10.1101/gr.1865504.
71. Streit K-S. Differentielle Expression und molekulare Evolution von Mollusken-Hämocyanin [Dissertation]. Mainz: Johannes Gutenberg-Universität Mainz; 2008.
72. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler AD. The Human Genome Browser at UCSC. Genome Res. 2002;12:996–1006. https://doi.org/10.1101/gr.229102.
73. Lieb B, Boisguerin V, Gebauer W, Markl J. cDNA sequence, protein structure, and evolution of the single hemocyanin from Aplysia californica, an opisthobranch gastropod. J Mol Evol. 2004;59:536–45. https://doi.org/10.1007/s00239-004-2646-3.
74. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10. https://doi.org/10.1016/S0022-2836(05)80360-2.
75. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33:1870–4. https://doi.org/10.1093/molbev/msw054.
76. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7. https://doi.org/10.1093/nar/gkh340.
77. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–12. https://doi.org/10.1002/jcc.20084.
78. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14:1188–90. https://doi.org/10.1101/gr.849004.
79. Xu L, Dong Z, Fang L, Luo Y, Wei Z, Guo H, et al. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res. 2019;47:W52–8. https://doi.org/10.1093/nar/gkz333.

Publisher's Note

2.4

The evolution of hemocyanin genes in Caenogastropoda: Gene duplications and intron accumulation in highly diverse gastropods

Gabriela Giannina Schäfer
Lukas Jörg Grebe
Robin Schinkel
Bernhard Lieb

Journal of Molecular Evolution
Under Review,
eingereicht am: 22.04.2021

Dieses Kapitel wurde nach Fertigstellung dieser Arbeit in leicht veränderter Form beim *Journal of Molecular Evolution* angenommen und unter der **DOI: 10.1007/s00239-021-10036-y** veröffentlicht.

Mein Beitrag:

Ich habe die Untersuchung der Hämocyanin-Genstrukturen der Caenogastropoda mitkonzipiert und die Probensammlung von *Nucella lapillus* als intertidal lebender Vertreter dieser Gruppe geplant und mit Kollegen durchgeführt. Außerdem isolierte ich die DNA und RNA für die Sequenzierung und führte die biochemischen Analysen durch. Ich habe gemeinsam mit Lukas Grebe und Robin Schinkel, deren Masterarbeiten ich mitbetreut habe, die NGS-Daten assembliert und die Nukleotidsequenzen analysiert. Weiterhin habe ich die Erstversion des Manuskripts verfasst und anschließend die Kommentare der Mitautoren eingearbeitet.

The evolution of hemocyanin genes in Caenogastropoda:
Gene duplications and intron accumulation in highly diverse gastropods

Gabriela Giannina Schäfer[a,b], Lukas Jörg Grebe[a], Robin Schinkel[a], Bernhard Lieb[a]

[a] Johannes Gutenberg-University of Mainz, Institute of Molecular Physiology, Johann-Joachim-Becher-Weg 7, D-55128 Mainz, Germany

[b] **corresponding author;** schaefga@uni-mainz.de; ORCID: 0000-0002-5214-9830

Abstract

Hemocyanin is the oxygen transport protein of most molluscs and represents an important physiological factor that has to be well-adapted to their environments because of strong influences of abiotic factors to its oxygen affinity. Multiple independent gene duplications and intron gains have been reported for hemocyanin genes of Tectipleura (Heterobranchia) and the caenogastropod species Pomacea canaliculata which contrast the consistency of hemocyanin genes observed in Vetigastropoda. The goal of this study was to analyze hemocyanin gene evolution within the diverse group of Caenogastropoda in more detail. Our findings reveal multiple gene duplications and intron gains within their hemocyanin genes and imply that these represent general features of Apogastropoda hemocyanins. Whereas hemocyanin exon-intron structures are identical within different Tectipleura lineages, they differ strongly within Caenogastropoda between different groups as well as between paralogous hemocyanin genes of the same species. Thus, intron accumulation took place more gradually within Caenogastropoda but finally led to a similar consequence namely a multitude of introns. Since both phenomena occurred independently within Heterobranchia and Caenogastropoda this corroborates the hypothesis that introns may contribute to adaptive radiations by offering new opportunities for genetic variability (multiple paralogs which may evolve differently) and regulation (multiple introns). Despite of many questions remaining open, our study indicates that adaptations of hemocyanins might be one among the many influencing factors which enabled the large diversity of Apogastropoda. Thus, it might be a starting point for further studies on hemocyanins and their genes which may elucidate possible correlations of adaptive radiation and hemocyanin diversity.

Keywords

Hemocyanin, adaptation, gene structure, intron accumulation, gene duplication, Caenogastropoda

Background

Mollusca is the second largest animal phylum and includes over 82,000 described species (for numbers cf. WoRMS Editorial Board 2020). The huge diversity of this phylum is represented best by the two large gastropod groups Heterobranchia and Caenogastropoda which together form the clade Apogastropoda. They comprise over 64,000 species living in various habitats including the sea, fresh waters and terrestrial ecosystems as well as all kinds of intermediate environments. The numerous habitat shifts, which were undergone multiple times independently by different groups of Apogastropoda, were enabled by a multitude of adaptations which resulted in an enormous diversification. Besides the evolution of a range of different lifestyles and morphological adaptations, modifications of their respiratory systems have been inevitable during habitat shifts. In addition to the evolution of new respiratory organs such as pneumostomes and lungs (Dayrat and Tillier 2002; Jörger et al. 2010; Kocot et al. 2013; Schrödl 2014), also molecular adaptations which influence respiration have been detected (e.g. adaptations of mitochondrial complexes of Panpulmonata to increase metabolic efficiency (Romero et al. 2016b) or the evolution of multiple metabolic states demanding different levels of oxygen (Schweizer et al. 2019)).

Another very important factor of gastropod respiration which has to be adapted to environmental conditions is the oxygen transporter hemocyanin. Previous studies showed that the oxygen affinity, which strongly influences the function of hemocyanin, is temperature dependent (Brix et al. 1989; Brix et al. 1995; Mangum 1990; Miller 1985; Miller and van Holde 1974). Thus, shifts to habitats with different temperatures must be accompanied by adaptations of these proteins to further sustain a sufficient oxygen supply. Especially environments with varying temperatures, e.g. land and intertidal

zones in contrast to solely marine habitats, require well-adapted oxygen-transport proteins. Different hemocyanin isoforms can have different oxygen affinities (Swerdlow et al. 1996) and differential expression helps to adapt to varying oxygen conditions (e.g. low oxygen pressure in eggs of *Sepia officinalis*; Gutowska and Melzner 2009; Strobel et al. 2012). We previously reported a multitude of hemocyanin gene duplications in different species of Tectipleura (Schäfer et al. 2018). This large group of Heterobranchia comprises very diverse snails that conquered land and freshwater several times independently in different lineages (Dinapoli and Klussmann-Kolb 2010; Jörger et al. 2010; Kano et al. 2016; Kocot et al. 2013; Romero et al. 2016a). Therefore, we hypothesized that hemocyanin duplications may have helped to increase genetic variability by leading to a multitude of hemocyanin isoforms with potentially different properties and/or varying expression patterns. Accordingly, they might represent one out of many factors which might have enabled the exploitation of new habitats and the extremely large adaptive radiation (Schäfer et al. 2018).

The overall shape of molluscan hemocyanins is a partly-hollow cylinder of 4 MDa formed by decamers of 35 nm in diameter which can assemble to di- or multidecamers (fig. 1A). These large oxygen transport molecules are freely floating within the hemolymph of most molluscs (van Holde and Miller 1995). The basic structure of one single 400 kDa polypeptide subunit encompasses eight paralogous domains called functional units a, b, …, h (short FU-a to FU-h) which are connected by short linker regions (fig. 1B). The FUs have similar tertiary structures forming 45 to 50 kDa large globular substructures of the polypeptides and comprising one oxygen binding site each. Thus, one didecamer, which is the most common hemocyanin molecule in gastropods, encompasses 160 oxygen binding sites (basic structure reviewed in Markl 2013 and Kato et al. 2018). The basic composition of hemocyanin subunits including multiple FU domains as well as the primary structure of these FUs are highly conserved over all different molluscan classes that have been analyzed so far (overview in Markl 2013).

The segmentation of molluscan hemocyanin subunits in multiple FU domains is also represented by the highly conserved basic exon-intron structure of their genes (Lieb et al. 2001). Gene segments that are coding for different functional units are separated from each other by phase 1 introns (fig. 1B). Within all FUs of molluscan hemocyanins that have been analyzed so far, they lie at an almost equivalent positions just upstream from linker peptide coding gene regions (Altenhein et al. 2002; Bergmann et al. 2006; Lieb et al. 2001; Schäfer et al. 2021b). Accordingly, these introns are termed *linker introns*, while those lying within FU-coding regions are termed *internal introns*. These internal introns are less conserved in their positions and differ between hemocyanins of different molluscan lineages (fig. 1B). Previous studies showed that the number of these internal introns vary greatly between Octopodoidea (5 internal introns) or Lepetellida (Vetigastropoda; 8 internal introns) and Tectipleura (Heterobranchia; 46 internal introns) but are conserved within these different groups of molluscs (Altenhein et al. 2000; Altenhein et al. 2002; Lieb et al. 2001; Yao et al. 2019).

In spite of the strong conservation of the cylindrical hemocyanin molecule structure, of their subunits and of their genes, deviations from these basic structures have been described for several molluscan groups. They mostly concern the number of functional units which changed due to domain duplications or losses, e.g. hemocyanins of Cephalopoda lost FU-h (van Holde and Miller 1995). While the basic structure of gastropod hemocyanins corresponds to the typical eight functional unit domains, multiple variations have been found for hemocyanins of Caenogastropoda: Within the extremely large and diverse Cerithioidea (Cerithiida, Caenogastropoda), the so-called mega-hemocyanin has been identified (Lieb et al. 2010). It represents a hemocyanin tridecamer which includes two typical decamers built from 400 kDa subunits and additionally one larger decamer which is located between the two typical decamers. This decamer is composed of subunits with a molecular mass of 550 kDa. These 550 kDa mega-hemocyanin subunits are lacking FU-g and FU-h but encompass six additional functional units which are paralogous to FU-f (FU-f_1, FU-f_2 …) (Gatsogiannis et al. 2015). These additional FUs reach within the center of the molecule and fill the mega-hemocyanin cylinders. Thereby, they increase the oxygen transport capacity from 80 to 120 oxygen molecules within the hemocyanin decamer due to further binding sites, resulting in 280 oxygen binding sites within a mega-hemocyanin molecule. The viscosity and the colloid-osmotic pressure of the hemolymph, however, remains the same as in a typical hemocyanin tridecamer, so that the oxygen transport efficiency is increased (Gatsogiannis et al. 2015). The ability to differentially express the 400 kDa and 550 kDa hemocyanin subunits most probably facilitates variable ratios of typical hemocyanins and mega-hemocyanins. This may further help to adapt to different living conditions and may have accelerated the adaptive radiation of the extremely diverse group of Cerithioidea (Lieb et al. 2010).

Recently, we reported a further deviation of one out of two hemocyanin subunits of *Rapana venosa* (RtH2 derived from the synonym *R. thomasiana*) and *Nucella lapillus* (NlH2) (Schäfer et al. 2021a). Both species belong to Muricidae which represents another main group of Caenogastropoda (fig. 2). In contrast to the mega-hemocyanin of Cerithioidea, these subunits comprise FU-a to FU-h which is typical for gastropod hemocyanins. They rather vary from the basic structure of hemocyanins within one particular functional unit (specifically FU-g) as it has not been observed before in any other group of molluscs (Schäfer et al. 2021a). Typically, the primary structure of hemocyanin functional units is highly conserved in hemocyanins over all molluscan classes.

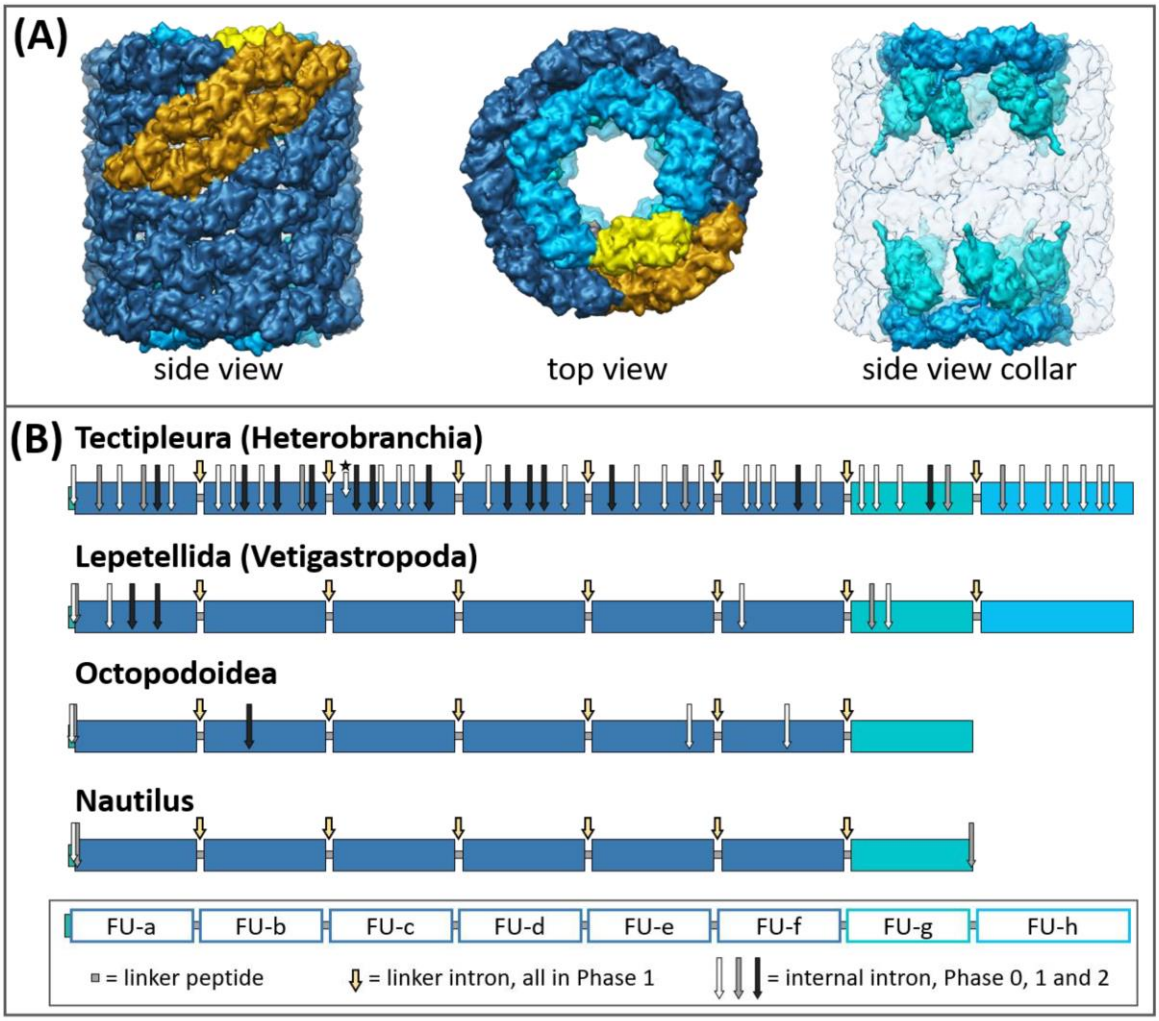


Fig. 1: Molluscan hemocyanin: molecules and genes. (A) Typical gastropod hemocyanin didecamer adapted from Schäfer et al. 2021a and based on the 9 Å model of KLH_1 (Gatsogiannis and Markl 2009, PDB: 4BED). The wall (FU-a – FU-f) is colored in dark blue. The collar is restricted to both sides of the didecamer and built by 10 FU-g (cyan) and 10 FU-h (light blue). Side and top view are depicted with one hemocyanin subunit dimer highlighted in gold (wall) and grey/yellow (collar: FU-g/h). (B) Exon-intron structure of molluscan hemocyanins. Shown are coding sequences of molluscan hemocyanins. Their genes typically contain ~10,200 – 10,300 nucleotides coding for the eight functional units (FU-a, FU-b, … FU-h) of one hemocyanin polypeptide subunit. FU-h is about 300 nucleotides longer than the other functional units. Hemocyanin genes of Cephalopoda do not comprise FU-h. Big boxes represent functional units (FU-a, FU-b, …), small boxes between them linker peptides. Arrows symbolize intron positions in hemocyanin genes with respect to the coding sequences. Linker introns are conserved in phase 1 within all known hemocyanin genes (yellow arrows). Internal introns occur in all intron phases and are color-coded accordingly: located before the first (phase 0, white), after the first (phase 1, grey) or after the second (phase 2, black) nucleotide of a codon. Hemocyanin genes of Tectipleura comprise a significantly larger number of internal introns than those of Lepetellida, Octopodoidea or Nautilus.

However, we have identified 118 highly hydrophilic amino acids within the N-terminal region of FU-g of one isoform expressed in *R. venosa* (RtH2) and 340 hydrophilic amino acids inserted within the same region of the orthologous gene of *Nucella lapillus* (NlH2) which are additional to the highly conserved amino acids within a typical molluscan hemocyanin. These additional amino acids seem to build an extra mass within the hemocyanin didecamer of up to 800 kDa and might influence the function of this hemocyanin molecule within these species (Schäfer et al. 2021a).

Chiumiento et al. (2020) analyzed hemocyanins of *Pomacea canaliculata*, a species that belongs to Ampullariida which represents a third main group of Caenogastropoda (fig. 2). They identified four hemocyanin subunits that correspond to the basic polypeptide structure of this oxygen transporter. Analyzing the gene structures of these hemocyanins, they observed variations from all hemocyanin exon-intron architectures that have been described, yet. Hemocyanin genes of *P. canaliculata* encompass a remarkable larger number of introns than hemocyanins of Vetigastropoda or Cephalopoda (27 -32 in *P. canaliculata*, only 9 – 15 in Vetigastropoda or Cephalopoda; Altenhein et al. 2002; Bergmann et al. 2006; Lieb et al. 2001; Yao et al. 2019). Since hemocyanin genes of Tectipleura include 53 introns each (Schäfer et al. 2021b), large numbers of introns in hemocyanin genes may be a phenomenon of Apogastropoda in general. In contrast to paralogous hemocyanins of other molluscs, however, the exon-intron architectures of the four hemocyanin genes of *P. canaliculata* differ from each other according to Chiumiento et al. (2020). The exon-intron structures of hemocyanins that have been analyzed for Vetigastropoda (Lepetellida), Heterobranchia (Tectipleura) and Octopodoidea, on the other hand, are highly conserved over different hemocyanin paralogs (Altenhein et al. 2002; Schäfer et al. 2021b).

Regarding these exceptional phenomena of caenogastropod hemocyanins, we started to analyze the evolution of the genes of this oxygen transporter within Caenogastropoda in more detail. Therefore, we inferred the evolutionary background of caenogastropod hemocyanin genes by reconstructing phylogenetic trees and compared their exon-intron structures. This enabled us to unravel a more detailed scenario of intron evolution in gastropod hemocyanins. Thereby, our analysis include the previously described hemocyanins of the Cerithioidea *Melanoides tuberculata* (MtH_{400} and MtH_{550} (Lieb et al. 2010), the Muricidae species *Rapana venosa* and *Nucella lapillus* (RtH and NlH; Gebauer et al. 1999; Schäfer et al. 2021b) and three hemocyanins of *Pomacaea canaliculata* (Ampullariida; Chiumiento et al. 2020)). Additionally, we investigated hemocyanins of *Littorina saxatilis*, a species that belongs to the large group of Hypsogastropoda as the Muricidae *Rapana venosa* and *Nucella lapillus* but, in contrast to them, does not belong to the siphonate clade (fig. 2; Ponder et al. 2019).

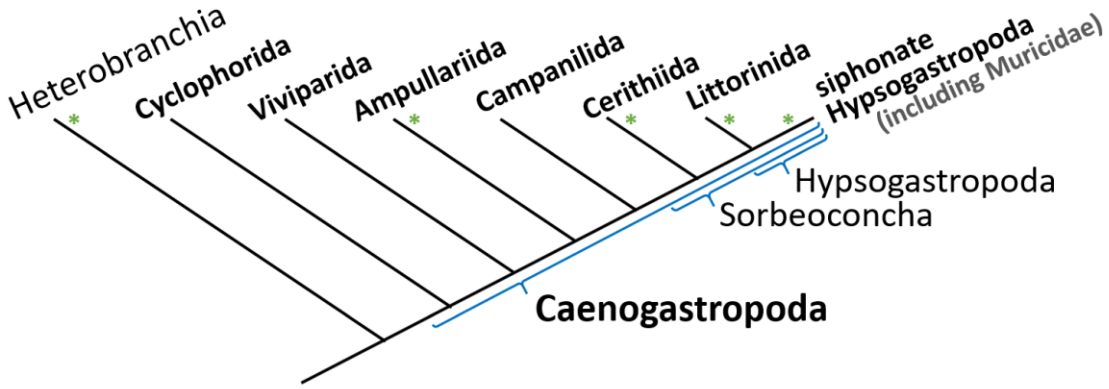


Fig. 2: Broad systematics of Caenogastropoda according to Ponder et al. (2019).

Despite a large number of phylogenetic studies on Caenogastropoda, many phylogenetic relationships within that large group of gastropods remained unresolved until today. The depicted systematics are combined from different studies by Ponder et al. (2019). The large groups included in this tree have often been recognized in systematic analysis but even though are debated controversially because their exact relationships and the classification of some lineages within these groups differ within various studies. Groups of which hemocyanin genes are included in this study are indicated by asterisks.

Methods

Animal sampling and DNA isolation

One individual of *Melanoides tuberculata* was taken from a freshwater aquarium in the Institute of Molecular Physiology in Mainz. Three specimens of *Nucella lapillus* were collected at the western Atlantic coast of Brittany, France (Schäfer et al. 2021a). DNA of one individual of both species was isolated from foot tissue using E.Z.N.A.® Mollusc DNA Kit (Omega Bio-Tek, Norcross, GA, USA). Via Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA), DNA was purity checked and quantified. Subsequently, it was sent to StarSeq in Mainz, Germany, for next-generation sequencing (NGS, Illumina Next Seq500) and library preparation.

In silico assembly of hemocyanin cDNAs of Littorina saxatilis and Pomacea canaliculata

Hemocyanin cDNA sequence assemblies were performed with Geneious 9.1.8 (Kearse et al. 2012) using publicly available transcriptomic data of *L. saxatilis* (SRR9651721, SRR9651722, SRR9651724) and *P. canaliculata* (SRR6429145, SRR6429146, SRR6429153). Paired-end reads were set, sequencing adapters removed and transcriptomic raw reads quality trimmed. Processed reads of *L. saxatilis* were mapped to the previously published cDNA sequence of the 400 kDa hemocyanin of *M. tuberculata* (KC405575, overlap identity: 70%). Those of *P. canaliculata* were mapped to the cDNA sequences which we deduced from the published hemocyanin gene structures (Chiumiento et al. 2020). Reads that mapped against the references were isolated and used as references for iterative mappings of the remaining reads to elongate cDNA fragments (minimum overlap: 60 nucleotides; minimum overlap identity: 99%; maximum mismatches: 1%). This mapping was reiterated until the isolated fragments ended up in complete coding sequences.
The existence of multiple hemocyanin genes per species and the repetitive structure of their cDNAs with sequences coding for functional units which share some highly conserved amino acid motifs, can result in wrong assemblies. To preclude such misassemblies, we verified the sequences by (i) low sensitive mappings which enable misassembly detection, (ii) analyzing the sequences for highly identical sections between hemocyanin sequences of a species to enable manually double checking them for right assemblies and (iii) using paired-end reads. For a more detailed description of sequence assembly and verification see Schäfer et al. 2021a.

Reconstructing exon-intron structures of hemocyanin genes

For the reconstruction of gene architectures, we used Geneious 9.1.8 as bioinformatic tool (Kearse et al. 2012). Genetic NGS data were processed as described for transcriptomic raw reads. Trimmed and paired-end reads were then mapped to coding sequences of *Rapana venosa* (BK014286, BK014287), *Nucella lapillus* (MT939254, MT939255), *Melanoides tuberculata* (KC405575, KC405576), *Littorina saxatilis* and *Pomacea canaliculata* (obtained in this study). Mapping results showed that some parts of the cDNA sequences were uncovered by genomic NGS data or showed inconsistencies to those. These sections were used to subdivide cDNA sequences in different sections representing hypothetical exons. These sections were extended by repetitive mappings of genomic NGS data until their 3' and 5' ends deviated in at least ten base pairs from cDNA sequences (procedure explained more detailed in Schäfer et al. 2021b.

Sequence alignment and phylogenetic tree

We used MEGA 7 (Kumar et al. 2016) to align amino acid sequences by applying the implemented Muscle algorithm and to determine LG+G+I+F as the best evolutionary model. It was used to compute the Maximum Likelihood tree including branch supports based on 100 bootstrap replicates using MEGA version 7 (Kumar et al. 2016). Hemocyanin sequences of the Cephalopoda *Nautilus pompilius* and *Enteroctopus dofleini* were used as outgroups. We used MrBayes 3.2.6 (Huelsenbeck and Ronquist 2001) which is implemented in Geneious 9.1.8 (Kearse et al. 2012) to conduct Bayesian inference based on two parallel runs of four Monte Carlo Markov Chains (MCMC) with one million generations, a subsampling frequency of 500 and a burn-in of 250,000.

Results

In order to reveal relationships of different hemocyanin paralogs and to identify hemocyanin gene duplication events within Caenogastropoda, we inferred the phylogeny of these proteins including eleven hemocyanin sequences of over 10,000 nucleotides each from species which belong to four different main groups of Caenogastropoda. To further investigate the evolution of hemocyanin genes of Caenogastropoda, we compiled exon-intron structures for all of these eleven caenogastropod hemocyanin genes. These analyses include the already published hemocyanin cDNA sequences of the Muricidae (siphonate Hypsogastropoda) *Nucella*

lapillus (MT939254, MT939255) and *Rapana venosa* (BK014286, BK014287) and of the Cerithioidea species *Melanoides tuberculata* (KC405575, KC405576). Additionally, we revised the hemocyanin sequences of the Ampullariida species *Pomacea canaliculata* (BK014379, BK014378, BK014377) and newly assembled the cDNA of the hemocyanins of *Littorina saxatilis* (BK014376, BK014375) which belongs to the asiphonate Hypsogastropoda (fig. 2).

Canonical hemocyanin coding sequences identified for *Littorina saxatilis* and *Pomacea canaliculata*

We obtained two complete hemocyanin coding sequences for *L. saxatilis* by assembling transcriptomic NGS data (LisaH1, LisaH2). For *P. canaliculata* four hemocyanin cDNAs have already been published by Chiumiento et al. (2020). As previously reported, these sequences contained some inconsistencies (cf. Schäfer et al. 2021a). By assembling transcriptomic NGS data, we identified and corrected three of those hemocyanin cDNA sequences (PcH I, IIb, III). This facilitated phylogenetic comparability and enabled us to include them in our sequence analysis as well as to elucidate their correct gene architectures (see below).
The obtained hemocyanin sequences of *L. saxatilis* and *P. canaliculata* include the eight canonical FUs (a, b, …, h) and comprise about 10,250 - 10,300 nucleotides as typical for gastropod hemocyanins (cf. Lieb et al. 2000; Lieb et al. 2004; Schäfer et al. 2018). Accession numbers, lengths and molecular weight for all hemocyanin cDNA sequences and primary structures of *L. saxatilis* and *P. canaliculata* are shown in tab. 1.

Phylogenetic analyses reveal multiple independent gene duplications in different caenogastropods

Molecular phylogenetic trees based on amino acid alignments and inferred by maximum likelihood and Bayesian analyses are well-supported (fig. 3, supplement 1). They are largely congruent with each other, sharing the same nodes except for minor differences concerning one hemocyanin of *Littorina saxatilis* (LisaH2) which is grouped together with hemocyanins of Muricidae in the Bayesian analysis. The obtained phylogenies of caenogastropod hemocyanins are in accordance with the currently accepted systematic relationships of the four groups Muricidae, Littorinida, Cerithiida and Ampullariida (fig. 2; Bouchet et al. 2017; Ponder et al. 2019). The phylogenies obtained by maximum likelihood and MrBayes both reveal that the multiple hemocyanin genes which we have identified for the different analyzed species resulted from independent gene duplications that occurred after the diversification of Caenogastropoda into Ampullariida, Cerithiida and Hypsogastropoda (orange arrows in fig. 3). Although the position of LisaH2 is uncertain within the hemocyanins of Hypsogastropoda, the results of maximum likelihood and Bayesian inferences both support multiple independent duplications. The gene duplication that led to the two hemocyanin isoforms of *R. venosa* and *N. lapillus* most probably took place in a common ancestor of both Muricidae species but after separation of siphonate and asiphonate Hypsogastropoda (fig. 3).

Strong variability of hemocyanin gene architectures in Caenogastropoda

We achieved the reconstruction of eleven hemocyanin gene structures of five Caenogastropoda (including corrected gene structures of *P. canaliculata*) and identified that nine of them vary from each other concerning the number of internal introns and their positions with respect to the coding sequences (fig. 4). Thereby, we obtained differences of exon-intron architectures between hemocyanin genes of different caenogastropod lineages (variations between Muricidae, Littorinida, Cerithioidea and Ampullariida) and between different hemocyanin genes within the same species (e.g. between RtH1 and RtH2 of *R. venosa* or LisaH1 and LisaH2 of *L. saxatilis*). Identical exon-intron structures are only present for hemocyanin gene 2 of the two Muricidae species (RtH2 and NlH2) as well as for the hemocyanin genes PcH IIb and PcH III of *Pomacea canaliculata*.

Tab. 1: Hemocyanins of *Littorina saxatilis* (LisaH) and *Pomacea canaliculata* (PcH).
Shown are accession numbers and lengths of coding sequences (CDS) in nucleotides (nt); number of amino acids (aa) for the deduced primary structure of the polypeptides and the calculated molecular weight in kDa.

	LisaH1	LisaH2	PcH I	PcH IIb	PcH III
accession number	BK014376	BK014375	BK014379	BK014378	BK014377
CDS (nt)	10,308	10,278	10,278	10,287	10,272
primary structure (aa)	3,436	3,426	3,426	3,429	3,424
deduced molecular mass (kDa)	392	391	391	391	390

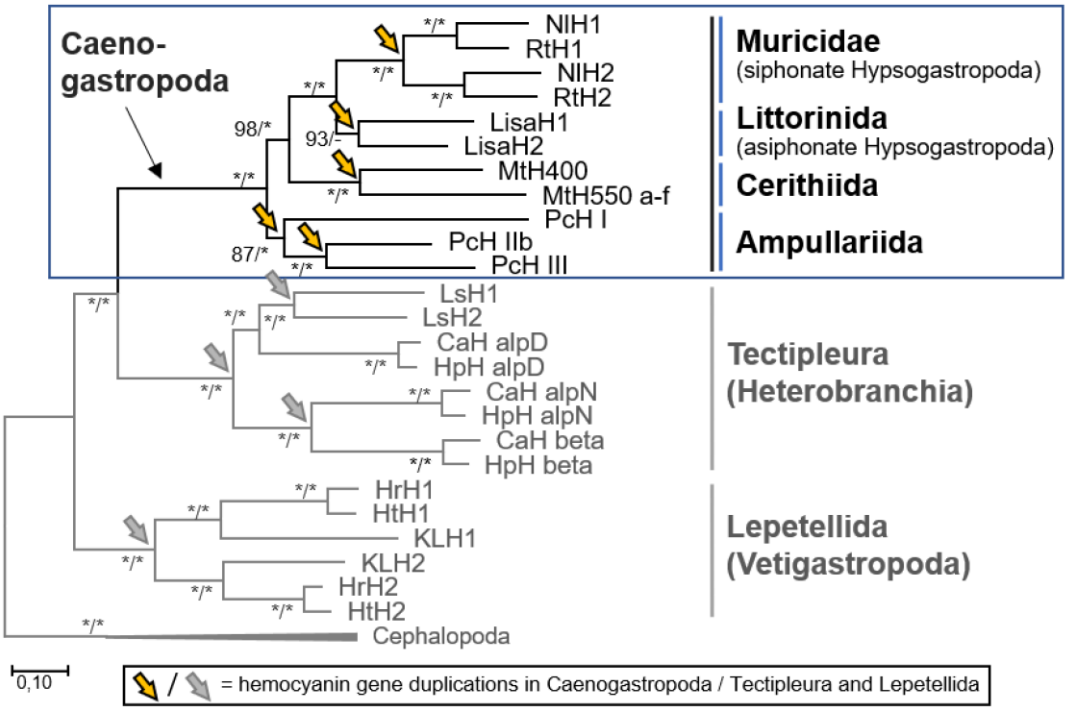


Fig. 3: Maximum likelihood (ML) tree of gastropod hemocyanins.
The phylogenetic tree is based on an amino acid sequence alignment and conducted with MUSCLE (Edgar 2004) implemented in MEGA7 (Kumar et al. 2016) using the LG+G+I+F model. It shows that independent hemocyanin gene duplications (symbolized by orange arrows) occurred within all in this study included main groups of Caenogastropoda, namely within Ampullariida, Ceritiida and within Hypsogastropoda. Thereby the tree includes hemocyanins of the following Caenogastropoda: *Pomacea canaliculata* (PcH I+IIb+III), *Melanoides tuberculata* ($MtH_{400+550}$), *Littorina saxatilis* (LisaH1+2), *Rapana venosa* (RtH1+2) and *Nucella lapillus* (NlH1+2). It further encompasses hemocyanins of Tectipleura (*Helix pomatia* HpHαD+αN+β; *Cornu aspersum* CaHαD+αN+β; *Lymnaea stagnalis* LsH1+2) and Lepetellida (*Haliotis tuberculata* HtH1+2; *H. rubra* HrH1+2; *Megathura crenulata* KLH1+2). Gene duplications within hemocyanins of Tectipleura and Lepetellida are represented by grey arrows. The tree was rooted with hemocyanins of the Cephalopoda *Nautilus pompilius* (NpH) and *Enteroctopus dofleini* (OdH_A and OdH_G). Nodes are congruent with those obtained by Bayesian inference except for the position of LisaH2 which is grouped together with hemocyanins of Muricidae in the Bayesian analysis tree (supplement 1). Hemocyanin gene duplications of Caenogastropoda are not affected by the described deviations of maximum likelihood or Bayesian inference. Nodes are labelled with bootstrap (BS) percentages based on 100 replicates from ML analyses and posterior probabilities (PP) computed by MrBayes (BS/PP). Asterisks indicate nodes supported by BS ≥ 99% / PP ≥ 0.99

Total numbers of internal introns of caenogastropod hemocyanins vary between 21 and 57 (tab. 2). Average numbers of internal introns per functional unit domain vary between 2.4 introns/FU (PcH IIb/III) and 6.5 introns/FU (MtH_{400}). The comparison with known gastropod hemocyanin genes shows that all analyzed hemocyanins of Caenogastropoda contain a larger number of internal introns than those of Vetigastropoda (Altenhein et al. 2002; Lieb et al. 2001). Hemocyanins of Caenogastropoda contain less internal introns per functional unit than those of Tectipleura (cf. HpH in tab. 2 and fig. 4; Schäfer et al. 2021b), with exception of MtH_{400} of *Melanoides tuberculata*. Hemocyanin genes of this cerithioid encompass the largest number of internal introns that have been identified so far. Linker introns are highly conserved across all molluscan hemocyanins that have been analyzed so far (Bergmann et al. 2006; Lieb et al. 2001; Schäfer et al. 2021b) and are also present in hemocyanins of Caenogastropoda (fig. 4). However, we did not obtain any intron within the linker peptide coding region between FU-f_1 and FU-f_2 of the mega-hemocyanin gene of *Melanoides tuberculata* (star in supplement 3). Thus, the MtH_{550} gene is not only the first mega-hemocyanin whose gene structure has been analyzed but also the first hemocyanin gene that has been detected to lack a linker intron between two FU-coding regions. All other analyzed hemocyanin genes of Caenogastropoda include typical phase 1 linker introns between all FU-coding regions (fig. 3, supplement 2+3), as characteristic for molluscan hemocyanins (Lieb et al. 2001)

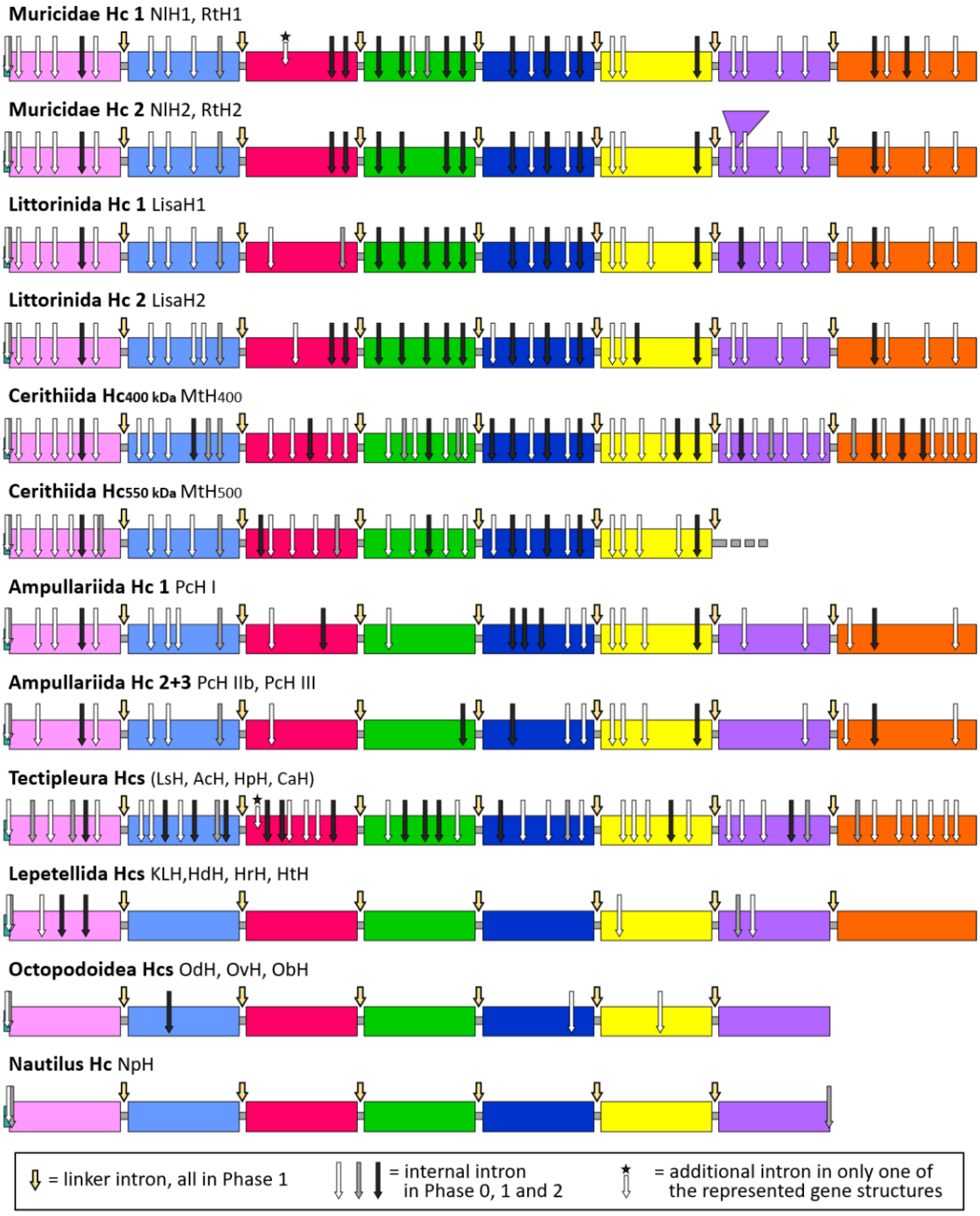


Fig. 4: Hemocyanin gene structures of Caenogastropoda, Tectipleura, Vetigastropoda and Cephalopoda. The comparison of exon-intron structures indicates that nine out of eleven analyzed hemocyanin genes of Caenogastropoda vary in their gene structures concerning number and positions of internal introns. Shown are coding sequences of hemocyanins (big boxes: coding for functional units; small and grey boxes: coding for linker peptides) and intron positions within their genes represented by arrows (cf. alignment in supplement 2). To enable comparability of gene structures, we did not enlarge the box representing the FU-g coding sequence of Muricidae Hc2 despite of 118/340 additional amino acids which have been identified for RtH2 and NlH2 (cf. supplement 2, Schäfer et al. 2021a). Instead, we included a violet triangular box on top of the box representing the additional amino acids. Introns are divided in linker introns which are highly conserved within all molluscan hemocyanins (bold yellow arrows) and internal introns which differ between various genes (thin arrows in phase 0: white; phase 1: grey or phase 2: black). Smaller arrows with a star on top represent internal introns which were obtained in only one hemocyanin out of a gene structure type (one intron within NlH1, Muricidae; and one within LsH1, Tectipleura). The comparison includes all in this study analyzed hemocyanins of

the following Caenogastropoda species: *Nucella lapillus* (NlH1+2), *Rapana venosa* (RtH1+2), *Littorina saxatilis* (LisaH1+2), *Melanoides tuberculata* (MtH_{400}; MtH_{550} FU-a – FU-f; exon-intron structures of the additional functional units f_1, f_2, …, f_6 of the mega-hemocyanin of *M. tuberculata* are shown in supplement 3) and *Pomacea canaliculata* (PcH I+IIb+III). Additionally, the conserved gene structures are included for hemocyanins of Tectipleura (*Lymnaea stagnalis* LsH, *Aplysia californica* AcH, *Helix pomatia* HpH and *Cornu aspersum*), Lepetellida (*Megathura crenulata* KLH, *Haliotis diversicolor* HdH, *H. rubra* HrH, *H. tuberculata* HtH), Octopodoidea (*Enteroctopus dofleini* OdH, *Octopus vulgaris* OvH and *O. bimaculoides*) and *Nautilus pompilius*.

Tab. 2: FU-internal introns of caenogastropod hemocyanins.

Internal introns vary between hemocyanin genes of different Caenogastropoda groups as well as between different genes within the same caenogastropod species. This contrasts the highly conserved exon-intron structures of Tectipleura (Heterobranchia) and Lepetellida (Vetigastropoda) represented in this table by the hemocyanins of *Helix pomatia* (HpH) and of *Haliotis tuberculata* (HtH). The table shows the numbers of introns lying within the functional units of hemocyanins of *Rapana venosa* (RtH), *Nucella lapillus* (NlH), *Littorina saxatilis* (LisaH), *Melanoides tuberculata* (MtH) and *Pomacea canaliculata* (PcH). Numbers of internal introns are depicted for the specific FUs. Additionally, the table includes the sum of internal introns of FU-a to FU-f (these FUs are included in all represented hemocyanins), as well as the total number of internal introns and the respective average number of internal introns per functional unit of each hemocyanin.

Hc / FU	RtH1/ NlH1	RtH2/ NlH2	LisaH1	LisaH2	MtH_{400}	MtH_{550}	PcH I	PcH IIb	PcH III	HpH	HtH
sign.	2	2	2	2	2	2	2	2	2	1	2
-a	5	5	5	5	6	7	4	3	3	5	3
-b	4	4	4	5	6	4	4	3	3	6	0
-c	2 (3)	2	2	3	5	5	2	1	1	6	0
-d	6	4	5	5	7	5	1	1	1	5	0
-e	5	5	5	6	6	6	5	3	3	5	0
-f	3	3	4	4	6	5	4	4	4	5	1
-g	4	4	4	4	7	-	2	1	1	5	2
-h	5	4	5	4	9	-	3	3	3	7	0
$-f_1$	-	-	-	-	-	2	-	-	-	-	-
$-f_2$	-	-	-	-	-	6	-	-	-	-	-
$-f_3$	-	-	-	-	-	4	-	-	-	-	-
$-f_4$	-	-	-	-	-	5	-	-	-	-	-
$-f_5$	-	-	-	-	-	3	-	-	-	-	-
$-f_6$	-	-	-	-	-	3	-	-	-	-	-
Σ(total, incl. sign.)	36 (37)	33	36	38	54	57	27	21	21	45	8
Σ(a-f)	25 (26)	23	25	28	36	32	20	15	15	32	4
Σ(FUs)	34 (35)	31	34	36	52	55	25	19	19	44	6
Ø Intron /FU	~ 4.3	~ 3.9	~ 4.3	~ 4.5	~ 6.5	~ 4.6	~ 3.1	~ 2.4	~ 2.4	~ 5.5	~ 0.8

Discussion

As most molluscs, Caenogastropoda use hemocyanin as oxygen transporter. This respiratory protein has already been identified within the Muricidae species *Rapana venosa* (Gebauer et al. 1999) and *Nucella lapillus* (Schäfer et al. 2021a), within the cerithioid *Melanoides tuberculata* (Lieb et al. 2010) and within the apple snail *Pomacea canaliculata* (Chiumiento et al. 2020). All of these Caenogastropoda species possess at least two hemocyanin genes like other Gastropoda (Gebauer et al. 1999; Markl et al. 1991; Schäfer et al. 2018). This may enable differential expression of several hemocyanin genes as shown for Cephalopoda (Oellermann et al. 2015a; Oellermann et al. 2015b; Thonig et al. 2014). Similar regulatory mechanisms may help Caenogastropoda to adapt to different living conditions by sustaining oxygen

supply despite changing oxygen partial pressure and temperature as hypothesized for Tectipleura (further discussed below under "Evolutionary constraints on hemocyanin genes in Apogastropoda?" and in Schäfer et al. 2018). Analyzing transcriptomic NGS data, we corrected inconsistencies within published hemocyanin cDNA sequences of *P. canaliculata* and obtained two hemocyanins of *Littorina saxatilis* which belong to asiphonate Hypsogastropoda and thus represent another lineage of the large group of Caenogastropoda (fig. 2). Finally, we conducted molecular phylogenies based on maximum likelihood (fig. 3; Felsenstein 1981; Kumar et al. 2016) and MrBayes (supplement 1; Huelsenbeck and Ronquist 2001; Mau and Newton 1997) and analyzed hemocyanin gene structures of the following species: *R. venosa, N. lapillus, L. saxatilis, M. tuberculata* and *P. canaliculata* (fig. 4). Our analyses revealed ongoing hemocyanin gene evolution within all main groups of Caenogastropoda which have been analyzed within this study: siphonate and asiphonate Hypsogastropoda (Muricidae and Littorinida), Cerithiida and Ampullariida (fig. 2).

Multiple independent hemocyanin gene duplications: a phenomenon of Apogastropoda

Both maximum likelihood analysis and Bayesian phylogenetic inferences (fig. 3, supplement 1) revealed that multiple hemocyanin isoforms identified for *P. canaliculata*, *M. tuberculata*, *L. saxatilis* and the two Muricidae species *R. venosa* and *N. lapillus* resulted from independent gene duplications. This corresponds to multiple gene duplications in Tectipleura (Heterobranchia) which took place independently from each other in different groups (e.g. Stylommatophora, Hygrophila, Anaspidea; Schäfer et al. 2018) and thus might be a general phenomenon for Apogastropoda. These multiple independent gene duplications contrast the much stronger conserved hemocyanin genes 1 and 2 of Lepetellida analyzed within *Haliotis tuberculata*, *Haliotis diversicolor* and *Megathura crenulata* (Lieb and Markl 2004; Yao et al. 2019). These paralogous hemocyanin genes resulted from one gene duplication which took place before the split of Lepetellida into Haliotoidea and Fissurelloidea ~ 343 million years ago (Lieb and Markl 2004) and maintained until today in both lineages. Thus, the comparison of hemocyanin genes of Apogastropoda and Vetigastropoda indicates strongly different evolutionary dynamics of hemocyanin gene duplications during the evolution of these gastropod lineages.

Strong variability of hemocyanin gene structures of Caenogastropoda suggests higher continuous intron gain rate than identified in any other mollusc

In addition to multiple independent gene duplications, our results on exon-intron architectures also reveal a high rate of evolutionary changes of hemocyanin genes in Caenogastropoda. We have previously reported that hemocyanin genes of Tectipleura encompass a significantly larger number of FU-internal introns than other analyzed hemocyanin genes (on average 5.6 introns per FU; Schäfer et al. 2021b). In hemocyanin genes of Caenogastropoda we have also identified a larger number of internal introns (2.4 – 6.5 introns per FU) compared to those of Vetigastropoda with only 0.8 introns per FU on average (fig. 4, tab. 2). Thereby, the distinct hemocyanin genes of Caenogastropoda differ strongly from each other in terms of number and position of internal introns. This corresponds to hemocyanins of different caenogastropod lineages (Muricidae, Littorinida, Cerithioidea and Ampullariida) as well as to paralogous genes within the same species (fig. 4, tab. 2). Largely varying gene architectures between different hemocyanin genes within one molluscan species have not been reported yet. So far, only for Hygrophila, a group of Tectipleura, two paralogous hemocyanin genes have been identified which vary in one out of 45/46 introns. Precisely, hemocyanin gene 1 of *Lymnaea stagnalis*, *Radix balthica* and *Biomphalaria glabrata* (LsH1, RbH1, BgHcl-2, all orthologous to each other) encompasses one intron additionally to those conserved within all other analyzed Tectipleura hemocyanins including hemocyanin gene 2 of those Hygrophila species (fig. 4, arrow with star and Schäfer et al. 2021b).

To analyze the evolution of hemocyanin gene structures of Gastropoda in more detail, we have derived the most parsimonious scenario of intron evolution within these genes (fig. 5). Including all known hemocyanin gene structures of gastropods, it is based on the most parsimony principle (Rogozin et al. 2006) assuming that an intron shared within a sister group has already been present within their common ancestor. The same applies to missing introns detected within sister groups that are to be assumed lost in an ancestor. Intron sliding was not considered because this phenomenon is difficult to identify by their location solely if introns vary in more than one nucleotide (Rogozin et al. 2000). Furthermore, it most likely contributes little to gene structure evolution (Poverennaya et al. 2020; Stoltzfus et al. 1997). Positions of nearby lying FU-internal introns within the analyzed hemocyanin

genes of Caenogastropoda vary in at least six nucleotides and with exception of six cases in more than 15 nucleotides (supplement 2). Thus, we do not assume that intron sliding has strongly contributed to hemocyanin gene evolution. It has to be noted that the presence of some introns within an ancestor cannot be assessed because two possible scenarios would have taken the same number of evolutionary steps (fig. 5). Intron origins which could not be determined exactly are marked with smaller arrows in a hypothetical gene precursor of a common ancestor in fig. 5. Nevertheless, the model of intron evolution shows the most parsimonious explanation for the revealed exon-intron structures and clearly indicates that ongoing changes within hemocyanin gene structures during the evolution of different Caenogastropoda are most likely. Although it represents only one possible hypothesis of hemocyanin gene evolution, the scenario strongly suggests a gradual accumulation of introns which led to gene structures with several internal introns, regardless of the exact origins of particular introns.

Considered more closely, the presented model of intron evolution reveals that the first gains of internal introns which are present in hemocyanins of Caenogastropoda most probably occurred within a common ancestor of Apogastropoda (fig. 5). All in all, we have detected four internal introns which are present within all analyzed species of Tectipleura and Caenogastropoda. The model of intron evolution shows that multiple FU-internal introns (12-13) and two introns within the signal peptide-coding sequences are likely to have evolved within a common ancestor of Caenogastropoda. The most parsimonious scenario is that the accumulation throughout the evolution of Sorbeoconcha, Hypsogastropoda and Muricidae as well as within hemocyanin genes of the particular lineages of the analyzed species continued gradually (fig. 5).

Our findings support the hypothesis that the accumulation of introns is a general phenomenon within hemocyanin genes of Apogastropoda and contrasts the hemocyanin gene evolution of Vetigastropoda and Cephalopoda (Chiumiento et al. 2020; Schäfer et al. 2021b). We were able to show that intron accumulation occurred in hemocyanin genes of several Apogastropoda lineages independently and started most likely within a hemocyanin gene of a common ancestor of Heterobranchia and Caenogastropoda. We have shown before that the high intron gain rate is hemocyanin gene-specific within Apogastropoda (Schäfer et al. 2021b). Such lineage- as well as gene-specific evolutionary rates have already been described for further genes (Carmel et al. 2007; Rogozin et al. 2003). As previously proposed (Schäfer et al. 2021b), a large number of introns might be accompanied by regulatory advantages which may cause this lineage- and gene-specific trend of intron accumulation in hemocyanins of Apogastropoda. Despite the enlarged number of introns, hemocyanin genes of the two large groups of Apogastropoda, precisely Heterobranchia (represented by Tectipleura) and Caenogastropoda, vary strongly in their rate of changes concerning their gene architecture. The strong variation of exon-intron structures of caenogastropod hemocyanin genes, which we have obtained in this study, contrasts the highly conserved exon-intron structures that have been detected in hemocyanins of Tectipleura. Hemocyanin gene architectures of Tectipleura evolved most probably within a common ancestor and thus before the multiple gene duplications occurred in several caenogastropod groups. Subsequently, the exon-intron structure maintained mostly conserved until today within all analyzed species of this large group of gastropods (Schäfer et al. 2021b). Hemocyanin gene structures of Caenogastropoda, in contrast, changed independently within the different lineages and within different paralogous genes.

Since the number of Tectipleura species corresponds to ~80% to the number of Caenogastropoda species and even ~96% to the species of Hypsogastropoda (fig. 5; WoRMS Editorial Board 2020), the larger number of differences within hemocyanin genes of Caenogastropoda cannot be explained by a higher species number. We rather suggest that the intron gain rate decreased within Tectipleura quite early because their hemocyanin genes became saturated by introns more quickly than those of Caenogastropoda. Gene- and lineage-specific saturation intron densities have already been described for plants (Basu et al. 2008) and mammals (Kordiš 2011). According to the most parsimonious scenario (fig. 5) of intron evolution in gastropod hemocyanins, 42 introns were gained during the evolution from an Apogastropoda ancestor to a precursor of Tectipleura and might have saturated the genes. This is strikingly more than the number of intron gains during the evolution from an Apogastropoda ancestor to the ancestors of Caenogastropoda (12 – 13 gains), Sorbeoconcha (15 – 18 gains) or Hypsogastropoda (22 – 30 gains, fig. 5). Schäfer et al. (2021b) showed that introns of hemocyanin genes of Tectipleura are evenly distributed within their coding sequences and consequently led to a regular size distribution of exons. Conservation of exon length, as we have discovered for Tectipleura hemocyanin genes, has been suggested to indicate evolutionary advantages (Davila-Velderrain et al. 2014; Fu and Lin 2012).

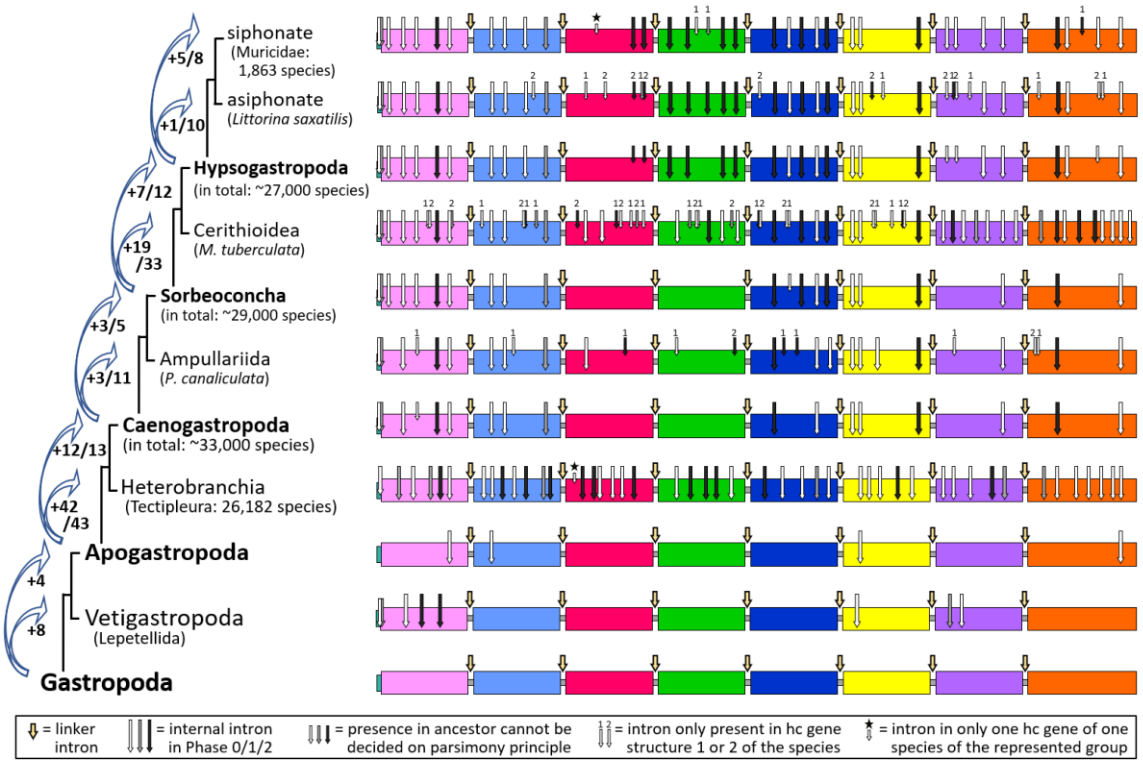


Fig. 5: Maximum parsimony scenario of exon-intron architecture evolution within gastropod hemocyanins. The phylogenetic tree on the left shows the relation of gastropod species with known hemocyanin gene architectures (Ponder et al. 2019). For groups of Apogastropoda also the number of including species according to WoRMS Editorial Board (2020) is shown. If gene structures are known for only one representative of a lineage, the specific species is given. On the right, hypothetical gene structures are shown for the common ancestors. They are based on the maximum parsimony principle, assuming that introns which are at the same position concerning the coding sequences of hemocyanin genes of sister groups have already been present in hemocyanin genes of a common ancestor. The same applies for intron losses if both descendants do not include a former present intron. The deduced model of gene structure evolution within gastropods shows a gradually gain of introns during the evolution of Caenogastropoda. The amount of intron gains varies between the different hemocyanin lineages and is numbered on the left side of the taxonomic tree. We were not able to exactly assess every intron origin because two independent intron gains within two descendant species represent the same number of evolutionary events as one gain within a common ancestor together with an intron loss within one gene of one of the descendants. In these cases we do not esteem one scenario to be more likely than the other one and highlighted them with smaller arrows in the ancestor genes. In contrast to hemocyanin gene structures of Tectipleura, those of Caenogastropoda do not only vary between different caenogastropod lineages but also within paralogous genes within the same species. We symbolized such introns which are present in the gene structure of only one hemocyanin gene of the same species with small arrows with the number of the gene in which it is located in on top. Arrows with a star on top represent exceptions from a conserved group-specific hemocyanin gene structure (e.g. additional Hygrophila-specific intron, see fig. 4 and cf. Schäfer et al. (2018)). As in all other molluscan hemocyanins, linker introns are located at the same position within the sequences coding for linker peptides between all canonical functional units.

Accordingly, our results of conserved exon-intron structures may imply that the large number of introns and the regular distribution of exons within hemocyanin genes of Tectipleura might bring evolutionary benefits (e.g. expanded possibilities of gene expression regulation) (Schäfer et al. 2021b). More elaborated explanations of possible evolutionary advantages are further discussed below under "Evolutionary constraints on hemocyanin genes in Apogastropoda?". If a similar evolutionary constraint is acting on hemocyanin genes of Caenogastropoda as well, this might explain the continuous and ongoing accumulation of introns in these genes during the evolution of Caenogastropoda as proposed in the maximum parsimony scenario (fig. 5). The hemocyanin gene architectures suggested for the common ancestors of Caenogastropoda, Sorbeoconcha or Hypsogastropoda include less introns than the gene structure of Tectipleura hemocyanins (fig. 5).

Therefore, they comprise relatively large exons which may represent possible targets for intron insertion (Hawkins 1988; Hwang and Cohen 1997; Roy and Irimia 2009). Due to lineage- and gene-specific intron gain/loss rates which may result from evolutionary constraints, this may lead to gene architectures with accumulating introns during the evolution of Caenogastropoda until saturation is reached. In contrast to that, introns in Tectipleura hemocyanins were gained in a common ancestor and maintained until today. The large number of introns led to relatively small exon lengths which may preserve further intron gains (Hawkins 1988; Hwang and Cohen 1997; Roy and Irimia 2009) and thus might have already caused the described saturation intron density (Basu et al. 2008; Kordiš 2011). This may be the reason for the conserved gene structure of Tectipleura hemocyanins which contrast the continuous intron gains in hemocyanins of Caenogastropoda. With exception of the rate of intron gains, however, intron accumulation seems to be common within hemocyanins of both groups of Apogastropoda and may be caused by evolutionary constraints.

Evolutionary constraints on hemocyanin genes in Apogastropoda?

Our results on the large number of gene duplications and the various intron gains identified for hemocyanin genes of Caenogastropoda support the hypothesis of high dynamics within hemocyanin gene evolution in Apogastropoda (Schäfer et al. 2018; 2021b). Both, Caenogastropoda and Heterobranchia represent the most diverse groups of the phylum Mollusca and encompass together over 64,000 species (WoRMS Editorial Board 2020) which are living in nearly all kinds of habitats from deep sea to deserts. Yet, no hemocyanin genes of other Heterobranchia groups than Tectipleura have been analyzed. However, we have analyzed representatives from the two different Tectipleura monophyla, Euopisthobranchia and Panpulmonata, and, since Tectipleura represent by far the largest group of Heterobranchia, our results of the conserved gene architecture refer to the major part of Heterobranchia (~ 26,000 Tectipleura species of a total number of about 30,000 analyzed Heterobranchia species, WoRMS Editorial Board 2020).

Radiation which led to the large diversity of Apogastropoda is based on an enormous number of adaptations, including the evolution of new respiratory organs as lungs or pneumostomes, changed abilities of osmoregulation and reproduction as well as new behavioral strategies (Vermeij and Dudley 2000; Vermeij and Wesselingh F. P. 2002). We have already suggested that strong evolutionary rates of hemocyanin genes may represent molecular adaptations of Apogastropoda and thus might be associated with the large diversity and the multiple habitat shifts within that group (Schäfer et al. 2018; Schäfer et al. 2021b). This hypothesis is based on the strong temperature dependence of the oxygen affinity of hemocyanins (Brix et al. 1989; Burnett et al. 1988; Mangum 1990). Various adaptations of this oxygen transport protein to different environmental conditions have been reported (González et al. 2017; Melzner et al. 2007; Oellermann et al. 2015a; Oellermann et al. 2015b; Strobel et al. 2012; Yesilyurt et al. 2008; Zielinski et al. 2001). They appear as necessary to ensure a sufficient supply of oxygen and therefore as fundamental for molluscs to survive. Consequently, they represent one essential precondition for habitat shifts between strongly different environments (e.g. from sea to land or freshwater). Strong variability as previously identified for hemocyanin genes of Tectipleura and now verified for five groups of Caenogastropoda, could enable these required adaptations (Schäfer et al. 2018; 2021b).

Gene duplications as we have identified for hemocyanins of Tectipleura and Caenogastropoda (fig. 3) play a major role in genomic complexity and evolution (Magadum et al. 2013; Ohno 1970). They are a driving force in organismal diversity (Lynch and Conery 2000) and can promote adaptations (Qian and Zhang 2014). Hemocyanin gene duplications could be followed by differential evolution of the different genes and could eventually lead to hemocyanins with, for example, different oxygen affinities, varying pH- or temperature sensitivities or differential expression patterns. This might represent an origin of genetic variability and adaptations as it has already been discovered for Cephalopoda (Oellermann et al. 2015a; Oellermann et al. 2015b; Strobel et al. 2012; Thonig et al. 2014). The squid *Sepia officinalis*, for example, possesses multiple hemocyanin genes that underlie differential expression (Strobel et al. 2012; Thonig et al. 2014). Thonig et al. (2014) identified ontogeny-specific expression patterns of hemocyanin genes of this squid species. Thereby, one hemocyanin gene is highly expressed within the egg and might help to sustain an adequate oxygen supply despite of hypoxic and hypercapnic conditions arising within the eggs during embryogenesis. After hatching, however, the expression of this gene is downregulated, whereas two other hemocyanin isoforms are upregulated (Thonig et al. 2014). In order to examine if the multitude of hemocyanin genes in Apogastropoda contributes to adaptive radiation and evolutionary benefits as well, further studies are needed which examine biochemical properties and physiological

implications of different hemocyanin isoforms of Heterobranchia and Caenogastropoda.

In addition to gene duplications, our here presented findings confirm that intron accumulation is a general phenomenon of hemocyanin genes in Apogastropoda (Chiumiento et al. 2020; Schäfer et al. 2021b). As already proposed, this may indicate evolutionary constraints on a large number of introns in gene structures of Apogastropoda. Introns, in general, can increase species diversity (Calarco and Ellis 2020) and thus may also contribute to adaptation. The extensive number of internal introns found within hemocyanin genes of Apogastropoda might facilitate regulation of differential expression (discussed more detailed in Schäfer et al. 2021b). Several regulatory mechanisms of introns have already been identified (e.g. Chorev and Carmel 2012; Rose 2008; Rose 2018) and include among others temperature-dependent gene-expression (Airoldi et al. 2015; Evantal et al. 2018; Gotic et al. 2016; James et al. 2018). Regulation of hemocyanin gene expression could help to control the availability of different hemocyanin isoforms which harbor different properties (e.g. varying oxygen affinities at different temperatures). Thus, regulatory functions of introns could represent evolutionary advantages that promote intron accumulation in hemocyanin genes of Apogastropoda (for more details see Schäfer et al. 2021b). According to this hypothesis, the enormous number of internal introns in hemocyanin genes in combination with the evolution of multiple hemocyanin genes which have been identified for Apogastropoda might be one of the many adaptations that finally enable the large diversity of Caenogastropoda and Heterobranchia.

Our results show that the accumulation rate of introns which maintained in hemocyanins of Caenogastropoda is different within various lineages and highest within Cerithioidea (fig. 5). In addition, the two paralogous hemocyanin genes of the cerithioid gastropod *Melanoides tuberculata* exhibit the largest variations between their exon-intron structures that have been found within one gastropod species (small arrows with numbers on top in fig. 5, cf. also supplement 3). Furthermore, the 550 kDa mega-hemocyanin subunit represents the only hemocyanin gene which lacks a linker intron between two different functional units. The loss of regions coding for FU-g and FU-h as well as the gain of six FU-f-coding sequence sections additionally led to strong variations of the protein structures which influence its physiological features (typical 400 kDa hemocyanin and 550 kDa mega-hemocyanin subunits which differ in their oxygen binding capacity, cf. introduction above and Lieb et al. 2010). These enormous differences on gene and simultaneously on protein level might represent combined adaptations of hemocyanins which enable the strong diversity of habitats Cerithioidea live in. The members of this superfamily of Caenogastropoda are distributed in freshwater, brackish water and marine ecosystems in mostly warm temperate regions (for an overview see Strong et al. 2011). Their habitats include some extreme biotopes as rocky intertidal shores, mudflats and mangrove forests which entail environmental conditions as strongly changing temperatures, humidity differences and little oxygen availability. The fact that the strongest differences of paralogous hemocyanins (including gene and protein level) have been identified in a group of gastropods with such extremely diverse habitats may represent a further hint that the evolution of hemocyanin plays a major role for the evolution of molluscs. Since both genes can be expressed in different rations (Lieb et al. 2010) they can help to adapt to different living conditions.

Conclusions

The oxygen affinity of molluscan hemocyanins is strongly influenced by abiotic factors, as for example temperature (Mangum 1990). A multitude of adaptations of these oxygen transport proteins have already been described (e.g. Strobel et al. 2012) and seem to be indispensable for a lot of molluscan species to ensure a sufficient oxygen supply. Our findings reveal a strong diversity of hemocyanin genes of Caenogastropoda including multiple independent gene duplications within different caenogastropod groups as well as a strong accumulation of FU-internal introns (21 - 57) within their genes. Since gene duplications and intron gains have also been discovered within hemocyanin genes of Tectipleura (Schäfer et al. 2018; 2021b), they most likely represent general phenomena of hemocyanin gene evolution within Apogastropoda. This contrasts hemocyanin evolution of Vetigastropoda and may support the hypothesis that diversity of this oxygen transporter may boost adaptation. Thereby gene duplications may provide new genes for evolutionary renovation by mutation and selection and the accumulation of introns may contribute to regulatory opportunities. Although, follow-up studies are required to analyze biochemical and physiological properties of apogastropod hemocyanins (e.g. varying oxygen affinities, different temperature dependencies of the isoforms and differential expression patterns), our results indicate that the evolution of hemocyanins in Heterobranchia and Apogastropoda might be one out of many influencing factors that enabled the large diversity of Apogastropoda by facilitating adaptive radiation.

Funding This study was supported by a grant from the German Research Foundation (DFG) to B. Lieb, project number LI998/15-1. Additional support came from the Wagner and the Feldbausch Foundation (both to B. Lieb, University of Mainz).

Authors' contributions Gabriela G. Schäfer was a major contributor to the conception of the analyses. She prepared genomic DNA and RNA for NGS. Additionally, she analyzed as well as interpreted nucleotide sequences and gene structures in cooperation with Lukas J. Grebe and Robin Schinkel who conducted large parts of the NGS analyses of MtH, RtH and LisaH. Bernhard Lieb conceived the project and supervised this study as principle investigator. Gabriela G. Schäfer wrote the first version of the manuscript that was commented, revised and approved by all authors.

Availability of data and materials The cDNA sequences and gene structures obtained during the current study are available in NCBI under the following accession numbers: LisaH1: BK014376; LisaH2: BK014375; PcH I: BK014379; PcH IIb: BK014378; PcH III: BK014377; NlH1: MT939254; NlH2: MT939255; RtH1: BK014286; RtH2: BK014287; MtH400 KC405575, MtH550 KC405576.

Competing interests The authors declare that they have no competing interests.

References

Airoldi CA, McKay M, Davies B (2015) MAF2 Is Regulated by Temperature-Dependent Splicing and Represses Flowering at Low Temperatures in Parallel with FLM. PLoS One 10:e0126516. https://doi.org/10.1371/journal.pone.0126516

Altenhein B, Lieb B, Awenius C, Markl J (2000) Gene Structure of Gastropod Hemocyanin. Zoology Suppl. III. Proceedings of the 93th Annual Meeting Bonn, Germany

Altenhein B, Markl J, Lieb B (2002) Gene structure and hemocyanin isoform HtH2 from the mollusc Haliotis tuberculata indicate early and late intron hot spots. Gene 301:53–60. https://doi.org/10.1016/S0378-1119(02)01081-8

Basu MK, Rogozin IB, Deusch O, Dagan T, Martin W, Koonin EV (2008) Evolutionary dynamics of introns in plastid-derived genes in plants: saturation nearly reached but slow intron gain continues. Mol Biol Evol 25:111–119. https://doi.org/10.1093/molbev/msm234

Bergmann S, Lieb B, Ruth P, Markl J (2006) The hemocyanin from a living fossil, the cephalopod Nautilus pompilius: protein structure, gene organization, and evolution. J Mol Evol 62:362–374. https://doi.org/10.1007/s00239-005-0160-x

Bouchet P, Rocroi J-P, Hausdorf B, Kaim A, Kano Y, Nützel A, Parkhaev P, Schrödl M, Strong EE (2017) Revised Classification, Nomenclator and Typification of Gastropod and Monoplacophoran Families. Malacologia 61:1–526. https://doi.org/10.4002/040.061.0201

Brix O, Bårdgard A, Cau A, Colosimo A, Condò SG, Giardina B (1989) Oxygen-binding properties of cephalopod blood with special reference to environmental temperatures and ecological distribution. J. Exp. Zool. 252:34–42. https://doi.org/10.1002/jez.1402520106

Brix O, Colosimo A, Giardina B (1995) Temperature dependence of oxygen binding to cephalopod haemocyanins: Ecological implications. Marine and Freshwater Behaviour and Physiology 25:149–162 . https://doi.org/10.1080/10236249409378914

Burnett LE, Scholnick DA, Mangum CP (1988) Temperature Sensitivity of Molluscan and Arthropod Hemocyanins. The Biological Bulletin 174:153–162. https://doi.org/10.2307/1541782

Calarco L, Ellis J (2020) Contribution of introns to the species diversity associated with the apicomplexan parasite, Neospora caninum. Parasitol Res 119:431–445. https://doi.org/10.1007/s00436-019-06561-x

Carmel L, Rogozin IB, Wolf YI, Koonin EV (2007) Evolutionarily conserved genes preferentially accumulate introns. Genome Res 17:1045–1050. https://doi.org/10.1101/gr.5978207

Chiumiento IR, Ituarte S, Sun J, Qiu JW, Heras H, Dreon MS (2020) Hemocyanin of the caenogastropod Pomacea canaliculata exhibits evolutionary differences among gastropod clades. PLoS One 15:e0228325. https://doi.org/10.1371/journal.pone.0228325

Chorev M, Carmel L (2012) The function of introns. Front Genet 3:55. https://doi.org/10.3389/fgene.2012.00055

Davila-Velderrain J, Servin-Marquez A, Alvarez-Buylla ER (2014) Molecular evolution constraints in the floral organ specification gene regulatory network module across 18 angiosperm genomes. Mol Biol Evol 31:560–573. https://doi.org/10.1093/molbev/mst223

Dayrat B, Tillier S (2002) Evolutionary relationships of euthyneuran gastropods (Mollusca): A cladistic re-evaluation of morphological characters. Zoological Journal of the Linnean Society 135:403–470. https://doi.org/10.1046/j.1096-3642.2002.00018.x

Dinapoli A, Klussmann-Kolb A (2010) The long way to diversity - phylogeny and evolution of the Heterobranchia (Mollusca: Gastropoda). Mol Phylogenet Evol 55:60–76. https://doi.org/10.1016/j.ympev.2009.09.019

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340

Evantal N, Anduaga AM, Bartok O, Patop IL, Weiss R, Kadener S (2018) Thermosensitive alternative splicing senses and mediates temperature adaptation in Drosophila. bioRxiv 503409. https://doi.org/10.7554/eLife.44642

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–376. https://doi.org/10.1007/BF01734359

Fu GC-L, Lin W-C (2012) Identification of gene-oriented exon orthology between human and mouse. BMC Genomics 13 Suppl 1:S10. https://doi.org/10.1186/1471-2164-13-S1-S10

Gatsogiannis C, Markl J (2009) Keyhole limpet hemocyanin: 9-A CryoEM structure and molecular model of the KLH1 didecamer reveal the interfaces and intricate topology of the 160 functional units. J Mol Biol 385:963–983. https://doi.org/10.1016/j.jmb.2008.10.080

Gatsogiannis C, Hofnagel O, Markl J, Raunser S (2015) Structure of mega-hemocyanin reveals protein origami in snails. Structure 23:93–103. https://doi.org/10.1016/j.str.2014.10.013

Gebauer W, Stoeva S, Voelter W, Dainese E, Salvato B, Beltramini M, Markl J (1999) Hemocyanin subunit organization of the gastropod Rapana thomasiana. Arch Biochem Biophys 372:128–134. https://doi.org/10.1006/abbi.1999.1478

González A, Nova E, Del Campo M, Manubens A, Ioannes A de, Ferreira J, Becker MI (2017) The oxygen-binding properties of hemocyanin from the mollusk Concholepas concholepas. Biochim Biophys Acta Proteins Proteom 1865:1746–1757. https://doi.org/10.1016/j.bbapap.2017.08.017

Gotic I, Omidi S, Fleury-Olela F, Molina N, Naef F, Schibler U (2016) Temperature regulates splicing efficiency of the cold-inducible RNA-binding protein gene Cirbp. Genes Dev 30:2005–2017. https://doi.org/10.1101/gad.287094.116.

Gutowska MA, Melzner F (2009) Abiotic conditions in cephalopod (Sepia officinalis) eggs: Embryonic development at low pH and high pCO2. Mar Biol 156:515–519. https://doi.org/10.1007/s00227-008-1096-7

Hawkins JD (1988) A survey on intron and exon lengths. Nucleic Acids Res 16:9893–9908

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Hwang DY, Cohen JB (1997) U1 small nuclear RNA-promoted exon selection requires a minimal distance between the position of U1 binding and the 3' splice site across the exon. Mol Cell Biol 17:7099–7107. https://doi.org/10.1128/MCB.17.12.7099

James AB, Calixto CPG, Tzioutziou NA, Guo W, Zhang R, Simpson CG, Jiang W, Nimmo GA, Brown JWS, Nimmo HG (2018) How does temperature affect splicing events? Isoform switching of splicing factors regulates splicing of LATE ELONGATED HYPOCOTYL (LHY). Plant Cell Environ 41:1539–1550. https://doi.org/10.1111/pce.13193

Jörger KM, Stöger I, Kano Y, Fukuda H, Knebelsberger T, Schrödl M (2010) On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. BMC Evol Biol 10:323. https://doi.org/10.1186/1471-2148-10-323

Kano Y, Brenzinger B, Nützel A, Wilson NG, Schrödl M (2016) Ringiculid bubble snails recovered as the sister group to sea slugs (Nudipleura). Sci Rep 6:30908. https://doi.org/10.1038/srep30908

Kato S, Matsui T, Gatsogiannis C, Tanaka Y (2018) Molluscan hemocyanin: structure, evolution, and physiology. Biophys Rev 10:191–202. https://doi.org/10.1007/s12551-017-0349-4

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (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

Kocot KM, Halanych KM, Krug PJ (2013) Phylogenomics supports Panpulmonata: Opisthobranch paraphyly and key evolutionary steps in a major radiation of gastropod molluscs. Mol Phylogenet Evol 69:764–771. https://doi.org/10.1016/j.ympev.2013.07.001

Kordiš D (2011) Extensive intron gain in the ancestor of placental mammals. Biol Direct 6:59. https://doi.org/10.1186/1745-6150-6-59

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054

Lieb B, Markl J (2004) Evolution of molluscan hemocyanins as deduced from DNA sequencing. Micron 35:117–119. https://doi.org/10.1016/j.micron.2003.10.035

Lieb B, Altenhein B, Markl J (2000) The sequence of a gastropod hemocyanin (HtH1 from Haliotis tuberculata). J Biol Chem 275:5675–5681. https://doi.org/10.1074/jbc.275.8.5675

Lieb B, Altenhein B, Markl J, Vincent A, van Olden E, van Holde KE, Miller KI (2001) Structures of two molluscan hemocyanin genes: significance for gene evolution. Proc Natl Acad Sci U S A 98:4546–4551. https://doi.org/10.1073/pnas.071049998

Lieb B, Boisguerin V, Gebauer W, Markl J (2004) cDNA sequence, protein structure, and evolution of the single hemocyanin from Aplysia californica, an opisthobranch gastropod. J Mol Evol 59:536–545. https://doi.org/10.1007/s00239-004-2646-3

Lieb B, Gebauer W, Gatsogiannis C, Depoix F, Hellmann N, Harasewych MG, Strong EE, Markl J (2010) Molluscan mega-hemocyanin: an ancient oxygen carrier tuned by a ~550 kDa polypeptide. Front Zool 7:14. https://doi.org/10.1186/1742-9994-7-14

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science 290:1151–1155 . https://doi.org/10.1126/science.290.5494.1151

Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R (2013) Gene duplication as a major force in evolution. J Genet 92:155–161. https://doi.org/10.1007/s12041-013-0212-8

Mangum CP (1990) Gas Transport in the Blood. In: Gilbert DL, Adelman WJ, Arnold JM (eds) Squid as Experimental Animals, vol 211. Springer US, Boston, MA, pp 443–468

Markl J, Savel-Niemann A, Wegener-Strake A, Sding M, Schneider A, Gebauer W, Harris JR (1991) The role of two distinct subunit types in the architecture of keyhole limpet hemocyanin (KLH). Naturwissenschaften 78:512–514. https://doi.org/10.1007/BF01131401

Markl J (2013) Evolution of molluscan hemocyanin structures. Biochim Biophys Acta 1834:1840–1852. https://doi.org/10.1016/j.bbapap.2013.02.020

Mau B, Newton MA (1997) Phylogenetic Inference for Binary Data on Dendograms Using Markov Chain Monte Carlo. Journal of Computational and Graphical Statistics 6:122–131. https://doi.org/10.1080/10618600.1997.10474731

Melzner F, Mark FC, Pörtner H-O (2007) Role of blood-oxygen transport in thermal tolerance of the cuttlefish, Sepia officinalis. Integr Comp Biol 47:645–655. https://doi.org/10.1093/icb/icm074

Miller K, van Holde KE (1974) Oxygen binding by callianassa californiensis hemocyanin. Biochemistry 13:1668–1674. https://doi.org/10.1021/bi00705a017

Miller KI (1985) Oxygen equilibria of Octopus dofleini hemocyanin. Biochemistry 24:4582–4586. https://doi.org/10.1021/bi00338a015

Oellermann M, Lieb B, Pörtner H-O, Semmens JM, Mark FC (2015a) Blue blood on ice: modulated blood oxygen transport facilitates cold compensation and eurythermy in an Antarctic octopod. Front Zool 12. https://doi.org/10.1186/s12983-015-0097-x

Oellermann M, Strugnell JM, Lieb B, Mark FC (2015b) Positive selection in octopus haemocyanin indicates functional links to temperature adaptation. BMC Evol Biol 15:133. https://doi.org/10.1186/s12862-015-0411-4

Ohno S (1970) Evolution by Gene Duplication. Springer Berlin Heidelberg, Berlin, Heidelberg

Ponder WF, Lindberg DR, Ponder JM (2019) Biology and Evolution of the Mollusca. CRC PRESS, [S.l.]

Poverennaya IV, Potapova NA, Spirin SA (2020) Is there any intron sliding in mammals? BMC Evol Biol 20:164. https://doi.org/10.1186/s12862-020-01726-0

Qian W, Zhang J (2014) Genomic evidence for adaptation by gene duplication. Genome Res 24:1356–1362. https://doi.org/10.1101/gr.172098.114

Rogozin IB, Lyons-Weiler J, Koonin EV (2000) Intron sliding in conserved gene families. Trends in Genetics 16:430–432. https://doi.org/10.1016/S0168-9525(00)02096-5

Rogozin IB, Wolf YI, Sorokin AV, Mirkin BG, Koonin EV (2003) Remarkable Interkingdom Conservation of Intron Positions and Massive, Lineage-Specific Intron Loss and Gain in Eukaryotic Evolution. Current Biology 13:1512–1517. https://doi.org/10.1016/S0960-9822(03)00558-X

Rogozin IB, Wolf YI, Babenko VN, Koonin EV (2006) Dollo parsimony and the reconstruction of genome evolution. In: Albert VA (ed) Parsimony, Phylogeny, and Genomics. Oxford University Press, pp 190–200

Romero PE, Pfenninger M, Kano Y, Klussmann-Kolb A (2016a) Molecular phylogeny of the Ellobiidae (Gastropoda: Panpulmonata) supports independent terrestrial invasions. Mol Phylogenet Evol 97:43–54. https://doi.org/10.1016/j.ympev.2015.12.014

Romero PE, Weigand AM, Pfenninger M (2016b) Positive selection on panpulmonate mitogenomes provide new clues on adaptations to terrestrial life. BMC Evol Biol 16:164. https://doi.org/10.1186/s12862-016-0735-8.

Rose AB (2008) Intron-mediated regulation of gene expression. Curr Top Microbiol Immunol 326:277–290. https://doi.org/10.1007/978-3-540-76776-3_15

Rose AB (2018) Introns as Gene Regulators: A Brick on the Accelerator. Front Genet 9:672 https://doi.org/10.3389/fgene.2018.00672

Roy SW, Irimia M (2009) Mystery of intron gain: new data and new models. Trends in Genetics 25:67–73. https://doi.org/10.1016/j.tig.2008.11.004

Schäfer GG, Pedrini-Martha V, Schnegg R, Dallinger R, Jackson DJ, Lieb B (2018) Hemocyanin genes as indicators of habitat shifts in Panpulmonata? Mol Phylogenet Evol 130:99–103. https://doi.org/10.1016/j.ympev.2018.10.014

Schäfer GG, Grebe LJ, Depoix F, Lieb B (2021a) Hemocyanins of Muricidae: New 'Insights' Unravel an Additional Highly Hydrophilic 800 kDa Mass Within the Molecule. J Mol Evol. https://doi.org/10.1007/s00239-020-09986-6

Schäfer GG, Pedrini-Martha V, Jackson DJ, Dallinger R, Lieb B (2021b) The evolution of hemocyanin genes in Tectipleura: a multitude of conserved introns in highly diverse gastropods. BMC Ecol Evol 21:36. https://doi.org/10.1186/s12862-021-01763-3

Schrödl M (2014) Opinion: Time to say "Bye-bye Pulmonata"? SPIXIANA 37:161–164

Schweizer M, Triebskorn R, Köhler H-R (2019) Snails in the sun: Strategies of terrestrial gastropods to cope with hot and dry conditions. Ecol Evol 9:12940–12960 . https://doi.org/10.1002/ece3.5607

Stoltzfus A, Logsdon JM, JR, Palmer JD, Doolittle WF (1997) Intron "sliding" and the diversity of intron positions. Proc Natl Acad Sci U S A 94:10739–10744. https://doi.org/10.1073/pnas.94.20.10739

Strobel A, Hu MYA, Gutowska MA, Lieb B, Lucassen M, Melzner F, Portner HO, Mark FC (2012) Influence of temperature, hypercapnia, and development on the relative expression of different hemocyanin isoforms in the common cuttlefish Sepia officinalis. J Exp Zool A Ecol Genet Physiol 317:511–523. https://doi.org/10.1002/jez.1743

Strong EE, Colgan DJ, Healy JM, Lydeard C, Ponder WF, GLAUBRECHT M (2011) Phylogeny of the gastropod superfamily Cerithioidea using morphology and molecules. Zoological Journal of the Linnean Society 162:43–89. https://doi.org/10.1111/j.1096-3642.2010.00670.x

Swerdlow RD, Ebert RF, Lee P, Bonaventura C, Miller KI (1996) Keyhole limpet hemocyanin: Structural and functional characterization of two different subunits and multimers. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 113:537–548. https://doi.org/10.1016/0305-0491(95)02091-8

Thonig A, Oellermann M, Lieb B, Mark FC (2014) A new haemocyanin in cuttlefish (Sepia officinalis) eggs: Sequence analysis and relevance during ontogeny. Evodevo 5:6. https://doi.org/10.1186/2041-9139-5-6

van Holde KE, Miller KI (1995) Hemocyanins. Adv Protein Chem 47:1–81

Vermeij GJ, Dudley R (2000) Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems. Biol. J. Linn. Soc. 70:541–554. https://doi.org/10.1111/j.1095-8312.2000.tb00216.x

Vermeij GJ, Wesselingh F. P. (2002) Neogastropod molluscs from the Miocene of western Amazonia, with comments on marine to freshwater transitions in molluscs. Journal of Paleontology 76:265–270. https://doi.org/10.1666/0022-3360(2002)076<0265:NMFTMO>2.0.CO;2

WoRMS Editorial Board (2020) World Register of Marine Species. http://www.marinespecies.org

Yao T, Zhao M-M, He J, Han T, Peng W, Zhang H, Wang J-Y, Jiang J-Z (2019) Gene expression and phenoloxidase activities of hemocyanin isoforms in response to pathogen infections in abalone Haliotis diversicolor. Int J Biol Macromol 129:538–551. https://doi.org/10.1016/j.ijbiomac.2019.02.013

Yesilyurt BT, Gielens C, Meersman F (2008) Thermal stability of homologous functional units of Helix pomatia hemocyanin does not correlate with carbohydrate content. FEBS J 275:3625–3632. https://doi.org/10.1111/j.1742-4658.2008.06507.x

Zielinski S, Sartoris FJ, Pörtner HO (2001) Temperature effects on hemocyanin oxygen binding in an antarctic cephalopod. The Biological Bulletin 200:67–76. https://doi.org/10.2307/1543086

2.5

Responsiveness of metallothionein and hemocyanin genes to cadmium and copper exposure in the garden snail *Cornu aspersum*

Veronika Pedrini-Martha
Raimund Schnegg
Gabriela Giannina Schäfer
Bernhard Lieb
Willi Salvenmoser
Reinhard Dallinger

Journal of Experimental Zoology
February 1, 2021; Volume 335, Issue 2, Pages:228–238.

Online: November 2020

https://doi.org/10.1002/jez.2425

Mein Beitrag:

Ich habe die bioinformatischen Analysen durchgeführt. Dabei verglich ich die Transkriptomdaten der Individuen, die mit Kupfer angereicherter Nahrung gefüttert wurden, mit denen der Kontrollgruppe hinsichtlich quantitativer Unterschiede der exprimierten Hämocyanin- und Metallothionein-cDNA-Sequenzen.

Received: 18 September 2020 | Accepted: 13 October 2020

DOI: 10.1002/jez.2425

RESEARCH PAPER

JEZ-A ECOLOGICAL AND INTEGRATIVE PHYSIOLOGY WILEY

Responsiveness of metallothionein and hemocyanin genes to cadmium and copper exposure in the garden snail *Cornu aspersum*

Veronika Pedrini-Martha[1] | Raimund Schnegg[1] | Gabriela Giannina Schäfer[2] | Bernhard Lieb[2] | Willi Salvenmoser[1] | Reinhard Dallinger[1]

[1]Department of Zoology and Center of Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria

[2]Institute of Molecular Physiology, Johannes Gutenberg-University of Mainz, Mainz, Germany

Correspondence
Veronika Pedrini-Martha and Reinhard Dallinger, Department of Zoology and Center of Molecular Biosciences Innsbruck, University of Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria.
Email: Veronika.Pedrini-Martha@uibk.ac.at and Reinhard.Dallinger@uibk.ac.at

Funding information
Austrian Science Fund, Grant/Award Number: I 3032-B21; Deutsche Forschungsgemeinschaft, Grant/Award Number: LI998/15-1

Abstract

Terrestrial gastropods express metal-selective metallothioneins (MTs) by which they handle metal ions such as Zn^{2+}, Cd^{2+}, and Cu^{+}/Cu^{2+} through separate metabolic pathways. At the same time, they depend on the availability of sufficient amounts of Cu as an essential constituent of their respiratory protein, hemocyanin (Hc). It was, therefore, suggested that in snails Cu-dependent MT and Hc pathways might be metabolically connected. In fact, the Cu-specific snail MT (CuMT) is exclusively expressed in rhogocytes, a particular molluscan cell type present in the hemocoel and connective tissues. Snail rhogocytes are also the sites of Hc synthesis. In the present study, possible interactions between the metal-regulatory and detoxifying activity of MTs and the Cu demand of Hc isoforms was explored in the edible snail *Cornu aspersum*, one of the most common European helicid land snails. This species possesses CdMT and CuMT isoforms involved in metal-selective physiological tasks. In addition, *C. aspersum* expresses three different Hc isoforms (CaH αD, CaH αN, CaH β). We have examined the effect of Cd^{2+} and Cu^{2+} exposure on metal accumulation in the midgut gland and mantle of *C. aspersum*, testing the impact of these metals on transcriptional upregulation of *CdMT*, *CuMT*, and the three *Hc* genes in the two organs. We found that the *CuMT* and *CaH αD* genes exhibit an organ-specific transcriptional upregulation in the midgut gland of Cu-exposed snails. These results are discussed in view of possible interrelationships between the metal-selective activity of snail MT isoforms and the synthesis and metabolism of Hc isoforms.

KEYWORDS
bioaccumulation, Gastropoda, metal metabolism, respiration, stress

1 | INTRODUCTION

Terrestrial gastropods possess sophisticated mechanisms for diverting essential and nonessential transition metal ions such as Zn^{2+}, Cd^{2+}, and Cu^{+}/Cu^{2+} through separate metabolic pathways (Dallinger, 1993; Dvorak et al., 2019). A crucial role in this regulation is attributed to snail metallothioneins (MTs; Dallinger et al., 1997) which can specifically bind either Cd^{2+} and Zn^{2+} or Cu^{+} through activation of metal-selective isoforms (Palacios et al., 2011). The Cu-specific MT (CuMT) of terrestrial helicid snails, in particular, has been

Veronika Pedrini-Martha and Raimund Schnegg contributed equally to this study.

suggested to play a role in the synthesis of hemocyanin (Hc; Dallinger et al., 2005), the Cu-binding respiratory protein of most snails and molluscs (Kato et al., 2018; Lieb et al., 2010). In fact, snail CuMT is exclusively expressed in rhogocytes (Chabicovsky et al., 2003), which represent a particular molluscan cell type that occurs in the primary body cavity, either freely floating in the hemocoel or embedded in connective tissues (Haszprunar, 1996; Stewart et al., 2014). Apart from CuMT, rhogocytes are also the sites of Hc synthesis, as demonstrated for several Hcs and Hc isoforms in different gastropod species (Albrecht et al., 2001; Martin et al., 2011; Sminia & Boer, 1973; Sminia & Vlugt-Van Daalen, 1977). In the mantle of *Haliotis laevigata*, for example, two morphologically distinct rhogocyte populations were identified which probably represent two separate phases of Hc synthesis (Sairi et al., 2015).

However, the hypothesis of CuMT involvement in gastropod Hc synthesis is compromised by the fact that many species of Gastropoda do not possess Cu-selective MTs (Dallinger et al., 2020), and some do not use Hc as a primary respiratory protein (Alyakrinskaya, 2002; Lieb et al., 2006). Yet, this does not preclude an interaction between CuMTs and Hc synthesis, in species which express both, Cu-selective MTs and Hcs, in the same rhogocyte cells. Apart from Hc synthesis, gastropod rhogocytes have long been known to be involved in homeostasis and/or detoxification of trace metals (Kokkinopoulou et al., 2014, 2015; Marigómez et al., 2002; Simkiss & Mason, 1983). Because of the indispensable role of copper (Cu^{+}/Cu^{2+}) as a central component of Hc in most snails, the homeostatic regulation of this metal may be crucial. Metal balance analyses of snail tissues indicate that Cu is accumulated, among other organs, in the midgut gland and mantle (Berger & Dallinger, 1989), in which a particularly high density of rhogocytes can be detected (Sairi et al., 2015). Chromatographic fractionations show, moreover, that snail tissue Cu is mainly allocated to Hc and MT fractions (Berger et al., 1997; Dallinger, 1996), even upon exposure to nonessential trace elements like cadmium (Cd; Dallinger et al., 2004). Yet, Cd exposure can significantly impact on Cu metabolism and homeostasis in tissues of snails (Gnatyshyna et al., 2020; Nica et al., 2019) and other molluscs (Wu & Wang, 2010). In mammals, Cd interacts with important proteins and enzymes involved in Cu homeostasis, including MT, Cu-chaperonins, and Cu-transporters (Moulis, 2010). Overall, this may also apply to gastropods and particularly to snail species that express both Hc and a Cu-selective MT isoform (Dallinger et al., 2005). Yet, information about possible interactions between metal accumulation (Cd^{2+}, Cu^{+}/Cu^{2+}) and CuMT or Hc expression in snails is still poor.

The edible snail *Cornu aspersum* (the escargot) is one of the most common helicid land snails in Western and Central Europe. This species possesses CdMT and CuMT isoforms involved in metal-selective physiological tasks such as Cd detoxification and homeostatic Cu regulation (Höckner et al., 2011; Palacios et al., 2011). In addition, *C. aspersum* expresses three different functional Hc isoforms (CaH αD, CaH αN, CaH β) (Schäfer et al., 2019). An involvement of CuMT in Hc synthesis of this species can therefore not be excluded. In the present study, we have examined the effect of Cd^{2+} and Cu^{2+} exposure on metal accumulation in the midgut gland and mantle of *C. aspersum*, along with the impact of these metals on transcriptional upregulation of *CdMT*, *CuMT*, and the three *Hc* genes in the two organs. Interestingly, the genes of CuMT and of CaH αD exhibit an organ-specific expression pattern in the midgut gland of Cu-exposed snails. These results are discussed in view of possible physiological interrelationships between the metal-selective activity of snail MT isoforms and the synthesis and metabolism of Hc isoforms.

2 | MATERIALS AND METHODS

2.1 | Animal rearing and metal exposure

Specimens of *C. aspersum* (garden snail) were obtained from a commercial dealer (Wiener Schneckenmanufaktur e.U., Vienna, Austria). Before exposure, snails were kept for several weeks in groups on garden soil complemented with lime powder ($CaCO_3$) under constant conditions (12:12 h photoperiod, 18°C ambient temperature) for acclimatization. They were fed with lettuce (*Lactuca sativa*) ad libitum. For exposure experiments, 30 snails were kept individually in octagonal plastic boxes (diameter 12 cm; height 6 cm) for 10 days under the same conditions (see also Pedrini-Martha et al., 2016). Lettuce was metal-enriched by soaking salad leaves for 1 h in Titrisol standard dilutions (Merck) of 2 mg/L $CdCl_2$ (actual concentration: 130.12 ± 57.43 µg/g dry weight [d.w.] Cd) or 10 mg/L $CuCl_2$ (actual concentration: 285.18 ± 77.28 µg/g d.w Cu), followed by 30 min of draining. Ten snails per treatment group were provided with metal-enriched lettuce whereas controls were fed with untreated lettuce (Cd: 0.72 ± 0.31 µg/g d.w.; Cu: 9.68 ± 2.10 µg/g d.w.), every second day. Feeding behavior was monitored and documented before every new food supply. At the end of the exposure period, eight snails of each group (controls, Cd, Cu) were killed and dissected on an ice-cooled aluminum plate, intermittently rinsed with RNase Away® Reagent (Ambion by Life Technologies, Thermo Fisher Scientific) and tissue samples of mantle edge and midgut gland were isolated. Snails which remained in a dormant condition during exposure and did not consume lettuce were excluded from further processing. Tissue aliquots for RNA isolation were stored in RNALater™ (Thermo Fisher Scientific) at −80°C until further processing. Tissue aliquots for metal analysis as well as leaves of *L. sativa* were immediately processed (see Section 2.2 and Pedrini-Martha et al., 2016).

2.2 | Metal measurement

All tissue ($n = 8$) and lettuce samples ($n = 4$) were oven-dried at 60°C in 12 ml screw-capped polyethylene rubes (Greiner). After d.w. determination, a 1:1 mixture of deionized water and nitric acid (65%) (Suprapur; Merck) was added and samples were heat-digested at 70°C through several days until a clear solution was obtained. Samples were diluted with deionized water to a final volume of 11.5 ml. Cd and Cu concentrations of tissue and lettuce samples

were analyzed by flame atomic absorption spectrophotometry (model 2380; Perkin Elmer). Samples with concentrations below the detection limits (e.g., Cd concentrations in mantle tissue) were re-measured by graphite furnace analysis (Z-8200 Polarized Zeeman Atomic Absorption Spectrophotometer with SSC-300 Auto Sampler; Hitachi). Instruments were calibrated with appropriate Cd and Cu solutions prepared from a Titrisol Cd or Cu standard solution (1000 mg/L; Merck). TORT-2 Lobster Midgut gland (NRC) and Polish Virginia Tobacco leaves (INCT-PVTL-6; INCT) were used as standard reference materials for validation of metal measurement accuracy. Metal concentrations of standard reference materials were found to be within the accepted range (±10%) of certified metal values.

2.3 | RNA extraction and complementary DNA (cDNA) synthesis

Total RNA extractions from six individuals of each exposure group were performed applying the RNeasy Plant Mini Kit (Qiagen) with On-Column DNA digestion (Qiagen) according to the manufacturer's instruction with following modifications: For tissue homogenization, samples were homogenized with a Precellys 24 ball mill (Bertin Corp.) in 450 µl Buffer RLT+ 4.5 µl 2-mercaptoethanol (Merck). Integrity of isolated RNA was verified by gel electrophoresis and quantified using the Quant-iT™ RiboGreen® RNA assay kit (Invitrogen, Thermo Fisher Scientific). For cDNA synthesis 200 ng total RNA were applied using the SuperScirpt™ IV First-Strand Synthesis System (Invitrogen, Thermo Fisher Scientific) following the manufacturer's instructions.

2.4 | Mapping of RNAseq data against known cDNA sequences of Hcs and MTs of *C. aspersum*

Total RNA from midgut gland of two controls and two Cu-exposed snails were sent to StarSeq for new generation sequencing (NGS) as described previously (Schäfer et al., 2019). Bioinformatic analyses of RNA sequences were performed using Geneious 9.1.8 (Kearse et al. 2012). Sequencing adapters of transcriptomic NGS data were removed and raw reads were quality-trimmed. Trimmed reads of all four datasets were then mapped together to the three known Hc sequences from *C. aspersum* (CaH αD: MH485355, CaH αN: MH485356, CaH β: MH485357), to detect single nucleotide polymorphisms (SNPs). For each Hc isoform 20 sequence sections with a length of 45 nucleotides each that do not contain any SNPs were isolated as references for the quantitative analysis. The sequence sections were distributed over the complete coding regions but they were abundant at the same position with respect to the coding sequences of all *hemocyanin* genes to enhance comparability. For MT sequences, the CuMT (EF178297) and the CdMT (EF152281) were chosen as references for a prescreening to identify the allelic sequence variations for each individual to use them as references for the quantitative analysis. Trimmed reads of NGS datasets from different individuals were mapped separately from each other to these reference sequences of Hcs and MTs to obtain their relative quantity (given as transcripts per million) within the transcriptome of cupper fed individuals and those of the control group. Minimum coverage was set to 45 nucleotides and overlap identity to 100%.

2.5 | Quantification of *MT* and *Hc* gene transcription via quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Gene-specific primers of the *MT* (Höckner et al., 2011) and *Hc* genes for qRT-PCR were designed using Primer Express 3.0 software (Applied Biosystems by Thermo Fisher Scientific; see Table 1). Primer dissociation curves were used to determine the optimal primer concentration (see Table 1). Calibration curves were generated by using cleaned PCR products for *Hc* genes (Qiagen PCR purification kit; Qiagen) or amplicon plasmids for *MT* genes. Cycle quantification (Cq) values were estimated as followed: CaH αN: $y = -3.2929x + 32.692$; CaH αD: $y = -3.4732x + 35.011$; CaH β: $y = -3.5882x + 37.551$; CdMT: $y = -3.2775x + 37.382$; CuMT: $y = -3.3551x + 36.058$. *MT* and *Hc* gene

TABLE 1 Characterization of gene-specific primers used for quantitative reverse transcription polymerase chain reaction

Primer	Sequence 5′–3′	Length (bp)	Conc. (nm)	Primer efficiency	Amplicon length (bp)
CdMT S	GCCGCCTGTAAGACTTGCA	19	900	101.89	56
CdMT AS	CACGCCTTGCCACACTTG	18	900		
CuMT S	AACAGCAACCCTTGCAACTGT	21	900	98.63	73
CuMT AS	CGAGCACTGCATTGATCACAA	21	900		
CaH αD S	CCCTGTCAGCAAGGACAATACC	22	900	94.05	62
CaH αD AS	CAATGCGGGTGCCTTTCTT	19	900		
CaH αN S	GCCCTGGTCCAATGAGATTCT	21	900	101.23	60
CaH αN AS	CCAGCTTGTCGGACTGCAT	19	900		
CaH β S	ATCCCAATTGGTGCTGAGAAA	21	900	89.97	60
CaH β AS	CGTGTCCTTGGGCACAATG	19	900		

Note: Sequences, lengths and primer concentrations are listed. Values for primer efficiency and amplicon length are also reported.

expression was quantified with the QuantStudio™ 3 (Applied Biosystems, Thermo Fisher Scientific) in a 10 µl approach applying the Power SYBR® Green PCR Master Mix (Applied Biosystems, Thermo Fisher Scientific). The respective transcripts were amplified using the following protocol: one initial denaturation step for 10 min at 95°C, 40 cycles denaturation for 15 s at 95°C and annealing/extension for 1 min at 60°C.

2.6 | Statistical analysis

For statistical analysis and graphical drawings, the software package GraphPad Prism (Version 6.01; GraphPad Software Inc.) was used. Significant outliers were removed from the data set (Grubb's test, $p < .05$). Data were tested for normality (Shapiro Wilks test). For normal-distributed data, one-way analysis of variance statistics were applied. For intergroup comparison of all treatments, a Holm-Sidak's multiple comparison test was performed. For comparison of control values with those of Cu- and Cd-treated snails within a respective data set (e.g., *CdMT* gene expression of untreated and metal-exposed snails in the midgut gland), a Dunnett's multiple comparison test was performed. For non-normal-distributed data, a Kruskal–Wallis test and the posthoc test Dunn's multiple comparison were applied. Significance level was set at $p < .05$.

3 | RESULTS

3.1 | Tissue-specific accumulation of Cd and Cu by *C. aspersum*

Adult snails of *C. aspersum* accumulated Cd and Cu in a tissue-specific manner (Figure 1). Cd concentrations in the midgut gland tissue of exposed animals (181.46 ± 47.47 µg/g d.w.) were significantly higher compared to unexposed individuals (10.78 ± 3.68 µg/g d.w; Figure 1a and Table 2). Generally, Cd tissue concentrations in the mantle edge were significantly lower than midgut gland values, even after Cd exposure (Figure 1a and Table 2).

In untreated snails, Cu concentrations in the mantle edge were considerably higher compared to those in the midgut gland (Figure 1b and Table 2). However, after Cu exposure, Cu concentrations of both tissue types were in an equal range (Table 2). Even though Cu was accumulated in both tissues types after Cu exposure, Cd showed a clearly higher bioaccumulation capacity than Cu, with bioaccumulation factor values of 17 and 23 in the midgut gland and mantle of Cd-exposed snails, respectively (Figure 1b and Table 2). Cu tissue concentrations in Cd-exposed animals and Cd tissue levels in Cu-treated snails did not change in comparison to the respective concentration values of control individuals (Figure 1 and Table 2). Data also revealed that Cu tissue concentrations were more variable in the midgut gland compared to the respective Cd values (Table 2).

3.2 | Basal and metal-dependent *MT* gene expression in *C. aspersum*

The basal transcription of both *MT* genes showed an isoform- and tissue-specific expression pattern in untreated snails (Figure 2).

Whereas the *CdMT* gene was markedly higher expressed in the midgut gland than in the mantle (46-fold), the *CuMT* transcripts appeared to be more abundant in the mantle, even though no significance could be detected. In addition, the *CdMT* gene in the midgut gland was 10-fold higher expressed than the *CuMT* gene, whereas the *CuMT* gene expression in the mantle seemed to be more pronounced in comparison to the *CdMT* isoform (Figure 2). Upon 10 days of Cd exposure, *CdMT* gene expression was significantly upregulated by a 12-fold increase in the midgut gland and an eightfold induction in the mantle tissue, when compared to controls (Figure 3).

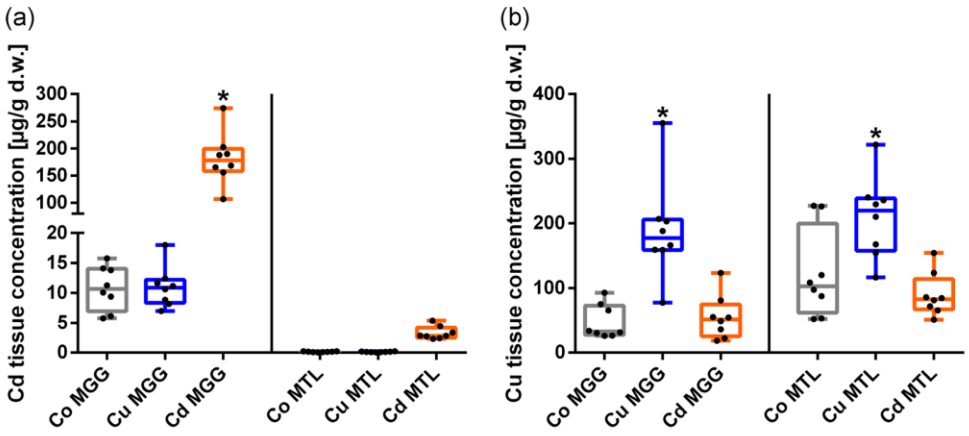


FIGURE 1 Metal accumulation in the midgut gland and mantle in *Cornu aspersum*. Whiskar box plots for Cd (a) and Cu (b) concentrations in the MGG and MTL ($n = 8$ each) of controls (Co; gray boxes), Cu-exposed (blue boxes) and Cd-treated snails (orange boxes) are shown. The respective boxes extend from the 25th to the 75th percentile, with square lines showing the medians. Single values of each group are represented by single black dots. *Indicate significance ($p < .05$) compared to control values (Holm-Sidak's multiple comparison). MGG, midgut gland; MTL, mantle tissue [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Accumulation values for Cd and Cu in the midgut gland and mantle of *Cornu aspersum*

Tissue	Group	Cd (µg/g)	SD (%)	BAF	Cu (µg/g)	SD (%)	BAF
Midgut gland	Controls	10.78	34		47.83	55	
	Cu-exp.	10.98	31	1	189.32	41	4
	Cd-exp.	181.46	26	17	54.95	62	1
Mantle	Controls	0.14	47		121.51	57	
	Cu-exp.	0.13	48	1	209.37	30	2
	Cd-exp.	3.30	32	23	89.76	37	1

Note: Mean Cd and Cu tissue concentrations (µg/g d.w.) ($n = 8$) for controls and metal-exposed snails are shown. For comparison of the variability of single values within each group, the respective standard deviation values are expressed in percent (*SD* [%]). In addition, the BAF reflects the metal accumulation in soft tissues of metal-treated snails referred to controls.

Abbreviation: BAF, bioaccumulation factor.

In contrast, after Cu exposure the *CuMT* gene transcription in the midgut gland was upregulated only threefold (Figure 3a), whereas no significant increase could be detected in the mantle, even though the visual pattern may suggest a stimulating tendency of Cu exposure on *CuMT* gene transcription (Figure 3b). Overall, gene expression data suggested that only exposure to the cognate metal that is bound by the expressed metal-selective MT isoforms can stimulate their respective mRNA transcription. In other words, only Cd exposure had a significant impact on *CdMT* gene expression, while Cu exposure lead to a significant increase of only *CuMT* gene transcription (Figure 3).

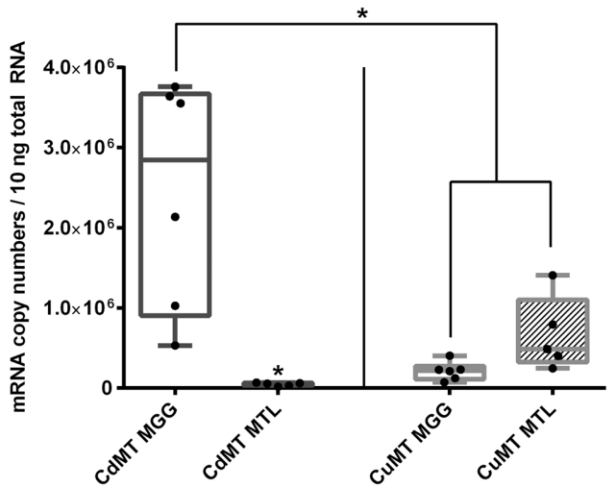


FIGURE 2 Basal expression of *MT* genes in the midgut gland and mantle of unexposed individuals of *Cornu aspersum*. Whiskar box plots show *CdMT* (left hand side) and *CuMT* (right hand side) gene transcription of untreated snails in MGG ($n = 6$) and MTL ($n = 5$). The respective boxes extend from the 25th to the 75th percentile, with square lines representing the medians. Single values are displayed as single black dots. For statistical analysis a Dunn's multiple comparison test was applied (see Section 2.6). *Indicate significance ($p < .05$). MGG, midgut gland; MTL, mantle tissue

3.3 | Basal and metal-dependent *hemocyanin* gene expression in *C. aspersum*

All three *Hc* genes, namely *CaH αD*, *αN*, and *β*, showed varying basal transcription rates in untreated snails. Whereas the genes of *CaH αD* and *αN* exhibited nearly identical basal transcription levels in the midgut gland and slightly different expression rates in the mantle, *CaH β* transcripts were most abundant in both tissue types. *CaH β* exhibits markedly higher basal gene transcription levels compared to *CaH αD* (with factors of 10 in the midgut gland, and of seven in the mantle) and *CaH αN* (with factors of 13 in the midgut gland, and of 21 in the mantle; Figure 4). Moreover, transcription levels of the *CaH β* gene showed the highest variability range of single values measured in both tissue types (Figure 4).

A tissue-specific view of metal-exposed snails unveils that a significant transcriptional upregulation was observed for the *CaH αD* gene expression after Cu exposure, whereas the transcription rates of the two other genes *CaH αN* and *β* did not change after either kind of metal exposure, compared to their respective control levels in the midgut gland (Figure 5). In the mantle, however, the transcription levels of all three *Hc* genes remained unaffected in Cd- and Cu-treated snails, when compared to their respective control levels (Figure 6).

3.4 | Mapping of RNAseq data against *hemocyanin* and *MT* reference gene sequences

Trimmed reads of RNA sequences from midgut gland of control and Cu-exposed snails ($n = 2$ each) were mapped against reference sequences of *CaH* genes *αD*, *αN*, and *β*, and the two *MT* genes *CdMT* and *CuMT*, to assess their relative quantity within the respective transcriptomes. As seen in Table 3, the transcriptomic mapping data seem to confirm the results derived from mRNA quantification by qRT-PCR (Figures 3 and 6), showing a clear trend of higher transcript abundances for the gene *CaH αD* and the *CuMT* gene in the midgut gland of Cu-exposed snails versus controls. In spite of the fact that

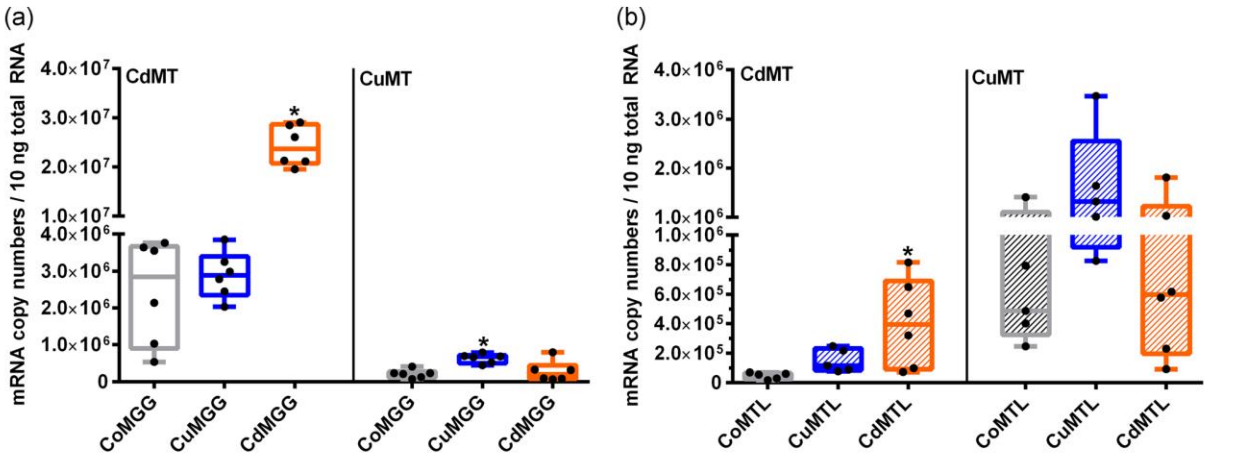


FIGURE 3 Metal-dependent expression of *MT* genes in the midgut gland (a) and mantle (b) of *Cornu aspersum*. Whiskar box plots show quantification of *CdMT* and *CuMT* gene transcription in controls (Co) (gray boxes), Cu-exposed (blue boxes) and Cd-treated snails (orange boxes) (n = 5–6 per treatment group). The respective boxes extend from the 25th to the 75th percentile, with square lines showing the medians. Single values of each group are reported as single black dots. To analyze the impact of metal exposure on *CdMT* and *CuMT* gene expression the two tissue types were compared separately. For *CdMT* gene expression in the mantle and in the midgut gland, and for *CuMT* gene expression in the mantle, a Dunn's multiple comparison test was applied. For *CuMT* gene expression in the midgut gland, a Dunnett's multiple comparison test was applied (see Section 2.6). *Indicate significance (p < .05) within the respective data set [Color figure can be viewed at wileyonlinelibrary.com]

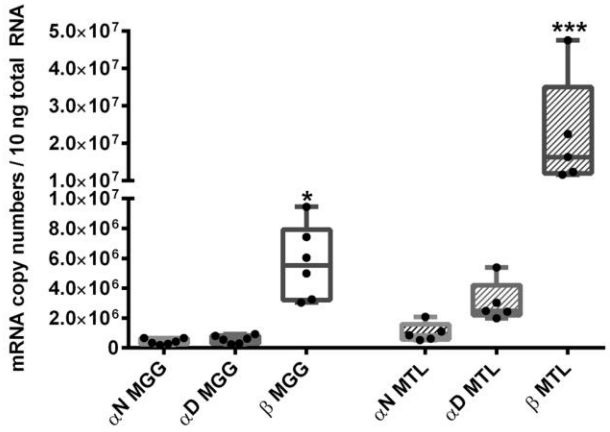


FIGURE 4 Basal expression of Hc genes in the midgut gland and mantle of unexposed individuals of *Cornu aspersum*. Whiskar box plots show the gene transcription of the three *Hc* genes *CaH* *αN*, *αD*, and *β* in the MGG (n = 6) (empty bars) and MTL (n = 5) (hatched bars) of untreated snails. The respective boxes extend from the 25th to the 75th percentile, with square lines showing the medians. Single values of each group are represented by single black dots. Data were not normal distributed, so a Dunn's multiple comparison test was applied (see Section 2.6). *Indicate significance (p < .05), with *meaning a significant difference between *β* MGG and *αN* MGG, and ***indicating a significant difference between gene transcription of *β* MTL compared to *αN* MGG, *αD* MGG and *αN* MTL. MGG, midgut gland; MTL, mantle tissue

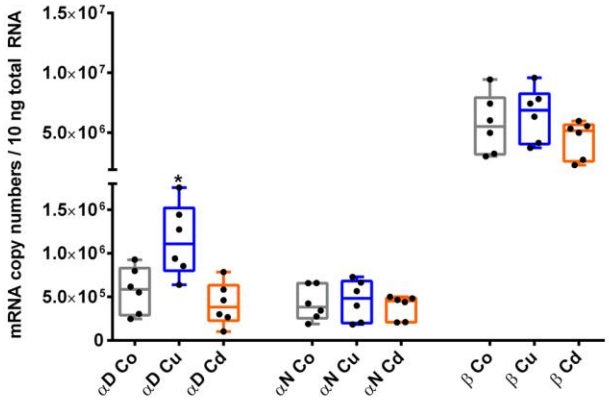


FIGURE 5 Metal-dependent expression of *Hc* genes in the midgut gland of *Cornu aspersum*. Whiskar box plots show gene transcription patterns of the three *Hc* genes *CaH* *αD*, *αN*, and *β* in controls (Co) (gray boxes), Cu-exposed (blue boxes) and Cd-treated snails (orange boxes) (n = 5–6 per treatment group). The boxes extend from the 25th to the 75th percentile, with square lines showing the medians. Single values of each group are shown as single black dots. Data for *αD* and *αN* gene expression were analyzed by the Dunnett's multiple comparison test (data normal-distributed), whereas metal-dependent gene expression of the *Hc* *β* gene was analyzed with a Dunn's multiple comparison procedure (data not normal-distributed) (see Section 2.6). *Indicate significance (p < .05) [Color figure can be viewed at wileyonlinelibrary.com]

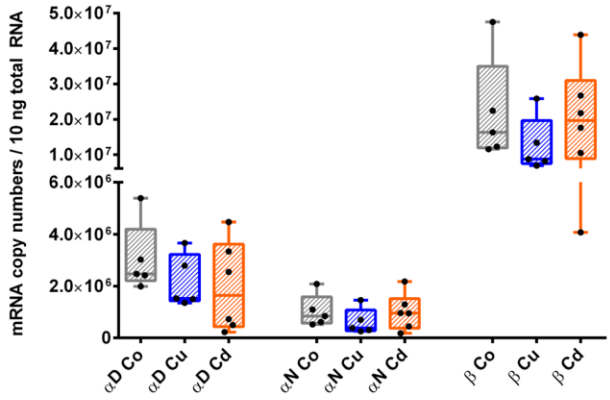


FIGURE 6 Metal-dependent expression of *Hc* genes in the mantle of *Cornu aspersum*. Whiskar box plots show gene transcription patterns of the three *Hc* genes *CaH αD*, *αN*, and *ß* in controls (Co) (gray boxes), Cu-exposed (blue boxes) and Cd-treated snails (orange boxes) (n = 5–6 per treatment group). The boxes extend from the 25th to the 75th percentile, with square lines showing the medians. Single values of each group are shown as single black dots. All data of the respective *Hc* genes were analyzed using the Dunnett's multiple comparison test (see Section 2.6). *Indicate significance (p < .05) [Color figure can be viewed at wileyonlinelibrary.com]

these results must be considered with caution because of the small sample size (n = 2 of each treatment group), they seem to confirm the much more robust real-time PCR data.

4 | DISCUSSION

4.1 | Organ-specific metal accumulation: Cd detoxification versus Cu regulation

Although in the present study only the midgut gland and the mantle of *C. aspersum* were considered, the observed patterns of Cd and Cu accumulation in these organs (see Figure 1 and Table 2) are consistent with previous findings in terrestrial helicid snails. Whereas Cd accumulates nearly exclusively in the midgut gland (Dallinger & Wieser, 1984; Dallinger et al., 1989), Cu is distributed more evenly among several snail tissues (Dallinger, 1993; Hispard et al., 2008). Elevated levels of this metal are normally found in the midgut gland, the mantle, the foot, and the kidney, especially after Cu exposure (Berger & Dallinger, 1989; Boshoff et al., 2013; Nowakowska et al., 2012). Apart from metal- and tissue-specific accumulation patterns, an important difference between Cd and Cu exists with respect to their kinetic and metabolic behavior. Although at low environmental Cd concentrations some helicid snail species are able to excrete the metal through excretion via their mucus and feces (Notten et al., 2006), higher Cd levels in the substrate lead to a strong accumulation of the metal in the midgut gland due to a high efficiency of metal absorption versus low excretion rates (Berger & Dallinger, 1989). This eventually results in a progressive increase of metal levels and a persistent Cd storage in this organ on a long-term scale (Dallinger & Wieser, 1984; Williamson, 1980). In contrast to Cd, Cu accumulation in snail organs occurs at lower levels and in a fluctuating manner, increasing during phases of Cu exposure, and decreasing down to basal levels after discontinuation of exposure (Dallinger & Wieser, 1984). These two different behavior patterns have been explained by storage detoxification for the nonessential metal Cd, and by physiological regulation for the essential trace element Cu (Dallinger, 1993; Dallinger et al., 2005). Many studies have shown that low external Cd concentrations can have a significant impact on internal tissue concentrations of essential metals like Zn or Cu, including plants (Borowska et al., 2017; Webster et al., 2011), vertebrates (Aydin et al., 2001; Kamunde & MacPhail, 2011; Pelgrom et al., 1995) and even snails (Nica et al., 2019). Although Nica et al. (2019) found an increase of internal Cu concentrations in the midgut gland of Cd-exposed *C. aspersum*, this effect could not be observed in the present study, probably due to the higher metal concentrations applied in our exposure experiments. Notably at moderately elevated and high concentrations, the pathways of Cd and Cu in helicid snails are largely controlled by two metal-selective MT isoforms, one detoxifying Cd in the midgut gland (called CdMT), the second one being associated with Cu (called CuMT) which is exclusively expressed in rhogocyte cells (Chabicovsky et al., 2003; Dallinger et al., 1997). As mentioned above, these cells can be scattered through connective tissues of all main snail organs (Haszprunar, 1996; Stewart et al., 2014). Consequently, the metabolic pathways of Cd and Cu in snail tissues remain strictly separated (Dallinger, 1996). This ensures that the molecular interactions between the two metals in snail cells and tissues can be minimized, a circumstance that might be important in view of the crucial role attributed to Cu in connection with Hc synthesis and metabolism (Dallinger et al., 2005).

TABLE 3 *Hc* and *MT* gene expression in controls and Cu-treated snails

	Co1	Co2	Cu1	Cu2
αD	3.6	2.9	11.1	19.0
αN	31.3	31.7	25.7	54.4
β	38.5	47.7	32.8	64.4
CuMT	38.4	12.9	176.9	174.3
CdMT	15.5	21.9	48.9	23.8

Note: Gene expression patterns in midgut gland of control (Co1, Co2) and Cu-exposed snails (Cu1, Cu2) as inferred from transcriptomic mappings against reference sequences for *MT* (*CdMT* and *CuMT*) and Hc genes (*αD*, *αN*, *β*), expressed as gene-specific transcripts per million.

4.2 | Transcriptional upregulation of Cd- and Cu-specific *MT* genes

The two metal-specific *MT* genes (*CdMT* and *CuMT*) of *C. aspersum* exhibit tissue-specific basal transcription patterns with high

expression levels of *CdMT* in the midgut gland and clearly elevated basal transcription rates of *CuMT* in the mantle (Figure 2). This is consistent with the perception that the expressed CdMT protein in the midgut gland plays an important role in Cd detoxification, whereas the elevated *CuMT* expression in the mantle may reflect the presence of the CuMT isoform in the highly abundant rhogocytes within this tissue (Chabicovsky et al., 2003; Sairi et al., 2015).

In accordance with the predicted metal-selective roles of the expressed CdMT and CuMT proteins (Höckner et al., 2011; Palacios et al., 2011), our results revealed also a metal and organ-specific transcriptional upregulation of the *CdMT* and *CuMT* genes in *C. aspersum* (Figure 3). Whereas *CdMT* gene expression was manifold upregulated in the midgut gland and mantle of Cd-exposed individuals, *CuMT* transcription was significantly upregulated only in the midgut gland of Cu-stressed snails. Cd-dependent *CdMT* gene upregulation was so far observed in all studies using terrestrial snail species of the clade of Stylommatophora (Palacios et al., 2011; Pedrini-Martha et al., 2020), including *C. aspersum* (Höckner et al., 2011). In contrast, only transient or no upregulation at all of snail *CuMT* genes can be detected after Cu exposure (Höckner et al., 2011; Palacios et al., 2011). This suggests that *CuMT* gene upregulation in snails may occur at low and not always detectable rates, depending perhaps on its transient variability and on sample-specific peculiarities such as the abundance of rhogocytes in the respective tissue preparations. On the other hand, our results confirm that the CuMT in snail rhogocytes may predominantly serve homeostatic regulation of intracellular Cu levels (Dallinger et al., 2005) with only transient upregulation peaks during phases of intracellular Cu excess. Rhogocytes are one of the most important cell types for Cu metabolism in gastropods for several reasons. First, they are the main sites of CuMT expression (Chabicovsky et al., 2003; Dallinger et al., 2005). Secondly, excessive amounts of Cu are detoxified in rhogocytes by storage compartmentalization into so-called granules which are quickly formed upon excessive Cu exposure (Dallinger et al., 2005). Thirdly, rhogocytes are the main sites for synthesis of the gastropod respiratory proteins, namely hemoglobin or Hc, as demonstrated for *Haliotis tuberculata* (Albrecht et al., 2001), *Megathura crenulata* (Martin et al., 2011), *Lymnaea stagnalis* (Kokkinopoulou et al., 2015), or *Biomphalaria glabrata* (Kokkinopoulou et al., 2014). It was, therefore, hypothesized that the CuMT in rhogocytes may act as a donor or acceptor of Cu^{+} ions during Hc synthesis or degradation (Dallinger et al., 2005). Alternatively, it can be speculated that CuMT may function as a Cu transporter involved in the generation of rhogocyte Cu-granules.

4.3 | Tissue-specific *hemocyanin* gene expression and its response to metal stress

Of the three *Hc* genes (*αD*, *αN*, and *β*) identified in *C. aspersum*, *CaH β* exhibits the highest basal transcription levels in both, midgut gland and mantle tissues (Figure 4). Overall, the expression of all three *Hc* genes appears to be higher in the mantle edge compared to the midgut gland. This is perhaps because the Hc demand of the mantle tissue may be particularly high, considering its involvement in oxygen uptake in pulmonate land snails. Cell and tissue-specific as well as developmental expression patterns of Hc isoform genes have been observed in other gastropod species too, including *Haliotis asinina* (Streit et al., 2005) and *H. laevigata* (Sairi et al., 2015). Multiplication of *Hc* genes with differential expression patterns may have fostered the adaptation capacity of gastropods to novel habitats during evolution, extending their repertoire to respond to a variety of physiological and environmental conditions (Schäfer et al., 2019). Also, in nongastropod molluscs such as cephalopods, multiple *Hc* genes are expressed that probably possess different oxygen affinities under varying environmental conditions (Melzner et al., 2007).

Cd exposure, especially over a long-term range, can induce an increased oxygen demand in exposed molluscs such as marine snails (Dalla Via et al., 1989) and oysters (Lannig et al., 2008). A possible response strategy to compensate for this higher oxygen demand would be an increased expression of the oxygen-carrying protein Hc. In the present study, however, the transcription of Hc genes in Cd-exposed *C. aspersum* was not upregulated at all. This may be the case in other mollusc species, too, which makes them particularly vulnerable to increased temperatures owing to climatic changes, especially in combination with environmental metal pollution (Lannig et al., 2008).

All the more so surprising was the observation that one of the *Hc* genes, specifically *CaH αD*, was upregulated in the midgut gland of Cu-exposed snails (Figure 6 and Table 3). Since rhogocytes have been shown to be the only cells that synthesize Hc in the midgut gland of gastropods (Albrecht et al., 2001; Martin et al., 2011; Streit et al., 2005), it is highly probable that upregulation of *CaH αD* occurs predominantly in rhogocytes. Apart from Cu regulation and Hc synthesis (Dallinger et al., 2005), these cells seem to also possess their own metal detoxification mechanisms which rely on intracellular compartmentalization of toxic metals within granular vesicles (Dallinger et al., 2005; Kokkinopoulou et al., 2015; Simkiss & Mason, 1983), rather than on metal complexation by CdMT, that seems to prevail in snail tissues outside of rhogocytes (Chabicovsky et al., 2003). Therefore, compartmentalization of toxic metals (including Cd) within granular rhogocyte vesicles represents an alternative detoxification pathway when Cd-specific MTs are not present. Hence, we assume that mollusc rhogocytes are, in addition to their specific functions in Hc synthesis, highly capable of metal stress resistance (Kokkinopoulou et al., 2014, 2015). This ability may be crucial to ensure nonimpairment of Cu pathways during Hc synthesis. Hence, the observed upregulation of the Hc isoform *CaH αD* in the midgut gland of Cu-treated snails may reflect a role of the isoform in stress resistance, rather than indicating a direct connection between an increased Cu supply and Hc synthesis. In fact, intraspecific multiplicity of Hc isoforms may translate into functional specification of single isoforms, including their involvement in stress resistance like innate immunity, as well as antiviral or antibacterial activity (Wu et al., 2016; Yao et al., 2019). In the cuttlefish *Sepia officinalis*, several *Hc* genes are expressed differentially depending on

the developmental stage and on adaptation to external stressors such as temperature of hypercapnia (Strobel et al., 2012). Also, in *Sepiella maindroni* Hc was upregulated after hypoxia or bacterial challenge (Li et al., 2017).

4.4 | CuMT expression and Hc metabolism: A direct connection?

In a previous study with the Roman snail, *Helix pomatia*, it has been shown that the essential trace element Cu is strictly regulated in rhogocytes by diverting excess amounts of the metal into an intracellular granular pool, whereas the level of the Cu associated with the expressed CuMT protein remained fairly stable (Dallinger et al., 2005). The selective Cu binding by CuMT, its Cu-buffering function, and its coexpression with a respiratory Cu-protein in snail rhogocytes gave rise to the hypothesis that in these animals CuMT may be involved in Hc synthesis (Dallinger et al., 1997), serving perhaps as a Cu donator to nascent Hc molecules (Dallinger et al., 2000; Dallinger et al., 2005). Such a hypothesis is supported by findings that in arthropods and particularly in crustaceans, too, Cu-specific MT isoforms can deliver Cu to arthropod Hcs (Brouwer et al., 1986; Brouwer et al., 2002), which in turn may also be involved in moulting and metal stress resistance (Engel & Brouwer, 1987; Engel et al., 2001).

The simultaneous upregulation of *CuMT* and Hc *CaH αD* genes in the midgut gland of Cu-exposed snails as shown in the present study (Figures 3 and 6) is not a convincing argument to support the hypothesis of a Cu donation process through presumed interactions between CuMT and Hc molecules. Rather, it is assumed that the transcriptional upregulation of both genes might be a response to a stressful situation induced by Cu exposure.

5 | CONCLUSIONS

The tissue specific accumulation pattern of Cd and Cu as well as the metal-selective upregulation of the respective *Cd* or *CuMT* genes in *C. aspersum* reflects separated metabolic and detoxification pathways of these two metals within helicid snails. Whereas Cd is detoxified by binding to the CdMT protein, excessive amounts of Cu are scavenged in Cu granule within rhogocytes. In this study, gene transcription of all three *Hc* genes (*CaH αD*, *CaH αN*, *CaH β*) in *C. aspersum* was quantified for the first time, showing tissue and isoform specific expression patterns. Interestingly, *CaH αD* gene expression is moderately upregulated due to Cu-exposure. However, a possible role in Cu detoxification still has to be elucidated.

ACKNOWLEDGEMENTS

This study was supported by a cooperative research project granted by the Austrian Science Foundation (FWF), project number I 3032-B21 to R.D. and the German Research Foundation (DFG), project number LI998/15-1 to L.B.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

DATA AVAILABILITY STATEMENT

The raw data for this study can be provided upon request by the corresponding authors through the server of the University of Innsbruck.

ORCID

Reinhard Dallinger https://orcid.org/0000-0001-6084-4895

REFERENCES

Albrecht, U., Keller, H., Gebauer, W., & Markl, J. (2001). Rhogocytes (pore cells) as the site of hemocyanin biosynthesis in the marine gastropod Haliotis tuberculata. *Cell and Tissue Research*, *304*, 455–462. https://doi.org/10.1007/s004410100368

Alyakrinskaya, I. O. (2002). Physiological and biochemical adaptations to respiration of hemoglobin-containing hydrobionts. *Biological Bulletin*, *29*, 269–283. https://doi.org/10.1023/A:1015438615417

Aydin, H. H., Coker, C., & Ersöz, B. (2001). In vivo interaction between cadmium and essential trace elements copper and zinc in rats. *Turkish Journal of Medical Sciences*, *31*, 127–129.

Berger, B., & Dallinger, R. (1989). Accumulation of cadmium and copper by the terrestrial snail *Arianta arbustorum* L.: Kinetics and budgets. *Oecologia*, *79*, 60–65. https://doi.org/10.1007/BF00378240

Berger, B., Dallinger, R., Gehrig, P., & Hunziker, P. E. (1997). Primary structure of a copper-binding metallothionein from mantle tissue of the terrestrial gastropod *Helix pomatia* L. *Journal of Biochemistry*, *328*, 219–224.

Borowska, S., Brzóska, M. M., Gałażyn-Sidorczuk, M., & Rogalska, J. (2017). Effect of an extract from aronia melanocarpa l. Berries on the body status of zinc and copper under chronic exposure to cadmium: An in vivo experimental study. *Nutrients*, *9*, 1374. https://doi.org/10.3390/nu9121374

Boshoff, M., Jordaens, K., Backeljau, T., Lettens, S., Tack, F., Vandecasteele, B., De Jonge, M., & Bervoets, L. (2013). Organ- and species-specific accumulation of metals in two land snail species (Gastropoda, Pulmonata. *Science of the Total Environment*, *449*, 470–481. https://doi.org/10.1016/j.scitotenv.2013.02.003

Brouwer, M., Syring, R., & Brouwer, T. H. (2002). Role of a copper-specific metallothionein of the blue crab, *Callinectes sapidus*, in copper metabolism associated with degradation and synthesis of hemocyanin. *Journal of Inorganic Biochemistry*, *88*, 228–239.

Brouwer, M., Whaling, P., & Engelt, D. W. (1986). Copper-Metallothioneins in the American lobster, Ho arus americanus: Potential role as Cu(I) donors to apohemocyanin. *Environmental Health Perspectives*, *65*, 93–100.

Chabicovsky, M., Niederstätter, H., Thaler, R., Hödl, E., Parson, W., Rossmanith, W., & Dallinger, R. (2003). Localization and quantification of Cd- and Cu-specific metallothionein isoform mRNA in cells and organs of the terrestrial gastropod *Helix pomatia*. *Toxicology and Applied Pharmacology*, *190*, 25–36. https://doi.org/10.1016/S0041-008X(03)00148-0

Dalla Via, G. J., Dallinger, R., & Carpen, E. (1989). Effects of cadmium on *Murex trunculus* from the adriatic sea. II. Oxygen consumption and acclimation effects. *Archives of Environmental Contamination and Toxicology*, *18*, 562–567.

Dallinger, R. (1993). Strategies of metal detoxification in terrestrial invertebrates. In R. Dallinger, & P. S. Rainbow (Eds.), *Ecotoxicology of metals in invertebrates* (pp. 245–289). CRC Press, Lewis Publishers.

Dallinger, R. (1996). Metallothionein research in terrestrial invertebrates: Synopsis and perspectives. *Comparative Biochemistry and Physiology*

Part C: Pharmacology, Toxicology and Endocrinology, *113*(2), 125–133. https://doi.org/10.1016/0742-8413(95)02078-0

Dallinger, R., Berger, B., & Bauer-Hilty, A. (1989). Purification of cadmium-binding proteins from related species of terrestrial helicidae (gastropoda, mollusca): A comparative study. *Molecular and Cellular Biochemistry*, *85*, 135–145. https://doi.org/10.1007/BF00577109

Dallinger, R., Berger, B., Gruber, C., Hunziker, P., & Stürzenbaum, S. (2000). Metallothioneins in terrestrial invertebrates: Structural aspects, biological significance and implications for their use as biomarkers. *Cell Mol. Biol.*

Dallinger, R., Berger, B., Hunziker, P., & Kägi, H. R. J. (1997). Metallothionein in snail Cd and Cu metabolism. *Nature*, *143*, 831–833. https://doi.org/10.1038/40785

Dallinger, R., Chabicovsky, M., & Berger, B. (2004). Isoform-specific quantification of metallothionein in the terrestrial gastropod Helix pomatia. I. Molecular, biochemical, and methodical background. *Environmental Toxicology and Chemistry*, *23*, 890–901. https://doi.org/10.1897/03-100

Dallinger, R., Chabicovsky, M., Hödl, E., Prem, C., Hunziker, P., & Manzl, C. (2005). Copper in Helix pomatia (Gastropoda) is regulated by one single cell type: Differently responsive metal pools in rhogocytes. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology*, *289*, R1185–R1195. https://doi.org/10.1152/ajpregu.00052.2005

Dallinger, R., & Wieser, W. (1984). Patterns of accumulation, distribution and liberation of Zn, Cu, Cd and Pb in different organs of the land snail Helix pomatia L. *Comparative Biochemistry and Physiology*, *79C*, 117–124.

Dallinger, R., Zerbe, O., Baumann, C., Egger, B., Capdevila, M., Palacios, Ò., Albalat, R., Calatayud, S., Ladurner, P., Schlick-Steiner, B. C., Steiner, F. M., Pedrini-Martha, V., Lackner, R., Lindner, H., Dvorak, M., Niederwanger, M., Schnegg, R., & Atrian, S. (2020). Metallomics reveals a persisting impact of cadmium on the evolution of metal-selective snail metallothioneins. *Metallomics*, *12*, 702–720. https://doi.org/10.1039/C9MT00259F

Dvorak, M., Schnegg, R., Salvenmoser, W., Palacios, Ò., Lindner, H., Zerbe, O., Hansel, A., Leiminger, M., Steiner, G., Dallinger, R., & Lackner, R. (2019). Distinct pathways for zinc metabolism in the terrestrial slug Arion vulgaris. *Scientific Reports*, *9*, 20089. https://doi.org/10.1038/s41598-019-56577-7

Engel, D. W., & Brouwer, M. (1987). Metal regulation and molting in the blue crab, callinectes sapidus: metallothionein function in metal metabolism. *The Biological Bulletin*, *173*(1), 239–251.

Engel, D. W., Brouwer, M., & Mercaldo-Allen, R. (2001). Effects of molting and environmental factors on trace metal body-burdens and hemocyanin concentrations in the American lobster, *Homarus americanus*. *Marine Environmental Research*, *52*, 257–269. https://doi.org/10.1016/S0141-1136(01)00098-8

Gnatyshyna, L., Falfushynska, H., Stoliar, O., & Dallinger, R. (2020). Preliminary study of multiple stress response reactions in the pond snail *Lymnaea stagnalis* exposed to trace metals and a thiocarbamate fungicide at environmentally relevant concentrations. *Archives of Environmental Contamination and Toxicology*, *79*, 89–100. https://doi.org/10.1007/s00244-020-00728-9

Haszprunar, G. (1996). The molluscan rhogocyte (pore-cell, blasenzelle, cellule nucale), and its significance for ideas on nephridial evolution. *The Journal of Molluscan Studies*, *62*, 185–211. https://doi.org/10.1093/mollus/62.2.185

Hispard, F., Schuler, D., de de Vaufleury, A., Scheifler, R., Badot, P. M., & Dallinger, R. (2008). Metal distribution and metallothionein induction after cadmium exposure in the terrestrial snail helix aspersa. *Environmental Toxicology*, *27*, 1533–1542. https://doi.org/10.1897/07-232.1

Höckner, M., Stefanon, K., De Vaufleury, A., Monteiro, F., Pérez-Rafael, S., Palacios, Ò., Capdevila, M., Atrian, S., & Dallinger, R. (2011). Physiological relevance and contribution to metal balance of specific and non-specific metallothionein isoforms in the garden snail, *Cantareus aspersus*. *BioMetals*, *24*, 1079–1092. https://doi.org/10.1007/s10534-011-9466-x

Kamunde, C., & MacPhail, R. (2011). Metal-metal interactions of dietary cadmium, copper and zinc in rainbow trout, *Oncorhynchus mykiss*. *Ecotoxicology and Environmental Safety*, *74*, 658–667. https://doi.org/10.1016/j.ecoenv.2010.10.016

Kato, S., Matsui, T., Gatsogiannis, C., & Tanaka, Y. (2018). Molluscan hemocyanin: Structure, evolution, and physiology. *Biophysical Reviews*, *10*, 191–202. https://doi.org/10.1007/s12551-017-0349-4

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kokkinopoulou, M., Güler, M. A., Lieb, B., Barbeck, M., Ghanaati, S., & Markl, J. (2014). 3D-ultrastructure, functions and stress responses of gastropod (Biomphalaria glabrata) rhogocytes. *PLOS One*, *9*, e101078. https://doi.org/10.1371/journal.pone.0101078

Kokkinopoulou, M., Spiecker, L., Messerschmidt, C., Barbeck, M., Ghanaati, S., Landfester, K., & Markl, J. (2015). On the ultrastructure and function of rhogocytes from the pond snail Lymnaea stagnalis. *PLOS One*, *10*, 1–23. https://doi.org/10.1371/journal.pone.0141195

Lannig, G., Cherkasov, A. S., Pörtner, H.-O., Bock, C., & Sokolova, I. M. (2008). Cadmium-dependent oxygen limitation affects temperature tolerance in eastern oysters (Crassostrea virginica Gmelin). *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology*, *294*, 1338–1346. https://doi.org/10.1152/ajpregu.00793.2007.-Marine

Li, R., Xu, Z., Mu, C., Song, W., & Wang, C. (2017). Molecular cloning and characterization of a hemocyanin from *Sepiella maindroni*. *Fish and Shellfish Immunology*, *67*, 228–243. https://doi.org/10.1016/j.fsi.2017.06.009

Lieb, B., Dimitrova, K., Kang, H.-S., Braun, S., Gebauer, W., Martin, A., Hanelt, B., Saenz, S. A., Adema, C. M., Rgen Markl, J., & Ayala, F. J. (2006). Red blood with blue-blood ancestry: Intriguing structure of a snail hemoglobin. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(32), 12011–12016.

Lieb, B., Gebauer, W., Gatsogiannis, C., Depoix, F., Hellmann, N., Harasewych, M. G., Strong, E. E., & Markl, J. (2010). Molluscan mega-hemocyanin: An ancient oxygen carrier tuned by a ~550 kDa polypeptide. *Frontiers in Zoology*, *7*, 14. https://doi.org/10.1186/1742-9994-7-14

Marigómez, I., Soto, M., Cajaraville, M. P., Angulo, E., & Giamberini, L. (2002). Cellular and subcellular distribution of metals in molluscs. *Microscopy Research and Technique*, *56*, 358–392. https://doi.org/10.1002/jemt.10040

Martin, A. M., Martin, G. G., Butler, R., & Goffredi, S. K. (2011). Synthesis of keyhole limpet hemocyanin by the rhogocytes of *Megathura crenulata*. *Invertebrate Biology*, *130*, 302–312. https://doi.org/10.1111/j.1744-7410.2011.00249.x

Melzner, F., Mark, F. C., & Prtner, H. O. (2007). Role of blood-oxygen transport in thermal tolerance of the cuttlefish, *Sepia officinalis*. *Integrative and Comparative Biology*, *47*, 645–655. https://doi.org/10.1093/icb/icm074

Moulis, J. M. (2010). Cellular mechanisms of cadmium toxicity related to the homeostasis of essential metals. *BioMetals*, *23*, 877–896. https://doi.org/10.1007/s10534-010-9336-y

Nica, D. V., Draghici, G. A., Andrica, F. M., Popescu, S., Coricovac, D. E., Dehelean, C. A., Gergen, I. I., Kovatsi, L., Coleman, M. D., & Tsatsakis, A. (2019). Short-term effects of very low dose cadmium feeding on copper, manganese and iron homeostasis: A gastropod perspective. *Environmental Toxicology and Pharmacology*, *65*, 9–13. https://doi.org/10.1016/j.etap.2018.10.005

Notten, M. J. M., Oosthoek, A. J. P., Rozema, J., & Aerts, R. (2006). Heavy metal pollution affects consumption and reproduction of the landsnail *Cepaea nemoralis* fed on naturally polluted *Urtica dioica* leaves. *Ecotoxicology*, *15*, 295–304. https://doi.org/10.1007/s10646-006-0059-3

Nowakowska, A., Łaciak, T., & Caputa, M. (2012). Heavy metals accumulation and antioxidant defence system in *Helix pomatia* (Pulmonata: Helicidae). *Molluscan Res*, *32*, 16–20.

Palacios, O., Pagani, A., Pérez-Rafael, S., Egg, M., Höckner, M., Brandstätter, A., Capdevila, M., Atrian, S., & Dallinger, R. (2011). Shaping mechanisms of metal specificity in a family of metazoan metallothioneins: Evolutionary differentiation of mollusc metallothioneins. *BMC Biology*, *9*, 4. https://doi.org/10.1186/1741-7007-9-4

Pedrini-Martha, V., Köll, S., Dvorak, M., & Dallinger, R. (2020). Cadmium uptake, MT gene activation and structure of large-sized multi-domain metallothioneins in the terrestrial door snail *Alinda biplicata* (Gastropoda, Clausiliidae). *International Journal of Molecular Sciences*, *21*, 1631. https://doi.org/10.3390/ijms21051631

Pedrini-Martha, V., Niederwanger, M., Kopp, R., Schnegg, R., & Dallinger, R. (2016). Physiological, diurnal and stress-related variability of cadmium-metallothionein gene expression in land snails. *PLOS One*, *11*, e0150442. https://doi.org/10.1371/journal.pone.0150442

Pelgrom, S. M. G. J., Lamers, L. P. M., Lock, R. A. C., Balm, P. H. M., & Wendelaar Bonga, S. E. (1995). Interactions between copper and cadmium modify metal organ distribution in mature tilapia, *Oreochromis mossambicus*. *Environmental Pollution*, *90*, 415–423.

Sairi, F., Valtchev, P., Gomes, V. G., & Dehghani, F. (2015). Distribution and characterization of rhogocyte cell types in the mantle tissue of *Haliotis laevigata*. *Marine Biotechnology*, *17*, 168–179. https://doi.org/10.1007/s10126-014-9605-9

Schäfer, G. G., Pedrini-Martha, V., Schnegg, R., Dallinger, R., Jackson, D. J., & Lieb, B. (2019). Hemocyanin genes as indicators of habitat shifts in Panpulmonata? *Molecular Phylogenetics and Evolution*, *130*, 99–103. https://doi.org/10.1016/j.ympev.2018.10.014

Simkiss, K., & Mason, A. Z. (1983). Metal ions: Metabolic and toxic effects. In P. W. Hochachka (Ed.), *The mollusca* (pp. 101–164). Academic Press. https://doi.org/10.1016/B978-0-12-751402-4.50011-X

Sminia, T., & Boer, H. H. (1973). Haemocyanin production in pore cells of the freshwater snail *Lymnaea stagnalis*. *Z. Zellforsch*, *145*, 443–445.

Sminia, T., & Vlugt-Van Daalen, L. (1977). Haemocyanin synthesis cell and tissue research in pore cells of the terrestrial snail *Helix aspersa*. *Cell Tiss. Res. 183*.

Stewart, H., Westlake, H. E., & Page, L. R. (2014). Rhogocytes in gastropod larvae: Developmental transformation from protonephridial terminal cells. *Invertebrate Biology*, *133*, 47–63. https://doi.org/10.1111/ivb.12041

Streit, K., Jackson, D., Degnan, B. M., & Lieb, B. (2005). Developmental expression of two *Haliotis asinina* hemocyanin isoforms. *Differentiation*, *73*, 341–349. https://doi.org/10.1111/j.1432-0436.2005.00035.x

Strobel, A., Hu, M. Y. A., Gutowska, M. A., Lieb, B., Lucassen, M., Melzner, F., Pörtner, H. O., & Mark, F. C. (2012). Influence of temperature, hypercapnia, and development on the relative expression of different hemocyanin isoforms in the common cuttlefish *Sepia officinalis*. *Journal of Experimental Zoology. Part A, Ecological Genetics and Physiology*, *317*, 511–523. https://doi.org/10.1002/jez.1743

Webster, R. E., Dean, A. P., & Pittman, J. K. (2011). Cadmium exposure and phosphorus limitation increases metal content in the freshwater alga *Chlamydomonas reinhardtii*. *Environmental Science and Technology*, *45*, 7489–7496. https://doi.org/10.1021/es200814c

Williamson, P. (1980). Variables affecting body burdens of lead, zinc and cadmium in a roadside population of the snail *Cepaea hortensis* MUller. *Oecologia (Berl.)*, *44*, 213–220.

Wu, H., & Wang, W. X. (2010). NMR-based metabolomic studies on the toxicological effects of cadmium and copper on green mussels Perna viridis. *Aquatic Toxicology*, *100*, 339–345. https://doi.org/10.1016/j.aquatox.2010.08.005

Wu, J., Cunningham, A. L., Dehghani, F., & Diefenbach, R. J. (2016). Comparison of *Haliotis rubra* hemocyanin isoforms 1 and 2. *Gene Reports*, *4*, 123–130. https://doi.org/10.1016/j.genrep.2016.04.011

Yao, T., Zhao, M. M., He, J., Han, T., Peng, W., Zhang, H., Wang, J. Y., & Jiang, J. Z. (2019). Gene expression and phenoloxidase activities of hemocyanin isoforms in response to pathogen infections in abalone *Haliotis diversicolor*. *International Journal of Biological Macromolecules*, *129*, 538–551. https://doi.org/10.1016/j.ijbiomac.2019.02.013

How to cite this article: Pedrini-Martha V, Schnegg R, Schäfer GG, Lieb B, Salvenmoser W, Dallinger R. Responsiveness of metallothionein and hemocyanin genes to cadmium and copper exposure in the garden snail *Cornu aspersum*. *J Exp Zool*. 2020;1–11. https://doi.org/10.1002/jez.2425

3

Fazit & Ausblick

Die gemeinsame Betrachtung der Ergebnisse zu den Hämocyaninen der Heterobranchia und der Caenogastropoda zeigt, dass die Hämocyaningene innerhalb der Großgruppe der Apogastropoda auf vielfältige Weise evolviert sind. In vielen verschiedenen Linien dieser artenreichsten aller Mollusken-Gruppen haben mehrfach unabhängig voneinander Hämocyanin-Genduplikationen stattgefunden und so zu paralogen Genen geführt. Mit dem Hämocyanin der Muricidae liegt neben dem bereits beschriebenen Mega-Hämocyanin der Cerithioidea eine weitere strukturelle Besonderheit eines Hämocyanins der Caenogastropoda vor. Auch die Genstruktur der Hämocyanine hat sich während der Evolution der Apogastropoda stark verändert. Die Ergebnisse meiner Arbeit zeigen, dass im Laufe der Evolution der Apogastropoda-Hämocyanine vermutlich immer mehr Introns entstanden und erhalten geblieben sind, was zu unterschiedlichen Genstrukturen mit 28 bis 61 Introns führte.

In den Hämocyaningenen verschiedener Tectipleura-Gruppen liegt mit Ausnahme eines zusätzlichen Introns, welches wir innerhalb eines Hämocyaningens der Hygrophila entdeckt haben, die gleiche Exon-Intron-Struktur vor. Diese ist demnach höchstwahrscheinlich bereits in einem gemeinsamen Vorfahren der Tectipleura evolviert und seitdem konserviert geblieben, was möglicherweise auf eine Sättigung der Gene mit Introns zurückzuführen ist. Um weitere Aussagen über die Evolution dieser Genstrukturen treffen zu können, sollten in zukünftigen Arbeiten die Hämocyaningene weiterer Heterobranchia, die nicht den Tectipleura angehören, analysiert werden. Durch die Untersuchung von Hämocyaningenen aus der Gruppe der basalen („lower“) Heterobranchia könnten so Rückschlüsse gezogen werden, in wie weit eine Ansammlung von Introns bereits in einem gemeinsamen Vorläufer der Heterobranchia stattgefunden hat. Auch das aufgestellte Szenario zur Evolution der Hämocyanin-Genstrukturen aus Kapitel 2.4 könnte dadurch weiter überprüft werden, um die Entstehung erster interner Introns in einem gemeinsamen Vorfahren aller Apogastropoda zu verifizieren.

Meine Ergebnisse aus Kapitel 2.4 zeigen, dass innerhalb der Caenogastropoda Veränderungen der Exon-Intron-Strukturen der Hämocyaningene kontinuierlich während der Evolution der einzelnen Schnecken-Linien stattgefunden haben. Anders als bei den Hämocyaningenen der Tectipleura sind hier verschiedene Introns erst nach den Genduplikationen, die zu den mehrfach heute vorhandenen Caenogastropoda-Hämocyaninen geführt haben, entstanden. Um zu überprüfen, ob es sich bei der

kontinuierlichen Ansammlung interner Introns tatsächlich um einen allgemeinen Trend innerhalb der Caenogastropoda handelt, sollten in Folgeuntersuchungen auch Hämocyaningene weiterer, bisher diesbezüglich nicht untersuchter Caenogastropoda-Gruppen (Cyclophorida, Viviparida, Campanilida; vgl. Abb. 3) analysiert werden.

Aufgrund der im Vergleich zu den Lepetellida (Vetigastropoda) deutlich stärkeren Evolutionsrate der Hämocyaningene in Apogastropoda habe ich schließlich die Hypothese aufgestellt, dass die starke Evolutionsrate mit der enormen Radiation und der großen Diversität der Apogastropoda zusammenhängen oder womöglich sogar eine Voraussetzung für diese darstellen könnte. Um diese Hypothesen weiter überprüfen zu können, ist es notwendig, dass viele weitere Untersuchungen der Hämocyanine bezüglich ihrer physiologischen Eigenschaften (z.B. Temperaturabhängigkeit, Sauerstoffaffinität) folgen, um mögliche funktionelle Veränderungen zu analysieren.

Zusammenfassend bietet meine Arbeit damit einen ersten Einblick in die vielfältige Evolution der Hämocyaningene in der Gruppe der Apogastropoda und beschreibt unterschiedliche evolutive Erneuerungen wie Genduplikationen und die Zunahme von Introns in Hämocyaningenen. Sie kann hiermit als Grundlage für weitere Studien zur Erforschung dieser riesigen Sauerstofftransport-Proteine und deren Einfluss auf die Evolution der Diversität der Apogastropoda und der Mollusken im Allgemeinen dienen.

4

Literaturverzeichnis

Aktipis, S. W., G. Giribet, D. R. Lindberg, W. F. Ponder (2008): Gastropoda:. An overview and analysis. In: Ponder, W. (Hrsg.): Phylogeny and Evolution of the Mollusca. University of California Press, S. 201–237.

Albrecht, U., H. Keller, W. Gebauer, J. Markl (2001): Rhogocytes (pore cells) as the site of hemocyanin biosynthesis in the marine gastropod *Haliotis tuberculata*. Cell and tissue research 304, 3/2001, S. 455–462.

Altenhein, B., J. Markl, B. Lieb (2002): Gene structure and hemocyanin isoform HtH2 from the mollusc *Haliotis tuberculata* indicate early and late intron hot spots. Gene 301, 1-2/2002, S. 53–60.

Bang, R., R. DeSalle, W. Wheeler (2000): Transformationalism, taxism, and developmental biology in systematics. Systematic Biology 49, 1/2000, S. 19–27.

Bergmann, S., B. Lieb, P. Ruth, J. Markl (2006): The hemocyanin from a living fossil, the cephalopod *Nautilus pompilius*: protein structure, gene organization, and evolution. Journal of molecular evolution 62, 3/2006, S. 362–374.

Bergmann, S., J. Markl, B. Lieb (2007): The first complete cDNA sequence of the hemocyanin from a bivalve, the protobranch *Nucula nucleus*. Journal of molecular evolution 64, 5/2007, S. 500–510.

Boettger, C. R. (1954): Die Systematik der euthyneuren Schnecken. Verhandlungen der Deutschen Zoologischen Gesellschaft Supplementband 18, 1954, S. 253–280.

Bouchet, P., J.-P. Rocroi, B. Hausdorf, A. Kaim, Y. Kano, A. Nützel, P. Parkhaev, M. Schrödl, E. E. Strong (2017): Revised Classification, Nomenclator and Typification of Gastropod and Monoplacophoran Families. Malacologia 61, 1-2/2017, S. 1–526.

Burmester, T. (2001): Molecular evolution of the arthropod hemocyanin superfamily. Molecular biology and evolution 18, 2/2001, S. 184–195.

Chabicovsky, M. (2003): Localization and quantification of Cd- and Cu-specific metallothionein isoform mRNA in cells and organs of the terrestrial gastropod *Helix pomatia*. Toxicology and Applied Pharmacology 190, 1/2003, S. 25–36.

Chao, Y.-C., M. Merritt, D. Schaefferkoetter, T. G. Evans (2020): High-throughput quantification of protein structural change reveals potential mechanisms of temperature adaptation in Mytilus mussels. BMC evolutionary biology 20, 2020.

Chiumiento, I. R., S. Ituarte, J. Sun, J. W. Qiu, H. Heras, M. S. Dreon (2020): Hemocyanin of the caenogastropod *Pomacea canaliculata* exhibits evolutionary differences among gastropod clades. PloS one 15, 1/2020, e0228325.

Colgan, D. J., W. F. Ponder, E. Beacham, J. Macaranas (2007): Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca). Molecular phylogenetics and evolution 42, 3/2007, S. 717–737.

Collin, R. (2005): Development, phylogeny, and taxonomy of Bostrycapulus (Caenogastropoda: Calyptraeidae), an ancient cryptic radiation. Zoological Journal of the Linnean Society 144, 1/2005, S. 75–101.

Cox, L. R. (1960): Thoughts on the classification of the Gastropoda. Journal of Molluscan Studies, 1960.

Cunha, T. J., G. Giribet (2019): A congruent topology for deep gastropod relationships. Proceedings. Biological sciences 286, 1898/2019, S. 20182776.

Dallinger, R., M. Chabicovsky, E. Hödl, C. Prem, P. Hunziker, C. Manzl (2005): Copper in *Helix pomatia* (Gastropoda) is regulated by one single cell type: differently responsive metal pools in rhogocytes. American journal of physiology. Regulatory, integrative and comparative physiology 289, 4/2005, R1185-95.

Dinapoli, A., A. Klussmann-Kolb (2010): The long way to diversity - phylogeny and evolution of the Heterobranchia (Mollusca. Gastropoda). Molecular phylogenetics and evolution 55, 1/2010, S. 60–76.

Finke, E., H. O. Pörtner, P. G. Lee, D. M. Webber (1996): Squid (*Lolliguncula brevis*) life in shallow waters: oxygen limitation of metabolism and swimming performance. The Journal of experimental biology 199, Pt 4/1996, S. 911–921.

Frýda, J. (2021): Gastropods. In: Alderton, D., S. A. Elias (Hrsg.): Encyclopedia of Geology. Elsevier, S. 299–310.

Gai, Z., A. Matsuno, K. Kato, S. Kato, M. R. I. Khan, T. Shimizu, T. Yoshioka, Y. Kato, H. Kishimura, G. Kanno, Y. Miyabe, T. Terada, Y. Tanaka, M. Yao (2015): Crystal Structure of the 3.8-MDa Respiratory Supermolecule Hemocyanin at 3.0 A Resolution. Structure (London, England : 1993) 23, 12/2015, S. 2204–2212.

Gatsogiannis, C., O. Hofnagel, J. Markl, S. Raunser (2015): Structure of mega-hemocyanin reveals protein origami in snails. Structure (London, England : 1993) 23, 1/2015, S. 93–103.

Gebauer, W., S. Stoeva, W. Voelter, E. Dainese, B. Salvato, M. Beltramini, J. Markl (1999): Hemocyanin subunit organization of the gastropod *Rapana thomasiana*. Archives of biochemistry and biophysics 372, 1/1999, S. 128–134.

Gosselck, F., Darr A., J. H. Jungbluth, M. Zettler (2009): Trivialnamen für Mollusken des Meeres und Brackwassers in Deutschland (Polyplacophora, Gastropoda, Bivalvia, Scaphopoda et Cephalopoda). In: Mollusca, S. 3–32.

Gray, J. E. (1840): Synopsis of the contents of the British Museum. Printed by G. Woodfall and Son, Angel Court, Skinner Street, London.

Harasewych, M. G., Adamkewicz S. L., M. Plassmeyer, P. M. Gillevet (1998): Phylogenetic relationships of the lower Caenogastropoda (Mollusca, Gastropoda, Architaenioglossa, Campaniloidea, Cerithioidea) as determined by partial 18s rDNA sequences. Zoologica Scripta 27, 4/1998, S. 361–372.

Haszprunar, G. (1988): On the origin and evolution of major gastropods group, with special reference to the streptoneura. Journal of Molluscan Studies 54, 4/1988, S. 367–441.

Haszprunar, G. (1985): The Heterobranchia - a new concept of the phylogeny of the higher Gastropoda. Journal of Zoological Systematics and Evolutionary Research 23, 1/1985, S. 15–37.

Hsia, C. C. W., A. Schmitz, M. Lambertz, S. F. Perry, J. N. Maina (2013): Evolution of air breathing: oxygen homeostasis and the transitions from water to land and sky. Comprehensive Physiology 3, 2/2013, S. 849–915.

Hyman L. H. (1967): The Invertebrates. Vol. 6, Mollusca I. McGraw-Hill, New York.

Jörger, K. M., I. Stöger, Y. Kano, H. Fukuda, T. Knebelsberger, M. Schrödl (2010): On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. BMC evolutionary biology 10, 2010, S. 323.

Kano, Y., B. Brenzinger, A. Nützel, N. G. Wilson, M. Schrödl (2016): Ringiculid bubble snails recovered as the sister group to sea slugs (Nudipleura). Scientific reports 6, 2016, S. 30908.

Kato, S., T. Matsui, C. Gatsogiannis, Y. Tanaka (2018): Molluscan hemocyanin: structure, evolution, and physiology. Biophysical reviews 10, 2/2018, S. 191–202.

Kocot, K. M., J. T. Cannon, C. Todt, M. R. Citarella, A. B. Kohn, A. Meyer, S. R. Santos, C. Schander, L. L. Moroz, B. Lieb, K. M. Halanych (2011): Phylogenomics reveals deep molluscan relationships. Nature 477, 7365/2011, S. 452–456.

Kocot, K. M., K. M. Halanych, P. J. Krug (2013): Phylogenomics supports Panpulmonata. Opisthobranch paraphyly and key evolutionary steps in a major radiation of gastropod molluscs. Molecular phylogenetics and evolution 69, 3/2013, S. 764–771.

Lang, W. H. (1988): cDNA cloning of the *Octopus dofleini* hemocyanin: sequence of the carboxyl-terminal domain. Biochemistry 27, 19/1988, S. 7276–7282.

Lieb, B., B. Altenhein, R. Lehnert, W. Gebauer, J. Markl (1999): Subunit organization of the abalone *Haliotis tuberculata* hemocyanin type 2 (HtH2), and the cDNA sequence encoding its functional units d, e, f, g and h. European journal of biochemistry 265, 1/1999, S. 134–144.

Lieb, B., B. Altenhein, J. Markl (2000): The sequence of a gastropod hemocyanin (HtH1 from *Haliotis tuberculata*). The Journal of biological chemistry 275, 8/2000, S. 5675–5681.

Lieb, B., B. Altenhein, J. Markl, A. Vincent, E. van Olden, K. E. van Holde, K. I. Miller (2001): Structures of two molluscan hemocyanin genes: significance for gene evolution. Proceedings of the National Academy of Sciences of the United States of America 98, 8/2001, S. 4546–4551.

Lieb, B., V. Boisguerin, W. Gebauer, J. Markl (2004): cDNA sequence, protein structure, and evolution of the single hemocyanin from *Aplysia californica*, an opisthobranch gastropod. Journal of molecular evolution 59, 4/2004, S. 536–545.

Lieb, B., K. Dimitrova, H.-S. Kang, S. Braun, W. Gebauer, A. Martin, B. Hanelt, S. A. Saenz, C. M. Adema, J. Markl (2006): Red blood with blue-blood ancestry. Intriguing structure of a snail hemoglobin. Proceedings of the National Academy of Sciences 103, 32/2006, S. 12011–12016.

Lieb, B., W. Gebauer, C. Gatsogiannis, F. Depoix, N. Hellmann, M. G. Harasewych, E. E. Strong, J. Markl (2010): Molluscan mega-hemocyanin: an ancient oxygen carrier tuned by a ~550 kDa polypeptide. Frontiers in zoology 7, 2010, S. 14.

Lieb, B., J. Markl (2004): Evolution of molluscan hemocyanins as deduced from DNA sequencing. Micron 35, 1-2/2004, S. 117–119.

Lieb, B., C. Todt (2008): Hemocyanin in mollusks - a molecular survey and new data on hemocyanin genes in Solenogastres and Caudofoveata. Molecular phylogenetics and evolution 49, 1/2008, S. 382–385.

Markl, J. (2013): Evolution of molluscan hemocyanin structures. Biochimica et biophysica acta 1834, 9/2013, S. 1840–1852.

Markl, J., A. Savel-Niemann, A. Wegener-Strake, M. Sding, A. Schneider, W. Gebauer, J. R. Harris (1991): The role of two distinct subunit types in the architecture of keyhole limpet hemocyanin (KLH). Die Naturwissenschaften 78, 11/1991, S. 512–514.

Martin, A. M., G. G. Martin, R. Butler, S. K. Goffredi (2011): Synthesis of keyhole limpet hemocyanin by the rhogocytes of *Megathura crenulata*. Invertebrate Biology 130, 4/2011, S. 302–312.

Medina, M., A. Collins (2003): The role of molecules in understanding molluscan evolution. In: Lydeard, C., D. R. Lindberg, G. J. Vermeij (Hrsg.): Molecular systematics and phylogeography of mollusks. Smithsonian Books, Washington, DC, S. 14–44.

Melzner, F., F. C. Mark, H.-O. Pörtner (2007): Role of blood-oxygen transport in thermal tolerance of the cuttlefish, *Sepia officinalis*. Integrative and comparative biology 47, 4/2007, S. 645–655.

Miller, K., K. E. van Holde (1974): Oxygen binding by callianassa californiensis hemocyanin. Biochemistry 13, 8/1974, S. 1668–1674.

Miller, K. I. (1985): Oxygen equilibria of *Octopus dofleini* hemocyanin. Biochemistry 24, 17/1985, S. 4582–4586.

Miller, K. I., M. E. Cuff, W. F. Lang, P. Varga-Weisz, K. G. Field, K. E. van Holde (1998): Sequence of the *Octopus dofleini* hemocyanin subunit: structural and evolutionary implications. Journal of molecular biology 278, 4/1998, S. 827–842.

Mordan, P., C. Wade (2008): Heterobranchia II. The Pulmonata. In: Ponder, W. (Hrsg.): Phylogeny and Evolution of the Mollusca. University of California Press, S. 409–426.

Oellermann, M., J. M. Strugnell, B. Lieb, F. C. Mark (2015): Positive selection in octopus haemocyanin indicates functional links to temperature adaptation. BMC evolutionary biology 15, 2015, S. 133.

Ponder, W. F., D. J. Colgan, J. M. Healy, N. Alexander, L. R. Simone, E. E. Mielke (2008): Caenogastropoda. In: Ponder, W. (Hrsg.): Phylogeny and Evolution of the Mollusca. University of California Press, S. 331–383.

Ponder, W. F., D. R. Lindberg (1997): Towards a phylogeny of gastropod molluscs: an analysis using morphological characters. Zoological Journal of the Linnean Society 119, 2/1997, S. 83–265.

Ponder, W. F., D. R. Lindberg, J. M. Ponder (2019): Biology and Evolution of the Mollusca. CRC PRESS, [S.l.].

Pörtner, H. O., D. M. Webber, R. G. Boutilier, R. K. O'Dor (1991): Acid-base regulation in exercising squid (*Illex illecebrosus, Loligo pealei*). The American journal of physiology 261, 1 Pt 2/1991, R239-46.

Romero, P. E., M. Pfenninger, Y. Kano, A. Klussmann-Kolb (2016): Molecular phylogeny of the Ellobiidae (Gastropoda. Panpulmonata) supports independent terrestrial invasions. Molecular phylogenetics and evolution 97, 2016, S. 43–54.

Schäfer, G. G. (2017): Bioinformatische und ontogenetische Analyse des Hämocyanin-Gens von *Lymnaea stagnalis*. Masterarbeit, Mainz.

Schäfer, G. G., L. J. Grebe, F. Depoix, B. Lieb (2021): Hemocyanins of Muricidae: New 'Insights' Unravel an Additional Highly Hydrophilic 800 kDa Mass Within the Molecule. Journal of molecular evolution, 2021.

Seibel, B. A., F. Chausson, F. H. Lallier, F. Zal, J. J. Childress (1999): Vampire blood: respiratory physiology of the vampire squid (Cephalopoda: Vampyromorpha) in relation to the oxygen minimum layer. Experimental Biology Online 4, 1/1999, S. 1–10.

Senozan, N., J. Landrum, J. Bonaventura, C. Bonaventura (1981): Hemocyanin of the giant keyhole limpet, *Megathura crenulata*. In: Lamy, J., J. Lamy (Hrsg.): Invertebrate oxygen-binding proteins. Structure, active site, and function : Proceedings of a workshop sponsored by the European Molecular Biology Organization, held in Tours, France, August 20-24 1979. Dekker, New York, S. 703–717.

Siezen, B. J., E. van Bruggen (1974): Structure and properties of hemocyanins. Journal of molecular biology 90, 1/1974, S. 77–89.

Sigwart, J. D., D. R. Lindberg (2015): Consensus and confusion in molluscan trees: evaluating morphological and molecular phylogenies. Systematic Biology 64, 3/2015, S. 384–395.

Simone, L. R. L. (2011): Phylogeny of the Caenogastropoda (Mollusca), based on comparative morphology. Arquivos de Zoologia 42, 4/2011, S. 161.

Simone, L. R. L. (2001): Phylogenetic analyses of Cerithioidea (Mollusca, Caenogastropoda) based on comparative morphology. Arquivos de Zoologia 36, 2/2001, S. 147.

Smith, S. A., N. G. Wilson, F. E. Goetz, C. Feehery, S. C. S. Andrade, G. W. Rouse, G. Giribet, C. W. Dunn (2011): Resolving the evolutionary relationships of molluscs with phylogenomic tools. Nature 480, 7377/2011, S. 364–367.

Steele, J. H. (1985): A comparison of terrestrial and marine ecological systems. Nature 313, 6001/1985, S. 355–358.

Streit, K., Geiger D. L., B. Lieb (2006): Molecular Phylogeny and the Geographic Origin of Haliotidae traced by Haemocyanin sequences. Journal of Molluscan Studies 72, 1/2006, S. 105–110.

Streit, K.-S. (2008): Differentielle Expression und molekulare Evolution von Mollusken-Hämocyanin. Dissertation, Mainz.

Strobel, A., M. Y. A. Hu, M. A. Gutowska, B. Lieb, M. Lucassen, F. Melzner, H. O. Portner, F. C. Mark (2012): Influence of temperature, hypercapnia, and development on the relative expression of different hemocyanin isoforms in the common cuttlefish *Sepia officinalis*. Journal of experimental zoology. Part A, Ecological genetics and physiology 317, 8/2012, S. 511–523.

Strong, E. E. (2003): Refining molluscan characters: morphology, character coding and a phylogeny of the Caenogastropoda. Zoological Journal of the Linnean Society 137, 4/2003, S. 447–554.

Swerdlow, R. D., R. F. Ebert, P. Lee, C. Bonaventura, K. I. Miller (1996): Keyhole limpet hemocyanin. Structural and functional characterization of two different subunits and multimers. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 113, 3/1996, S. 537–548.

Symanowski, F., J.-P. Hildebrandt (2010): Differences in osmotolerance in freshwater and brackish water populations of *Theodoxus fluviatilis* (Gastropoda: Neritidae) are associated with differential protein expression. Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology 180, 3/2010, S. 337–346.

Thonig, A., M. Oellermann, B. Lieb, F. C. Mark (2014): A new haemocyanin in cuttlefish (*Sepia officinalis*) eggs. Sequence analysis and relevance during ontogeny. EvoDevo 5, 1/2014, S. 6.

van Holde, K. E., K. I. Miller (1995): Hemocyanins. Advances in protein chemistry 47, 1995, S. 1–81.

Velkova, L., I. Dimitrov, H. Schwarz, S. Stevanovic, W. Voelter, B. Salvato, P. Dolashka-Angelova (2010): Structure of hemocyanin from garden snail *Helix lucorum*. Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology 157, 1/2010, S. 16–25.

Vermeij, G. J., R. Dudley (2000): Why are there so few evolutionary transitions between aquatic and terrestrial ecosystems? Biological Journal of the Linnean Society 70, 4/2000, S. 541–554.

Vermeij, G. J., Wesselingh F. P. (2002): Neogastropod molluscs from the Miocene of western Amazonia, with comments on marine to freshwater transitions in molluscs. Journal of Paleontology 76, 2/2002, S. 265–270.

Wägele, H., A. Klussmann-Kolb, E. Verbeek, M. Schrödl (2014): Flashback and foreshadowing - a review of the taxon Opisthobranchia. Organisms Diversity & Evolution 14, 1/2014, S. 133–149.

Wägele, H., A. K. Kolb, V. Vonnemann, M. Medina (2008): Heterobranchia I. In: Ponder, W. (Hrsg.): Phylogeny and Evolution of the Mollusca. University of California Press, S. 384–408.

Warnke, K. M., A. Meyer, B. Ebner, B. Lieb (2011): Assessing divergence time of Spirulida and Sepiida (Cephalopoda) based on hemocyanin sequences. Molecular phylogenetics and evolution 58, 2/2011, S. 390–394.

WoRMS Editorial Board (2020): World Register of Marine Species. Abrufbar unter http://www.marinespecies.org.

Yao, T., M.-M. Zhao, J. He, T. Han, W. Peng, H. Zhang, J.-Y. Wang, J.-Z. Jiang (2019): Gene expression and phenoloxidase activities of hemocyanin isoforms in response to pathogen infections in abalone *Haliotis diversicolor*. International journal of biological macromolecules 129, 2019, S. 538–551.

Zapata, F., N. G. Wilson, M. Howison, S. C. S. Andrade, K. M. Jörger, M. Schrödl, F. E. Goetz, G. Giribet, C. W. Dunn (2014): Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proceedings. Biological sciences 281, 1794/2014, S. 20141739.

5

Quellenverzeichnis

Seite	#	Bild/Spezies	Autor	Creative Common Version
6	A1	*Aplysia californica*	Chad King	public domain (gemeinfrei)
6	A2	*Clione limacina*	Kevin Raskoff	public domain (gemeinfrei)
6	A3	*Tylodina perversa*	Parent Géry	https://creativecommons.org/licenses/by-sa/3.0/deed.de
6	A4	*Placida cremoniana*	Sike Rohrlach	https://creativecommons.org/licenses/by-sa/2.0/deed.de
6	A5	*Oxynoe viridis*	Izuzuki Diver	https://creativecommons.org/licenses/by-sa/2.5/deed.de
6	A6	*Rissoella opalina*	Yasunori Kano	https://creativecommons.org/licenses/by-sa/4.0/deed.de
6	A7	*Architectonica perspectiva*	Nhobgood Nick Hobgood	https://creativecommons.org/licenses/by-sa/3.0/deed.de
6	A8	*Melibe leonina*	Cabrillo Marine Aquarium	public domain (gemeinfrei)
6	A9	*Aplysiopsis elegans*	Sike Rohrlach	https://creativecommons.org/licenses/by-sa/2.0/deed.de
6	B1	*Siphonaria spec.*	Layton, Diana	https://creativecommons.org/licenses/by-sa/4.0/deed.de
6	B2	*Onchidella celtica*	Eigenes Werk	----------
6	B3	*Trimusculus mauritianus*	Philippe Bourjon	https://creativecommons.org/licenses/by-sa/4.0/deed.de
6	B4	*Otina ovata*	Eigenes Werk	----------
6	C1	*Acochlidium fijiiensis*	Katharina M. Jörger	https://creativecommons.org/licenses/by-sa/2.0/deed.de
6	C2	*Stagnicola palustris*	Alexander Mrkvicka	https://creativecommons.org/licenses/by-sa/3.0/deed.de
6	C3	*Ancylus fluviatilis*	Alexander Mrkvicka	https://creativecommons.org/licenses/by-sa/3.0/deed.de
6	C4	*Lymnaea stagnalis*	Bj. Schoenmakers	public domain (gemeinfrei)
6	D1	*Isognomostoma isognomostomos*	Francesco Welter Schultes	public domain (gemeinfrei)
6	D2	*Edentulina spec.*	Bernard DUPONT	https://creativecommons.org/licenses/by-sa/2.0/deed.de
6	D3	*Limax cinereoniger*	Holger Krisp	https://creativecommons.org/licenses/by-sa/3.0/deed.de
6	D4	*Cepaea nemoralis*	Mad Max	https://creativecommons.org/licenses/by-sa/3.0/deed.de
6	D5	*Euhadra peliomphala*	Takahashi	public domain (gemeinfrei)
6	D6	*Helix pomatia*	Waugsberg	https://creativecommons.org/licenses/by-sa/3.0/deed.de
6	D7	*Liguus virigineus*	Anders Sandberg	https://creativecommons.org/licenses/by-sa/2.0/deed.de
6	D8	*Achatina fulica*	Charles J. Sharp	https://creativecommons.org/licenses/by-sa/4.0/deed.de
6	D9	*Limax maximus*	Michal Manas	https://creativecommons.org/licenses/by-sa/2.5/deed.de
8	A1	*Lentigo lentiginosus*	Philippe Bourjon	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	A2	*Cyphoma gibbosum*	Laszlo Ilyes	https://creativecommons.org/licenses/by-sa/2.0/deed.de
8	A3	*Phenacovolva rosea*	Nhobgood Nick Hobgood	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	A4	*Chicoreus palmarosae*	Bricktop	https://creativecommons.org/licenses/by-sa/2.5/deed.de
8	A5	*Lambis scorpius*	Frédéric Ducarme	https://creativecommons.org/licenses/by-sa/4.0/deed.de
8	A6	*Dentiovula dorsuosa*	Nick Hobgood	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	A7	*Arestorides argus*	Elisabeth Morcel	https://creativecommons.org/licenses/by-sa/4.0/deed.de
8	A8	*Rapana venosa*	George Chernilevsky	public domain (gemeinfrei)
8	B1	*Telescopium telescopium*	Toni Wöhrl	https://creativecommons.org/licenses/by-sa/4.0/deed.de
8	B2	*Batillaria estuarina*	Jan Delsing	public domain (gemeinfrei)
8	B3	*Nodilittorina pyramidalis*	Peter Woodard	https://creativecommons.org/licenses/by-sa/4.0/deed.de
8	B4	*Crepidula fornicata*	Rudolphous	https://creativecommons.org/licenses/by-sa/4.0/deed.de
8	B5	*Nucella lapillus*	Eigenes Werk	----------
8	C1	*Pomacea bridgesii*	Stijn Ghesquiere	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	C2	*Sadleriana fluminensis*	Alexander Mrkvicka	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	C3	*Oncomelania hupensis*	Fred A. Lewis	https://creativecommons.org/licenses/by-sa/4.0/deed.de
8	C4	*Melanoides tuberculata*	Dennis L.	https://creativecommons.org/licenses/by-sa/2.0/deed.de
8	C5	*Anentome helena*	RSX	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	C6	*Melanopsis praemorsa*	Alexander Mrkvicka	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	C7	*Emmericia patula*	Alexander Mrkvicka	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	C8	*Elimia clara*	Whelan NV	https://creativecommons.org/licenses/by-sa/2.5/deed.de
8	D1	*Pupinella rufa*	Takahashi	public domain (gemeinfrei)
8	D2	*Rhiostoma huberi*	Manuel Caballer	https://creativecommons.org/licenses/by-sa/4.0/deed.de
8	D3	*Cochlostoma auritum*	Francesco Welter Schultes	public domain (gemeinfrei)
8	D4	*Terebralia palustris*	Christoph Kühne	https://creativecommons.org/licenses/by-sa/2.0/deed.de
8	D5	*Pomatias elegans*	Hectonichus	https://creativecommons.org/licenses/by-sa/3.0/deed.de
8	D6	*Cochlostoma septemspirale*	Céreric Audibert	public domain (gemeinfrei)
8	D7	*Cyclophorus perdix*	H. Zell	https://creativecommons.org/licenses/by-sa/3.0/deed.de

6

Danksagung

Ich danke herzlichst…

… [redacted] für die Bereitstellung und Betreuung meines Promotionsthemas. Ein großes Dankeschön für das entgegengebrachte Vertrauen, die Unterstützung während der gesamten Zeit meiner Doktorarbeit, die Ermöglichung vieler gewinnbringender Konferenzteilnahmen und erfolgreicher Exkursionen, viele konstruktive Diskussionen und die kritische Durchsicht und Kommentierung der einzelnen Veröffentlichungen sowie dieser Arbeit. Der wertschätzende Umgang und die angenehme Zusammenarbeit haben mich die vergangenen vier Jahre immer wieder motiviert und diese Arbeit möglich gemacht.

… [redacted] dafür, dass er bereitwillig das Zweitgutachten meiner Doktorarbeit übernommen hat.

… [redacted] und [redacted] sowie [redacted] für die angenehme Kooperation und die vielen hilfreichen Kommentare und Denkanstöße beim Überarbeiten gemeinsamer Manuskripte. Vielen Dank außerdem an [redacted] und [redacted] für die nette Zusammenarbeit.

… [redacted] und [redacted] für die Unterstützung im Labor sowie [redacted] für die Durchführung der elektronenmikroskopischen Untersuchungen.

… [redacted] und [redacted] für hilfreiche Ratschläge.

… [redacted] und [redacted] für ihre Unterstützung im alltäglichen Uniwahnsinn.

… den ehemaligen Kandidaten der Arbeitsgruppe [redacted], die auch nachdem sie die Uni verlassen haben immer wieder für mich da waren. Vielen Dank [redacted], [redacted] und [redacted] für eure Unterstützung, die interessanten Gespräche, tollen Exkursionen, eine wunderbare Konferenz und unheimlich viel Spaß!

… [redacted] für ihre stetige Hilfsbereitschaft und Unterstützung, für viele angenehme Unterhaltungen, ein tolles Arbeitsklima und die schöne gemeinsame Zeit in der Arbeitsgruppe.

… [redacted] und [redacted] für eine wunderbare Exkursion in die Bretagne und die enorme Hilfe beim erfolgreichen Schneckensammeln, für die stetige Unterstützung, viele lustige Unternehmungen, tolle Gespräche und die einzigartige Arbeitsatmosphäre.

… meinen Bachelor- und Master-Kandidaten [redacted], [redacted], [redacted], [redacted], [redacted] und [redacted] für die tolle und erfolgreiche Zusammenarbeit und eure Unterstützung!

… [redacted]

… meiner gesamten Familie! Ein besonderer Dank gilt meiner Mutter, die mich immer unterstützt hat, für mich da war und der ich das für diese Arbeit nötige Selbstvertrauen verdanke. Ein großes Dankeschön geht außerdem an meinen Vater für seine Hilfe und Unterstützung und dafür, dass er mir die Naturwissenschaften nähergebracht hat. Außerdem bedanke ich mich bei meinen Schwestern, auf die ich immer zählen kann, sowie bei Lotti, die mich mehrfach in der Uni besucht und selbst im Urlaub Ausschau nach Schnecken gehalten hat.

7

Lebenslauf

Persönliche Daten

Name:	Gabriela Giannina Schäfer
Wohnort:	[redacted]
Email:	[redacted]
Geburtstag:	[redacted]
Geburtsort:	[redacted]

Ausbildung & berufliche Erfahrung

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

Stipendien & Auszeichnungen:

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

8

Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig und nur mit Hilfe der angegebenen Personen und Mittel (Literatur, Methoden) angefertigt zu haben. Bei den von mir durchgeführten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der Satzung der Johannes Gutenberg-Universität Mainz festgelegt sind, eingehalten.

Ort, Datum	Gabriela Schäfer