KEEP AN EYE ON ARGONAUTA!

A comparative evolutionary analysis of key genes for vision in littoral versus deep-sea octopuses

Giacinto De Vivo

Thesis submitted for the degree of **Doctor of Philosophy -** XXIII Cycle
Life and Biomolecular Sciences
September 2024

From Comingio Merculiano (1845–1915) in Jatta Giuseppe
I Cefalopodi viventi nel Golfo di Napoli (sistematica) : monografia

Supervision Panel
Director of Studies: **Dr Salvatore D'Aniello** (Stazione Zoologica Anton Dohrn, Italy)
Internal Supervisor: **Dr Fabio Crocetta** (Stazione Zoologica Anton Dohrn, Italy)
External Supervisor: **Prof Rafael Zardoya** (Museo Nacional de Ciencias Naturales, Spain)

Examiner Panel
External examiner: **Prof Davide Pisani** (University of Bristol, United Kingdom)
External examiner: **Dr Fabio Cortesi** (University of Queensland, Australia)
Chair: **Dr Ylenia Carotenuto** (Stazione Zoologica Anton Dohrn, Italy)

The Open University
School of Life, Health and Chemical Sciences Milton Keynes (UK)
Stazione Zoologica Anton Dohrn
Department of Biology and Evolution of Marine Organisms – Naples (Italy)

“Here, O bhikkhus, a bhikkhu understands the eye and material forms and the fetter that arises dependent on both (eye and forms); he understands how the arising of the non-arisen fetter comes to be; he understands how the abandoning of the arisen fetter comes to be; and he understands how the non-arising in the future of the abandoned fetter comes to be.”

Satipaṭṭhāna Sutta

ABSTRACT

The ability to detect light and process visual information is crucial to octopods' success, allowing environment exploration, food detection, predator avoidance, and exceptional camouflaging. These is made possible by their complex camera-like eyes and highly developed optic lobes. However, data on octopod photoreception, particularly in deep-sea species, remain scarce. To address this gap, adaptations to various light conditions among different octopod species living in diverse environments were investigated by examining differences in their molecular light-sensing machinery, the opsins.

To begin, the ancestral opsin toolkit of octopods was reconstructed, revealing that their common ancestor possessed at least five opsins: r-opsin1, r-opsin2, xenopsin, retinochrome, and peropsin. Additionally, a novel group of opsin-like molecules, termed pseudopsins, was identified. These molecules lack the retinal-binding domain, which is essential for photoreception.

Furthermore, RT-qPCR analysis was conducted to examine the expression patterns of these various opsins in two distantly related octopods, *Argonauta argo* and *Octopus vulgaris,* revealing a complex expression pattern: r-opsin1 is expressed in the retina, optic lobe, and suckers; r-opsin2 in the optic lobe; xenopsin in the retina; retinochrome in all tissues; peropsin in the optic lobe; and pseudopsin in all tissues. These findings raise new questions about the role of opsins in non-photoreceptive tissues and the function of xenopsin, which is capable of activating a signalling cascade in cephalopod eyes.

Next, low-coverage RNA sequencing of the visual system (retina and optic lobe) was performed on five octopod species inhabiting different light environments. These data were used to perform a positive selection analysis on the two most highly

expressed opsins in the eye, r-opsin1 and retinochrome. Among the species studied, positive selection was observed in r-opsin1 only in *Pteroctopus tetracirrhus* living in the bathypelagic environment. In contrast, retinochrome exhibited positive selection sites in all deep-sea species, providing new insights into the evolution of these molecules.

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, Dr Salvatore D'Aniello, who gave me the opportunity to embark on this incredible scientific journey. He mentored me throughout these three years, providing trust, support, and valuable advice.

I would also like to thank my two co-supervisors, Dr Fabio Crocetta and Prof Rafael Zardoya. The help of Dr Fabio Crocetta was essential for his encyclopaedic knowledge of the Gulf of Naples fauna and for instructing me on the various species. Prof Rafael Zardoya provided insightful suggestions and guidance throughout the research.

Beyond my supervising team, a special mention goes to Dr Roberto Feuda for hosting me in his lab in Leicester and for being an outstanding collaborator and scientific mentor.

I would like to extend my thanks to Dr Filomena Caccavale, Dr Pasquale De Luca, Dr. Giovanna Ponte, and Dr Giovanni Annona. They introduced me to many molecular biology and wet lab techniques, and their support was essential to the success of this thesis. The same gratitude goes to the RAF and lab technicians of the Stazione Zoologica Anton Dohrn: Dr Pamela Imperatore, Dr Mara Francone, Dr Luca Ambrosino, Elvira Mauriello, and Raimondo Pannone, who provided critical technical assistance during the experiments.

Many thanks also go to Clifton Lewis, Mathew Goutily, and Dr Alessandra Aleotti from Dr Roberto Feuda's lab for their help with the phylogeny of opsins in Lophotrochozoa; to Dr Giobbe Forni from Prof Andrea Luchetti's lab for contributing to the selection analysis of opsin genes; and to Dr Lorena Buono for her assistance with transcriptome analysis.

I will always remember Francesca and Margaux, two young students whose interest and enthusiasm were truly inspiring.

Over the course of this year, I have met many extraordinary people and outstanding young researchers who have also become close friends beyond the lab walls, making this journey more bearable: Maria, Isabella, Annalisa, Giulia, Valentina, Riccardo, Tanya, Antonio, Lorenza, Andrea, Alessandra, and John. Eva provided constant support throughout my PhD, particularly during challenging moments, and their help greatly improved the quality of this dissertation.

I would also like to thank Giuseppe for being great friend and his family for their help and care during my stay in Naples. Finally, my deepest thanks go to my family for always believing in me and their support. This dissertation is dedicated to my sister Emanuela.

Table of Contents

CHAPTER 1 - Introduction

Characterized by high intelligence and a diversified lifestyle, octopods represent one of the most distinctive cephalopod groups, as indicated by their distribution across a differentiated environment, including pelagic, benthic, deep-sea and littoral (Albertin et al., 2015; Destanović et al., 2023; Hanke & Kelber, 2020; Yoshida et al., 2015; Young, 1991). Octopods, like all the other living cephalopods, are renowned as active predators, relying on a multitude of senses to locate and capture their prey. Notably, their exceptional visual acuity, facilitated by well-developed camera eyes (eyes with lens, iris, cornea, and retina), and remarkable camouflage abilities underscore the importance of studying their sophisticated light-sensing molecular machinery, the opsins (Hanke & Kelber, 2020; Yoshida et al., 2015).

The convergent and rapid evolution of camera eyes in different metazoan phyla, including cephalopods, is one of the most interesting phenomena in natural history (Ogura et al., 2004). It is strictly connected with the astonishing ability of organisms to adapt and conquer different environments and to find similar solutions in species separated by millions of years, such as cnidarians (cubozoans), arthropods (spiders), and vertebrates (Fernald, 2006; Kozmik et al., 2008; Ogura et al., 2004; Serb & Eernisse, 2008; Yoshida & Ogura, 2011). Indeed, the presence of complex eyes is indicative of the importance that vision has to the ecology of the organism.

In the case of cephalopods, it is not by chance that they evolved camera eyes. This becomes particularly clear when the wide variety of eye structures evolved in molluscs, such as aesthetes in polyplacophorans, or the lens eyes present in gastropods and bivalves, are taken into consideration (Serb & Eernisse, 2008). Eyes are so important to cephalopods that even the simple pinhole eye of

Nautilus pompilius, one of the earliest divergent living cephalopods, is an adaptation to fast daily excursions in the water column and not just an ancestral character retained by these animals (Zhang et al., 2021).

Cephalopods rely on high-resolution image-forming not only to detect prey or avoid predators but also to communicate and orient themselves (Chung & Marshall, 2016; Hanke & Kelber, 2020; Young, 1991). Despite the importance of light in these animals, exhaustive knowledge on adaptation to different light conditions in different octopod species remains largely missing; therefore, the main aim of this PhD thesis is to try to fill this gap. This first chapter will introduce the reader to the main theoretical concepts, and the state of the art in research is illustrated, including a more detailed explanation of the research aims that led to this final dissertation.

1.1 Light in the water column

Compared to land, light behaves differently in water, being absorbed and scattered depending on several factors such as latitude, turbidity, flickery and, depth (Maximov, 2000; Warrant & Locket, 2004). These changes are both quantitative and qualitative (Figure 1.1). Quantitatively, light intensity decreases with depth, meaning that at a certain level of the water column, in the bathypelagic zone (~1000 m in depth in open oceans), sunlight is almost completely absent, and bioluminescence becomes the only available light source. Qualitatively, only some wavelengths reach certain depths, as others are absorbed or scattered in the upper part of the water column. In particular, the longer and shorter wavelengths, the two extremes of the light spectrum (violet and red), are respectively scattered and absorbed earlier, while only blue-green light penetrates to deeper levels of the water column and becomes the only available light after ~200 m in open and clear oceans. Therefore, depending on the extent of light absorption, the water column can be divided into

three main zones: the euphotic zone (Epipelagic 0-200 m in depth), where sunlight is gradually scattered and absorbed; the disphotic zone (Mesopelagic 200-1000 m in depth), where wavelengths from violet to blue and from red to blue are completely absent; and the aphotic zone (Bathypelagic 1000-4000 m in depth), in which no sunlight reaches, and bioluminescence is the sole source of light (Warrant & Locket, 2004). UV-light does not penetrate more than 50 m. Conditions like latitude, turbidity and seasonality change this pattern, making some wavelength less available earlier in the water column (https://www.noaa.gov/jetstream/ocean/layers-of-ocean).

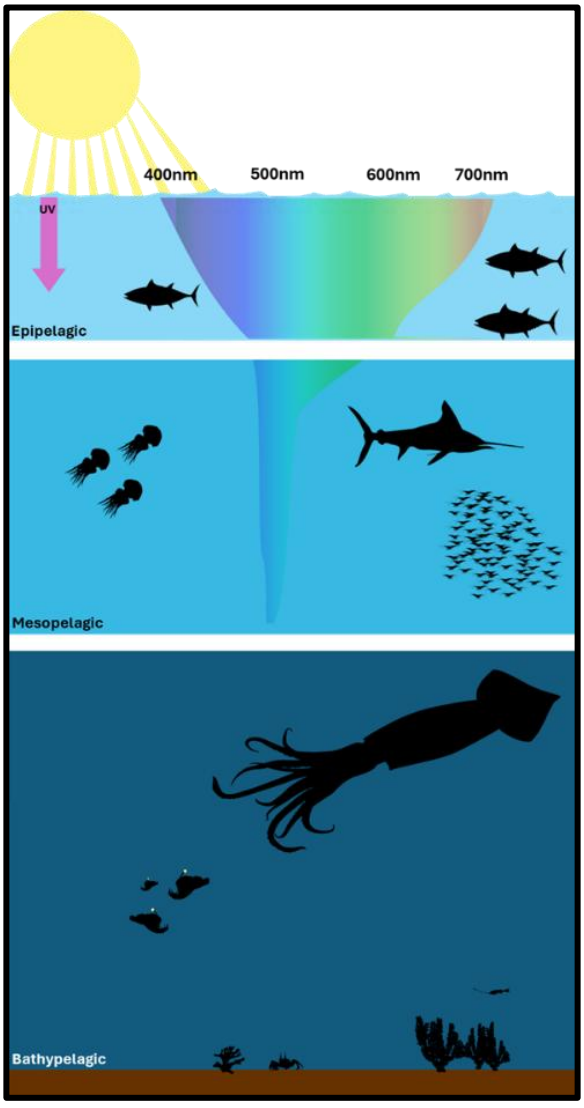


Figure 1.1 – Light absorption in open ocean. The image illustrates how sunlight changes through the water column in the different oceanic zones. Most wavelengths are absorbed by epipelagic zone (up to ~200) and only blue light reaches the mesopelagic (up to ~1000 m in open oceans). Sunlight is absent in the bathypelagic zone (up to ~4000 meters in open oceans) and only bioluminescence becomes available as main light source.

1.2 General adaptation to light environments in the deep sea

To accommodate quantitative changes, low light level requires morphological adaptation to capture more photons and increase sensitivity (Warrant & Locket, 2004). Larger photoreceptive cells (or photoreceptive cells grouped in larger units) are common evolutionary solutions. This means that bigger eyes are often convergently evolved in deep sea animals, as they are required to increase the focal length and accommodate larger photoreceptors by keeping the same resolution. An additional mirror cell layer can be used to reflect the light back to the photoreceptors and increase light sensitivity, such as the vertebrate *tapetum lucidum* or the analogous acellular crystalline deposits in the eyes of nocturnal spiders (Warrant & Locket, 2004). Bigger lens can also help the animal to capture more sunlight and to focus it on the retina, leading to convergent evolution of dorsally directed tubular eyes (e.g., the big eye of the squid *Histiotheuthis bonnelli* and the eyes of many abyssal fishes, the most iconic being the barreleye fish *Macropinna microstoma*) (Thomas et al., 2017; Warrant & Locket, 2004). In contrast, to detect fast punctiform bioluminescent emission from the prey or potential predators, eyes with a wide parabolic shaped retina, placed laterally, are common solutions; this morphology is called the typical eye (e.g., the small eye of the squid *Histiotheuthis bonnelli* and the eyes of the lantern fish *Myctophum punctatum*). Sometimes, both eye morphologies can be found in the same animal: for example, the squid *Histiotheuthis bonnelli* possesses two morphological distinct eyes, one tubular and one typical. Another example are the highly specialized eyes of the brownsnout spookfish *Dolichopteryx longipes*, possessing tubular eyes pointing upward, bearing an additional structure (mirrow eye) pointing downward that is functionally analogous to a typical eye (Robison & Reisenbichler, 2008; Warrant & Locket, 2004). Other adaptations to low

light intensity require molecular changes. For example, a way is to increase sensitivity trough temporal summation, in which a photoreceptor collects photons over a longer period of time, prolongating the signal (Warrant, 1999; Warrant & Locket, 2004). Temporal summation can also be used to extend the visual range as it allows capturing photons that come from longer distances, taking more travelling time. These mutations involve changes in the sequence of the opsins, the main molecular photoreceptors in metazoans, and the photoisomerases (Hagen et al., 2023; Musilova et al., 2019, 2021). Opsins detect the change in configuration of its ligand, the retinal, from *cis* to *trans*, while the photoisomerases revert the retinal back to *cis*, making the opsin quickly available to capture another photon (Zhang et al., 2021). A prolongated opsin signalling activity might contribute to temporal summation, while more efficient photoisomerases can contribute to a quick restoration of the retinal even in low light environments. In cephalopods, the photoisomerases the retinochrome (Hara & Hara, 1972; Vöcking et al., 2021; Zhang et al., 2021).

On the other hand, adapting to qualitative changes in light within the water column involves mutations that enable the retinal attached to opsin to more effectively absorb the blue-green wavelengths present in the deep-sea environment (Hagen et al., 2023; Musilova et al., 2019, 2021). This is due to amino acid modifications leading to changes in the electromagnetic field, making a particular retinal configuration (11-*cis* retinal or 11-*trans* retinal) more or less stable. Generally, changes in the opsin sequence leading to mutations involving amino acids surrounding retinal contribute to this process. Although not well studied in molluscs, shift in light absorption has been extensively explored in vertebrates (Hagen et al., 2023; Musilova et al., 2019, 2021). Given the remarkable similarity of visual organs with vertebrates, we expect that those adaptations might also occur in octopods. Therefore, by studying opsin structure and evolution in octopods, we aim to extend

the knowledge on the evolution of these peculiar and important molecules, which will be explored further in the subsequent paragraphs.

1.3 The Opsins

Opsins are important molecules involved in animal visual processes and other photoreceptive functions such as circadian entrainment (e.g., melanopsin in humans), seasonality, or orientation in the water column (Davies et al., 2010; Gühmann et al., 2015; Terakita, 2005). Opsins are G-protein-coupled receptors with seven transmembrane domains that covalently bind a vitamin A derivative, the retinal (often referred to as chromophore), through a protonated Schiff base in a specific lysine site (K296 in cow rhodopsin used as reference). This binding confers to opsins the photoreceptive properties: under dark conditions, the retinal is attached to the opsin in its 11-*cis* form, while, when hit by a photon, the retinal changes its configuration to all-*trans* (Figure 1.2). The change in the retinal configuration cause a conformational change of the opsin, which enables it to bind the G-protein starting the signalling cascade (Terakita, 2005). A photoactive opsin is referred as to be in its meta state.

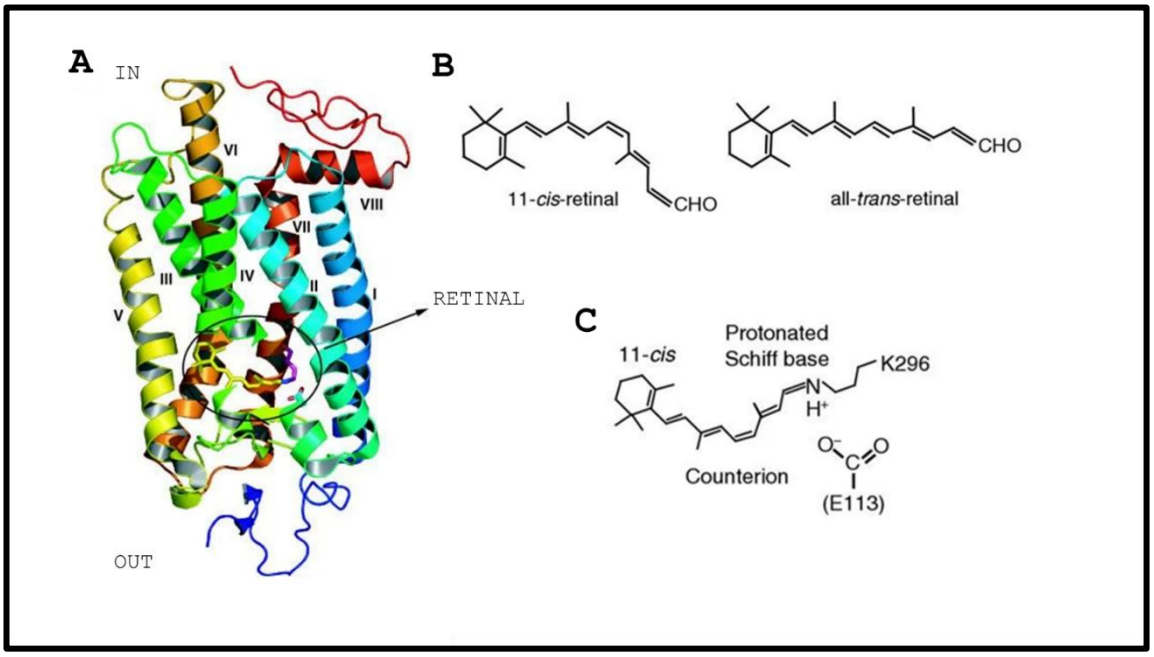


Figure 1.2 – Opsin structure and retinal interaction. 3-D model of the cow rhodopsin (A), showing its seven transmembrane domain and the cytoplasmic helix VIII. Opsins bind specific ligand, the retinal (B), in specific conserved site, the lysin K296 in cow rhodopsin. In dark conditions, the retinal is in its 11-*cis* configuration, while when interact with a photon it changes to *all-trans.* The retinal bind the K296 thorough a protonated Schiff base (C), that is stabilized by a negative charged counterion, in this example E113. Modified from Terakita (2005).

Opsins emerged in the animal kingdom at the Eumetazoa node, but opsin-like sequences that lack the retinal binding site have been found in Placozoa (named Placopsins) and Bilateria (named Pseudopsins) (De Vivo et al., 2023; Feuda et al., 2012; Fleming et al., 2020). Opsins are divided into seven subfamilies: rhabdomeric opsins (r-opsins), ciliary opsins (c-opsins), xenopsins, and those belonging to group-4, such as peropsins/retinochromes, Go opsins, RGRs (RPE-retinal G protein-coupled), and neuropsins. In general, there is about 25% amino acid similarity between subfamilies, which increases to over 40% within the members of the same subfamily (Terakita, 2005).

Opsins are traditionally categorized using the G-protein that they bind, but this can be misleading. For example, xenopsins might activate a Gt/Gi signalling cascade, and therefore, the definition of Gt-coupled opsins to indicate only c-opsins

might be ambiguous (Döring et al., 2020). To overcome this problem, in this dissertation terms like Gt-coupled opsins, referring to c-opsins, and Gq-coupled opsins, used to refer to r-opsins, are abandoned, since it is more important to highlight their phylogenetic relationship and thus the homology rather than the protein that they activate. The term Go opsin is maintained since there are no alternative names in the literature to refer to this group and it is important to not create further confusion. On the other hand, also the terms ciliary and rhabdomeric opsins might be misleading since they imply the opsin is expressed in a certain type of photoreceptors or photoreceptive structures, which is not always the case (Matsuo et al., 2023; Vöcking et al., 2017). Therefore, here we use the abbreviated forms r-opsin and c-opsin to indicate these groups.

1.4 Opsin gene structure

The gene structure of opsins depends on the opsin subfamily, but it can also vary among taxa, with many shared introns within the same subfamily being lost over time (Döring et al., 2020) (Figure 1.3). C-opsins retain three introns in common, while xenopsins retain two (Döring et al., 2020). Members of group-4 opsins possess six shared introns, which are conserved among them. Specifically, two peropsin introns are conserved in RGRs, and three are conserved in neuropsins. The number of introns in r-opsins can vary more widely. For example, r-opsin 1 in *Octopus vulgaris* (usually referred to as *Octopus* rhodopsin) possesses two introns, while r-opsin 2 (usually referred to as non-canonical *Octopus* rhabdomeric opsin) shows three introns that are not shared with r-opsin 1. To further highlight this diversity, the structure within *Octopus* r-opsin 1 homologous genes reveals that vertebrate r-opsins (usually referred to as melanopsins) possess nine introns, while arthropods UV-opsin five (Döring et al., 2020; Terakita, 2005).

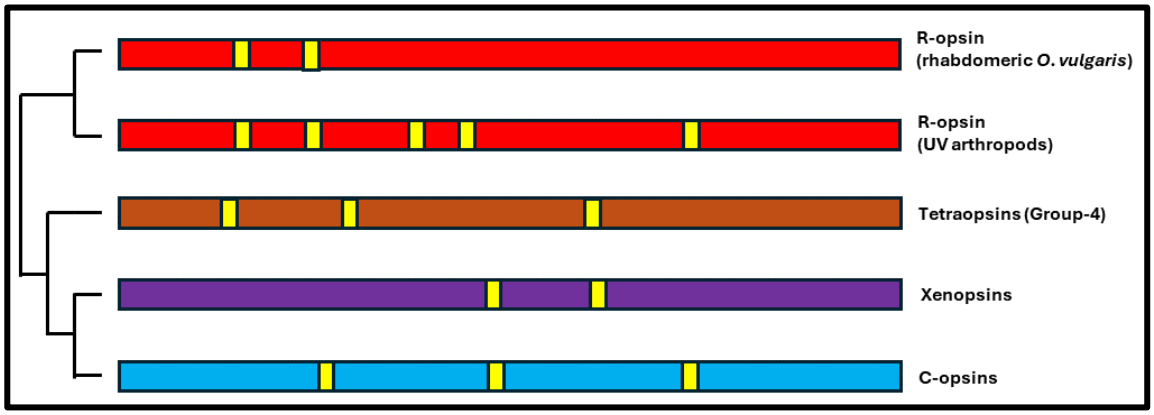


Figure 1.3 – Opsin gene structure. Schematic gene structure of the different opsin gene subfamilies in Bilateria showing conserved introns (yellow boxes). Intron position taken from Döring et al. (2020).

1.5 Opsin taxonomic distribution and phylogeny

Opsins are missing in poriferans and placozoans but are present in cnidarians, ctenophores, and bilaterians, indicating a deep origin in Eumetazoa (Feuda et al., 2012; Fleming et al., 2020). In particular, cnidarians possess, c-opsins, and xenopsins but lack members of group-4. Furthermore there is a group of cnidarian opsin that might have diverged early during opsin phylogeny (McCulloch et al., 2023). In some phylogenetic studies the same group lies within r-opsin, and their position within the opsin tree remain uncertain (Feuda et al., 2012; Fleming et al., 2020). Similarly, ctenophore opsins were previously positioned at the base of the opsin tree; however, this is now considered a phylogenetic artifact, and they are better recognized as c-opsins (Feuda et al., 2012, 2014; Fleming et al., 2020; McCulloch et al., 2023).

With few specific exceptions in cnidarians, all the different opsin subfamilies are distributed among all the main Bilateria branches, both in protostomes and deuterostomes, indicating they were present in their common ancestor and were only subsequently lost in different phyla or entire evolutionary lineages. For

example, deuterostomes lack xenopsins (Arendt, 2017). Given their ancient and complex evolutionary history, it is evident that different models of evolution and datasets produce different phylogenetic hypotheses. Furthermore, new groups have been described, such as bathyopsins (or echinopsin B) and chaopsins (or echinopsin A) (D'Aniello et al., 2015; Ramirez et al., 2016). These groups are better recognized in this thesis as early divergent c-opsins and r-opsins (Rawlinson et al., 2019; Vöcking et al., 2022).

In general, when performing opsin phylogeny, melatonin receptors and placopsins are often used as outgroups (Feuda et al., 2012). Many works put c-opsins in a closer relationship with Group-4 opsins and consider r-opsins early divergent (Bonadè et al., 2020; Davies et al., 2010; Feuda et al., 2012, 2014; Fleming et al., 2020; McCulloch et al., 2023; Rawlinson et al., 2019; Yoshida et al., 2015). Xenopsins, on the other hand, are sometimes found to be closer to c-opsins (McCulloch et al., 2023; Vöcking et al., 2017, 2022; Yoshida et al., 2015) and other times closer to Group-4 (Bonadè et al., 2020; Ramirez et al., 2016; Rawlinson et al., 2019) (Figure 1.4).

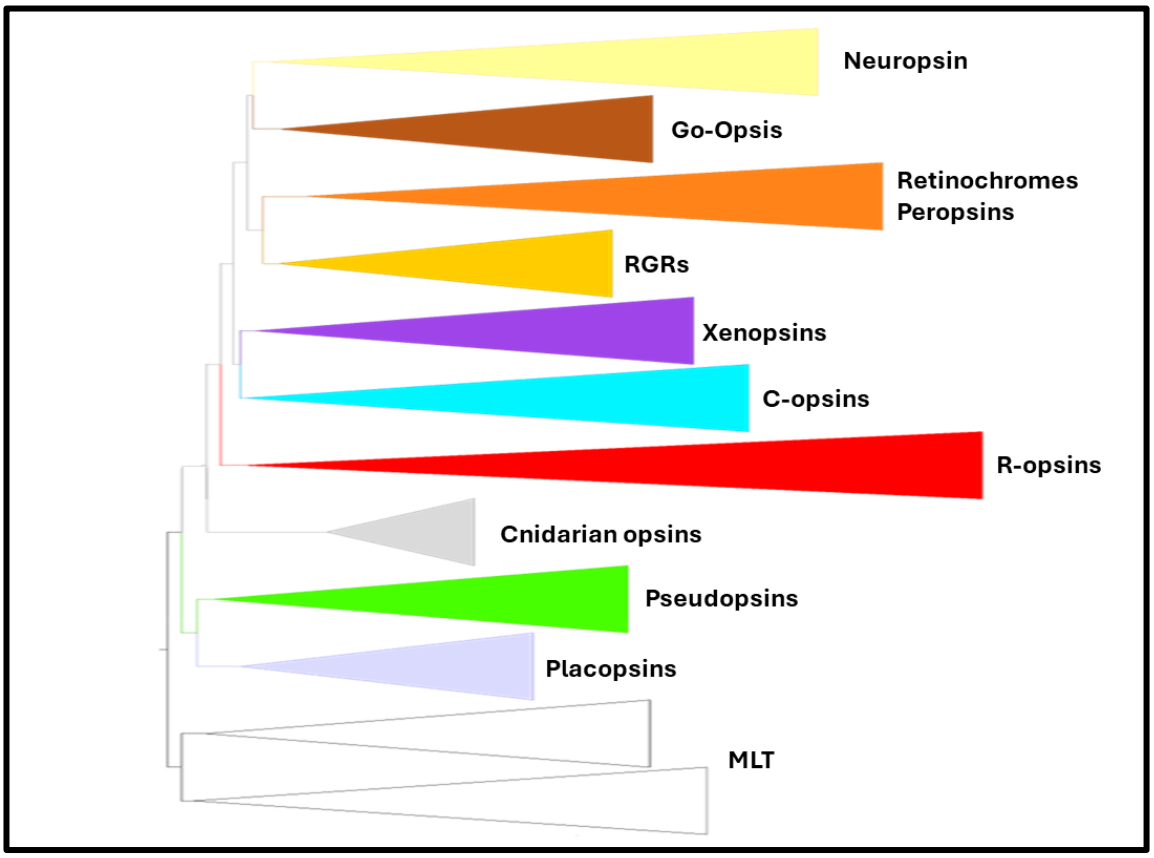


Figure 1.4 – Opsin phylogenetic relationship. Opsin subfamilies and their phylogenetic relationship according to De Vivo et al. (2023). Melatonin receptors (MLT), placopsins and pseudopsins are used as outgroup.

Considering the previously cited studies, it is possible to go deeper and identify distinct clades within opsins (Bonadè et al., 2020; Davies et al., 2010; Feuda et al., 2014; Fleming et al., 2020; McCulloch et al., 2023; M. Ramirez et al., 2016; Rawlinson et al., 2019; Vöcking et al., 2022; Yoshida et al., 2022). Within r-opsins, there is a clade often referred to as canonical r-opsins, containing the visual opsins of arthropods (UV/SW, MW and LW), cephalopods, and other invertebrates that have rhabdomeric receptors, and vertebrate melanopsins. A second r-opsin group, the non-canonical r-opsins, is composed of one or various clades of functionally uncharacterized r-opsins found in protostomes. A third group might be represented by cnidarian r-opsins, of which phylogenetic remain uncertain. Group-4 is divided into four different clades: RGR in a sister relationship with peropsin/retinochrome, and neuropsins in a sister relationship with Go-opsins. Xenopsin is composed of

four clades: xenopsin 1, xenopsin 2, and possibly two cnidarian opsin clades. C-opsins contain one cnidarian clade, one clade of ctenophore opsins, one or more clades of protostome c-opsins (which might include vertebrate tmt and opn3), four clades of non-visual vertebrate c-opsins (parietopsin, parapinopsin, vertebrate ancient opsin, and pinopsin), and one clade of visual opsins (LWS/MWS, SWS1, SWS2, RH1, and RH2).

1.6 Opsin protein structure

As previously mentioned, opsins are GPCR-coupled receptors with the typical seven transmembrane domains (TM) shared among the group (Terakita, 2005). To be more precise, opsins belong to the rhodopsin superfamily (PFAM00001), which includes other gene families containing more than 670 different receptors only in humans, such as histamine receptors, β-adrenergic receptors, and melatonin receptors (Mickael et al., 2016). Being members of the same superfamily, they share the same three-dimensional structure as well as some amino acid residues. These shared amino acid residues constitute about half of those conserved among opsins (Terakita, 2005). For convenience, as previously done for the retinal binding site, all the conserved residue positions are mapped using cow rhodopsin as a reference (Figure 1.5).

Regarding the three-dimensional structure, opsins generally present an extracellular amino terminus and a cytoplasmic carboxyl terminus (Terakita, 2005). There are three cytoplasmic loops (CL) and three extracellular loops (EC). On the N terminus chain, there are two amino-terminal glycosylation sites at residues 2 and 15. On the opposite side of the protein, on the C-terminus chain, there is the cytoplasmic helix HVIII and two carboxyl-terminal palmitoylation sites at residues 322 and 323. Three conserved cysteine (C) residues at positions 110 (TMIII), 185

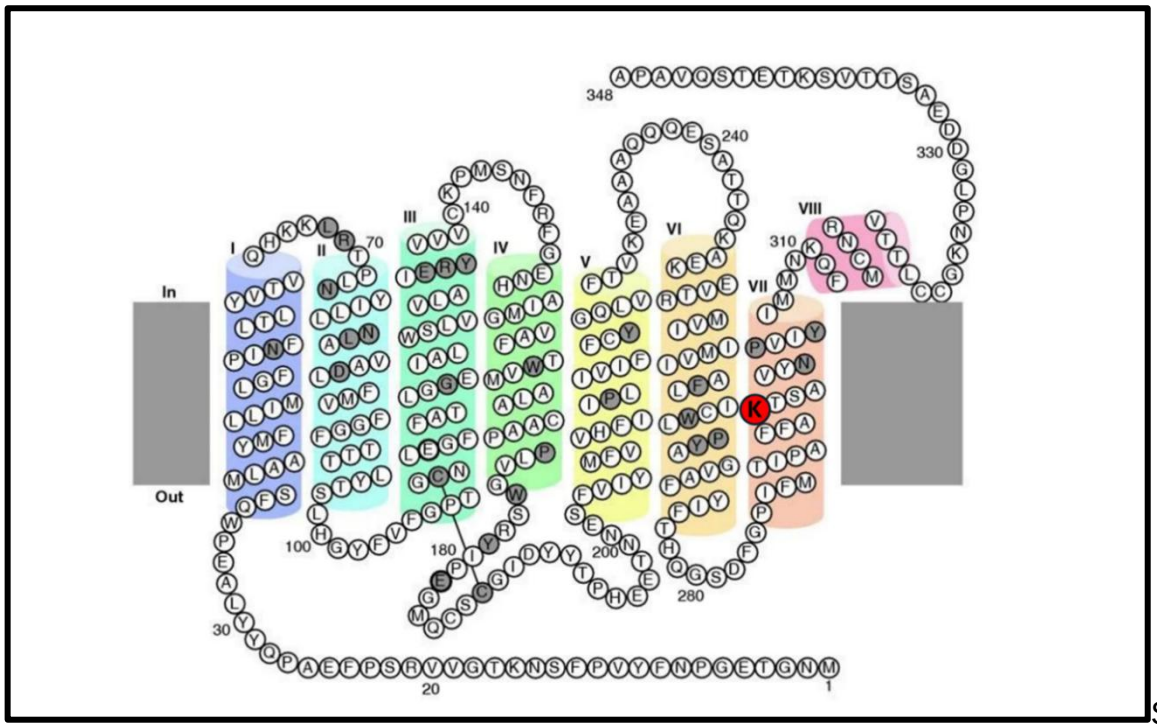
(ECII), and 187 (ECII) are involved in disulfide bond formation (Terakita, 2005). Opsins might function in dimer, and the dimer interaction between two opsins is supposed to occur between TMI and HVIII (Zhang et al., 2016).

The most important residue shared by opsins is the retinal binding domain K296, located on TMVII, which is crucial for their photoreceptive activity (Terakita, 2005). The retinal is affected by severe steric constraints due to the structure of the retinal binding site. Despite this, opsins might have roles beyond photoreception, such as mechanoreception, thermoception, or chemoreception (Feuda et al., 2022). In some opsin orthologues, such as gluopsin (Gühmann et al., 2022) and possibly placopsin, the retinal binding site might have been lost or shifted elsewhere in the sequence, though the function of placopsin and gluopsin remains uncovered (Feuda et al., 2012; Gühmann et al., 2022).

Another important residue is the counterion, which stabilizes the protonated Schiff base by interacting with a negatively charged amino acid residue. This is generally the glutamic acid E181 (ECII), which in c-opsins is replaced by E113 (TMIII). In other opsins, position 113 is often occupied by a tyrosine (Y113). There is an ERY/DRY motif at position 134 (TMIII), residues 134 to 136, which stabilizes the inactive opsin molecule by providing a negative charge (Terakita, 2005). Finally, there is a conserved NPxxY(x)5,6F motif at residues 302 to 313, which contains an NKQ/HPK motif (310 to 312) that maintains the structural integrity of the protein after photopigment activation (Davies et al., 2010; Terakita, 2005).

Additionally, sites surrounding the retinal binding pocket have the potential to change the distribution of the protonated Schiff base positive charge along the chromophore polyene chain in both the unexcited (11-*cis*) and excited (all-*trans*) states (Vöcking et al., 2022). These changes significantly affect the amount of energy necessary to electronically excite the retinal and cause the conformational change, thereby altering the absorption peak (Hagen et al., 2023a; Liénard et al.,

2022). Due to this crucial ability, these sites are known as key tuning sites. Their location can vary between different opsins, and it is unknown to what extent this is conserved among different taxa or closely related paralogues. Key tuning sites are well characterized in the visual c-opsin of vertebrates, but not much is known in other animal groups.



S

Figure 1.5 – Conserved opsin structural sites mapped on cow rhodopsin. In red, the key tuning site K296. Modified from Terakita (2005).

1.7 Opsin visual cycle and signalling cascade

Opsins can be bistable or monostable. Visual opsins of vertebrates are monostable, meaning that the retinal bleaches and is released from the protein upon the conformational change (Tejero et al., 2024; Terakita, 2005). On the other hand, cephalopod and arthropod visual opsins, as well as their homologs in vertebrates, the non-visual melanopsin, are bistable, meaning that the retinal does not bleach and reverts its configuration from all-*trans* to 11-*cis* after a second photon interaction

(Koyanagi et al., 2008; Koyanagi & Terakita, 2014; Terakita et al., 2012). This occurs because, in bistable opsins, the Schiff base is not deprotonated and hydrolysed. This mechanism is prevented by the presence of the tyrosine residue Y113 instead of glutamine, E113, in bistable opsins (Tejero et al., 2024).

Once the conformational change occurs, the opsin TMV and TMVI are pushed toward the cytoplasmic side, exposing their hydrophobic chains to interact with the G-protein (Scheerer et al., 2008). This particular conformation is known as metarhodopsin. Depending on the opsin type, it will bind a different G-protein, leading to a different signalling cascade. Binding of the opsin with arrestins will interrupt the signalling cascade. A list of G-protein/opsin interactions taken from Vöcking et al. (2022), including the final outcome (depolarization and hyperpolarization of the membrane), can be found in the table below.

OPSIN	G-protein	Enzyme	Mechanism	Ion Channel	Leads to	Stability	Photoisomerases
C-opsin	Gt/Gi/Go	PDE/GC	cGMP/cAMP	CNG	Hyperpolarization	monostable	
Xenopsin	Go/Gt?	AC?	cAMP?	CNG?	Hyperpolarization ?	Monostable ?	
RGR							Yes
Peropsin	Gi	AC	cAMP	CNG	Hyperpolarization ?	Monostable ?	
Retinochrome							Yes
Go Opsin	Go	GC	cGMP	CNG	Hyperpolarization	bistable	
Neuropsin	Gi	AC	cAMP	CNG?	Hyperpolarization	bistable	
R-opsin	Gq	PLC	PIP	TRP	Depolarization	bistable	

Table 1.1 – The diversity of opsin signalling cascade. The table shows the different opsin subfamilies, including the known G-protein that they bind, their function their signalling cascade leading to hyperpolarization or depolarization of the membrane. Some opsins do not show a signalling activity but are well known restore the retinal configuration (photoisomerases). Question mark (?) indicates uncertainty.

Since the signalling cascade is well-characterized in the visual opsins of vertebrates, such as the c-opsins (Gt-signalling cascade), and the visual opsins of

many invertebrates, such as the canonical r-opsins (Gq-signalling cascade), these processes will be illustrated in more detail (Vöcking et al., 2022; Yau & Hardie, 2009).

Focusing on the rod cells of vertebrates, the c-opsin signalling cascade has been exhaustively described (Vöcking et al., 2022; Yau & Hardie, 2009). In its metarhodopsin state, rhodopsin activates the subunit Gα of the Gt-protein (transducin), which in turn activates a cGMP Phosphodiesterase (PDE). By hydrolysing cGMP, PDE causes a decrease in cGMP levels in the cell. This leads to the closure of nucleotide-gated channels (CNGs), inhibiting the entrance of Na^+ and $Ca2^+$ ions into the cell, thereby leading to hyperpolarization.

To understand the canonical r-opsin phototransduction cascade, *Drosophila* visual opsins have been extensively studied (Hardie & Juusola, 2015; Vöcking et al., 2022). Gq is activated by the opsin in its active state metarhodopsin. This leads to the activation of the enzyme phospholipase C (PLC), which interacts with a membrane phospholipid, PIP2. PIP2 is then cleaved, generating diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3). Despite the function of DAG and IP3 still not being well understood, this leads to the opening of TRPC and TRPL channels. It is possible that PIP2 cleavage, by altering the membrane structure, causes the opening of the channels (Hardie & Juusola, 2015). This allows Ca^{2+} and Na^+ ions to enter the cytoplasm, leading to the depolarization of the membrane and finally, protein kinase C (PKC) activation that generate negative feedback by inactivating PLC. In both c-opsins and r-opsins, arrestins are involved in blocking the signalling (Terakita et al., 2012), a similar mechanism has been described in cephalopods (Figure 1.6).

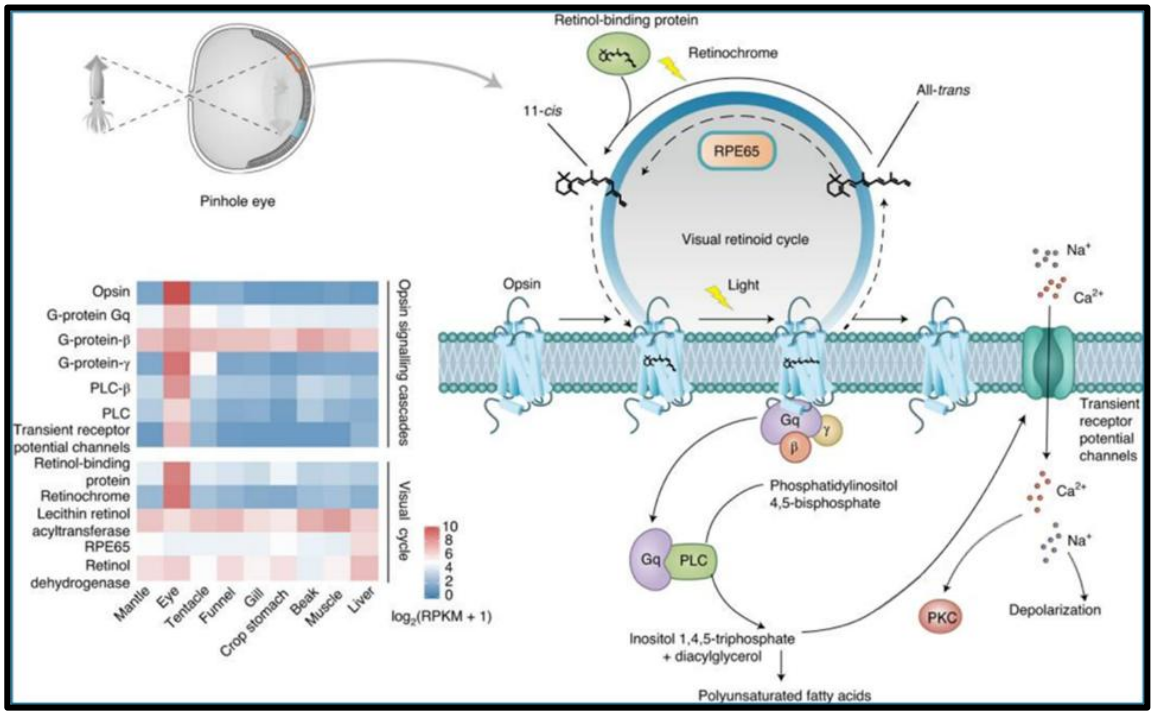


Figure 1.6 – Visual r-opsin signalling cascade in cephalopods. Modified from Zhang *et al.* (2021). Creative Commons Attribution 4.0 International License.

A special case are retinochromes and RGRs, which do not interact with G-proteins (Terakita et al., 2012; Vöcking et al., 2021, 2022); these opsins are known as photoisomerases and bind the retinal in all-*trans* to restore its configuration to 11-*cis*. This mechanism is well-known in the *Octopus* retina, which expresses retinochrome. RGRs, on the other hand, is expressed in vertebrates' Müller cells, which send the 11-*cis* retinal to cone and rod photoreceptor cells (Terakita, 2005; Yau & Hardie, 2009) (Figure 1.6). Another mechanism for retinal regeneration in vertebrates is an enzymatic pathway that involves retinal pigment epithelium cells and are well-described (Vöcking et al., 2022). Similar enzymatic mechanisms are not known in cephalopods, which are the main focus of this dissertation.

1.8 Opsin spectral tuning mechanism

Different opsin molecules absorb more effectively specific wavelengths depending on their molecular structure and their amino acidic sequence (Liénard et al., 2022; Maximov, 2000; Terakita et al., 2012). In other words, different opsins show a different absorption peak or λmax and changes in the protein sequence determine a shift of the λmax, causing what is known as spectral tuning. This is because the λmax depends on the energy required to isomerize the chromophore, which is more or less stable depending on the amino acids surrounding it (Hagen et al., 2023).

Amino acid sites known to be involved in shifting the λmax are known as key tuning sites (Jacobs, 2009). They are often identified through *in silico* evolutionary comparative analysis (for example, selection analysis) and confirmed using *in vitro* or transgenic-induced mutagenesis experiments (Dungan et al., 2016; Feuda et al., 2016; Hagen et al., 2023b; Liénard et al., 2022; Ricci et al., 2022, 2023). Adding an intermediate step and mapping the detected mutations on the protein's 3-D structure can also be useful to understand the molecular mechanism and make predictions before proceeding with mutagenesis experiments (Dungan et al., 2016; Liénard et al., 2022).

In general, the spectral sensitivity is often affected by the addition or removal of hydroxyl-bearing amino acids, such as serine (S), tyrosine (Y), and threonine (T) (Hagen et al., 2023). These amino acids change the distribution of the protonated Schiff base positive charge of the retinal on the polyene chain. Mutations that add an OH-bearing amino acid near the beta-ionone ring distribute the positive charge on the distal side of the retinal, maintaining its excited state. Therefore, less energy (higher wavelengths) is required for isomerization to all-trans, causing a shift of the absorption peak toward red light. Conversely, mutations that add an OH-bearing amino acid near the retinal binding sites stabilize the retinal in its 11-*cis* configuration. More energy (lower wavelengths) is required for isomerization, causing a shift toward UV-blue light. This is a general and common phenomenon in

many species with a clear and well-understood mechanistic basis, although other mutations involving non-OH-bearing amino acids are known. For example, mutations causing the loss of electrically charged amino acids, such as D83N and E199Q, shift the λmax of vertebrate RH1 towards lower wavelengths (Hagen et al., 2023). Mutations involving changes between valine, leucine, and isoleucine have also been observed in species living in different light environments, but it is unclear if they are involved in changes of λmax (Chung & Marshall, 2016; Hagen et al., 2023; Ricci et al., 2022). Through secondary interactions, distant sites might also contribute to changes in the absorption peak, such as the loss of disulfide bridges between two amino acids (Musilova et al., 2019).

Other mechanisms of spectral tuning that do not involve changes in the amino acid sequence have been described in vertebrates (Hagen et al., 2023). These often involve pre-filtering of light before it reaches the molecular photoreceptors, such as the involvement of pigments in the tapetum or in the lens, or the presence of oil droplets within the photoreceptor. It must be noted that in vertebrate eyes, light crosses the photoreceptor cell body before reaching the cone or rod (reverse eye), whereas in invertebrates such as arthropods and cephalopods, the eye is not reversed, and the rhabdomere is first reached by the light, so intracellular oil droplets cannot act as effective light filters (Hanke & Kelber, 2020). Although colour vision in octopods has yet to be proven, it has been speculated that octopuses might distinguish distinct colours using chromatic aberrations created by the shape of their pupil, otherwise they might rely on discrimination of polarized light (Hanke & Kelber, 2020; Stubbs & Stubbs, 2016; Temple et al., 2021).

Another interesting mechanism to change the λmax requires the use of a more or less stable chromophore. For example, the use of 11-*cis*-3,4-Didehydroretinal, which is less stable and isomerizes more easily to all-*trans*,

instead of the classical 11-*cis* retinal, is a mechanism described in crocodiles to accommodate light in freshwater environments (Hagen et al., 2023).

1.9 Rhodopsin evolution in fishes

Vertebrate rhodopsin (RH1) is the first G-coupled receptor (GPCR) protein with a fully resolved amino acid sequence and crystal structure, it is well characterized and used in multiple occasions as an experimental model to understand key tuning mechanisms (Tejero et al., 2024; Terakita, 2005; Terakita et al., 2012). In RH1 more than 13 key tuning sites are known, some of which are shared with other vertebrate visual opsins (Hagen et al., 2023b). Therefore, it is an excellent example to thoroughly comprehend opsins' evolution in varying light environments, and to provide a parallel to the study of octopods' adaptation to the deep sea.

The majority of light wavelengths are filtered or scattered in the water column, with only blue light reaching the deep sea (Maximov, 2000; Warrant & Locket, 2004). Consequently, RH1 has shifted from green to blue in deep sea fishes, with a λmax change from ~500 to ~470 nm (Musilova et al., 2019; Ricci et al., 2022). In some groups, gene duplications occurred, and the absorption peak of each duplicated gene diversified to better cover and discriminate between all the spectrum available in the environment (λ = 460 to 490 nm) (Musilova et al., 2019; Ricci et al., 2022). This adaptation has convergently emerged in other marine vertebrate groups (Hagen et al., 2023; Musilova et al., 2019; Ricci et al., 2022). In sunlight-depleted zones, bioluminescence becomes the major source of selective pressure. Diretmidae, a family of abyssal fish, possess 37 RH1 opsin paralogs that can detect and discriminate different peaks between a range of λmax fitting the emission range of bioluminescent organisms (λ = 420 to 520 nm) (Musilova et al., 2019)).

Additionally, within vertebrates, opsin expression of one or the other paralogue might also depend on environmental conditions. In turbid waters, such as those of rivers and lakes, even shallow environments are deprived of different wavelengths (Zhang et al., 2000). To cope with that, migratory fishes such as the common eel (*Anguilla anguilla*) usually retain two copies of RH1, and express one or the other opsin depending on the developmental stage. The young stage that lives in freshwater expresses the hight wavelength (λmax ~501 nm) paralogue, while the adult stage migrating to the open sea expresses the low wavelength one (λmax ~482 nm) (Zhang et al., 2000).

It is important to highlight that teleost fishes, due to a third round of whole genome duplication, possess and retain more gene copies (Hagen et al., 2023). Tandem duplications, retrotransposition events, and evolutionary phenomena like pseudogenization, gene loss, and gene conversion have further increased the complexity of this evolutionary scenario, leading some species to possess more than 30 copies of RH1, as well illustrated in the Diretmidae reported above (Musilova et al., 2019). In contrast, duplications of opsin genes in cephalopods are rare, and there are no reported duplications involving the main visual opsin group. Therefore, while studying visual opsins in fishes provides a well-established foundation to understand deep sea adaptations in other groups, it must be noted that vertebrate visual opsins are more evolutionary dynamic than those of other groups, such as cephalopods.

1.10 Anatomy of cephalopods visual system

Cephalopods have one of the most complex and well-developed visual systems in animals, comparable only to those of vertebrates and arthropods (Hanke & Kelber, 2020). Their eyes are of the camera type, a structure that has convergently evolved

and resemble that of vertebrates (Yoshida et al., 2015). They can be very large, as in the giant squid *Architeuthis dux*, which eyes can reach up ~ 270 mm in diameter, making them the largest eyes ever measured in living organisms, approximately the size of a football ball. In littoral octopod species, such as *Octopus vulgaris*, the eye diameter can be up to ~20 mm in an average-sized animal (~200-1000 g), while in some deep-sea species the eyes are larger, for example in *Pteroctopus tetracirrhus* (~500 g) it is ~35 mm. Despite not being record numbers, these sizes match those of visual animals such as humans (~23 mm) or owls (~23-29 mm), highlighting the importance that these organs have in cephalopods (Hanke & Kelber, 2020). Each eye moves independently and have a large monocular visual field. *Octopus* can perform elevate rotating movement and possesses extra-ocular muscles independently innervated. Eyelid and corneas (pseudo-corneal) are constituted by the skin folding and surrounding the eye (Hanke & Kelber, 2020).

Going deeper into the anatomy (Figure 1,6), cephalopods pupils are one of the most peculiar traits, as they can change in shape, from rounded to square or peanut-like, depending on light conditions (Soto et al., 2020). The lens is divided into an anterior and a posterior part by a septum, and can be accommodated to adjust the focus by a circular or ring-shaped muscle (Hanke & Kelber, 2020). The *Octopus* lens does not seem to correct for chromatic aberration and this, together with the shape of the pupil, have been interpreted as possible mechanism for colour discrimination.

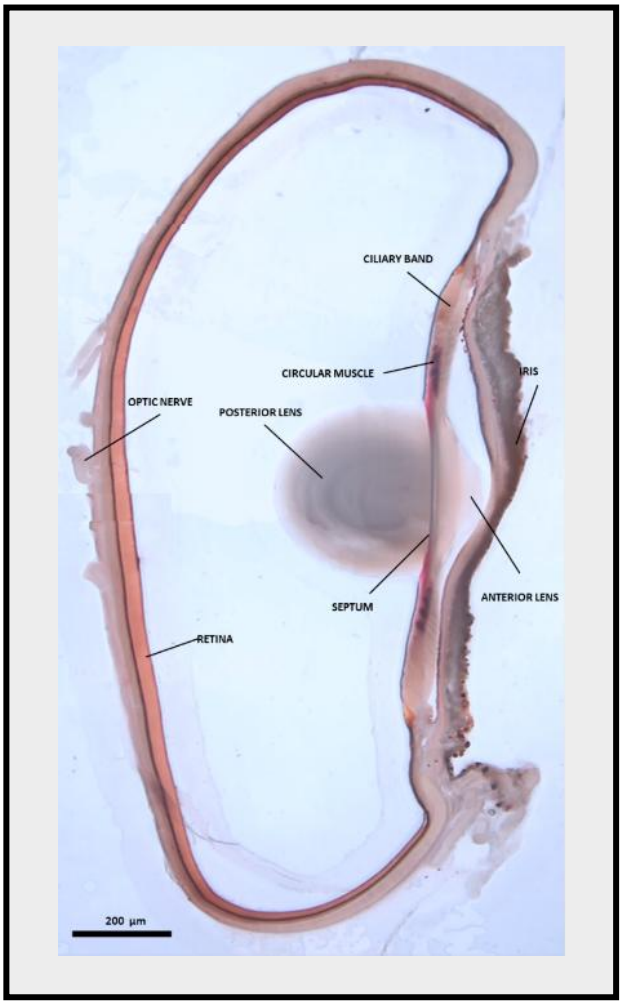


Figure 1.7 - Section of *Octopus vulgaris* eye. The picture shows the different structures present in the *Octopus* eye. A pigmented iris, a lens divided in posterior and anterior by a septum, a circular muscle in the ciliary band involved in accommodation processes, and a single layer retina with axons forming the optic nerves.

In the case of cephalopods, and therefore octopods, the eye is not inverted, meaning that the light reaches directly the photoreceptive side of the photoreceptor cell, without crossing its cellular body as it occurs in vertebrates (Figure 1.8). In other words, while the rhabdomeres of the octopus eye point toward the light, the rod and cones of the vertebrate eyes point to the opposite direction (Figure 1.8). In addition, while rod and cones are cells with well-developed and modified cilia, rhabdomeric cells are characterized by an expansion of the cellular membrane (rhabdomere) hosting various microvilli. Other differences are that, in vertebrate eyes, the signal departing from the photoreceptor is first transmitted to an intermediate layer of

retinal cells (e.g., the bipolar cells) and then to the neurons (or ganglion cells) that form the optic nerve. Since in vertebrates the eye is inverted, the optic nerve originates on the side of the vitreous body (the eye inner space) and needs to cross the retina to reach the brain, generating a retina depleted area known as blind spot. In contrast, not being inverted, the axon of cephalopods eyes generates directly from the photoreceptor cells and depart from the outer portion of the retina to form optic nerves that reach directly the optic lobe. Therefore, in cephalopods there are no cells intermediating the signal between the photoreceptors and the brain, and this role is covered directly by the optic lobe. Indeed, both the cephalopod optic lobe and the vertebrate inner retinal cells (e.g., bipolar) express opsins (Yoshida et al., 2015). Furthemore, in vertebrates the optic nerve reaches the opposite hemisphere (contralateral), while the octopus optic nerve is connected to the optic lobe of the same side (ipsilateral hemisphere).

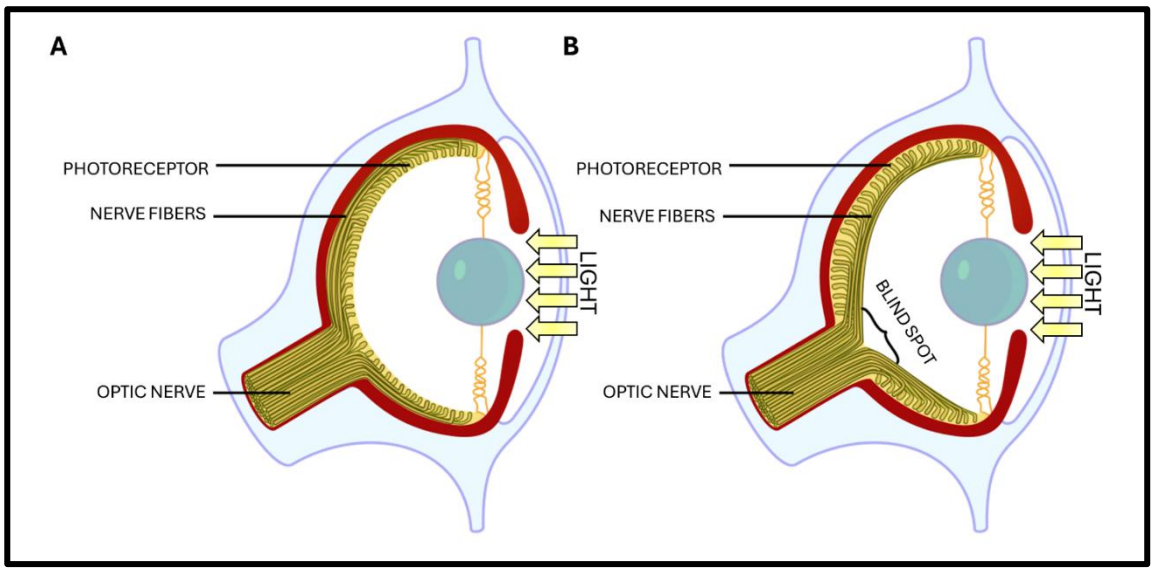


Figure 1.8 – Cephalopods and vertebrate camera eyes. Difference between the invertebrate camera eye (A) and the vertebrate reverse camera eye (B). In the vertebrate reverse camera eye, the nerve fibres are positioned internally, meaning that light must pass through them before reaching the photoreceptors, which are oriented toward the outside. To reach the brain, the optic nerves must cross the retina, creating a blind spot. This does not occur in invertebrate camera eyes, where the photoreceptors are oriented internally,

and the nerve fibres are on the external side of the retina. Modified from @wikisource CC BY-SA 3.0.

The cephalopod retina is internally delimited by an acellular membrane shell, made by the secretion of supporting cells (Figure 1.9) (Hanke & Kelber, 2020). Underneath that, there are densely packed photoreceptor cells carrying one rhabdomere each with a clear geometric disposition: each rhabdomere is divided in two rows facing opposite sides separated by a pigmented layer. Four rows of four different rhabdomeres are grouped in quartets with a supporting cell in the centre. Rhabdomeres are also displaced perpendicular to each other, indicating detection of polarized light. Pigment migration in supporting cells or in the rhabdomere can vary depending on light intensity and are an adaptation of dark/light conditions. Moving toward the external side of the retina, there are the rhabdomeric cells' bodies with nuclei and a supporting glia (Hanke & Kelber, 2020). Axons depart from the cell body forming the retinal plexus, where they interact with other axons originating in the centrifugal cell of the optic lobe. The rhabdomeric cell axons then join to form the optic nerves. Cephalopods lack of fovea, a depleted region of the vertebrates' retina where light is projected by the lens; instead, they show an increase in photoreceptors density in the horizontal region. The nerves originating from the ventral part of the eye interact with the dorsal part of the optic lobe, while those originating from the dorsal part of the eye interact with the ventral part of the optic lobe. Altogether, the optic nerves interact with more than four-fifths of the optic lobe's surface, penetrating its outermost layers. Because of this peculiar structure, the most dorsal and ventral optic nerves are longer (up to 20 mm) and cross in the centre, forming a medial chiasma.

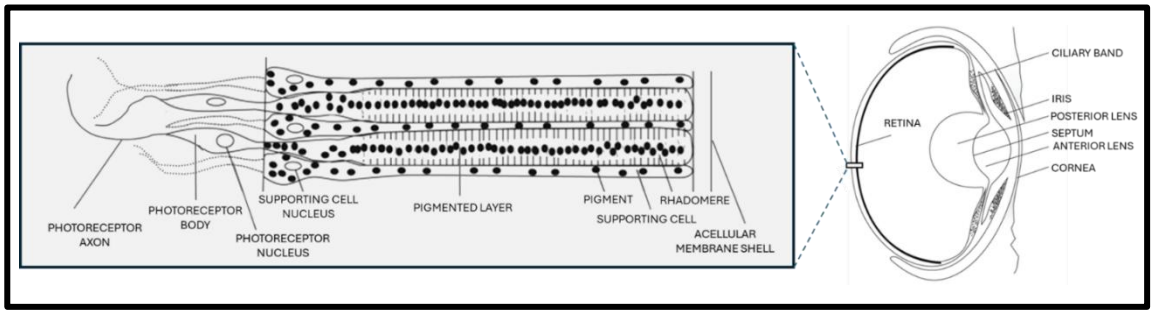


Figure 1.9 - Detail of *Octopus vulgaris* retina. The picture shows the different structures in the *O. vulgaris* retina and a scheme of the eye anatomy. The single layer retina is made by photoreceptive cells bearing one rhabdomere divided in 2 rows separated by a pigmented layer, between the rhabdomeres there are supporting cells, and an acellular membrane shell protects the structure from the inside. The cell body of the photoreceptive cells is placed underneath the rhabdomere and bear axons. Modified from Hanke and Kelber (2020).

The cephalopod optic lobe is an important organ of the central nervous system, as highlighted by its size (it is the largest nervous structure in the cephalopod) and the number of neurons it contains, 128,940,000 neurons in *O. vulgaris* (Young, 1962). Optic lobes lie on the orbit-shaped lateral structures of the cranial cartilage which protect the brain. It is generally bean-shaped, though in some species the anterior and posterior extremities are so developed that they confer what some authors call a ‘croissant shape’ (Pungor et al., 2023) (Figure 1.10). The optic lobe is differentiated into two layers: the cortex (or deep retina) and the medulla. In the outer area of the cortex, there is a granular layer, where the amacrine cell bodies are present, and more internally the plexiform layer, in which only axons are present including the nerves coming from the eye that penetrate and interact with the processes of small amacrine cells (Young, 1962). Then, the optic nerve terminations reach the inner part of the cortex, the inner granular layer (Pungor et al., 2023; Styfhals et al., 2022; Young, 1962). Here, the nerve terminations transmit the signal to the bipolar cells, which finally send it to medulla. Finally, in the medulla the visual signal is integrated and transmitted to the vertical lobe of the brain, where the sensory signals are processed.

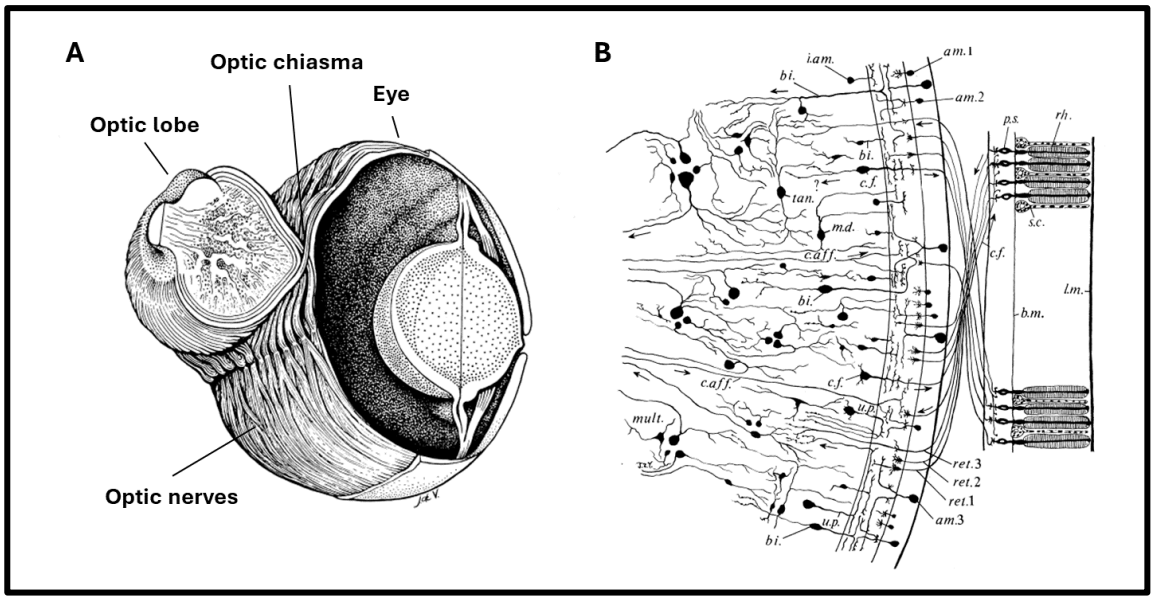


Figure 1.10 - Optic lobe and eye in *Octopus vulgaris*. A illustrates the interaction between the eye and the optic lobe, where the optic nerves cross to form a chiasma before reaching the opposite sides of the optic lobe. **B** shows these interactions in detail: the axons of the photoreceptive cells depart from the eye and penetrate the first layer (the plexiform layer) of the optic lobe, where they interact with the small amacrine cells and transmit the signal to the bipolar cells. The bipolar cells, located in the deepest part of the cortex, then transmit the signal to the innermost part of the optic lobe, the medulla. Amacrine cell bodies are mostly in the plexiform layer, while the bipolar cells originate in the inner part of the cortex. Centrifugal cells also originate here, and their axons cross the plexiform layer to reach the retina.

Legend: am. - amacrine cell; b.m. - basal membrane; bi. - bipolar cell; caff. - axon afferent from the medulla to the plexiform layer; cf. - centrifugal cell; i. am. - internal amacrine cell; lm - acellular membrane layer; md. - multi-dendrite cells; mult. - multipolar cell; p.s. - (Uribe & Zardoya, 2017) supporting cell; tan - tangential cell of outer medulla; up. - unipolar cell. Modified from Young (1961).

1.11 Colour vision, brightness contrast, and polarized vision in cephalopods

Cephalopoda are known for their exceptional camouflaging abilities, which enable them to match the colour of their surrounding environment ((Hanlon, 2007). Despite this, based on the presence of a single type of photoreceptor expressing one visual

opsin paralogue (often referred to as Octops rhodopsin, here r-opsin1) (Hara et al., 1967) and behavioural studies (Hanke & Kelber, 2020; Hanlon, 2007; Stubbs & Stubbs, 2016), it has become evident that they rely on monochromatic vision and, therefore, are considered colour blind (Hanlon, 2007). On the other hand, evidence suggests that their camouflaging ability is so effective that it can easily deceive trichromatic predators, leading to the hypothesis that cephalopods might indeed possess some form of colour vision (Chiao et al., 2011). Some authors have argued that chromatic aberration could, at least in part, explain their ability to perceive color and match their environment (Stubbs & Stubbs, 2016). IAccording to this model, by altering the shape of the pupil, light is refracted and split into different wavelengths, which then hit the retina in different areas. This model would also explain why cephalopods fail to discriminate colours in tests where brightness is uniform, as they cannot gain spectral information from a flat-field background or other similar assays (Mäthger et al., 2006; Stubbs & Stubbs, 2016). Despite this model, the ability of cephalopods to discriminate colours remains an open question.

On the other hand, it is universally accepted that octopods rely on polarized light to navigate and obtain visual cues from their environment (Hanke & Kelber, 2020; Temple et al., 2021). Sunlight is unpolarized, consisting of a mixture of wave packets with different polarizations (Marshall & Cronin, 2011). When sunlight encounter an object, the situation changes: some wavelengths are absorbed while others are scattered and reflected. When this occurs, light travels parallel to a single plane (linear polarization), becoming polarized (Marshall & Cronin, 2011). Water selectively modifies incident light wavelengths (see above) but does not affect polarization, meaning that polarized light can travel longer distances and remains a more reliable signal for aquatic animals, as it is available at any depth (Marshall & Cronin, 2011). Indeed, the parallel arrangement of rhabdomeres in the cephalopod eye allows them to be precisely organized into two channels of polarization

sensitivity, creating a highly sensitive polarization vision system(Temple et al., 2021). This has been extensively demonstrated in previous studies, where experimental assays have shown that cephalopods can respond even to small contrasts in polarization (Temple et al., 2021).

1.12 Aim of the project

The PhD project aims to check for the presence of opsin molecular adaptation in seven different incirrate octopod species living at different depts. The project focus on the species of the Gulf of Naples (Italy), which, given its differences in bathymetric levels (Passaro et al., 2016) and the various species of octopods inhabiting it, represent an excellent opportunity for evolutionary and comparative studies. Incirrate octopods are a group of specialized coleoid cephalopods with nektobenthic lifestyle. The sole exception are argonautids, represented by *Argonauta argo*, which are pelagic and have been chosen for this study as phylogenetic and ecological outlier (Uribe & Zardoya, 2017).The other species are littoral (*Octopus vulgaris*, *Callistoctopus macropus*, and *Eledone moschata*) and deep sea (*Eledone cirrhosa*, *Scaeurgus unicirrhus*, and *Pteroctopus tetracirrhus*) (Figure 1.11).

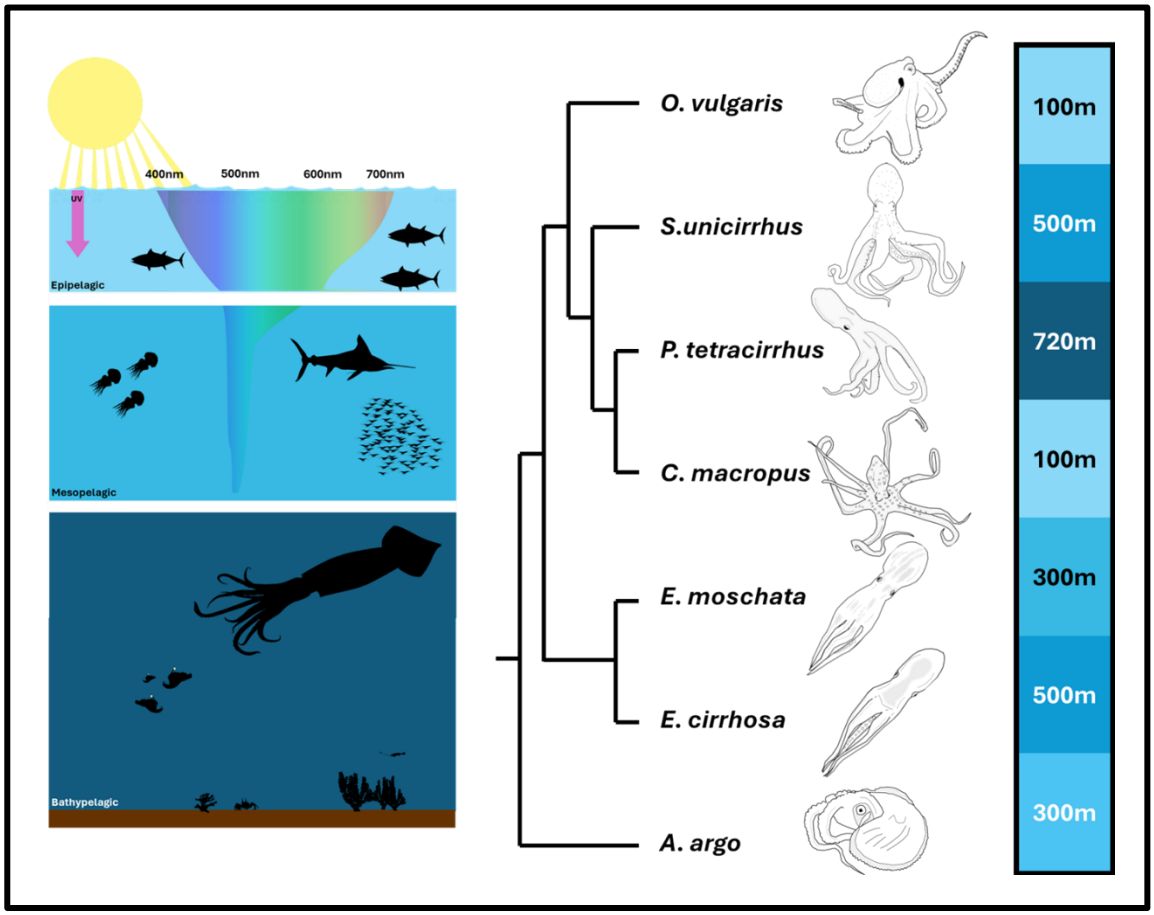


Figure 1.11 - **List of octopod species selected for the analysis**. The picture shows the different octopods species involved in the study including the maximum depth in which they have been reported.

While pelagic animals might move easily through the entire water column and experience different light realms that suits their need, predominantly benthic animals, like octopods, are subjected to the light that reaches the sea floor (Jereb and Roper, 2010), and therefore might experience a stronger selective pressure due to light conditions. Studies on cephalopod r-opsin1 adaptation to dim light environments have already been performed and while founding adaptations to light environment in other coleoid, no peculiar change in the peak of absorption of octopod r-opsin1 have been found. On the other hand, it must be noted that deep-sea octopod species (>200 m in dept) have not been included (Chung & Marshall, 2016), except for *Enteroctopus dofleini* which only migrate in the deep for reproductive reasons.

Further, while the expression of Octopus r-opsin1 and retinochrome is well-characterized in cephalopods' visual system (Bonadè et al., 2020; Chung & Marshall, 2016; Hara & Hara, 1972; Yoshida et al., 2015), other opsins are not, and this project also aim to reconstruct the number of opsins that the ancestral octopod possessed and if they are expressed in the visual organs, the eyes and the optic lobe.

CHAPTER 2 – Material and methods

2.1 Ancestral octopod opsin toolkit reconstruction

To assess the number of opsins potentially expressed in cephalopods, including octopods, we reconstructed the ancestral opsin toolkit for the group. During this process, it became evident that there was a lack of knowledge regarding the opsin content across different taxa within the superphylum Lophotrochozoa. This gap might lead to incorrect classification of opsins in phylogenetic analyses. Therefore, the analysis was extended to include the reconstruction of ancestral opsin toolkits in other taxa, such as other molluscs, annelids, bryozoans, brachiopods, platyhelminths, phoronids, and nemerteans, thus addressing a gap in the literature.

This paragraph of material and methods section is a more detailed version of what is reported in a previously published paper and all the analysis were conducted under the supervision of Dr Roberto Feuda from the University of Leicester, UK (De Vivo et al., 2023).

2.1.1 Data Collection

A total of 59 proteomes and 15 genomes from species representative of the main phyla within the Lophotrochozoa clade were downloaded from NCBI and UniProt. Augustus (training sequence set: *Homo sapiens*) (Stanke et al., 2004) was used to predict the coding sequences from the genome and obtain the relative proteome, making a total of 74 proteomes analysed. These included 22 molluscs, one

brachiopod, one phoronid, one nemertean, 13 annelids, two bryozoans, 32 platyhelminths, one orthonectid, and one acanthocephalan, with the last two taxa having more uncertain phylogenetic positions. To each proteome, genome and coding sequences was assigned the name following their species and taxonomy. For example, the fasta file of the *Octopus vulgaris* proteome was renamed "Octopus_vulgaris.aa," and all sequences were renamed by adding a reference indicating the taxon they belong to, followed by a number in the order they appear in the proteome. Thus, the first sequence listed was renamed ">Mol_Cephalopoda_Octopus_vulgaris_1".

2.1.2 Species Tree

From the 74 translated genomes, BUSCO analysis was performed. The single-copy orthologs dataset was extracted with BUSCO (-m protein -l metazoa_odb10 -c 20) (Seppey et al., 2019), and the computed BUSCO values were used to assess the completeness of each proteome. We used single copy orthologs to compute the species tree. For each species, single-copy orthologs were extracted and aligned using MAFFT (--retree 2) and assembled with FASconCAT (Katoh & Standley, 2013; Kück & Longo, 2014; Seppey et al., 2019). Phylogeny was computed using IQ-TREE (-m GTR20+G4 -B 1000 -redo -wbtl). The phylogenetic positions of Dinophilidae, Oweniidae, and Nemerteans were manually corrected using Mesquite3.31, and the tree was rooted using different roots (Bryozoa/Platyhelminthes or Orthonectids/Platyhelminthes (Carrillo-Baltodano et al., 2021; Maddison & Maddison, 2017; Martín-Durán et al., 2021). *Danio rerio* and *Trichoplax adherents* were also manually added to serve as outgroups.

2.1.3 Opsin Mining

To collect opsins, we used BLASTp (-evalue 1e-10) with the dataset from Feuda et al. (2012) as query. However, this BLAST result was insufficient to discriminate opsins within a genome, as many members of the Rhodopsin superfamily (e.g., melatonin receptors, glutamate receptors, histamine receptors, adrenergic receptors) were also collected. Usually, these sequences are filtered out by alignment, followed by the selection of only sequences with a lysine in the key tuning site. However, the goal of this research was not only to characterize opsins known in the ancestor of all main lophotrochozoan taxa but also to identify potential opsins in this group that might have lost the lysine in the retinal binding domain. Therefore, we implemented a pipeline that is not biased toward this specific opsin characteristic. First, TOPCONS (Bernsel et al., 2009) was used to identify the number of transmembrane domains, discarding sequences with fewer than three transmembrane domains. Secondly, all remaining sequences were assigned to different ortholog groups (orthogroups) using Broccoli (-phylogenies ml -threads 18) (Derelle et al., 2020), and each sequence within the group was annotated using EggNog (Huerta-Cepas et al., 2017). Broccoli performs ultra-fast phylogenetic analyses (Price et al., 2009) and creates a network of orthologous relationships. The different orthologous groups are then identified and separated using a parameter-free machine learning algorithm (Derelle et al., 2020). This method was used to cluster the different opsins into the same orthogroups. Subsequent annotation with EggNog helped identify non-opsin sequences by excluding orthogroups lacking at least one annotated opsin. To further confirm their identity, a CLANS (CLuster ANalysis of Sequences) analysis of all sequences in the orthogroups was performed (Frickey & Lupas, 2004). CLANS performs an all-vs-all BLAST and calculates

pairwise attraction values based on High-Scoring Segment Pair's (HSP) P-values (Frickey & Lupas, 2004).

From this, two different datasets were created: 1) "Opsin Dataset 1," for the opsin gene phylogenetic analysis, which included sequences assigned to melatonin receptors (MLT), a well-established opsin outgroup, along with a precompiled list of already annotated opsins from the literature, including placopsins; and 2) "Opsin Dataset 2," used for the reconciliation analysis, using only sequences found in the lophotrochozoan proteomes, with unstable sequences and the opsin outgroups removed.

2.1.4 Opsin Gene Phylogenetic Analysis

Opsin dataset 1 was analysed using IQ-TREE.2, with two separate phylogenetic analyses conducted (Minh et al., 2020). The first analysis was an ultrafast bootstrap (UFB) analysis (UFB1, -m GTR20+G4 -B 1000), and the second was a transfer bootstrap estimation (TBE) analysis (TBE, -m GTR20+G4 -b 100 --tbe). The TBE analysis was specifically used to identify potential rogue taxa. Rogue taxa are sequences that, by altering their position among otherwise well-supported branches, can reduce the bootstrap values. All sequences with an estimated transfer bootstrap index exceeding the 95_{th} percentile were removed, following the guidelines by Lemoine et al. (2018) and Minh et al. (2020). Subsequently, a second phylogenetic analysis was conducted with the unstable sequences removed (UFB2, -m GTR20+G4 -B 1000).

The GTR20+G4 model was chosen as it is considered optimal for opsin phylogenetic analyses, supported by multiple studies (D'Aniello et al., 2015; Feuda et al., 2012; Fleming et al., 2020). This model was confirmed to be the best fit through a best-fitting model analysis performed on our initial dataset, showing the lowest Bayesian Information Criterion (BIC) score of 743,801.83, compared to the

second-best model VT+R10 with a BIC of 743,825.94. The phylogenetic tree was rooted using melatonin receptors (MLT) and placopsin as outgroups. To account for the potential influence of outgroups, which might attract certain opsin groups and affect phylogenetic relationships, a third additional phylogenetic analysis was performed without MLT and placopsin (UFB3, -m GTR20+G4 -B 1000).

2.1.5 Reconciliation Analysis

Reconciliation analysis consists in mapping the gene tree topology onto the species tree and was used to understand the pattern of gene losses and duplications across the evolutionary history of opsins in these taxa. For the reconciliation analysis, *Danio rerio* opsin sequences and *Trichoplax adhaerens* placopsins were included as outgroups to Opsin dataset 2. A phylogenetic analysis was again performed using IQ-TREE.2 (UFB4, -m GTR20+G4 -B 1000), and the resulting phylogeny was reconciled with the species tree using GeneRax. The reconciliation was executed with the strategy of --max-spr- SPR 5 and using the UndatedDL reconciliation model (Morel et al., 2020). Each node of the species tree was annotated with the corresponding opsin subfamilies to elucidate the evolutionary dynamics, such as gains and losses, of these photoreceptive proteins.

2.3 Sample collection

Once we reconstructed the cephalopod opsin toolkit, with the aim of characterize those that are expressed in their visual system the eyes and the optic lobes of 6 species (*O. vulgaris*, *C. macropus*, *E. moschata*, *E. cirrhosa*, *P. tetracirrhus*, *S. unicirrhus*, and *A. argo*) have been collected. Additionally, proximal suckers of the right dorsal tentacle and retractor muscles were also collected, since suckers might

potentially show a particular opsin expression as suggested by the presence of rhodopsin-kinase gene (Al-Soudy et al., 2021), and the retractor muscle was intended to be used as control. Tissue from by-catch animals were dissected on an RNase-zap cleaned surface covered with aluminium foil and kept refrigerated using pieces of dry ice. Once a tissue sample was collected, dissection instruments were cleaned by immersing them in sodium hypochlorite, followed by Milli-Q water, and finally 75% ethanol before proceeding to collect other tissues to prevent potential cross-contamination. Once collected, tissues were immediately snap-frozen using liquid nitrogen, subsequently stored on dry ice, and then at -80°C.

2.4 RNA extraction

To perform RNA extraction, we developed a protocol that combines the Trizol/Trazol RNA extraction protocol, skipping the precipitation step, and adding an adapted cleanup protocol using the Sigma-Aldrich NucleoSpin Gel & PCR Clean-up kit. To completely remove possible DNA contamination, add DNase steps (see 2.4.3) using the material from Qiagen RNeasy Kits for RNA Purification, but other in column DNase protocols are suitable. RNA samples can be sent to a sequencing service for transcriptome library creation or alternatively, used for RT-PCR (reverse transcriptase PCR) and RT-qPCR analyses.

2.4.1 Trizol/Trazol Extraction

1. Collect 0.1-0.2 g of each tissue in separate 1.5 mL Eppendorf tubes along with 2 metal beads and 1 mL of TrAzol or Trizol under a chemical hood.
2. Homogenize using a tissue homogenizer (25 oscillations per second for 15 minutes) - the adapter should be pre-cooled at -20°C for at least 2 hours to

maintain a low temperature during the lysis procedure. Alternatively, homogenization can be done using an Ultra Turrax, although this was found to be less efficient.

3. Transfer the homogenized tissues in Trizol/Trizol to a new Eppendorf tube and add 200 µL of chloroform.
4. Gently mix for 3 minutes.
5. Incubate the sample on ice for 20 minutes.
6. Centrifuge for 15 minutes at 4°C at 1400 rcf to separate the solution into three phases: the aqueous phase, the interphase, and the lower phase. Collect the aqueous phase while avoiding the interphase into another Eppendorf tube.
7. Proceed to the cleanup step.

NB: Clean the beads using 75% ethanol, followed by washing with soapy water, then again with 75% ethanol twice, and finally autoclave them.

2.4.2 Cleanup (material from NucleoSpin Gel & PCR Clean-up kit)

1. For every 100 µL of the aqueous phase, add 350 µL of buffer NTI and gently mix by pipetting.
2. Add 250 µL of 100% ethanol and gently mix by pipetting.
3. Transfer up to 700 µL into a cleanup column placed on a collection tube and centrifuge for 15 seconds at 8000 rcf. Discard the flow-through.
4. Repeat the operation for the remaining aqueous phase using the same column to collect all the RNA. Add DNase steps (2.4.3 below) if needed.
5. Add 500 µL of RPE to the cleanup column and centrifuge for 15 seconds at 8000 rcf. Discard the flow-through.

6. Add another 500 µL of RPE and centrifuge for 2 minutes at 8000 rcf. Discard the flow-through.
7. Centrifuge the cleanup column for 1 minute to remove all ethanol.
8. Transfer the cleanup column to a new collection tube.
9. Elute RNA by adding 30 µL of DEPC water, wait for a minute, then centrifuge for 30 seconds at 8000 rcf.
10. Repeat the previous step to obtain a higher yield.
11. Collect 4 µL to measure it on a NanoDrop and run it on a gel to check the quality, more if further analyses are required.
12. Store at -80°C.

2.4.3 DNase Treatment – After Step 4 in Cleanup (2.4.2)

1. Add 350 µL Buffer RW1 to the cleanup column and centrifuge for 15 seconds at 8000 rcf. Discard the flow-through.
2. Add 10 µL DNase I stock solution to 70 µL Buffer RDD (incubation mix) and mix gently. Add 80 µL incubation mix to the cleanup column, leave on the benchtop for 15 minutes.
3. Add 350 µL Buffer RW1 to the cleanup column and centrifuge for 15 seconds at 8000 rcf. Discard the flow-through.
4. Proceed with Cleanup (2.4.2) step 5.

2.4.4 RNA Quantification and Quality Assessment (material from Qiagen RNeasy Kits)

To assess the integrity of RNA, the extracted RNA was run on agarose gel to detect the presence of ribosomal RNA (rRNA) bands and eventual degradation signals

(smear or lower bands). Purity and RNA concentration were evaluated using NanoDrop Spectrophotometers. The ratio A260/230 indicates contamination by phenol or carbohydrates, and the ratio A260/280 indicates possible contamination by phenol, proteins or might be the result of low concentration of nucleic acids. Purity values between 1.8 and 2.2 are considered acceptable. Lower values of A260/280 in the RNA extracted from the eyes have been attributed to the presence of pigment.

2.5 RT-qPCR

An explorative analysis on the expression of opsin genes in octopods was conducted via reverse transcriptase quantitative polymerase chain reaction (RT-qPCR). RNA was extracted from the eyes, optic lobes, suckers, and muscles of one specimen of *O. vulgaris* and *A. argo*, and then reverse transcribed into cDNA using the SuperScript™ VILO™ cDNA Synthesis Kit from Thermo Fisher Scientific following the manufacturer's instructions. Subsequently, RT-qPCR was conducted. These two species were chosen due to the availability of their fully published genomes and their phylogenetic distance, encompassing all octopuses relevant to this study (Destanović et al., 2023; Yoshida et al., 2022). Primers were meticulously designed to span exon-exon boundaries to mitigate potential amplification resulting from DNA contamination. Additionally, a prior DNase treatment of the RNA was performed, and gel electrophoresis did not reveal any DNA bands. The presence of potential opsin pseudogenes was ruled out through genome analysis. Primer self-complementarity was also verified. Refer to the Appendix for details regarding the primers used in the analysis, including the target gene, sequence, and efficiency.

Unfortunately, out of the three selected control genes that performed well with *O. vulgaris* (Actin, Elongation Factor 1-Alpha, Alpha-Tubulin), only Elongation Factor

1-Alpha worked effectively with *A. argo*. Each reaction was performed in duplicate using cDNA synthesized twice from the same RNA extraction to mimic biological replicates, addressing the limited availability of *A. argo* specimens with adequate RNA quality. Additionally, three technical replicates were performed for each experiment. Primer efficiency was evaluated by conducting the analysis on serial dilutions of cDNA ($10^{0.5}$, 10^0.2, 10^-1, 10^-2, 10^-3).

Each qPCR reaction contained:

4 µL TaqMan™ Fast Advanced Master Mix

3 µL MilliQ H2O

1 µL forward primer

1 µL reverse primer

1 µL cDNA

During qPCR, DNA amplification was continuously monitored using an intercalating dye that emits fluorescence upon binding to double-stranded DNA. This fluorescence signal enables the qPCR device, in this case, the Thermo Fisher Scientific QuantStudio™ 5 Real-Time PCR system, to detect changes in fluorescence at each cycle. The fluorescence intensity increases exponentially as the DNA amplifies until it reaches a plateau due to high concentrations of amplicons and decreased availability of nucleotides, primers, and polymerase efficiency. Consequently, three phases can be discerned: the baseline, where the signal is too faint to be detected; the exponential phase, where the curve begins to grow exponentially due to DNA duplication in each cycle; and the plateau phase, where the signal stabilizes. The quantification cycle (Cq) marks the cycle at which the signal becomes detectable, and the exponential curve begins (Dymond, 2013). This

parameter is crucial as it provides an estimate of the amplicon copy number in the original cDNA sample. Normalization has been computed using the $Ln2^{-\Delta Cq}$, where ΔCq is the difference between the Cq mean of target gene and the Cq mean of the reference gene. The first intention was to use the muscle as control tissue, but we found retinochrome and pseudopsin being expressed in this tissue, as confirmed also by a simple reverse transcriptase RT-PCR amplification. Therefore, to be very conservative, the $Ln2^{-\Delta Cq}$ value of r-opsin1 expressed in the muscle was considered as 0 to normalize the data.

2.6 Library preparation and mRNA sequencing

Data and previous analysis conducted on *O. vulgaris* and *A. argo*, including gene cloning and sequencing of RT-PCR products and RT-qPCR (see Chapter 3), have shown that all the opsins known in this species are expressed in the visual system (eye and optic lobe). Therefore, with the aim of sequencing opsins in other octopods, total RNA was extracted from the right eye and the optic lobe of *E. moschata*, *E. cirrhosa*, *C. macropus*, *P. tetracirrhus*, and *S. unicirrhus*. Their mRNA was sequenced using Ion Torrent next-generation sequencing technology. This was performed in the Molecular Biology and Sequencing Service at Stazione Zoologica Anton Dohrn, and all procedures, including library preparation, were carried out by the technicians with my assistance. The sequencing procedure involves several steps, including the preparation of a cDNA library for each species containing the transcribed genes (messenger RNA) of both the eye and optic lobe.

2.6.1 RNA Quality Assessment and Quantification

The RNA quality was evaluated on agarose gel and quantified using a NanoDrop. Additional quality control was performed using the Agilent™ 2100 Bioanalyzer™ and Agilent™ RNA 6000 Nano kit (see Appendix SuppFigure 1). In cephalopod RNA samples, the 28S RNA band (about 4000 bp) is easily degraded and splits into two bands of about 2000 bp each. This is also evident on agarose gel, where usually only a single band is visible, with 18S and 28S RNA subunits mixed (Adema, 2021). The program used by the Agilent™ 2100 Bioanalyzer™ does not account for this difference, as it is set for general eukaryotic RNA in which both bands are generally more stable and visible (Mueller et al., 2004). Therefore, in cephalopods, the RNA integrity number (RIN) is usually underscored. Despite being more subjective, an assessment by eye must be performed. We sequenced all the extracted RNAs with an estimated RIN value > 6, evaluated by the height of the 18S band and the presence of smears and peaks in the fast region.

2.6.2 RNA pooling

Eye and optic lobe total RNA were diluted to reach the same concentration and mixed together. They were then diluted again to reach a concentration of 2500 ng of total RNA in nuclease-free water (see table below). Spike-in control was added to ensure the correct outcome of the subsequent library preparation procedure and sequencing (Table1).

Sample ID	ng/µL	µL to collect 2,5 µg	µl spike-in 1/100	RNase-free water to 150 µL
CME/CMB	600	4,2	2	143,8
EME/EMB	225	12,0	2	136,0
ECE/ECB	1160	2,5	2	145,5
PTE/PTB	1200	2,0	2	146,8
SUNE/SUNB	466	5,4	2	142,6

Table 2.1 – RNA dilution. The table shows hoe the different extracted RNAs were diluted to reach a concentration of 2,5 µg in 150 µL and proceed with the library preparation. CME/CMB states for *C. macropus*, EME/EMB for *E. moschata*, ECE/ECB for *E. cirrhosa*, PTE/PTB for *P. tetracirrhus* and SUNE/SUNB *S. unicirrhus* visual system (Eye plus optic lobe) total RNA.

2.6.3 Purification from non-coding RNA

To prepare a library containing only messenger RNAs, the Thermo Fisher Invitrogen™ Dynabeads™ mRNA DIRECT™ Micro Purification Kit, which is specific for selecting only sequences with Poly(A) tails, was used following the manufacturer's protocol. The absence of 18S and 28S rRNA was checked using the Agilent™ 2100 Bioanalyzer™ instrument.

2.6.4 RNA Fragmentation

RNA was subsequently fragmented by adding RNase III, following the manufacturer's protocol, in a total reaction volume of 12 µL (10 µL of RNA in nuclease-free water and 2 µL containing the enzyme and the buffer). The reaction was assembled on ice and incubated for 10 minutes. There is no stop solution, and 20 µL of nuclease-free water must be immediately added to slow down the reaction, which can be kept on ice for up to one hour. Purification of fragmented RNA was

performed using the Magnetic Bead Cleanup Module, following the manufacturer's procedures. Assessment was done using the RNA 6000 Pico Kit with the Agilent™ 2100 Bioanalyzer™, where the graph shown a size distribution with a peak towards 200 bp, indicating the presence of mRNA fragments of the correct size.

2.6.5 Adaptor Ligation, Hybridization, and cDNA Synthesis

Next, additional sequences, named adaptors, were added to the extremities of each RNA fragment through an initial hybridization and a subsequent ligation step. After ligation, the RNA was reverse transcribed into cDNA and purified using the Magnetic Bead Cleanup Module. All these steps were performed using the materials in the Ion Total-RNA Seq Kit v2 and were carried out following the manufacturer's protocol. Once the cDNA is synthesized, the sequences are composed of retrotranscribed mRNA fragments with two adapter sequences on their extremities that are identical for each sample and can be used for amplification.

2.6.6 PCR

The cDNA needs to be amplified, which is achieved through a high-fidelity PCR reaction. Since more than one sample can be loaded on a single Ion Torrent chip, this PCR reaction can be used to add an additional marker sequence (barcode) to the adapter, helping to distinguish between sequences from different libraries. To do this, primers from the Ion Xpress™ RNA 3' Barcode Primer Platinum™ PCR SuperMix High Fidelity kit were used. The Ion Xpress™ RNA 3' Barcode primers contain different 3' primers with specific barcode sequences that can be assigned to different libraries. In contrast, the forward 5' primer of the kit is complementary to

the adapter sequence. The product was purified using the Magnetic Bead Cleanup Module, and the result was then assessed using the Agilent™ 2100 Bioanalyzer™.

2.6.7 Sequencing

For each sample, the barcoded cDNA libraries were diluted to achieve a similar concentration. Subsequently, an equal volume was collected and pooled together from each diluted library. Loading of the template on the Ion PI v3 Chip was performed using the Ion Chef™ System. Sequencing was performed using the Ion Torrent™ Ion S5™ System following the manufacturer's instruction.

2.7 Transcriptome assembly and opsin search

2.7.1 Transcriptome assembly

Transcriptome assembly was performed in the Bioinforma lab of the Stazione Zoologica Anton Dohrn. Row data quality of the sequenced libraries have been assessed using FastQC v0.12 and Trimmomatic v.0.38 have been used to remove all the low-quality sequences, low-quality bases at the start and the end of a read, and sequences with a length inferior of 50bp (LEADING:5 TRAILING:5 SLIDINGWINDOW:4:15 MINLEN:50). *De novo* assembly have been performed using Trinity v2.15 (min_kmer_cov 2 --normalize_reads). Results can be seen in the table below (Table 2.2).

Species	Number reads raw	Number reads cleaned	Total reads after cleaning
C. macropus	13,432,780	10,806,527	22,245,294
	14,437,024	11,438,767	
P. tetracirrhus	20,897,511	17,471,623	36,078,858
	22,418,243	18,607,235	
S. unicirrhus	14,109,261	11,837,710	24,229,506
	14,952,285	12,391,796	
E. cirrhosa	13,386,356	10,685,853	21,777,575
	14,048,844	11,091,722	
E. moschata	14,520,817	11,730,783	24,205,386
	15,632,508	12,474,603	

Table 2.2 – Number of total reads for each species. The table shows the number of total reads obtained before and after cleaning

2.7.2 Opsin mining

Using annotated opsins collected from the genome of *O. vulgaris*, we used BLAST search to look for them, using the script made by Dr. Lorena Buono.

blastn -query query -db $i -word_size 4 -evalue 10 -dust no -soft_masking false -gapopen 0 -gapextend 0 -out

2.8 Gene cloning

Gene cloning involves integrating a specific gene, amplified through RT-PCR, into a plasmid via a ligation reaction. This plasmid is then inserted into a cell through transformation, where it undergoes subsequent amplification during cell division. This procedure has the advantage of producing more reliable copies of the gene compared to conventional PCR, leveraging the repair mechanisms present in organisms, which helps to minimize mutations during the copying process.

Additionally, cells can be stored at low temperatures for extended periods. We used gene cloning to confirm that the sequences obtained from transcriptomic and genomic data were correct using Sanger sequencing in the Molecular Biology and Sequencing Service at Stazione Zoologica Anton Dohrn. Furthermore, by integrating the gene into a vector, it can be utilized for subsequent expression analysis experiments, such as the creation of *in situ* hybridization probes. The process involves the following steps:

2.8.1 PCR Extraction

After PCR amplification of the gene to be cloned, the PCR product was run on a 1% agarose gel. The amplified band was cut and extracted using the NucleoSpin Gel Extraction Kit, following the manufacturer's protocol, and eluted in 30 µL. A portion of the elution (10 µL) was run on an agarose gel again to verify the presence of the extracted band and the success of the extraction procedure. Additionally, 1 µL was used for Nanodrop measurement to evaluate the concentration in ng/µL. If the concentration is less than approximately 8 ng/µL, the elution can be placed in the Concentrator 5301, keeping the Eppendorf tube open at 30°C for 10 minutes to check if it reaches the desired concentration before proceeding with the ligation reaction and subsequent transformation protocol.

2.8.2 Ligation Reaction

The insert was integrated into the pGEM®-T Easy Vector System (Promega) according to the manufacturer's protocol. The pGEM®-T Easy Vector contains two regions, SP6 and T7, along with their respective binding sites for the insert, as well

as a region conferring resistance to ampicillin (Figure 2.1). The T4 DNA ligase enzyme included in this kit was used for the ligation reaction, which was incubated with the vector, insert, and buffer overnight at 4°C or for 3 hours at room temperature. During this step, the linearized plasmid becomes circular and incorporates the insert. If the insert fails to integrate but the plasmid circularizes, the SP6 and T7 regions will merge, forming a complete lacZ gene. When this insert-depleted plasmid is introduced into bacterial cells, the resulting colonies will express the enzyme β-Galactosidase, producing a blue coloration on agar plates supplemented with X-gal. Omitting lacZ expression as a control for successful insert integration can be justified if the presence of the insert is confirmed by colony PCR, as will be detailed later.

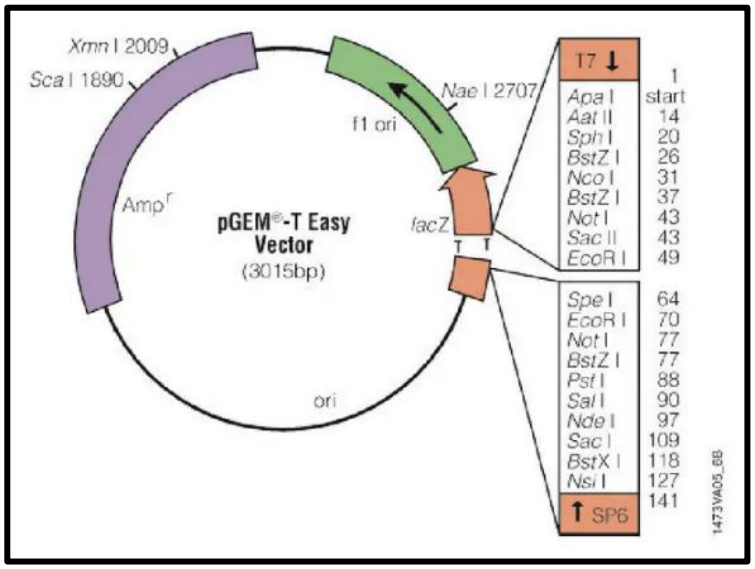


Figure 2.1 – **The pGEM®-T Easy Vector.** Picture taken from the manufacturer website showing the pGEM®-T Easy Vector plasmid structure, including the gene insertion sites in the lacZ region.

2.8.3 Transformation

Transformation was performed using chemically competent *E. coli* cells. In this process, 3 µL of the ligation product was combined with 40 µL of competent bacteria and kept on ice for 20 minutes. Subsequently, a thermal shock was applied by

incubating the mixture at 42°C for 2 minutes, followed by placing it on ice for 2 minutes. This thermal shock facilitates vector integration (transformation). The bacteria were then transferred to a Falcon tube and allowed to grow in 800 µL of liquid Lysogeny Broth (LB) for one hour at 37°C. They were subsequently plated on two LB + Ampicillin agar plates and incubated overnight at 37°C. One plate received 100 µL of cultured bacteria, while the other received 500 µL. All materials (LB, culture medium, and competent bacteria) were prepared by the Molecular Biology and Sequencing Service at Stazione Zoologica Anton Dohrn. Colonies formed on the LB + Ampicillin agar plates were selected for colony PCR. Single clean colonies showing the insert band were allowed to grow overnight (maximum 16 hours) in a Falcon tube containing 3 mL of LB plus 3 µL of ampicillin, placed in the incubator at 37°C.2.3.4 Colony PCR

Colony PCR consists of a standard PCR reaction in which, instead of using eluted DNA, a selected colony collected with a pipet tip is immersed in a PCR reaction tube. The tip is then stored in the fridge and assigned a specific colony number. M13 forward and reverse primers were used for amplification. Positive colonies on the respective tips can be allowed to grow as described above.

- **M13 Forward:** (5'-CAGGAAACAGCTATGAC-3')
- **M13 Reverse:** (5'-CGCCAGGGTTTTCCCAGTCACGAC-3')

2.8.5 Glycerol Stock

For glycerol stock preparation, 250 µL of grown bacteria were mixed with 700 µL of glycerol in an Eppendorf tube. The mixture was left on ice and subsequently stored at -20°C or -80°C. If necessary, the bacterial culture can be regrown in 3 mL of LB plus 3 µL of ampicillin.

2.8.6 Plasmid DNA Isolation

Minipreps were prepared and eluted using the EasyPure Plasmid Miniprep Kit, starting from bacterial pellets according to the manufacturer's instructions. The eluted minipreps were quantified, and the necessary quantity was sent for Sanger sequencing using the M13 forward and reverse primers. This sequencing is crucial not only to verify that the correct gene was cloned but also to determine the orientation of the gene for eventual probe preparation.

2.9 Characterization of key tuning site's selective pressures.

To characterize the selective pressures that shaped key tuning sites in visual opsins of octopod species (*O. vulgaris*, *E. moschata*, *E. cirrhosa*, *C. macropus*, *S. unicirrhus*, *P. tetracirrhus*, and *A. argo*) living in various light conditions, we focused on the most expressed and characterized opsins in visual systems: r-opsin1 and retinochromes. These sequences were obtained through RNA sequencing and Sanger sequencing of the cloned genes, as explained above. These analyses were performed in collaboration with Dr Giobbe Forni from Prof. Andrea Luchetti's lab at the University Alma Mater Studiorum of Bologna, that provided the scripts and run the first explorative analysis.

The first step was to perform ancestral state reconstruction in order to understand if the ancestral incirrate octopods lived in deep or shallow water, and thus if evolutionary changes in response to this environmental variable should be searched for in species living in the photic or deep zone.

Subsequently, codon models were used to infer synonymous and non-synonymous mutation fixation rates (dn and ds) in protein-coding gene sequences, and to retrieve putative codons under positive selection (Álvarez-Carretero et al., 2023). Since synonymous mutations do not change the sequence and structure of proteins, they are less likely to have a phenotypic effect and are consequently fixed at a rate that can indicate neutral evolution. On the other hand, non-synonymous mutations change the protein structure, which might have positive, negative, or neutral effects on the organism. Thus, a dn/ds ratio equal to one (the same fixation rate as under neutrality) can be considered as neutral evolution. A dn/ds ratio greater than one (higher fixation rate than neutrality) indicates positive selection, while a dn/ds ratio less than one (lower fixation rate than neutrality) suggests stabilization by purifying selection. Once the sites under selection were identified, they were mapped onto the 3-D model of the protein structure to check their distance from the retinal and estimate their potential role as key tuning sites.

2.9.1 Ancestral State Reconstruction

To reconstruct the environment of the common ancestors of incirrate octopuses, categoric and continuous data on maximum depth were collected from FAO cephalopods of the word using the lowest depth estimated (Jereb et al., 2005), while the time tree was attained from López-Córdova et al. (2022). Then, Phytools (Revell, 2024) library in R have been used to test between different discrete models (ER, Equal Rates, where all changes between states occur at the same rate; ARD, All Rates Different, where changes between states occur at the different rate; and SYM, Symmetrical, where reversions occur at the same rates). The test showed that ARD model shown the best fit. For the categoric analysis, 5 categories have been

established: Pelagic (P, for pelagic species such as *A. argo*); Euphotic (E, up to 100m); Disphotic 1 (D1, up to 500m); Disphotic 2 (D2, up to 1000m) and Aphotic (N, more than 1000m). *Vampyroteuthis infernalis* (Vampyromorphida), *Stauroteuthis gilchristi* (Stauroteuthidae) and *Opisthoteuthis massyae* (Opisthoteuthidae) were used as outgroups. The analysis indicated that the ancestral incirrate octopods lived most likely in disphotic environments.

2.9.2 Dataset Creation

Two fasta files (seqfiles) were created for the analysis. The first contained all the r-opsin1 nucleotide sequences from the selected species; the second contained all the retinochrome nucleotide sequences of the selected species. The alignment of the codons in the sequences was done using the MUSCLE algorithm implemented in MEGA (Sohpal et al., 2010). Subsequently, it was manually checked that the triplets were properly aligned, and saved in PHYLIP format.

2.9.3 Species Phylogeny

Due to a lack of information in the literature regarding the phylogenetic relationships of the analysed species, the species phylogeny was computed using collected r-opsin sequences on IQTREE2 (-m MF-B 1000) (Minh et al., 2020). Bootstrap values and branch lengths were removed, and the tree was saved in Newick (treefile) format. The tree was rooted with *A. argo*.

2.9.4 Molecular evolutionary analysis

Positive selection analysis was conducted using the CodeML program in the PAML package for all the alignments (Álvarez-Carretero et al., 2023). To begin, the analysis started under the one-ratio model which assumes a constant dn/ds rate across branch in the phylogeny and each site in the alignment (M0; noisy = 0; verbose = 0; runmode = 0; seqtype = 1; CodonFreq = 3; model = 0; NSsites = 0; icode = 1; getSE = 0; fix_blength = 2). Finding positive selection using this model is unrealistic, as it serves as a null hypothesis to test against the fit of other models and understand what might better explain the data. Then, the M0 analysis was compared to two different branch models: a model where each branch in the tree is allowed to have a different dn/ds (M1; noisy = 0; verbose = 0; runmode = 0; seqtype = 1; CodonFreq = 3; model = 1; NSsites = 0; icode = 1; getSE = 0; fix_blength = 2) and a model assuming ω is different between two groups of branches (M2; noisy = 0; verbose = 0; runmode = 0; seqtype = 1; CodonFreq = 3; model = 0; NSsites = 0; icode = 1; getSE = 0; fix_blength = 2). M2 models are used to test if specific groups of branches sharing a particular characteristic, such as environmental conditions, also share similar dN/dS rates.

Subsequently, the analysis has been performed under branch-site model for all the possible scenarios (littoral as foreground, deep as foreground, and if suggested by data, a single species as foreground). Foreground branches are those under test for positive selection. With this information in mind, two branch-site models were considered: one alternative (BSM1 assuming ω is different between different branches and sites; noisy = 0; verbose = 0; runmode = 0; seqtype = 1; CodonFreq = 3; model = 2; NSsites = 2; icode = 1; getSE = 1; fix_blength = 2; Small_Diff = $0.1e^{-6}$); and one general, the null hypothesis (BSM0 in which assuming ω assuming ω is different between different branches and never higher than 1 in the

different sites; noisy = 0; verbose = 0; runmode = 0; seqtype = 1; CodonFreq = 3; model = 2; NSsites = 2; icode = 1; getSE = 1; fix_blength = 2; fix_omega = 1; omega = 1). Likelihood ratio test (LRT) statistics have been used to compare the different models and results are reported in table, including likelihood ratio test (LRT, 2*(lnL1-lnL2)) Degrees of Freedom (DoF) and the p.value.

2.9.5 Mapping of the potential key tuning site

3-D models of the Retinochrome and R-opsin1 have been produced using the most recent version of the AlphaFold server (AlphaFold 3.0) available (https://golgi.sandbox.google.com/). Subsequently, the retinal have been manually included by overlapping it with *Todarodes pacificus* r-opsin1 and the detected mutation mapped on the protein 3-D structure using Chimera X.

CHAPTER 3 – Reconstruction of the ancestral opsin toolkit

When this investigation began, the opsin toolkit in octopods was largely unknown. The only available data came from a study conducted by Albertin and colleagues (2015), that detected the presence of 4 opsin genes in *Octopus bimaculoides*. Additionally, while opsins were well studied within bilaterians in vertebrates, echinoderms, and ecdysozoans, they were poorly investigated in lophotrochozoans, including molluscs, despite the importance of this group, which exhibits vast disparity in body plans (bauplans) and ecology. The clade Lophotrochozoa encompasses a large number of phyla, such as molluscs, annelids, brachiopods, phoronids, nemerteans, platyhelminths, and bryozoans, which successfully colonized a wide range of ecological niches, both marine and terrestrial. To conquer these diverse environments, they evolved a plethora of sensory organs, including a variety of photoreceptive structures, ranging from simple pigmented light-sensing areas to sophisticated camera-type eyes (Bok et al., 2017; Rawlinson et al., 2019; Serb & Eernisse, 2008), underscoring the significance of photoreception in these organisms and the need to study the opsin genes in these animals.

To address the gap in knowledge, in collaboration with Dr Roberto Feuda from the University of Leicester (UK), 74 translated genomes from 11 phyla were mined to investigate opsin evolution in lophotrochozoans, computed their phylogeny and subsequently reconciled opsin gene phylogeny with species phylogeny. Our results shed light on the patterns of opsin gene loss and gain within the different lophotrochozoan groups, including the reconstruction of the ancestral opsin toolkit of cephalopods. These findings provide a solid foundation for continuing our study of opsins in octopods by clarifying how many opsins were present in their common

ancestors. This results were published in *Molecular Biology and Evolution* (De Vivo et al., 2023), from which this chapter have been taken, including figures, text has been rephrased without changing the intended meaning of the original article. Supplementary material can be found in the online version of the article (https://doi.org/10.1093/molbev/msad066).

3.1 Results

3.1.1 Opsin phylogeny

Opsins were identified using BLASTp and classified into different orthogroups using Broccoli and CLANS (Derelle et al., 2020). After removing sequences with fewer than three transmembrane domains (see Methods chapter for details), 392 putative opsin genes were identified and integrated with a dataset containing known opsin sequences from various metazoans for phylogenetic analysis using ultrafast bootstrap (UFB) implemented in IQ-TREE and transfer bootstrap expectation (TBE), a method specifically designed for single-gene phylogeny (Lemoine et al., 2018; Minh et al., 2020). Melatonin receptors (MLT) and placopsins (Feuda et al., 2012) were used as outgroups. Initially, phylogenetic analysis using UFB (UFB1) and TBE was performed (Figure 3.1 A, Suppl. File S1, and Suppl. File S2). Then the t-index computed from TBE was used as criterion to remove unstable sequences (rogue lineages) that could decrease the bootstrap values and affect the reconstruction of duplication and loss events. This resulted in a new dataset of 380 lophotrochozoan sequences, which was used for a second UFB phylogenetic analysis (UFB2) (Figure 3.1 B, Suppl. File S3).

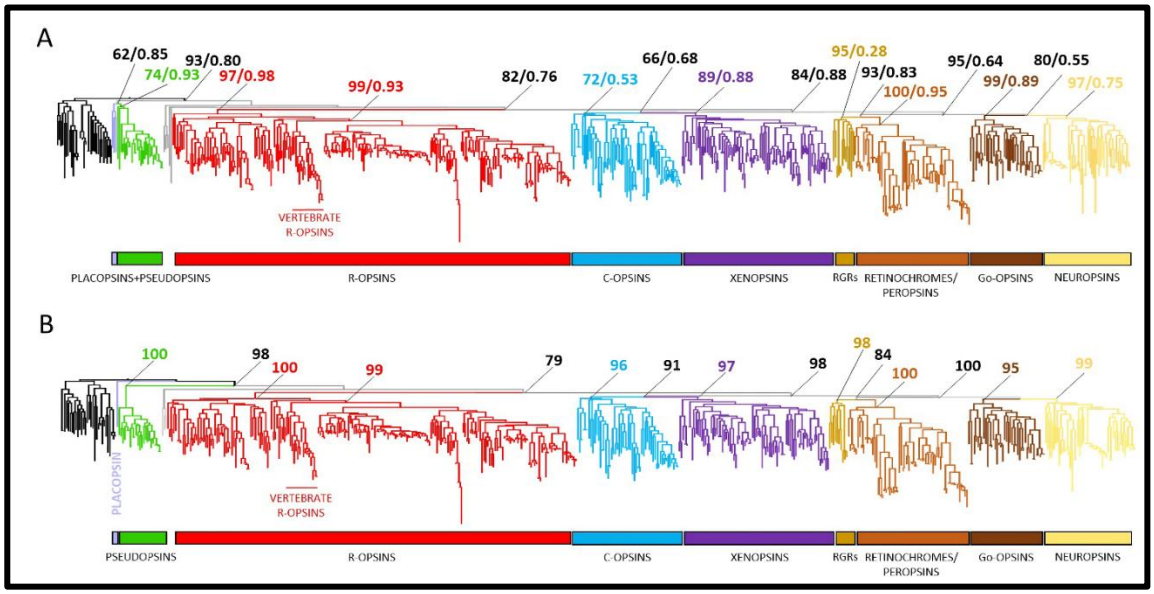


Figure 3.1 - Opsin of phylogeny in Lophotrochozoa. Maximum likelihood trees of lophotrochozoan opsins including non-lophotrochozoan sequences already annotated. The UFB1 and TBE trees are shown in (A), and the UFB2 tree in (B). UFB2 was computed after removing rogue lineages. Seven opsin subfamilies have been found in lophotrochozoans, with r-opsins being the sister group to all the other opsins and xenopsins related to c-opsins. The pseudopsins, a clade of opsin-like molecules, group alternatively with placopsins (A) or opsins (B). Non-bilaterian opsins are represented in grey, and the outgroup, MLT receptors, in black. Figure from De Vivo et al. (2023).

The monophyly of Go-opsins, neuropsins, peropsins, and RGR opsins (collectively referred to as Group-4) was supported by all trees (Figure 3.1) (UFB1=95; TBE=0.76; UFB2=100), as was the monophyly between c-opsins and xenopsins (UFB1=66; TBE=0.68; UFB2=91). Furthermore, our phylogenies resolved c-opsins/xenopsins and Group-4 as sister groups (UFB1=84; TBE=0.86; UFB2=98). R-opsins emerged as the earliest divergent group among all bilaterian opsins (UFB1=82; TBE=0.76; UFB2=79). Additionally, a monophyletic clade of opsin-like molecules, named ‘pseudopsins,’ consisting of 19 sequences from molluscs, nemerteans, and annelids was described. Pseudopsins formed a sister group to all other opsins (UFB1=93; TBE=0.80; UFB2=98; Figure 3.2A-B). Notably, when rogue taxa were not removed, pseudopsins formed a monophyletic clade with

placopsins (UFB1=62, TBE=0.85; Figure 3.2A). However, when rogue lineages were excluded, placopsins and pseudopsins no longer formed a monophyletic group; instead, placopsins were resolved as the sister group to the opsin/pseudopsin group (UFB2=98; Figure 3.1B). The position of pseudopsins might result from a potential Long Branch Attraction caused by the outgroup, leading us to perform an additional phylogenetic analysis with the outgroups (Placopsins + MLT) removed from the dataset. The sister relationship between pseudopsins and canonical opsins was again maintained (UFB3=100; Suppl. File S4). Like placopsins, pseudopsins do not possess a canonical retinal binding domain but retain conserved motifs indicating G-protein binding capability (Table 3.1).

Molecule/Sites	110	113	134-136	181	185	187	296	302-306	310-312
C-opsins (*Bos taurus* rhodopsin)	C	E	ERY	E	T	C	K	NPxxY	NKQ
R-opsins (*Octopis vulgaris* rhodopsin)	C	R	DRY	E	C	C	K	NPxxY	HPK
Mizuhopectenyessoensis_2164	C	Y	DKY	S	G	C	Q	NSxxY	NKI
Candidulaunifasciata_278	C	Y	DKY	H	G	C	L	HAxxY	NQI
Gigantopeltaaegis_1215	C	Y	DKY	V	G	C	Q	NAxxY	NKI
Octopusvulgaris_2483	C	Y	DKY	D	A	C	Q	NAxxY	NKI
Candidulaunifasciata_22	C	Y	DKY	F	G	C	L	HAxxY	NPI
Mercenariamercenaria_1853	C	Y	DKY	H	G	C	L	HSxxY	NKI
Crassostreavirginica_878	C	Y	DKY	H	G	C	Q	NGxxY	NKI
Pectenmaximus_2863	C	Y	DKY	S	G	C	Q	NSxxY	NKI
Oweniafusiformis_2741	C	Y	DKY	H	T	C	Y	NAxxY	NRI
Mytilusgalloprovincialis_2322	C	Y	DKY	H	D	C	Q	NGxxY	NKI
Octopusbimaculoides_2422	C	Y	DKY	D	A	C	Q	NAxxY	NKI
Nautiluspompilius_786466	C	Y	DKY	H	A	C	Q	NAxxY	NKI
Elysiachlorotica_1045	C	Y	DKY	H	G	C	L	HAxxY	NQI
Architeuthisdux_201604	C	Y	DKY	D	A	C	Q	NAxxY	NKI
Elysiamarginata_1135	C	Y	DKY	H	G	C	L	HAxxY	NQI
Plakobranchusocellatus_2988	C	Y	DKY	H	G	C	L	HAxxY	NQI
Sepiapharaonis_3310	C	Y	DKY	H	A	C	Q	NAxxY	NKI
Biomphalariaglabrata_157	C	Y	DKY	H	A	C	Q	HAxxY	NQI
Lineuslongissimus_618617	C	Y	DRY	D	G	C	T	NPxxY	NKQ

Table 3.1 - Comparison between opsins and pseudopsins in conserved motifs. The table shows the key residues and motifs highlight differences between representative *Bos taurus* c-opsins, *O. vulgaris* r-opsins and pseudopsins (all the other). Number indicating the aminoacidic position on the molecule are referred with respect to *B. taurus* rhodopsin. Sequences are reported with the species name and the numeric code as shown in Suppl. File S6. Pale yellow indicates structural residues, pink the counterion sites, red the retinal

binding site, and green the residues involved in the G-protein binding activation. Modified from De Vivo et al. (2023) supplementary material.

Furthermore, pseudopsins are different from canonical opsins, they cluster in a single orthogroup separated by the other known opsins, and behave like placopsins in CLANS analysis, where by increasing the e-value they first cluster with opsins and only subsequently, together with opsins, with other GCPR groups (Frickey & Lupas, 2004) (Figure 3.2). In summary, our analysis suggests that lophotrochozoans share canonical opsin genes and retain a putative clade of opsin-like sequences potentially related to placopsins.

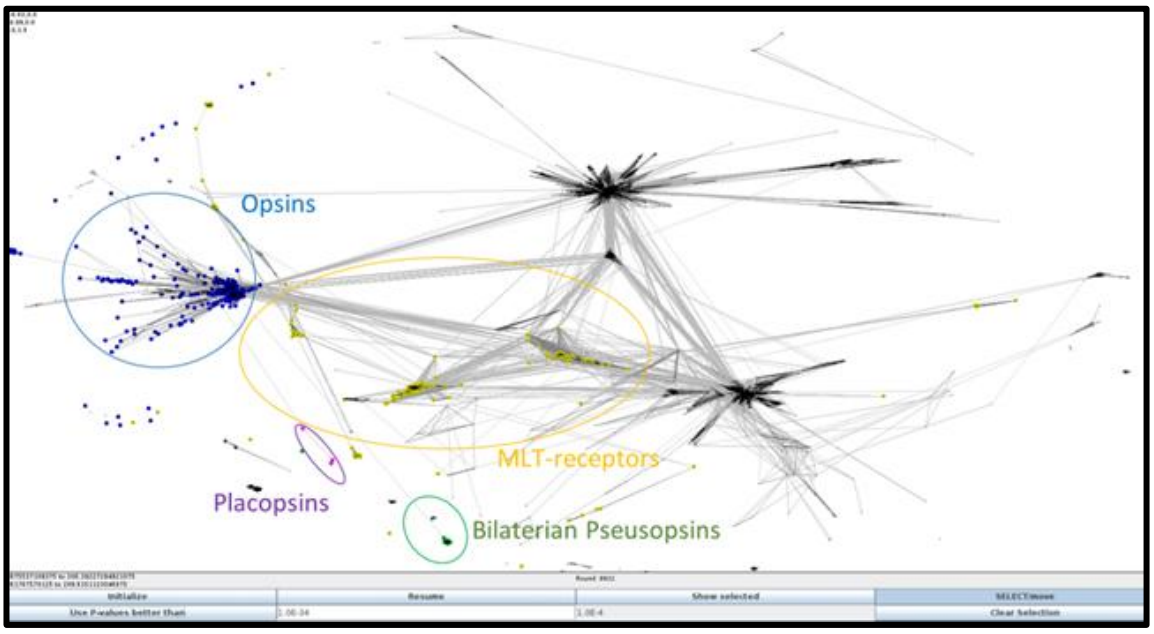


Figure 3.2 - CLANS analysis. The placopsin group (purple) and the pseudopsin group (green) connect simultaneously to the opsin group (blue) at E^{-34}. Melatonin receptors (MLT) are indicated in yellow, while other non-opsin sequences are indicated in black and. Picture from the supplementary data in De Vivo et al. (2023).

3.1.2 Loss and duplication of opsin genes in Lophotrochozoa

After removing unstable sequences, the pattern of opsin duplication and loss, including ancestral node content, was reconstructed using GeneRax, a maximum

likelihood gene-tree to species-tree reconciliation method (Morel et al., 2020). Overall, opsin evolution in lophotrochozoans appears highly dynamic, with numerous lineage-specific duplications and losses (Figure 3.3 and Figure 3.4).

GeneRax estimated that the common ancestor of lophotrochozoans possessed at least 13 opsins from seven subfamilies: four r-opsins; two c-opsins, retinochromes/peropsins, and xenopsins; one RGR, Go-opsin, and neuropsin (Figure 3.4). However, bryozoans and platyhelminths radically reduced their opsin gene complement (Figure 3.3 and Figure 3.4). Modern bryozoans retained only xenopsins, while all platyhelminths in our analysis possessed only r-opsins and retinochromes. However, the expression of a xenopsin in the larval eyes and ciliary structures (phaosomes) of polycladids has been demonstrated, showing that two xenopsin clades were already present in the platyhelminth common ancestor (Rawlinson et al., 2019). Among the cestodes, a group of parasitic worms, only aquatic forms (e.g., pseudophyllideans) possess opsins, while parasites of humans and terrestrial animals (cyclophyllideans) have lost all opsins (Kikuchi et al., 2021). Similarly, phoronids retain only two opsins: a neuropsin and a putative RGR, the latter recognized as an unstable sequence by TBE analysis and subsequently discarded (Figure 3.3 and Figure 3.4).

All other clades (annelids, molluscs, nemerteans, and brachiopods) did not experience a dramatic reduction in opsin genes. They retain a complement similar to that of the lophotrochozoan ancestor, followed by lineage-specific gene duplications and independent opsin losses in many clades (Figure 3.3 and Figure 3.4). Specifically, annelids lack RGR opsins, and some clades, such as clitellates, have drastically reduced their opsin content, preserving only r-opsins. Data from the annelid *Platynereis dumerilii* include one Go-opsin, four r-opsins, and two c-opsins (Gühmann et al., 2015). Brachiopods retain the complete opsin set, while

nemerteans exhibit the presence of RGRs, retinochromes, Go-opsins, c-opsins, and r-opsins. The common ancestor of molluscs retained the entire opsin complement; however, only polyplacophorans (chitons) possess a c-opsin, which has been lost in all other groups. Additionally, bivalves and gastropods have experienced xenopsin duplication events, retaining multiple paralogues. Surprisingly, cephalopods have the lowest number of opsins among molluscs, retaining only two r-opsins, one xenopsin, and two peropsin/retinochromes (Figure 3.3 and Figure 3.4). This is unexpected, given the morphological complexity of their eyes and visual system.

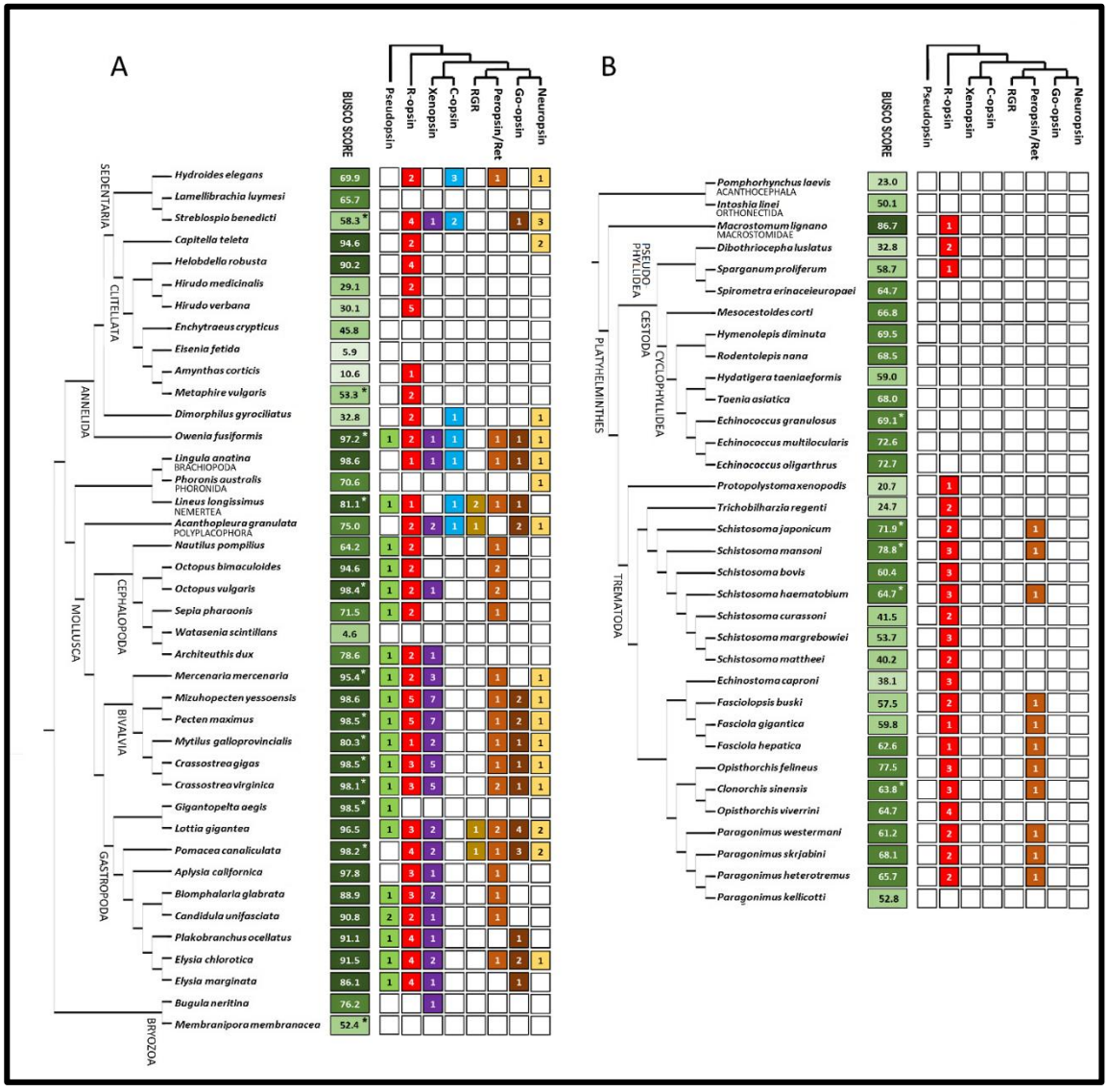


Figure 3.3 - Number and opsin subfamilies within lophotrochozoan species. The figure illustrates the species tree and the minimal number of opsins and pseudopsins found per species (excluding species-specific duplications) are shown respectively in the Bryozoa–Annelida/Mollusca node (A) and the Orthonectida/Acanthocephalan–Platyhelminthes node (B). The asterisk (*) indicates species with chromosome-level–assembled genomes. Figure from De Vivo et al. (2023).

3.2 Discussion

Our results reveal the evolution of opsin genes spanning 600 million years in one of the largest metazoan clades, Lophotrochozoans, showing a diverging evolutionary history within this group (Figure 3.3). Overall, all computed phylogenetic trees support the sister relationship between c-opsins and Group-4 opsins, with r-opsins being sister to this clade, consistent with previous works (Bonadè et al., 2020; Feuda et al., 2012, 2014; Fleming et al., 2020; Rawlinson et al., 2019; Vöcking et al., 2017; Yoshida et al., 2015). Furthermore, the tree supports the monophyly of xenopsins and c-opsins, similar to some studies (Vöcking et al., 2017; Yoshida et al., 2015) but different from others (Bonadè et al., 2020; M. Ramirez et al., 2016; Rawlinson et al., 2019).

Remarkably, a novel bilaterian opsin-like clade in lophotrochozoans named pseudopsins has been identified, closely related to placopsins or opsins. Pseudopsins do not possess the retinal binding domain, and therefore they might lack a photoreceptive function. However, the presence of G-binding conserved domains suggests that they might still be receptors involved in an unknown G-signalling mechanism, such as chemoreception. Despite this, the possibility that they can still bind retinal without the canonical retinal binding domain (K296 in cow rhodopsin) cannot be excluded. The occurrence of sequences possibly related to placopsins in bilaterians might indicate a wide distribution of these genes in

metazoans, suggesting the need for a re-evaluation of opsin evolutionary history. New studies investigating the expression patterns and functions of these cryptic receptors in the future must be encouraged, as they can potentially clarify the function of placopsins. In an alternative scenario, pseudopsins might just be highly divergent lophotrochozoan-specific opsins that lost the retinal binding domain.

Furthermore, our results suggest the presence of a rich opsin repertoire composed of 13 genes from 7 already known opsin subfamilies and 1 pseudopsin in the last common ancestor of lophotrochozoans. This is in agreement with other opsin analysis that included lophotrochozoa clades, such as Ramirez et al. (2016). The Lophotrochozoan common ancestor (LOCA) appeared in the Ediacaran, but information about its anatomy and ecology is lacking. By comparing the trochophore larvae (mollusc-like larvae) and the muller larvae (polyclad-like larvae), it has recently been speculated that LOCA possessed a swimming larval stage (Piovani et al., 2023). The adult, on the other hand, may have had a more benthic lifestyle, resembling animals like Kimberella and Wuffengella (Fedonkin & Waggoner, 1997; Guo et al., 2022).

The observation of the opsin complement in the species tree (Figure 3.4) shows two opposite tendencies: (i) the reduction of opsins in platyhelminths and bryozoans, and (ii) a stable number of opsins in the ancestors of molluscs, brachiopods, nemerteans, and annelids. This might highlight the different evolutionary forces that shaped opsin diversity in lophotrochozoans, but many questions remain unanswered. The parasitic lifestyle in some platyhelminth lineages might partly explain why the opsin complement was reduced in this clade. In other cases, environmental factors, such as terrestrialization in clitellates or life in extreme habitats, such as hydrothermal vents in the deep sea for *Lamellibrachia luymesi* and *Gigantopelta aegis*, may have caused the loss of opsins (Lan et al., 2021). Similarly,

events of genome compaction, such as in the case of *Dimorphilus gyrociliatus*, could result in a reduced opsin number (Martín-Durán et al., 2021). It is important to highlight that the opsin complement is not always correlated with the morphological complexity of the visual system, including the complexity of the eye. This is clearly shown by cephalopods, which have a very low number of opsins compared to other molluscs, yet possess complex camera-type eyes and well-known visual capabilities. Within cephalopods, and excluding species of molluscs living in light-depleted environments, *Nautilus pompilius* has been considered the one with the lowest numbers of opsins, retaining only two r-opsins and one retinochrome, but this could be due to the low genome quality (BUSCO score 64.2) (Zhang et al., 2021).

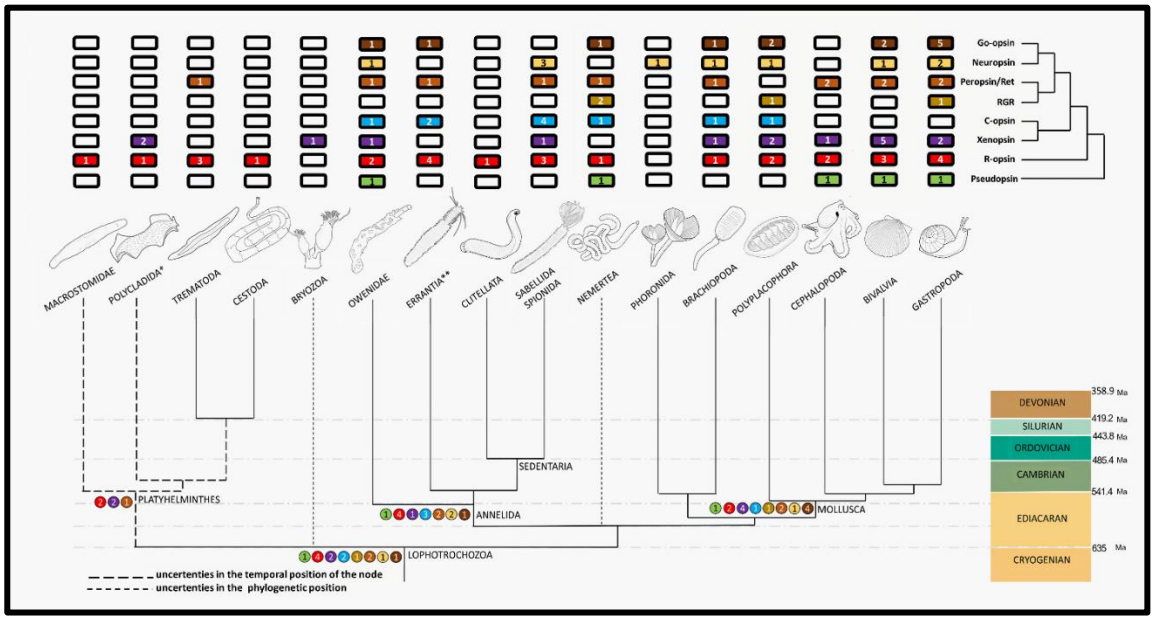


Figure 3.4 - Opsin duplications and losses in lophotrochozoans. Rectangles show the estimated number of opsins per taxa, while circles the number of opsins at the nodes (circles). The common ancestor of lophotrochozoans possessed 13 opsins from 7 opsin subfamilies plus 1 pseudopsin. A pattern of radical opsin loss can be found in the lineages leading to platyhelminths, bryozoans, clitellates, and phoronids. The inference of the ancestral opsin toolkits shown in the nodes have been made by integrating the GeneRax

results with data from (Rawlinson et al., 2019) (*) and Gühmann et al. (2015) (**). Figure from De Vivo et al. (2023).

Differences in opsin number among different taxa might be explained by the presence of complex life cycles or the number of photoreceptive structures. However, vision is not the only function of opsins; the rich repertoire of opsins can be used for other biological functions, such as light avoidance, circadian entrainment, melatonin release, growth control, phototactic behaviour, and non-light-dependent functions (Feuda et al., 2022).

To sum up, based on these results, due to their plethora of opsins, visual structures, life cycles, and complex opsin gene evolution, Lophotrochozoans can be considered excellent models for studying the evolution of photoreception in animals. However, these results still suffer from the limited genomic information available for certain groups (such as bryozoans, nemerteans, brachiopods, and phoronids) and the lack of chromosomal scale assembly for most lophotrochozoan species. More high-quality genomes will help to refine the details of opsin gene duplications and losses. Despite these limitations, the data provided here offer a solid foundation for clarifying the expression patterns and functions of different opsin genes in the phyla investigated.

CHAPTER 4 – Opsin expression in Octopoda

The ancestral opsin toolkit reconstruction in the previous chapter shows that cephalopods originally possessed five orthologous genes (r-opsin1, r-opsin2, xenopsin, retinochrome, and peropsin) and a putative opsin-like molecule, that was baptized pseudopsin. Despite that, little is known about the expression of these opsins in octopods, since the majority of studies focused their attention on opsin expression in the other branches of cephalopods: decapodiformes and nautiloids.

To fill this gap, the pattern of opsin expression in octopuses was explored through RT-qPCR in two distantly related incirrate octopod species: *Argonauta argo* and *Octopus vulgaris*. The genome of these two species was already published (Albertin et al., 2015; Destanović et al., 2023; Yoshida et al., 2022) and opsin sequences were easily mined through BLAST search, making possible to design species specific primers. Furthermore, these two species were chosen due to their phylogenetic distance, encompassing all the incirrate octopods, and making possible to produce hypothesis based on phylogenetic inference.

RT-qPCR results showed that all the opsins are expressed in the eye (retina) and/or brain (optic lobe). Therefore, with the aim to collect more opsin sequences from different octopods species, low coverage RNA sequencing was conducted combining extracted RNA from eye and optic lobe in five species: *Eledone moschata*, *Eledone cirrhosa*, *Callistoctopus macropus*, *Pteroctopus tetracirrhus*, and *Scaeurgus unicirrhus*. When necessary, to confirm the validity of these results, opsin gene cloning was performed. Our results indicate that most of the opsin gene

are present and expressed in these animals, and that additional investigation on their role is required.

4.1 Results

4.1.1 Opsin expression in Argonauta argo and Octopus vulgaris

Our RT-qPCR shows that the visual systems of *O. vulgaris* express all the opsins present in their genomes (Figure 4.1). Additionally, we found opsins to be expressed in the proximal suckers and funnel retractor muscle, which were initially intended to serve as control tissues for background expression. Specifically, the eye of *O. vulgaris* shows high expression of r-opsin1, moderate expression of xenopsin, and significant expression of retinochrome. Furthermore, exceptionally low levels of r-opsin2 and peropsin expression have been found in the eye. Similarly, the optic lobe shows high expression of r-opsin1 and retinochrome but, in contrast to the eye, also shows moderate expression of r-opsin2 and peropsin. Regarding the non-visual organs, the suckers show moderate expression levels of r-opsin1 and retinochrome, and the retractor muscle shows moderate expression levels of retinochrome. Pseudopsin is the only molecule that results to be expressed in all the tissues. To confirm these data, it was possible to amplify via RT-PCR all the sequences mentioned above from the respective tissues, with the exception of those showing low levels of expression, such as r-opsin2 and peropsin from the eye. Sequences expressed in the eye and optic lobe were also cloned.

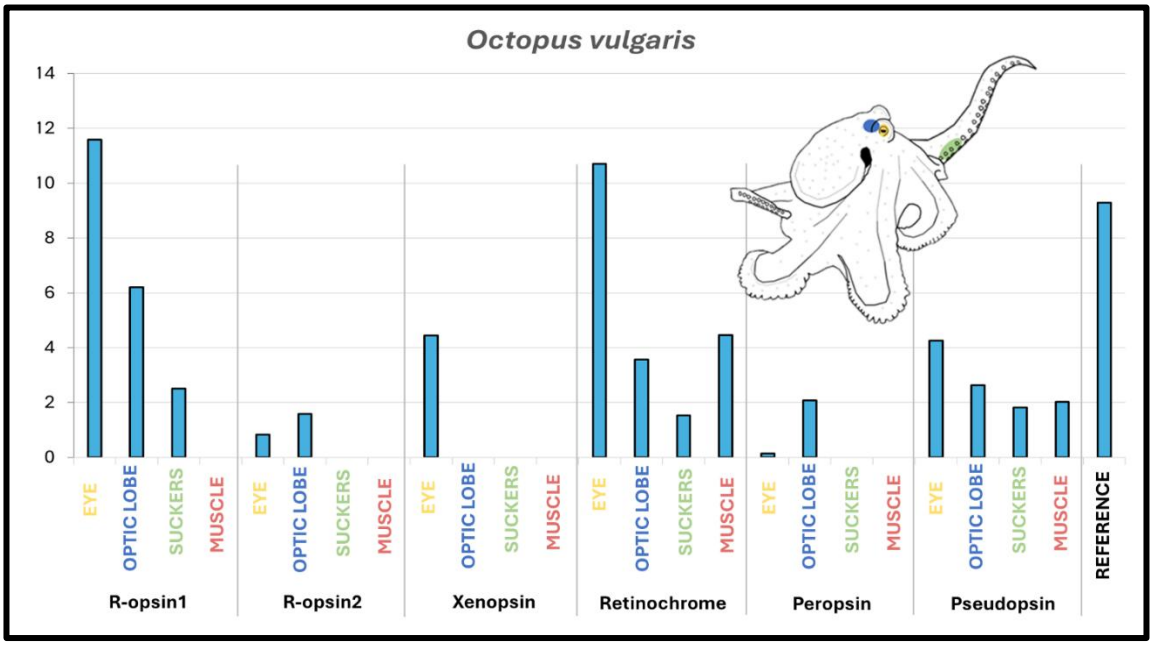


Figure 4.1 - Opsin expression in *O. vulgaris*. The bar graph shows the RT-qPCR results for opsin gene expression in *O. vulgaris*. The y-axis indicates the natural logarithm (Ln) of the average $2^{-\Delta Cq}$, normalized so that the value of zero corresponds to the expression of r-opsin1 in the muscles. The reference gene is Elongation Factor 1.

A similar outcome was observed in *A. argo*, where most of the opsins are expressed in a manner similar to *O. vulgaris*, except for r-opsin2 and peropsin, which were detected only in a single run of RT-qPCR in the optic lobe and their expression has not be confirmed by a second analysis (Figure 4.2). Our results confirm the presence of xenopsins exclusively in the eye also in this species. Non-visual tissues, such as the suckers and retractor muscle, showed the presence of r-opsin1 and retinochrome in the suckers, while only retinochrome was detected in the retractor muscle. Pseudopsin is moderately expressed in all tissues. In *A. argo*, it was possible to amplify, clone, and sequence only r-opsin1, retinochrome, and the two xenopsin paralogues via RT-PCR from the tissues mentioned above.

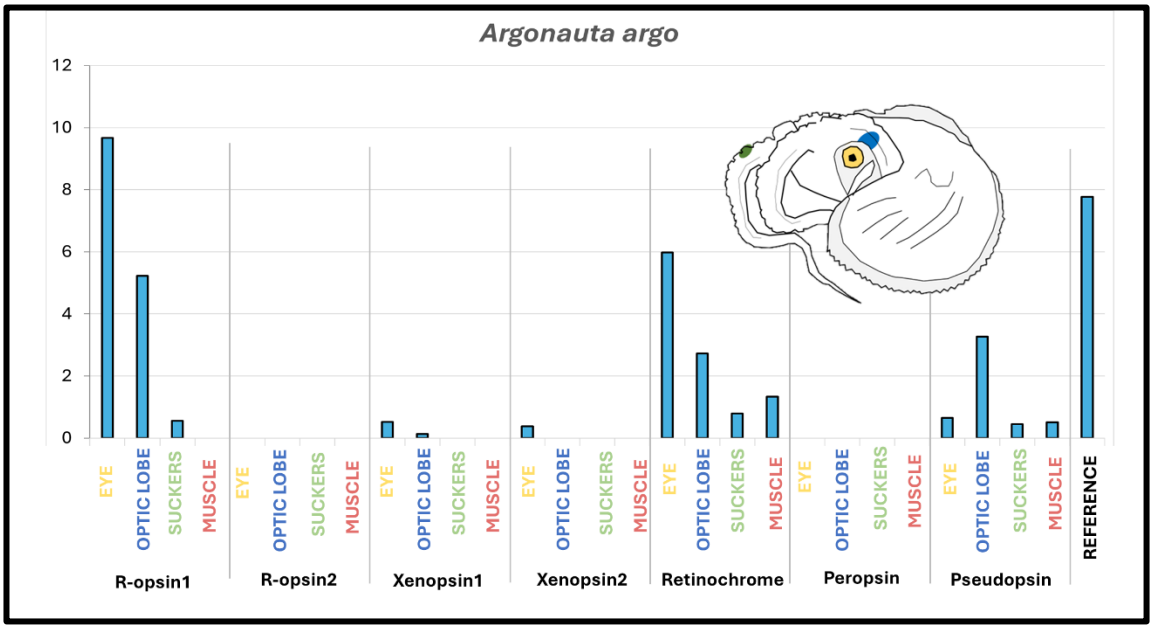


Figure 4.2 - Opsin expression in *A. argo*. The bar graph shows the RT-qPCR results for opsin gene expression in *A. argo*. The y-axis indicates the natural logarithm (Ln) of the average $2^{-\Delta Cq}$, normalized so that a value of zero corresponds to the expression of r-opsin1 in the muscles. The reference gene is Elongation factor 1.

4.1.2 Opsin search in low coverage transcriptomes

Low coverage (about 20 million reads) mRNA sequencing of the visual system (eye and optic lobe) in 5 species of octopods was performed with the aim of finding opsins. The transcriptomes were mined using BLASTn using *O. vulgaris* opsins as a query. It was not possible to sequence all the opsins, and only full-length sequences of r-opsin1, retinochrome and almost complete sequences of pseudopsins in all the samples were obtained in the majority of the species (Figure 4.3). Exception were *C. macropus*, in which the retinochrome resulted partial (the full length has been obtained via gene cloning) and *S. unicirrhus,* were r-opsin1 lacks the initial 29 bp. Fragments of peropsins (shorter than 500 bp) have been found in *S. unicirrhus*, *P. tetracirrhus*, and the two *Eledone species*. In *E. moschata*, fragment of all the opsins have been found, including xenopsin and r-opsin2. Recently, the genome of *E. cirrhosa* (GCA_964016885.1 and GCA_964016925.1)

has become publicly available on NCBI. We performed BLAST search on it to look for opsins, and we only retrieved full r-opsin1, partial sequences of r-opsin2, xenopsin, and peropsin (the last two are likely not correctly assembled). Pseudopsins seems still to be missing, but a complete sequence was found in its transcriptome.

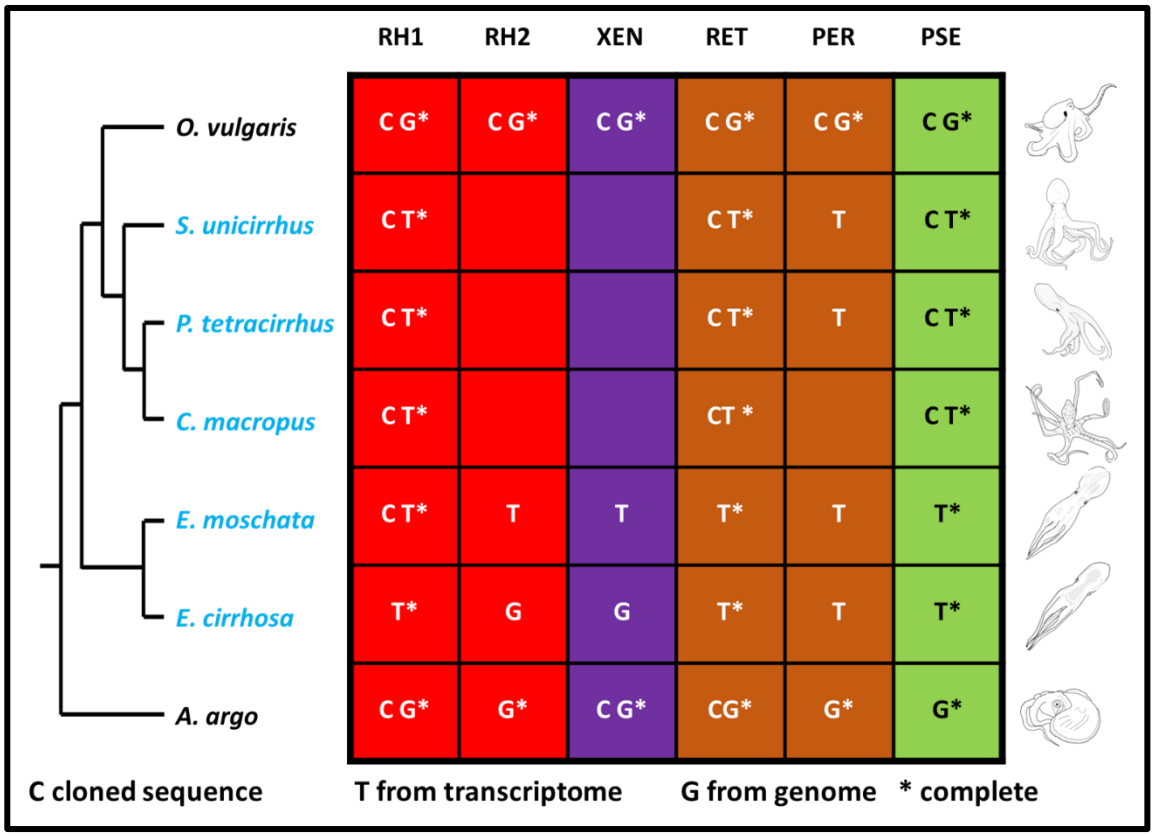


Figure 4.3 – Opsin found in the different species. The pictures show all the opsin found in the different species studies in this dissertation, in light blue the species of which the transcriptome of the visual system have been sequenced. Notice how r-opsin2 and xenopsins are missing in the majority of the transcripts, while peropsin is present only as partial sequences. Despite *S. unicirrhus* r-opsin1 lacks 29 bp it was still considered as complete, since its sequence has been covered for almost the 98% of the total nucleotides.

4.2 Discussion

4.2.1 Pattern of opsin expression in cephalopods

Studies on cephalopod opsins gene expression have been mostly conducted on decapodiform and nautiloids cephalopods. In particular, RNA sequencing and transcriptomic analysis, combined with reverse transcriptase quantitative PCR in the bobtail squid *Idiosepius paradoxus*, revealed the expression of r-opsin1, xenopsin (referred to as 'c-opsin'), retinochrome1, and retinochrome2 in the eye, while the central nervous system expresses r-opsin1, r-opsin2, and retinochrome2 (Yoshida et al., 2015). In this species, as in many other decapodiformes, peropsin has been lost and retinochrome duplicated (Bonadè et al., 2020; De Vivo et al., 2023; Yoshida et al., 2015). Surprisingly, xenopsin is also expressed in the gonads and retinochrome2 in the gut, indicating a possible function of these molecules outside of visual processes (e.g., circadian clock regulation). The same study conducted a similar analysis on *Nautilus pompilius*, revealing the expression of two opsins known at the time in this species: r-opsin1 and the retinochrome. They showed r-opsin1 to be expressed only in *Nautilus* eye, while the retinochrome in the eye, the central nervous system, and in two non-visual organs: the arm and the siphuncle. *Nautilus pompilius* also possesses r-opsin2, but evidences of its presence came after and the expression of this molecule in *Nautilus* remains uninvestigated (De Vivo et al., 2023). Another analysis conducted on the cuttlefish *Sepia officinalis* showed a different scenario: the experiment was restricted to the eye, the central nervous system and the skin, but only r-opsin1 and retinochromes were shown to be expressed in these organs. R-opsin1 and retinochrome were all expressed in later developmental stages mostly in all the tissues investigated except for retinochrome that resulted not expressed in the central nervous system of this species. This discrepancy could be explained by species-specific differences or can be related to

the different developmental stages examined. Indeed, while the *Idiosepius paradoxus* study investigated opsin expression in embryos and adult specimens, those conducted on *Sepia officinalis* focused on embryos and hatchlings. Furthermore, studies conducting *in situ* hybridization on *O. vulgaris* and *S. officinalis* only focused on the expression of r-opsin1 and retinochromes in the eye, but misses all the data regarding the other opsins (Bonadè et al., 2020; Hara et al., 1967). Therefore, opsin expression in octopod cephalopods remains largely missing in literature.

To fill this gap, I performed RT-qPCR to explore the level of opsin genes in octopod RNA extracted from different tissues. By doing that I discovered that it well matches with the expression pattern previously described in other cephalopods, especially the findings in the adult bobtail squid. In particular, despite being more expressed in this organ, r-opsin1 transcripts are not only present in the eye of *O. vulgaris* and *A. argo* but can be found in other organs as well, such as suckers and the optic lobe. The presence of r-opsin1 in the suckers can be indicative of light sensing ability of these organs, but it is unclear if this is due by the presence of a skin with special photoreceptive structures such as chromatophores, which is known to express opsins (Ramirez & Oakley, 2015), or it is a specific characteristic of these organs that, being cup shaped, can potentially discriminate the direction of the shadows.

The expression of retinochrome in the muscle tissue was a surprise. The reason for that remains unclear; it might be that in these organs retinochrome performs tasks unrelated to vision or photoreception in general. It is known that the retinochrome does not perform a signalling activity, but otherwise function as photoisomerases, restoring the original retinal configuration, and this has been linked to the r-opsin1 intense activity during vision, which might potentially exhaust

11-*cis* retinal reserves (Vöcking et al., 2021; Y. Zhang et al., 2021). Despite this well-known mechanism, other function can be hypothesised: since the retinal (or retinaldehyde) is involved in the production of retinoic acid, a compound linked to muscle signalling and metabolism (Belyaeva et al., 2019; Treves et al., 2012). Retinochrome might have, for example, a part in this process. This would explain its presence in other organs that are commonly considered non photoreceptive. Furthermore, retinochrome is not the only molecule able to reverse the retinal: the well-known Octopus r-opsin1, being a bistable opsin, can convert retinal between *cis* and *trans* configurations and *vice versa* (Tejero et al., 2024; Terakita, 2005). The key difference is that, in this molecule, the process occurs without bleaching the chromophore, which remains bound to the opsin. Since phylogenetically distant opsins are bistable and this characteristic is shared across different opsin subfamilies, it is likely that the ancestral opsin was bistable, with monostable opsins evolving later. Therefore, from an ancestral bistable state, retinochrome may have lost the ability to convert retinal from *cis* to *trans*, retaining only the ability to restore retinal upon a second photon interaction, ultimately functioning only as photoisomerases. The simple fact that r-opsin1 is able to photo reverse makes more difficult to establish the exact role of retinochrome, considered in general only a photoisomerases connected to the activity of r-opsin1 (Hara et al., 1967; Hara, Tomiyuki et al., 1972; Vöcking et al., 2021; Zhang et al., 2021). Being a molecule highly expressed in many tissues, the interpretation that retinochrome only to serve as accessory should be questioned and more investigation on the role of this molecule are required.

The function of less expressed opsins, such as r-opsin2 and peropsin, in the optic lobe of *O. vulgaris* remains enigmatic. However, transcriptomes have revealed their presence in other octopod species. Since the optic lobe works in a way

analogous to the vertebrate inner retina, it is possible that those molecules are used to integrate the visual signal coming from the retina. Indeed, both r-opsin1 and peropsin possess G-binding regions, indicating the ability to start a signalling cascade (Nagata et al., 2018). Furthermore, among cephalopods, peropsin has only been retained by octopods, and this is the first time its expression has been shown in cephalopod since it was discovered by Albertin et al. (2015). The same can be said for the newly discovered pseudopsin. The absence of r-opsin2 and peropsin expression in *A. argo* might be due to stringent normalization, which may have excluded opsins transcribed at very low levels in this species, or to the loss of these molecules due to RNA degradation. Pseudopsins show similar expression levels across all tissues, indicating that they might not be involved in photoreception but rather in a different signalling activity. On the other hand, the presence of xenopsin in the eyes of three different cephalopod species, as well as in the transcriptome of the visual system of *E. moschata*, might open new avenues for understanding the role of this opsin in molluscan visual photoreceptors. The low expression of some opsins might potentially also be caused by the restriction of our analyses to adult stages, and the pattern of opsin expression might change during development.

4.2.2 Xenopsin expression in the eye might be common in Lophotrochozoa

Xenopsins, one of the newly discovered major opsin groups, are considered either the sister group to c-opsins or to group-4. Xenopsins are absent in ecdysozoans and deuterostomes but are well conserved in lophotrochozoans and possibly in cnidarians (referred to as cnidopsins) (Arendt, 2017; De Vivo et al., 2023; Ramirez et al., 2016; Rawlinson et al., 2019; Vöcking et al., 2017). The expression of xenopsins has been investigated across many lophotrochozoan phyla. In bryozoans, xenopsins are the only opsins found, and in the larva of *Trycellaria*

inopinata they are localized in the cilia of the paired lateral eye photoreceptors (Döring et al., 2020). Furthermore, in the annelid *Malacoceros fuliginosus*, xenopsins are expressed in the ventral and dorsal larval eyes, where they are co-expressed with r-opsin3 in photoreceptors that bear both microvilli and cilia (Döring et al., 2020). Xenopsins are known to be expressed in the larvae and adults of the flatworm *Maritigrella crozieri* (Platyhelminthes), specifically in the larval epidermal eye and in the adult phaosome, a structure composed of extraocular cells with an intracellular vacuole containing multiple cilia (Rawlinson et al., 2019). Xenopsins and peropsins are also the only opsins retained in chaetognaths, a phylum belonging to Gnatifera, the clade sister of lophotrochozoa; in *Spadella cephaloptera*, xenopsins are expressed exclusively in the eye, while peropsins are mainly localized in the nervous system (Wollesen et al., 2023). This is intriguing since Cambrian stem-chaetognaths were larger predators that might have relied on a well-developed camera-like eye for vision (Park et al., 2024). In molluscs, xenopsins have been found to be expressed in rhabdomeric photoreceptors bearing cilia in the larva of the chiton *Leptochiton asellus*. These cilia are not well developed, but the authors highlight the presence of a gene that helps to localize opsins in cilia and speculate that r-opsins might be localized in microvilli and xenopsins in the cilia of these peculiar photoreceptive cells (Vöcking et al., 2017). Xenopsins have also been shown to be expressed in the adult eyes of bivalves, where they have largely duplicated, but only one paralogue shows high levels of expression (Hasan et al., 2024). Furthermore, xenopsins are expressed with r-opsin1, retinochrome, and peropsin in the photoreceptive cells of the terrestrial slug *Lymax* (Matsuo et al., 2023). The role of xenopsins in these organisms is largely unknown, but considering their expression patterns, mostly located in photoreceptors and visual organs, it seems likely that they play an active role in photoreception and visual information processing.

Our RT-qPCR results and the transcriptome of *E. moschata* reveal that xenopsins are expressed in the visual system, specifically in the eyes of these animals. Additionally, by phylogenetic inference it can be suggested that xenopsins are potentially expressed in the eyes of all octopods or, considering also the xenopsin expression in the bobtail squid eye, all cephalopods. Furthermore, the fact that xenopsins are not only duplicated but that both genes are expressed in the eyes of *A. argo*, a species subjected to water column excursions and different light environments, might reinforce the xenopsins role in visual processing. Xenopsins are related to c-opsins and, like them, possess a conserved NKQ motif, meaning they may bind the same proteins as c-opsins to initiate a signalling cascade that, in the case of c-opsins, leads to hyperpolarization. If this is confirmed and xenopsins activate a similar signalling cascade, this would mean that the effect of xenopsins opposes the depolarizing effect of the active r-opsin1, which is well known to be the main visual opsin in *Octopus* and other cephalopods. Since only r-opsin1 has been described in their retina, octopuses are said to be colour blind, but the low-level expression of a second opsin might indicate a different scenario. Indeed, since xenopsin is potentially capable of activating a signalling cascade in the eye, it could potentially lead to colour discrimination. For example, if a few retinal photoreceptive cells express both xenopsins and r-opsin1, a light spectrum that activates xenopsins might have an antagonistic effect (hyperpolarization) on the r-opsin1 signalling cascade (depolarization). This would mean that under certain wavelength activating both r-opsin ad xenopsin, cells expressing both opsins might not emit any signal, unlike those that only possess r-opsin1, providing the nervous system with the ability to potentially discriminate colour. This could partially explain their incredible camouflaging abilities since they are controlled by the central nervous system (Hanlon & Messenger, 1988). Furthermore, the ability to discriminate different wavelength in littoral environments can be essential to distinguish between the

random rapid light-intensity changes generated by the waves, causing flickering, and those generated a predator movement (Maximov, 2000). Unfortunately, this is only speculative since it was impossible to show the expression of xenopsins via colorimetric *in situ* hybridization and the location of this protein in the *O. vulgaris* eye remains enigmatic. In the meantime, high-coverage RNA sequencing of *C. macropus*, separating the eye and optic lobe, is being performed to see if these observations can be extended to other octopods and HCR *in situ* hybridization of xenopsins in *O. vulgaris* might overcome the limits of classic *in situ* hybridization colorimetric techniques. Additionally, immunohistochemical staining might contribute to localize these molecules in the retina. However, our observation is consistent with previously described data in other cephalopods (Yoshida & Ogura, 2011) and considering that photoreceptors expressing xenopsins and r-opsins have also been described in other lophotrochozoans, it might be that this is a shared characteristic within this superphylum. This suggests a scenario where the ancestral lophotrochozoan photoreceptor possessed both microvilli and cilia, expressing r-opsins and xenopsins together (Vöcking et al., 2017). From this, different photoreceptor types specialized, sometimes preserving the original opsin expression and other times maintaining just a single opsin and losing the other. The presence of xenopsin in the gonads of the bobtail squid might also suggest a different role for this opsin beyond visual processing, possibly related to seasonality and reproduction.

CHAPTER 5 – R-opsin1 and retinochrome sites under selection

As shown by their complex camera eyes and well-developed large optic lobes, coleoid cephalopods (cuttlefish, squid, and octopuses) rely heavily on their visual system to survive in their environment. Visual sensitivity in this group is used to perform a wide range of tasks, such as predator avoidance, social behaviours, mating, and locating prey or other potential food sources (Hanke & Kelber, 2020). Given the importance of vision to cephalopods, it is expected that different adaptations will be found in species living under various light conditions, including changes involving the main visual molecules, the opsins, mirroring a well-known and documented phenomenon in deep-sea fishes and other groups (Feuda et al., 2016; Hagen et al., 2023b; Musilova et al., 2019, 2021; Ricci et al., 2022).

Unlike vertebrates and arthropods, which use different opsin orthologues to discriminate colours, cephalopods classically rely on only two opsins for their visual tasks: the visual opsin r-opsin1 and the photoisomerases retinochrome (Bonadè et al., 2020; Chung & Marshall, 2016; Hanke & Kelber, 2020; Hara & Hara, 1972; Hara, Tomiyuki et al., 1972; Yoshida et al., 2015). Indeed, as previously shown, only these two opsins have a sufficient level of expression in the cephalopod retina, underscoring their key role in vision. The r-opsin1 and retinochrome visual cycle has been extensively studied and traced back to the molluscan common ancestor (Vöcking et al., 2021).

A previous study was conducted to detect differences in r-opsin1 among different cephalopods living at various depths, but it did not include any deep-sea

species of octopods, except for *Enteroctopus dofleini*, that migrates to the deep sea only for reproduction and usually lives in littoral environments (Chung & Marshall, 2016; Jereb et al., 2005). Furthermore, selective pressure on retinochrome has never been investigated. To address these gaps, a positive selection analysis of r-opsin1 and retinochrome in six incirrate octopod species from the Gulf of Naples spanning all across the phylogeny of the group (Uribe & Zardoya, 2017), including three deep-sea species (reaching depths >500 m), was performed. The advantage of using species from the same geographical area is that the impact of other factors (such as turbidity or latitude) is minimized since they affect all the species evenly, making depth the primary parameter that can be used as a proxy for light.

5.1 Results

5.1.1 Ancestral State Reconstruction

Before conducting selection analysis, the ancestral state of the different incirrate octopod nodes was reconstructed using Phytools package on R (Revell, 2024) by mapping the maximal depth at which they live onto the species phylogeny. Our results show that the common ancestor of incirrate octopods most likely inhabited the disphotic zone, living in a mesopelagic environment (Figure 5.1). Subsequently, certain branches independently adapted to more superficial or deeper levels of the water column. With this results in mind, positive selection analyses were performed using CodeML implemented in PAML.

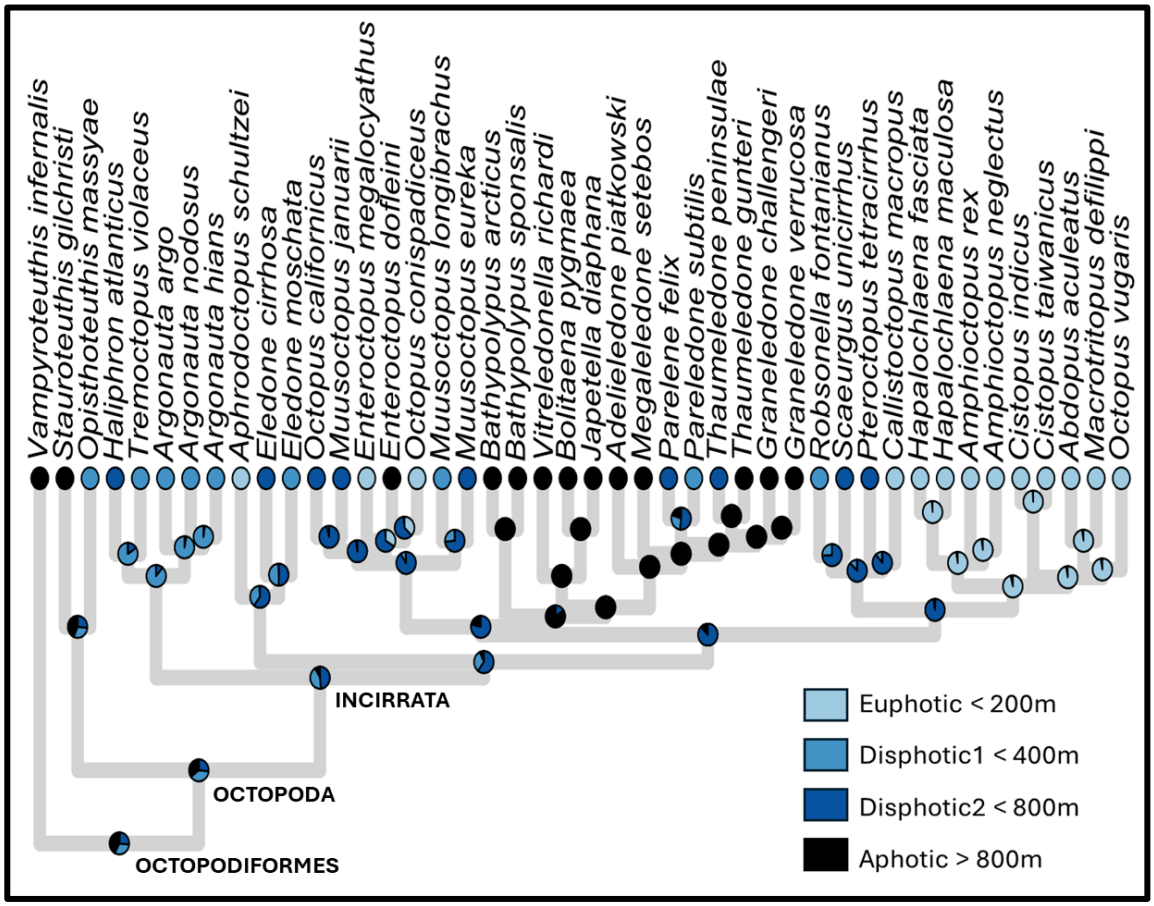


Figure 5.1 - Ancestral state reconstruction of Octopodiformes ancestral photic ecology. The phylogenetic tree illustrates the light ecology of different octopodiformes species (tips), with the corresponding ancestral states indicated by pie charts (nodes). Depth has been used as a proxy for estimating light ecology. Our analysis suggests that the littoral/euphotic and aphotic environments were most likely independently conquered by many lineages, originating from an ancestral disphotic condition.

5.1.2 R-opsin1 sites under positive selection

For r-opsin1, the best-fitting model indicates that only *Pteroctopus tetracirrhus* is under positive selection (according to M1: dN/dS ratio in *P. tetracirrhus* 0.4579, *O. vulgaris* 0.0704, *C. macropus* 0.0219, *S. unicirrhus* 0.0447, *E. moschata* 0.2055, *E. cirrhosa* 0.1191, *A. argo* 0.0981; and M2: *P. tetracirrhus* foreground dN/dS 0.4457, background branches 0.0803). This is also supported by the higher Likelihood ratio test (LRT) and lower p-value when compared to the group selection (deep-sea, littoral) alternative hypothesis (see table 5.1). Additionally, no sites under selection

were identified when the branch-site model was applied to the deep-sea and littoral groups as foregrounds. This was surprising but could be explained by the fact that *P. tetracirrhus* is the only analysed species inhabiting depths of up to 700 m, likely corresponding to a light-depleted bathybenthic environment in coastal areas. Therefore, among the species under investigation, it is the only one living in a markedly different environment, which may have led to positive selection.

Model comparison	**LRT**	**DOF**	**p. value**
M0 **VS** M1 R-opsin1	44.19269	10	3.039419e-06
M0 **VS** M2 R-opsin1 deep-sea	5.689126	1	0.01707035
M0 **VS** M2 R-opsin1 littoral	0.021866	1	0.8824441
M0 **VS** M2 R-opsin1 *P. tetracirrhus*	20.33408	1	6.503183e-06
BSM0 **VS** BSM1 R-opsin1 deep-sea	2.614638	1	0.1058817
BSM0 **VS** BSM1 R-opsin1 littoral	6.109282	1	0.01344737
BSM0 **VS** BSM1 R-opsin1 *P. teracirrhus*	40.27283	1	1 2.208588 e^{-10}

Table 5.1 – Comparison between different positive selection models on r-opsin1. The table shows the comparison between the different positive selection models in PAML. LRT indicates Likelihood Ratio Test, DOF difference in degrees of freedom. *Argonauta argo* was used as outgroup for the analysis.

A total of 14 codon mutations were found to be under selection (probability >0.8) in *P. tetracirrhus*, according to Bayes Empirical Bayes (BEB) analysis (see table 5.2)

***Pteroctopus tetracirrhus* r-opsin1**

Mutation	Distance from the retinal (Å)	Effect	BEB	Notes
Y58F	>20	Loss of OH-	0.838	
I87V	8	Unknown	0.978*	
V166I	15	Unknown	0,854	Already known in other cephalopods - green shift
F201N	10	Unknown	0,987*	
L212T	12	Gain of OH- nearby the β-ionione ring	0.981*	
A218G	17	Unknown	0,851	
T269S	17	Change of OH-bearing ammino acid	0.865	In EL
M272L	8	Unknown	0.838	
I281V	11	Unknown	0.842	
T336S	>20	Addition of OH-	0.866	
A388S	C-Terminal	Addition of OH-	0.985*	
Q452E	C-Terminal	Unknown	0.838	

Table 5.2 – Sites under selection in *P. tetracirrhus* r-opsin1. The table shows the different sites detected to be under selection in *P. tetracirrhus*, their distance from the retinal and the potential effect of the mutation. * Indicates sites with BEB >0.95. EL is for extracellular loop. BEB= Bayes Empirical Bayes; EL= Extracellular Loop.

The 3-D molecular structure of *P. tetracirrhus* r-opsin1 was subsequently generated using the AlphaFold server, and the amino acids under positive selection were mapped onto the 3-D structure (Figure 5.1). Of these, 8 were found to be in proximity (< 18 Å) of the retinal when mapped onto the protein 3-D model (see Table 5.1). In particular, there are two mutations that might suggest a shift towards the red end of the absorption spectrum: V166I, a mutation already known in other cephalopods with a λmax of 500 nm, and L212T, which introduces the OH-bearing amino acid threonine near the β-ionone ring of the retinal and is known as a potential tuning site (Chung & Marshall, 2016; Hagen et al., 2023b).

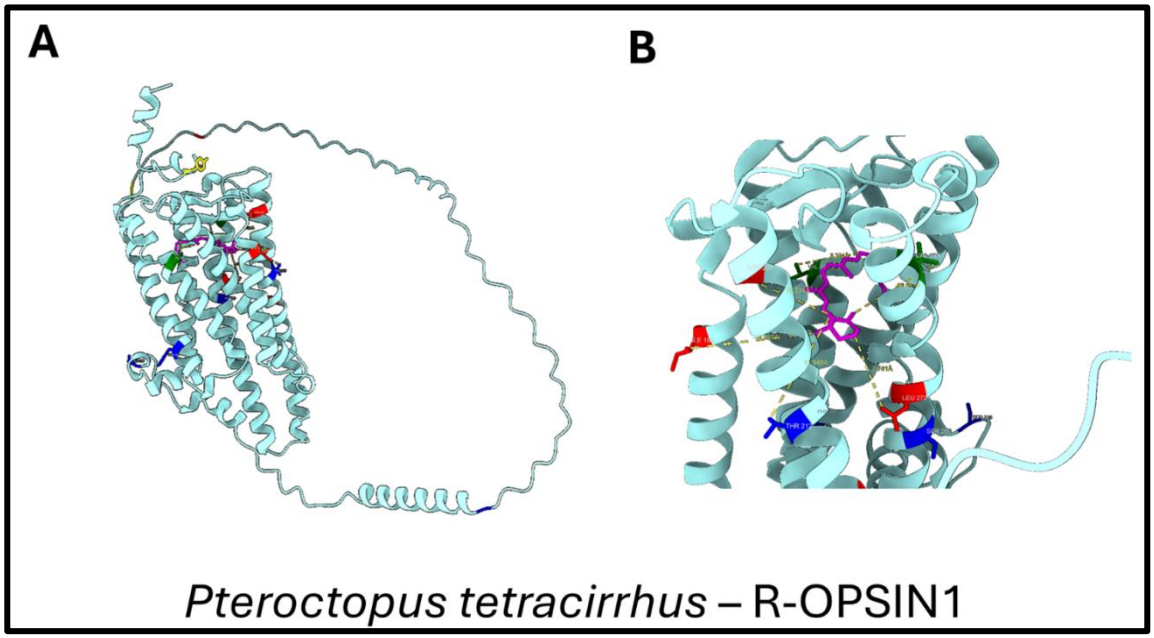


Figure 5.1 – *Pteroctopus tetracirrhus* r-opsin1 key tuning sites mapped on the 3-D protein structure. (A) Shows all the sites found to be under positive selection. (B) Shows sites near the retinal binding pocket (less than 18 Å from the retinal). Yellow indicates mutation sites that led to amino acids with a different electric charge; green indicates mutations involving the interchange between isoleucine (I) and valine (V); blue indicates mutations involving OH-bearing amino acids; red indicates mutations with an unknown effect.

5.1.3 Retinochrome sites under positive selection

The analysis of the photoisomerases retinochrome indicates that deep-sea species are more likely to have undergone positive selective pressure compared to littoral species (foreground ω ratio 0.3015, background ω ratio 0.1720). This is supported by the differences in LRT and p-values when compared to the null hypothesis (Table 5.3).

Model comparison	**LRT**	**DOF**	**p. value**
M0 **VS** M1 Retinochrome	24.84318	10	0.005650809
M0 **VS** M2 Retinochrome deep-sea	4.506174	1	0.0337727
M0 **VS** M2 Retinochrome littoral	0.670426	1	0.4129031
BSM0 **VS** BSM1 Retinochrome deep-sea	5.5413	1	0.01857278
BSM0 **VS** BSM1 Retinochrome littoral	0	1	1

Table 5.3 – Comparison between different positive selection models on retinochrome. The table shows the comparison between the different positive selection models in PAML. LRT indicates Likelihood Ratio Test, DOF difference in degrees of freedom. *Argonauta argo* was used as outgroup for the analysis.

A total of 14 codon mutations were found to be under selection according to the Bayes Empirical Bayes (BEB) analysis. Out of these, 8 codons were under positive selection, with a probability >50% for the positive selection class (ω > 1). Notably, only 2 sites (190 and 289) showed a probability >80% and no one >95% (Table 5.4).

Scaeurgus unicirrhus retinochrome

Mutation	Distance from the retinal (Å)	Effect	BEB	Notes
H107Q	11	Loss of positive charge	0.624	Shared with *C. macropus*
V152I	14	Unknown	0.658	
F181Y	7	Gain of OH- nearby the ring	0.743	Many
I190C	10	Unknown	0.856	
T263M	18	Loss of OH-	0.545	In EL
L289M	18	Unknown	0.942	Shared with *E. moschata*
G292S	>20	Gain of OH-	0.736	

Pteroctopus tetracirrhus retinochrome

Mutation	Distance from the retinal (Å)	Effect	BEB	Notes
H107D	11	Loss of positive charge; gain of negative charge	0.624	
V152I	16	Unknown	0.658	
I190V	9	Unknown	0.856	
G292A	>20	Unknown	0.736	

Eledone cirrhosa retinochrome

Mutation	Distance from the retinal (Å)	Effect	BEB	Notes

F181Y	6	Gain of OH- nearby the ring	0,743	
T263A	18	Loss of OH-	0.545	In EL, shared with *E. moschata*

Table 5.4 – Sites under selection in *S. unicirrhus*, *P. tetracirrhus* and *E. cirrhosa* retinochrome. The table shows the different sites detected to be under selection in the retinochrome of the three deep sea species, their distance from the retinal and the potential effect of the mutation. BEB= Bayes Empirical Bayes; EL= Extracellular Loop.

The 3-D molecular structures of *S. unicirrhus*, *P. tetracirrhus*, and *E. cirrhosa* were produced using the AlphaFold server, with amino acids under positive selection mapped onto the 3-D models and the species phylogeny (Figure 5.5 and Figure 5.6). Among these, five amino acids were found to be in proximity (<20 Å) to the retinal when mapped onto the protein 3-D model: five in *S. unicirrhus*, four in *P. tetracirrhus*, and two in *E. cirrhosa* (see the table above). Interestingly, mutations involving changes in OH-bearing amino acids are commonly shared between *S. unicirrhus* and *E. cirrhosa*. These include T263M and T263A (the latter also shared with *E. moschata*), which lead to the loss of an OH-bearing amino acid, and F181Y, which leads to the gain of an OH-bearing amino acid. This mutation occurred near the β-ionone ring.

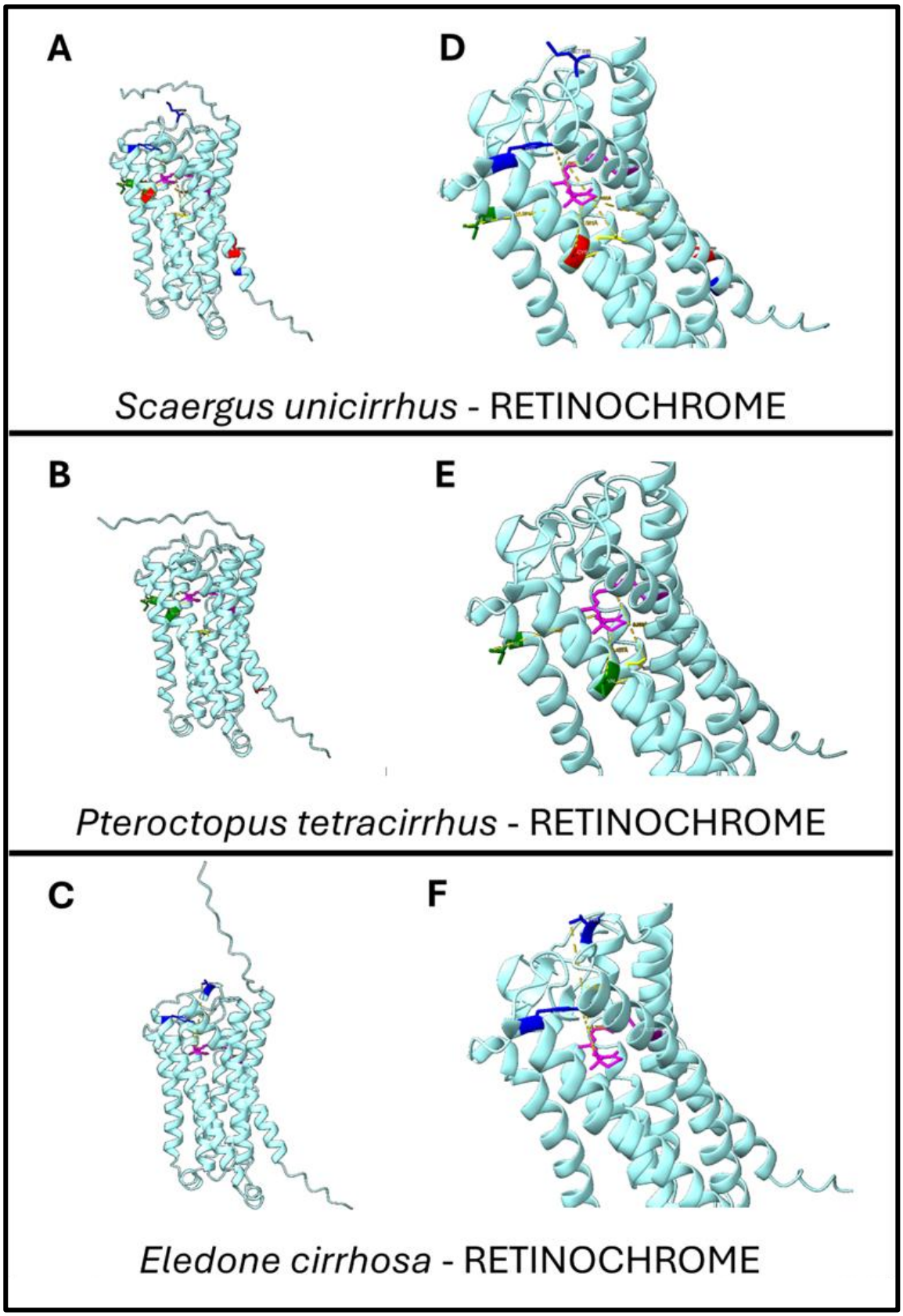


Figure 5.6 - Retinochrome key tuning sites mapped on the 3-D protein structure of the 3 deep sea species. Panels A, B, and C display all the sites found to be under positive selection in the three deep-sea species *S. unicirrhus*, *P. tetracirrhus*, and *E. cirrhosa*, respectively. Panels D, F, and G highlight sites near the retinal binding pocket (less than 18

Å from the retinal) in these three species. Yellow indicates mutation sites that resulted in amino acids with a different electric charge, green indicates mutations involving the interchange between isoleucine (I) and valine (V), blue indicates mutations involving OH-bearing amino acids, and red indicates mutations with an unknown effect.

Furthermore, *P. tetracirrhus* and *S. unicirrhus* share three mutations, most of which have unknown effects. The sole exception is the mutation at site 107 (H107Q and H107D), which involves the loss of a positively charged amino acid in *S. unicirrhus* and the gain of a negative charge in *P. tetracirrhus*. H107Q is also shared with *C. macropus*, a littoral nocturnal species related to *S. unicirrhus* and *P. tetracirrhus*. Changing the electrical charges surrounding the retinal might potentially influence isomerization. Additionally, the two other shared mutations (V152I and I190V) between the two species involve the interchange between isoleucine and valine, which, in other opsins, has been indicated to have a potential key tuning effect, though the mechanism remains unclear.

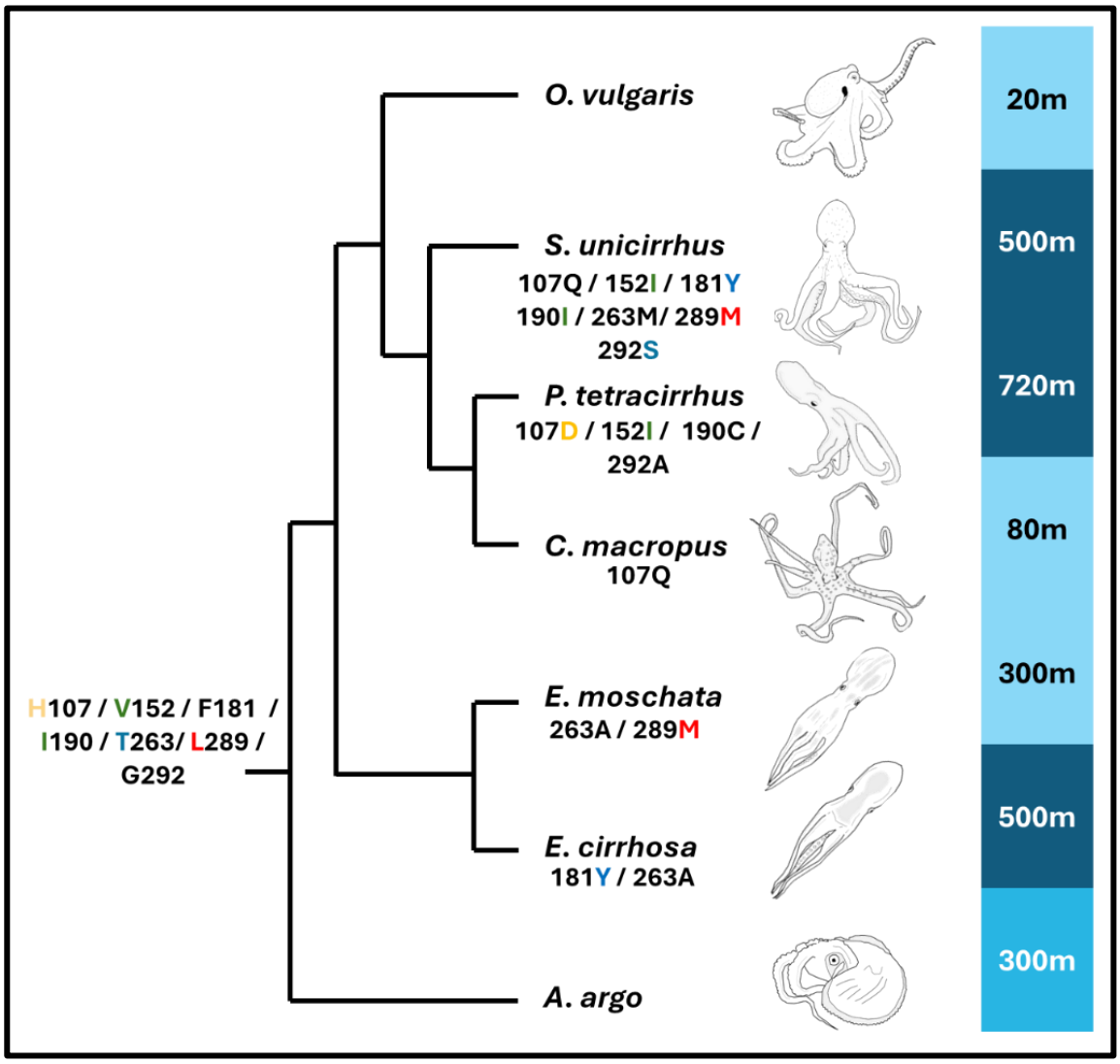


Figure 5.6 - Retinochrome sites under positive selection – The figure illustrates the variations of detected positively selected sites across all the species analysed. Yellow indicates mutation sites that resulted in amino acids with a different electric charge, green indicates mutations involving the interchange between isoleucine (I) and valine (V), blue indicates mutations involving OH-bearing amino acids, and red indicates mutations with an unknown effect.

5.2 Discussion

The ancestral photic ecology estimation in the main octopodiformes node, obtained through ancestral state reconstruction with Phytools, indicates secondary migrations into euphotic and aphotic environments among modern lineages and a deep-sea disphotic common ancestor. This scenario is conform to previous results (Klug et al., 2023) and coherent with the emergence of octopodiformes in the Cretaceous, where they might have evolved as benthic agile predators in the deep-sea, while decapodiformes conquered the nektonic euphotic realm during the Mesozoic marine revolution (Tanner et al., 2017). Despite that, the retinochrome has been found to be under selection only in deep-sea species and not in littoral species. This discrepancy might be due to inability to detect convergent evolution and reconstruct the ancestral state of the molecule when there is a lack of opsin data of the outgroup, such as the absence of sequences of cirrate octopods in the dataset, which are mostly deep sea and might support a different ancestral condition (Shea et al., 2018).

The analysis conducted in r-opsin1 indicates that there are no signals of positive selection when deep-sea and littoral species are confronted. This can have a clear explanation when the secondary epipelagic migration scenario is considered: while epipelagic species can rely on a wide light spectrum and adapt their absorption peaks to different light wavelengths, deep-sea species must adapt to the limited number of wavelengths available. Consequently, a species living in a euphotic environment might exploit wavelengths that are not present in the deep sea (e.g., red light) and must adapt accordingly when migrating into the deep sea (by shifting towards blue light). On the contrary, the wavelengths that are already available in the deep sea are already present in the surfaces and, consequently, deep sea species living in a disphotic environment are already adapted to exploit sunlight.

Furthermore, light wavelengths that can more efficiently penetrate the water column can also travel longer distances horizontally, enhancing the ability to detect distant objects without necessarily justifying any evolutionary change. This is definitely important for monochromatic organisms that rely more on long distance polarized light than colour discrimination such as octopods (Temple et al., 2021). For example, the littoral species *O. vulgaris* maintain a λmax of ~475 nm that is compatible with light spectra that reach deep-sea environments (Figure 5.7) (Chung & Marshall, 2016; Inoue et al., 2007). In this aspect, the migration of octopods from disphotic to euphotic areas might resemble the mammal nocturnal bottleneck occurred in the Mesozoic, an event that caused the loss SWS2 and RH2 opsins, and that resulted in a reduced opsin complement in placental mammals, which only conserved those essential for nocturnal vision, RH1, and the discrimination of blue and red wavelengths SWS1 and LWS (Jacobs, 2009).

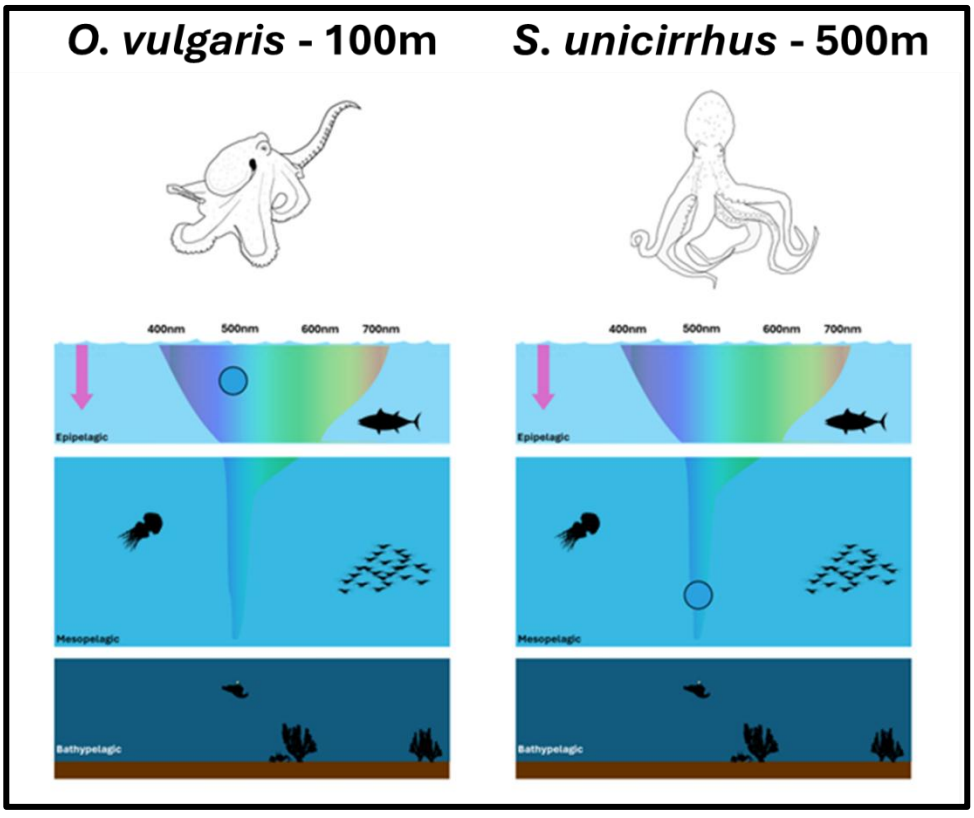


Figure 5.7 – No differences in λmax between littoral and mesopelagic deep-sea species. Despite inhabiting different depths, the λmax (~485 nm) of *O. vulgaris* r-opsin1

aligns with the light spectrum available in the disphotic zone. Therefore, no significant differences in λmax are expected between *O. vulgaris* and the deep-sea species *S. unicirrhus*.

On the other hand, bathypelagic octopod species and those migrating into deeper, sunlight-depleted, areas of the water column may experience different selective pressures. This scenario could explain why *Pteroctopus tetracirrhus* is the only species among those analysed found with positively selected sites. Indeed, living in depths of up to 700 meters, *P. tetracirrhus* represents the only bathybenthic octopod species in this study, revealing important adaptations to this environment. Unfortunately, little is known about key tuning sites in cephalopods, but two mutations in *P. tetracirrhus* may indicate a shift in the peak of absorption: V166I, which is present in species with a λmax shifted towards the red relative to other octopods (500 nm); and L212T, which introduces an additional OH-bearing amino acid (threonine) near the β-ionone ring of the retinal (Chung & Marshall, 2016). By adding a negative charge far from the binding site, these mutations help to distribute the free positive charge from the protonated Schiff base to the polyene chain of the retinal, reducing the energy required (shorter wavelength) for configurational changes. In short, they contribute to a red shift in the peak of absorption, potentially making the molecule more sensitive to green light in this species (Hagen et al., 2023b; Sekharan et al., 2012).

The eyes of *P. tetracirrhus* are notably larger compared to those of other littoral and epibenthic species (up to 3 cm, whereas in *S. unicirrhus* and *O. vulgaris*, they are not larger than 2 cm, approximately the size of a human eye) (Hanke & Kelber, 2020). Eyes are energetically expensive tissues, and because modification in the eye size occurs to accommodate low light levels and enhance vision in low-light environments, it can be deduced that vision is an important sense in this

species. Furthermore, living in a sunlight-depleted environment, *P. tetracirrhus* must rely on other light sources, the most abundant of which in the ocean is bioluminescence. Notably, while bioluminescence is generally blue in pelagic animals, in benthic species, the emission shifts toward green. This shift is significant for benthic predators like octopods, as green bioluminescence could be emitted by potential prey, such as benthic decapods, or by the organisms that host the prey, such as corals (Johnsen et al., 2012). Therefore, all these data combined suggest that *P. tetracirrhus* r-opsin1 might reveal particular adaptation to benthic bioluminescence in this species (Figure 5.8).

Although this scenario is hypothetical, it could be tested *in vitro* in the future and try to express the full r-opsin1 sequence of both *O. vulgaris* and *P. tetracirrhus* to measure their absorption peaks using the parallel sensitive heterologous expression methods (Liénard et al., 2022). However, challenges remain, particularly with the binding between the opsin and the retinal molecules.

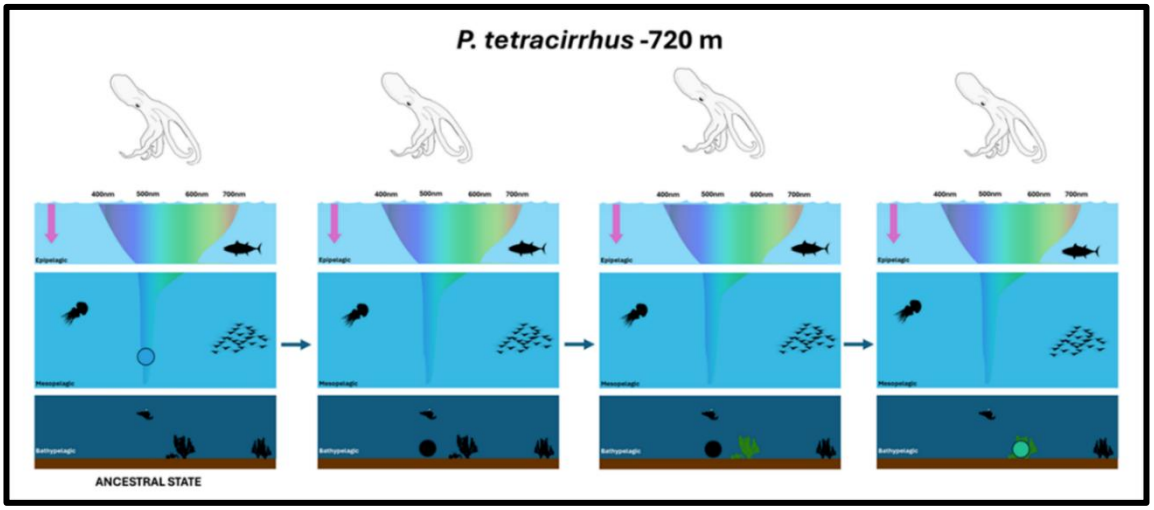


Figure 5.8 – Hypothetical changes in λmax of *P. tetracirrhus* r-opsin1. Living at sunlight-depleted depths, *P. tetracirrhus* must rely on other light sources, such as the green-light emission from bioluminescent deep-sea species. The picture illustrates the hypothetical evolutionary transition in *P. tetracirrhus*, transitioning from a disphotic (mesopelagic) environment to an aphotic (bathypelagic) environment. Not relying on blue

light reaching the deep sea, *P. tetracirrhus* might have adapted its absorption spectra to the green bioluminescent emission of bathypelagic prey.

Less is known about the potential effects of mutations on retinochrome. Unlike most visual opsins, such as c-opsins, retinochrome does not produce a signalling cascade, and is unique in that it only binds to the retinal in its all-*trans* form and converts it back to 11-*cis*, thereby restoring the retinal to its initial configuration and only subsequently bleaches the retinal to make it available in the cell (Hara, Tomiyuki et al., 1972; Inoue et al., 2007; Terakita, 2005; Vöcking et al., 2021; Zhang et al., 2021). While this mechanism is well-documented, retinochrome is not the only opsin capable of photo reversing the retinal in cephalopods (e.g., being bistable, r-opsin one can also reverse the retinal) and the exact retinochrome role remains to be discovered. If this is taken into consideration together with the absence of known key tuning sites in this molecule, predicting the potential effects of mutations becomes very challenging.

Some of the positively selected sites detected, particularly those involving changes in charge, may suggest a shift in the retinal's absorption maximum (λmax), potentially altering the energy required to revert the retinal. Similar considerations apply to amino acid changes involving OH groups. However, the nature and consequences of these changes are highly speculative, and a more detailed assessment of retinochrome's absorption spectra and retinal binding proprieties is necessary. Mutation can also be the result of adaptations needed to better perform under different physical conditions, such as pressure, or to perform different tasks other than photoreception that are unknown.

Three of the shared mutations on sites under selective pressure found in the retinochrome are also present in littoral species, making it difficult to definitively

identify them as key tuning sites. It is possible that certain energy levels, which are not available in the deep sea, are preferred for restoring the retinal. Conversely, the peak of absorption may be maintained between deep-sea and littoral conditions, but mutations might enhance the ability to restore the retinal configuration under low light levels in deep-sea species.

In addition, to adapting to varying light conditions encountered at different depths within the water column, deep-sea organisms must also contend with high pressures that can compress protein structures (Somero, 1992). This compression can impact essential physiological processes. To manage high-pressure environments, animals have developed both extrinsic and intrinsic adaptations. Extrinsic adaptations include modifications to cellular and membrane composition to reduce protein compressibility, while intrinsic adaptations involve amino acid substitutions in proteins and changes in adiabatic compressibility (Porter et al., 2016; Somero, 1992). Notably, some of these mutations have been identified in cephalopods primarily occurring in the outer regions of the protein (Porter et al., 2016). Despite some amino acid substitutions shared among related littoral and deep-sea species, none of the sites identified by Porter (2016) are under selection or present ammino acidic substitutions in deep sea branches in both retinochrome and r-opsins. However, it is still possible that the sites here identified under selection may show adaptation to high hydrostatic pressure environments.

CHAPTER 6 – Conclusions

Octopods are fascinating animals. Despite diverging from vertebrates 600 million years ago at the Bilateria node, they independently evolved complex traits, including a well-developed nervous system and sophisticated behaviour. Remarkably, like vertebrates, octopods extensively rely on vision, as evidenced by their complex camera-like eyes, often cited as an extraordinary example of evolutionary convergence (Hanke & Kelber, 2020; Serb & Eernisse, 2008; Yoshida et al., 2015; Yoshida & Ogura, 2011). However, studies on octopod adaptation to different light conditions are scarce, with most focusing on species that predominantly inhabit littoral environments (Chung & Marshall, 2016). To understand octopod adaptation to different light conditions, our aim was to study the opsins — the main visual molecules — in deep-sea and littoral octopod species to determine whether adaptations have occurred.

The original aim of the project was to characterize all the opsins across various littoral and deep-sea species, utilizing the diversity of species found in the Gulf of Naples. These included several incirrate octopods: the littoral species *Octopus vulgaris*, *Callistoctopus macropus*, and *Eledone moschata*; the deep-sea species *Scaeurgus unicirrhus*, *Pteroctopus tetracirrhus*, and *Eledone cirrhosa*; and the pelagic *Argonauta argo*.

From the outset, the project faced several challenges, primarily due to the lack of a clear understanding of opsins in octopods. Few data existed in the literature. For instance, only four opsins were known in octopods based on the genome of *Octopus bimaculoides*: r-opsin1 (previously known as rhodopsin), r-opsin2 (previously known as rhabdomeric opsin), retinochrome, and peropsin (Albertin et al., 2015). A partial sequence of xenopsin from the same genome was

also present in the NCBI database as a result of automatic genome prediction. Interestingly, while the *O. bimaculoides* sequence lacked the 3’ region, a sequence of *O. sisensis* (the Asian *Octopus vulgaris* variety) published on NCBI lacked the 5’ region. This combination enabled the design of 5’ and 3’ primers, allowing for the cloning of an almost complete sequence of xenopsin from *O. vulgaris* cDNA. Xenopsin was amplified via RT-PCR specifically from cDNA synthesized from extracted eye RNA, suggesting the presence of a second opsin in the eye and highlighting that many opsins remained uncharacterized in octopods.

Studies conducted in other cephalopod groups showed that most opsins are expressed, but the results varied due to different developmental stages being examined (Bonadè et al., 2020; Yoshida et al., 2015). These studies also demonstrated xenopsin expression in the eye, although it was initially classified as "c-opsin," as the xenopsin had not yet been described as a separate opsin group. Furthermore, peropsin expression was absent, and two clades of retinochrome were present, which was later clarified as a characteristic of decapodiform cephalopods.

Taken together, it became clear that a clarification of the cephalopod and octopod opsin complement was necessary before proceeding further (De Vivo et al., 2023). Since data were also missing for other lophotrochozoan groups, the analysis was expanded to include the entire clade. Translated genomes from various species were mined for opsins, and modern phylogenetic techniques for gene-tree species-tree reconciliation were used. The results were remarkable, showing that the common ancestor of lophotrochozoans possessed all the opsins found in Bilateria, a condition similar to that in early deuterostomes. Many modern species still retain this complex photoreceptive molecular machinery. The analysis also revealed that the complexity of visual structures in molluscs led to a reduction in the opsin complement. Cephalopods, in particular, have the lowest number of

opsins compared to their relatives. The cephalopod common ancestor possessed five opsins (r-opsin1, r-opsin2, xenopsin, retinochrome, and peropsin) and an opsin-like molecule, which was named pseudopsin. The function of pseudopsin remains largely unknown, though it may be related to placopsins. Despite the uncertain nature of this new protein group, it represents a new opsin outgroup found in bilaterians that requires further investigations. From the ancestral opsin toolkit, decapodiforms lost peropsins and duplicated retinochromes, while octopodiforms maintained the ancestral set, with some divergences, such as *Argonauta argo*, which duplicated xenopsins (Yoshida et al., 2022). The implementation of *in situ* hybridization or immunofluorescence techniques might provide new clues about the position of these molecules within the different tissues.

Once the ancestral opsin toolkit was reconstructed, a preliminary RT-qPCR analysis was conducted to understand the expression patterns of the identified opsin genes in octopods. Opsin sequences were collected from the two recently published genomes of *A. argo* and *O. vulgaris*, which, along with the genomes of two other *Octopus* species and *Eledone cirrhosa* (the last published on NCBI only recently, in March 2024), represent the only available octopod genomes. Primers were designed based on these sequences. Fortunately, *A. argo* and *O. vulgaris* are phylogenetically distant, with their common ancestry tracing back to the incirrate octopod common ancestor, allowing for inferences on other octopods based on their conserved expression. In *O. vulgaris*, all opsins were expressed in the visual system (eye and optic lobe), with xenopsin specifically expressed in the eye and r-opsin2 and peropsin in the optic lobe. Opsins were also present in the suckers, though it is unclear whether this is due to the presence of skin or a higher concentration of opsins in these organs —specifically, r-opsin1 and retinochrome. Surprisingly, retinochrome was expressed in the retractor muscle, which was intended to serve

as a reference tissue, suggesting that this opsin may have functions beyond its role as an accessory photoisomerases to r-opsin1 and may be involved in other mechanisms. Pseudopsin was expressed in all tissues, and its nature remains under investigation. Similar results were obtained in *A. argo*, although the presence of r-opsin2 and peropsin could not be confirmed in all the tissues analysed. Overall, the findings were consistent with those of a previous study on the decapodiform bobtail squid, indicating a good level of conservation in opsin expression among cephalopods.

With these results in mind, low-coverage RNA sequencing of the visual system of all species under investigation, for which genomes were not available, was performed to obtain opsin sequences. Based on our knowledge of *O. vulgaris*, we expected to find sequences for all five opsin genes. However, this was overly optimistic, as low coverage likely hindered full gene identification, and in cases where all five genes were found, such as in *E. moschata*, they were often partial sequences. On the other hand, through a combination of transcriptome data and gene cloning, full-length sequences of r-opsin1 and retinochrome were obtained for all the 5 species. These opsins are likely to play key roles in vision and are more likely to experience selective pressure due to varying light conditions.

Finally, we performed selection analysis. Various models were tested, particularly to check for positive selection signals across different groups (littoral vs. deep-sea, with *A. argo* considered littoral) or on individual branches. Our analysis revealed that r-opsin1 is under selection only in *P. tetracirrhus*, a species living at greater depths than the others in this study. Based on the mutation sites, it is hypothesized that the absorption peak shifted toward the green spectrum, likely to match the bioluminescence patterns of deep-sea benthic prey. Conversely, retinochrome showed signs of positive selection in all deep-sea species, but the

impact on its absorption peak remains unclear, as no previous studies have investigated the effects of mutations in these molecules. To address these limitations and validate our observations regarding *P. tetracirrhus* r-opsin1, next step will be to synthesize these molecules and measure their absorption peaks.

In sum, octopods use the visual opsin r-opsin1 (known as rhodopsin) and retinochrome in their visual cycle, a trait shared with other cephalopods and molluscs (Vöcking et al., 2021). Additionally, they retain other opsins, including r-opsin2, xenopsin, peropsin, and potentially pseudopsins with currently uncharacterized functions. Despite this, all of these opsins were found to be expressed, suggesting they play a role in octopod photoreception and photophysiology. Interestingly, r-opsin1 does not appear to be under selective pressure in deep-sea species inhabiting mesophotic environments. In contrast, positive selection was found in the only species under investigation that potentially lives in a completely light-depleted environment. The reason for this remains speculative and may be related to the detection on bioluminescent preys. Intriguingly, retinochrome shows signs of positive selection in all deep-sea species but the explanations remain enigmatic. To make new hypothesis and testing them, in vitro expression assays is currently being performed, as little is known about key tuning sites in cephalopods. Our analysis highlights several limitations due to gaps in knowledge, emphasizing the need for further sequencing efforts and *in vitro* tests. This will not only help fill these gaps but also shed light on the evolution of photoreception in these animals. For instance, further research could focus on studying outgroups, such as the opsins of cirrate octopods and vampyromorphids, or exploring the role of non-visual opsins in these animals. We have only just opened the door to this fascinating field, and future efforts will help us better understand one

of nature’s most intriguing phenomena: the endless strategies living organisms use to perceive the world around them.

References

Adema, C. M. (2021). Sticky problems: Extraction of nucleic acids from molluscs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1825), 20200162. https://doi.org/10.1098/rstb.2020.0162

Albertin, C. B., Simakov, O., Mitros, T., Wang, Z. Y., Pungor, J. R., Edsinger-Gonzales, E., Brenner, S., Ragsdale, C. W., & Rokhsar, D. S. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature*, *524*(7564), 220–224. https://doi.org/10.1038/nature14668

Al-Soudy, A.-S., Maselli, V., Galdiero, S., Kuba, M. J., Polese, G., & Di Cosmo, A. (2021). Identification and Characterization of a Rhodopsin Kinase Gene in the Suckers of Octopus vulgaris: Looking around Using Arms? *Biology*, *10*(9), 936. https://doi.org/10.3390/biology10090936

Álvarez-Carretero, S., Kapli, P., & Yang, Z. (2023). Beginner's Guide on the Use of PAML to Detect Positive Selection. *Molecular Biology and Evolution*, *40*(4), msad041. https://doi.org/10.1093/molbev/msad041

Arendt, D. (2017). The enigmatic xenopsins. *eLife*, *6*, e31781. https://doi.org/10.7554/eLife.31781

Belyaeva, O. V., Adams, M. K., Popov, K. M., & Kedishvili, N. Y. (2019). Generation of Retinaldehyde for Retinoic Acid Biosynthesis. *Biomolecules*, *10*(1), 5. https://doi.org/10.3390/biom10010005

Bernsel, A., Viklund, H., Hennerdal, A., & Elofsson, A. (2009). TOPCONS: Consensus prediction of membrane protein topology. *Nucleic Acids Research*, *37*(Web Server issue), W465-468. https://doi.org/10.1093/nar/gkp363

Bok, M. J., Porter, M. L., Ten Hove, H. A., Smith, R., & Nilsson, D.-E. (2017). Radiolar Eyes of Serpulid Worms (Annelida, Serpulidae): Structures, Function, and Phototransduction. *The Biological Bulletin*, *233*(1), 39–57. https://doi.org/10.1086/694735

Bonadè, M., Ogura, A., Corre, E., Bassaglia, Y., & Bonnaud-Ponticelli, L. (2020). Diversity of Light Sensing Molecules and Their Expression During the Embryogenesis of the Cuttlefish (Sepia officinalis). *Frontiers in Physiology*, *11*, 521989. https://doi.org/10.3389/fphys.2020.521989

Carrillo-Baltodano, A. M., Seudre, O., Guynes, K., & Martín-Durán, J. M. (2021). Early embryogenesis and organogenesis in the annelid Owenia fusiformis. *EvoDevo*, *12*(1), 5. https://doi.org/10.1186/s13227-021-00176-z

Chiao, C.-C., Wickiser, J. K., Allen, J. J., Genter, B., & Hanlon, R. T. (2011). Hyperspectral imaging of cuttlefish camouflage indicates good color match in the eyes of fish predators. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(22), 9148–9153. https://doi.org/10.1073/pnas.1019090108

Chung, W.-S., & Marshall, N. J. (2016). Comparative visual ecology of cephalopods from different habitats. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1838), 20161346. https://doi.org/10.1098/rspb.2016.1346

D'Aniello, S., Delroisse, J., Valero-Gracia, A., Lowe, E. K., Byrne, M., Cannon, J. T., Halanych, K. M., Elphick, M. R., Mallefet, J., Kaul-Strehlow, S., Lowe, C. J., Flammang, P., Ullrich-Lüter, E., Wanninger, A., & Arnone, M. I. (2015). Opsin evolution in the Ambulacraria. *Marine Genomics*, *24*, 177–183. https://doi.org/10.1016/j.margen.2015.10.001

Davies, W. L., Hankins, M. W., & Foster, R. G. (2010). Vertebrate ancient opsin and melanopsin: Divergent irradiance detectors. *Photochemical &*

Photobiological Sciences, *9*(11), 1444–1457. https://doi.org/10.1039/c0pp00203h

De Vivo, G., Crocetta, F., Ferretti, M., Feuda, R., & D'Aniello, S. (2023). Duplication and Losses of Opsin Genes in Lophotrochozoan Evolution. *Molecular Biology and Evolution*, *40*(4), msad066. https://doi.org/10.1093/molbev/msad066

Derelle, R., Philippe, H., & Colbourne, J. K. (2020). Broccoli: Combining Phylogenetic and Network Analyses for Orthology Assignment. *Molecular Biology and Evolution*, *37*(11), 3389–3396. https://doi.org/10.1093/molbev/msaa159

Destanović, D., Schultz, D. T., Styfhals, R., Cruz, F., Gómez-Garrido, J., Gut, M., Gut, I., Fiorito, G., Simakov, O., Alioto, T. S., Ponte, G., & Seuntjens, E. (2023). A chromosome-level reference genome for the common octopus, *Octopus vulgaris* (Cuvier, 1797). *G3: Genes, Genomes, Genetics*, *13*(12), jkad220. https://doi.org/10.1093/g3journal/jkad220

Döring, C. C., Kumar, S., Tumu, S. C., Kourtesis, I., & Hausen, H. (2020). The visual pigment xenopsin is widespread in protostome eyes and impacts the view on eye evolution. *eLife*, *9*, e55193. https://doi.org/10.7554/eLife.55193

Dungan, S. Z., Kosyakov, A., & Chang, B. S. W. (2016). Spectral Tuning of Killer Whale (*Orcinus orca*) Rhodopsin: Evidence for Positive Selection and Functional Adaptation in a Cetacean Visual Pigment. *Molecular Biology and Evolution*, *33*(2), 323–336. https://doi.org/10.1093/molbev/msv217

Dymond, J. S. (2013). Explanatory Chapter. In *Methods in Enzymology* (Vol. 529, pp. 279–289). https://doi.org/10.1016/B978-0-12-418687-3.00023-9

Fedonkin, M. A., & Waggoner, B. M. (1997). The Late Precambrian fossil Kimberella is a mollusc-like bilaterian organism. *Nature*, *388*(6645), 868–871. https://doi.org/10.1038/42242

Fernald, R. D. (2006). Casting a Genetic Light on the Evolution of Eyes. *Science*, *313*(5795), 1914–1918. https://doi.org/10.1126/science.1127889

Feuda, R., Hamilton, S. C., McInerney, J. O., & Pisani, D. (2012). Metazoan opsin evolution reveals a simple route to animal vision. *Proceedings of the National Academy of Sciences*, *109*(46), 18868–18872. https://doi.org/10.1073/pnas.1204609109

Feuda, R., Marlétaz, F., Bentley, M. A., & Holland, P. W. H. (2016). Conservation, Duplication, and Divergence of Five Opsin Genes in Insect Evolution. *Genome Biology and Evolution*, *8*(3), 579–587. https://doi.org/10.1093/gbe/evw015

Feuda, R., Menon, A. K., & Göpfert, M. C. (2022). Rethinking Opsins. *Molecular Biology and Evolution*, *39*(3), msac033. https://doi.org/10.1093/molbev/msac033

Feuda, R., Rota-Stabelli, O., Oakley, T. H., & Pisani, D. (2014). The comb jelly opsins and the origins of animal phototransduction. *Genome Biology and Evolution*, *6*(8), 1964–1971. https://doi.org/10.1093/gbe/evu154

Fleming, J. F., Feuda, R., Roberts, N. W., & Pisani, D. (2020). A Novel Approach to Investigate the Effect of Tree Reconstruction Artifacts in Single-Gene Analysis Clarifies Opsin Evolution in Nonbilaterian Metazoans. *Genome Biology and Evolution*, *12*(2), 3906–3916. https://doi.org/10.1093/gbe/evaa015

Frickey, T., & Lupas, A. (2004). CLANS: A Java application for visualizing protein families based on pairwise similarity. *Bioinformatics*, *20*(18), 3702–3704. https://doi.org/10.1093/bioinformatics/bth444

Gühmann, M., Jia, H., Randel, N., Verasztó, C., Bezares-Calderón, L. A., Michiels, N. K., Yokoyama, S., & Jékely, G. (2015). Spectral Tuning of Phototaxis by a

Go-Opsin in the Rhabdomeric Eyes of Platynereis. *Current Biology*, *25*(17), 2265–2271. https://doi.org/10.1016/j.cub.2015.07.017

Gühmann, M., Porter, M. L., & Bok, M. J. (2022). The Gluopsins: Opsins without the Retinal Binding Lysine. *Cells*, *11*(15), 2441. https://doi.org/10.3390/cells11152441

Guo, J., Parry, L. A., Vinther, J., Edgecombe, G. D., Wei, F., Zhao, J., Zhao, Y., Béthoux, O., Lei, X., Chen, A., Hou, X., Chen, T., & Cong, P. (2022). A Cambrian tommotiid preserving soft tissues reveals the metameric ancestry of lophophorates. *Current Biology*, *32*(21), 4769-4778.e2. https://doi.org/10.1016/j.cub.2022.09.011

Hagen, J. F. D., Roberts, N. S., & Johnston, R. J. (2023a). The evolutionary history and spectral tuning of vertebrate visual opsins. *Developmental Biology*, *493*, 40–66. https://doi.org/10.1016/j.ydbio.2022.10.014

Hagen, J. F. D., Roberts, N. S., & Johnston, R. J. (2023b). The evolutionary history and spectral tuning of vertebrate visual opsins. *Developmental Biology*, *493*, 40–66. https://doi.org/10.1016/j.ydbio.2022.10.014

Hanke, F. D., & Kelber, A. (2020). The Eye of the Common Octopus (Octopus vulgaris). *Frontiers in Physiology*, *10*, 1637. https://doi.org/10.3389/fphys.2019.01637

Hanlon, R. (2007). Cephalopod dynamic camouflage. *Current Biology: CB*, *17*(11), R400-404. https://doi.org/10.1016/j.cub.2007.03.034

Hanlon, R. T., & Messenger, J. B. (1988). Adaptive coloration in young cuttlefish (*Sepia officinalis* L.): The morphology and development of body patterns and their relation to behaviour. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *320*(1200), 437–487. https://doi.org/10.1098/rstb.1988.0087

Hara, T., & Hara, R. (1972). Cephalopod Retinochrome. In E. W. Abrahamson, Ch. Baumann, C. D. B. Bridges, F. Crescitelli, H. J. A. Dartnall, R. M. Eakin, G. Falk, P. Fatt, T. H. Goldsmith, R. Hara, T. Hara, S. M. Japar, P. A. Liebman, J. N. Lythgoe, R. A. Morton, W. R. A. Muntz, W. A. H. Rushton, T. I. Shaw, J. R. Wiesenfeld, … H. J. A. Dartnall (Eds.), *Photochemistry of Vision* (pp. 720–746). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-65066-6_18

Hara, T., Hara, R., & Takeuchi, J. (1967). Vision in Octopus and Squid: Rhodopsin and Retinochrome in the Octopus Retina. *Nature*, *214*(5088), 572–573. https://doi.org/10.1038/214572a0

Hara, Tomiyuki, Hara, Reiko, & Nara, Kashihara. (1972). *Cephalopod Retinochrome*.

Hardie, R. C., & Juusola, M. (2015). Phototransduction in Drosophila. *Current Opinion in Neurobiology*, *34*, 37–45. https://doi.org/10.1016/j.conb.2015.01.008

Hasan, M. S., McElroy, K. E., Audino, J. A., & Serb, J. M. (2024). Opsin expression varies across larval development and taxa in pteriomorphian bivalves. *Frontiers in Neuroscience*, *18*, 1357873. https://doi.org/10.3389/fnins.2024.1357873

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., & Bork, P. (2017). Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Molecular Biology and Evolution*, *34*(8), 2115–2122. https://doi.org/10.1093/molbev/msx148

Inoue, K., Tsuda, M., & Terazima, M. (2007). Photoreverse Reaction Dynamics of Octopus Rhodopsin. *Biophysical Journal*, *92*(10), 3643–3651. https://doi.org/10.1529/biophysj.106.101741

Jacobs, G. H. (2009). Evolution of colour vision in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1531), 2957–2967. https://doi.org/10.1098/rstb.2009.0039

Jereb, P., Roper, C. F. E., Norman, M. D., Finn, J. K., & FAO (Eds.). (2005). *Octopods and vampire squids* (Entirely rewritten, revised and updated version). Food and Agriculture Organization of the United Nations.

Johnsen, S., Frank, T. M., Haddock, S. H. D., Widder, E. A., & Messing, C. G. (2012). Light and vision in the deep-sea benthos: I. Bioluminescence at 500–1000 m depth in the Bahamian Islands. *Journal of Experimental Biology*, *215*(19), 3335–3343. https://doi.org/10.1242/jeb.072009

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kikuchi, T., Dayi, M., Hunt, V. L., Ishiwata, K., Toyoda, A., Kounosu, A., Sun, S., Maeda, Y., Kondo, Y., De Noya, B. A., Noya, O., Kojima, S., Kuramochi, T., & Maruyama, H. (2021). Genome of the fatal tapeworm Sparganum proliferum uncovers mechanisms for cryptic life cycle and aberrant larval proliferation. *Communications Biology*, *4*(1), 649. https://doi.org/10.1038/s42003-021-02160-8

Klug, C., Hoffmann, R., Tischlinger, H., Fuchs, D., Pohle, A., Rowe, A., Rouget, I., & Kruta, I. (2023). 'Arm brains' (axial nerves) of Jurassic coleoids and the evolution of coleoid neuroanatomy. *Swiss Journal of Palaeontology*, *142*(1), 22. https://doi.org/10.1186/s13358-023-00285-3

Koyanagi, M., Takano, K., Tsukamoto, H., Ohtsu, K., Tokunaga, F., & Terakita, A. (2008). Jellyfish vision starts with cAMP signaling mediated by opsin-G $_s$ cascade. *Proceedings of the National Academy of Sciences*, *105*(40), 15576–15580. https://doi.org/10.1073/pnas.0806215105

Koyanagi, M., & Terakita, A. (2014). Diversity of animal opsin-based pigments and their optogenetic potential. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1837*(5), 710–716. https://doi.org/10.1016/j.bbabio.2013.09.003

Kozmik, Z., Ruzickova, J., Jonasova, K., Matsumoto, Y., Vopalensky, P., Kozmikova, I., Strnad, H., Kawamura, S., Piatigorsky, J., Paces, V., & Vlcek, C. (2008). Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proceedings of the National Academy of Sciences*, *105*(26), 8989–8993. https://doi.org/10.1073/pnas.0800388105

Kück, P., & Longo, G. C. (2014). FASconCAT-G: Extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Frontiers in Zoology*, *11*(1), 81. https://doi.org/10.1186/s12983-014-0081-x

Lan, Y., Sun, J., Chen, C., Sun, Y., Zhou, Y., Yang, Y., Zhang, W., Li, R., Zhou, K., Wong, W. C., Kwan, Y. H., Cheng, A., Bougouffa, S., Van Dover, C. L., Qiu, J.-W., & Qian, P.-Y. (2021). Hologenome analysis reveals dual symbiosis in the deep-sea hydrothermal vent snail Gigantopelta aegis. *Nature Communications*, *12*(1), 1165. https://doi.org/10.1038/s41467-021-21450-7

Lemoine, F., Domelevo Entfellner, J.-B., Wilkinson, E., Correia, D., Dávila Felipe, M., De Oliveira, T., & Gascuel, O. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature*, *556*(7702), 452–456. https://doi.org/10.1038/s41586-018-0043-0

Liénard, M. A., Valencia-Montoya, W. A., & Pierce, N. E. (2022). Molecular advances to study the function, evolution and spectral tuning of arthropod visual opsins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1862), 20210279. https://doi.org/10.1098/rstb.2021.0279

López-Córdova, D. A., Avaria-Llautureo, J., Ulloa, P. M., Braid, H. E., Revell, L. J., Fuchs, D., & Ibáñez, C. M. (2022). Mesozoic origin of coleoid cephalopods

and their abrupt shifts of diversification patterns. *Molecular Phylogenetics and Evolution*, *166*, 107331. https://doi.org/10.1016/j.ympev.2021.107331

Maddison, W., & Maddison, D. (2017). *Mesquite: A modular system for evolutionary analysis, v. 3.61*. http://www.mesquiteproject.org

Marshall, J., & Cronin, T. W. (2011). Polarisation vision. *Current Biology: CB*, *21*(3), R101-105. https://doi.org/10.1016/j.cub.2010.12.012

Martín-Durán, J. M., Vellutini, B. C., Marlétaz, F., Cetrangolo, V., Cvetesic, N., Thiel, D., Henriet, S., Grau-Bové, X., Carrillo-Baltodano, A. M., Gu, W., Kerbl, A., Marquez, Y., Bekkouche, N., Chourrout, D., Gómez-Skarmeta, J. L., Irimia, M., Lenhard, B., Worsaae, K., & Hejnol, A. (2021). Conservative route to genome compaction in a miniature annelid. *Nature Ecology & Evolution*, *5*(2), 231–242. https://doi.org/10.1038/s41559-020-01327-6

Mäthger, L. M., Barbosa, A., Miner, S., & Hanlon, R. T. (2006). Color blindness and contrast perception in cuttlefish (Sepia officinalis) determined by a visual sensorimotor assay. *Vision Research*, *46*(11), 1746–1753. https://doi.org/10.1016/j.visres.2005.09.035

Matsuo, R., Koyanagi, M., Sugihara, T., Shirata, T., Nagata, T., Inoue, K., Matsuo, Y., & Terakita, A. (2023). Functional characterization of four opsins and two G alpha subtypes co-expressed in the molluscan rhabdomeric photoreceptor. *BMC Biology*, *21*(1), 291. https://doi.org/10.1186/s12915-023-01789-7

Maximov, V. V. (2000). Environmental factors which may have led to the appearance of colour vision. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1401), 1239–1242. https://doi.org/10.1098/rstb.2000.0675

McCulloch, K. J., Babonis, L. S., Liu, A., Daly, C. M., Martindale, M. Q., & Koenig, K. M. (2023). Nematostella vectensis exemplifies the exceptional expansion

and diversity of opsins in the eyeless Hexacorallia. *EvoDevo*, *14*(1), 14. https://doi.org/10.1186/s13227-023-00218-8

Mickael, M. E., Rajput, A., Steyn, J., Wiemerslage, L., & Bürglin, T. (2016). An optimised phylogenetic method sheds more light on the main branching events of rhodopsin-like superfamily. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *20*, 85–94. https://doi.org/10.1016/j.cbd.2016.08.005

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Morel, B., Kozlov, A. M., Stamatakis, A., & Szöllősi, G. J. (2020). GeneRax: A Tool for Species-Tree-Aware Maximum Likelihood-Based Gene Family Tree Inference under Gene Duplication, Transfer, and Loss. *Molecular Biology and Evolution*, *37*(9), 2763–2774. https://doi.org/10.1093/molbev/msaa141

Mueller, O., Lightfoot, S., & Schroeder, A. (2004). RNA integrity number (RIN)–standardization of RNA quality control. *Agilent Application Note, Publication*, *1*, 1–8.

Musilova, Z., Cortesi, F., Matschiner, M., Davies, W. I. L., Patel, J. S., Stieb, S. M., de Busserolles, F., Malmstrøm, M., Tørresen, O. K., Brown, C. J., Mountford, J. K., Hanel, R., Stenkamp, D. L., Jakobsen, K. S., Carleton, K. L., Jentoft, S., Marshall, J., & Salzburger, W. (2019). Vision using multiple distinct rod opsins in deep-sea fishes. *Science*, *364*, 588–592. https://doi.org/10.1126/science.aav4632

Musilova, Z., Salzburger, W., & Cortesi, F. (2021). The Visual Opsin Gene Repertoires of Teleost Fishes: Evolution, Ecology, and Function. *Annual*

Review of Cell and Developmental Biology, *37*(1), 441–468. https://doi.org/10.1146/annurev-cellbio-120219-024915

Nagata, T., Koyanagi, M., Lucas, R., & Terakita, A. (2018). An all-trans-retinal-binding opsin peropsin as a potential dark-active and light-inactivated G protein-coupled receptor. *Scientific Reports*, *8*(1), 3535. https://doi.org/10.1038/s41598-018-21946-1

Ogura, A., Ikeo, K., & Gojobori, T. (2004). Comparative Analysis of Gene Expression for Convergent Evolution of Camera Eye Between Octopus and Human. *Genome Research*, *14*(8), 1555–1561. https://doi.org/10.1101/gr.2268104

Park, T.-Y. S., Nielsen, M. L., Parry, L. A., Sørensen, M. V., Lee, M., Kihm, J.-H., Ahn, I., Park, C., De Vivo, G., Smith, M. P., Harper, D. A. T., Nielsen, A. T., & Vinther, J. (2024). A giant stem-group chaetognath. *Science Advances*, *10*(1), eadi6678. https://doi.org/10.1126/sciadv.adi6678

Piovani, L., Leite, D. J., Yañez Guerra, L. A., Simpson, F., Musser, J. M., Salvador-Martínez, I., Marlétaz, F., Jékely, G., & Telford, M. J. (2023). Single-cell atlases of two lophotrochozoan larvae highlight their complex evolutionary histories. *Science Advances*, *9*(31), eadg6034. https://doi.org/10.1126/sciadv.adg6034

Porter, M. L., Roberts, N. W., & Partridge, J. C. (2016). Evolution under pressure and the adaptation of visual pigment compressibility in deep-sea environments. *Molecular Phylogenetics and Evolution*, *105*, 160–165. https://doi.org/10.1016/j.ympev.2016.08.007

Pungor, J. R., Allen, V. A., Songco-Casey, J. O., & Niell, C. M. (2023). Functional organization of visual responses in the octopus optic lobe. *Current Biology*, *33*(13), 2784-2793.e3. https://doi.org/10.1016/j.cub.2023.05.069

Ramirez, M. D., & Oakley, T. H. (2015). Eye-independent, light-activated chromatophore expansion (LACE) and expression of phototransduction

genes in the skin of *Octopus bimaculoides*. *Journal of Experimental Biology*, *218*(10), 1513–1520. https://doi.org/10.1242/jeb.110908

Ramirez, M., Pairett, A., Pankey, M., Serb, J., Speiser, D., Swafford, A., & Oakley, T. (2016). The last common ancestor of most bilaterian animals possessed at least 9 opsins. *Genome Biology and Evolution*, evw248. https://doi.org/10.1093/gbe/evw248

Rawlinson, K. A., Lapraz, F., Ballister, E. R., Terasaki, M., Rodgers, J., McDowell, R. J., Girstmair, J., Criswell, K. E., Boldogkoi, M., Simpson, F., Goulding, D., Cormie, C., Hall, B., Lucas, R. J., & Telford, M. J. (2019). Extraocular, rod-like photoreceptors in a flatworm express xenopsin photopigment. *eLife*, *8*, e45465. https://doi.org/10.7554/eLife.45465

Revell, L. J. (2024). phytools 2.0: An updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, *12*, e16505. https://doi.org/10.7717/peerj.16505

Ricci, V., Ronco, F., Boileau, N., & Salzburger, W. (2023). Visual opsin gene expression evolution in the adaptive radiation of cichlid fishes of Lake Tanganyika. *Science Advances*, *9*(36), eadg6568. https://doi.org/10.1126/sciadv.adg6568

Ricci, V., Ronco, F., Musilova, Z., & Salzburger, W. (2022). Molecular evolution and depth-related adaptations of rhodopsin in the adaptive radiation of cichlid fishes in Lake Tanganyika. *Molecular Ecology*, *31*(10), 2882–2897. https://doi.org/10.1111/mec.16429

Robison, B. H., & Reisenbichler, K. R. (2008). Macropinna microstoma and the Paradox of Its Tubular Eyes. *Copeia*, *2008*(4), 780–784. https://doi.org/10.1643/CG-07-082

Scheerer, P., Park, J. H., Hildebrand, P. W., Kim, Y. J., Krauß, N., Choe, H.-W., Hofmann, K. P., & Ernst, O. P. (2008). Crystal structure of opsin in its G-

protein-interacting conformation. *Nature*, *455*(7212), 497–502. https://doi.org/10.1038/nature07330

Sekharan, S., Katayama, K., Kandori, H., & Morokuma, K. (2012). Color vision: 'OH-site' rule for seeing red and green. *Journal of the American Chemical Society*, *134*(25), 10706–10712. https://doi.org/10.1021/ja304820p

Seppey, M., Manni, M., & Zdobnov, E. M. (2019). BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods in Molecular Biology (Clifton, N.J.)*, *1962*, 227–245. https://doi.org/10.1007/978-1-4939-9173-0_14

Serb, J. M., & Eernisse, D. J. (2008). Charting Evolution's Trajectory: Using Molluscan Eye Diversity to Understand Parallel and Convergent Evolution. *Evolution: Education and Outreach*, *1*(4), 439–447. https://doi.org/10.1007/s12052-008-0084-1

Shea, E. K., Ziegler, A., Faber, C., & Shank, T. M. (2018). Dumbo octopod hatchling provides insight into early cirrate life cycle. *Current Biology*, *28*(4), R144–R145. https://doi.org/10.1016/j.cub.2018.01.032

Sohpal, V. K., Dey, A., & Singh, A. (2010). MEGA biocentric software for sequence and phylogenetic analysis: A review. *International Journal of Bioinformatics Research and Applications*, *6*(3), 230. https://doi.org/10.1504/IJBRA.2010.034072

Somero, G. N. (1992). Adaptations to High Hydrostatic Pressure. *Annual Review of Physiology*, *54*(1), 557–577. https://doi.org/10.1146/annurev.ph.54.030192.003013

Songco-Casey, J. O., Coffing, G. C., Piscopo, D. M., Pungor, J. R., Kern, A. D., Miller, A. C., & Niell, C. M. (2022). Cell types and molecular architecture of the Octopus bimaculoides visual system. *Current Biology*, *32*(23), 5031-5044.e4. https://doi.org/10.1016/j.cub.2022.10.015

Soto, C., Kelber, A., & Hanke, F. D. (2020). The Pupillary Response of the Common Octopus (Octopus vulgaris). *Frontiers in Physiology*, *11*, 1112. https://doi.org/10.3389/fphys.2020.01112

Stanke, M., Steinkamp, R., Waack, S., & Morgenstern, B. (2004). AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Research*, *32*(Web Server issue), W309-312. https://doi.org/10.1093/nar/gkh379

Stubbs, A. L., & Stubbs, C. W. (2016). Spectral discrimination in color blind animals via chromatic aberration and pupil shape. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(29), 8206–8211. https://doi.org/10.1073/pnas.1524578113

Styfhals, R., Zolotarov, G., Hulselmans, G., Spanier, K. I., Poovathingal, S., Elagoz, A. M., De Winter, S., Deryckere, A., Rajewsky, N., Ponte, G., Fiorito, G., Aerts, S., & Seuntjens, E. (2022). Cell type diversity in a developing octopus brain. *Nature Communications*, *13*(1), 7392. https://doi.org/10.1038/s41467-022-35198-1

Tanner, A. R., Fuchs, D., Winkelmann, I. E., Gilbert, M. T. P., Pankey, M. S., Ribeiro, Â. M., Kocot, K. M., Halanych, K. M., Oakley, T. H., Da Fonseca, R. R., Pisani, D., & Vinther, J. (2017). Molecular clocks indicate turnover and diversification of modern coleoid cephalopods during the Mesozoic Marine Revolution. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1850), 20162818. https://doi.org/10.1098/rspb.2016.2818

Tejero, O., Pamula, F., Koyanagi, M., Nagata, T., Afanasyev, P., Das, I., Deupi, X., Sheves, M., Terakita, A., Schertler, G. F. X., Rodrigues, M. J., & Tsai, C.-J. (2024). *Active state structures of a bistable visual opsin bound to G proteins*. https://doi.org/10.1101/2024.04.09.588704

Temple, S. E., How, M. J., Powell, S. B., Gruev, V., Marshall, N. J., & Roberts, N. W. (2021). Thresholds of polarization vision in octopuses. *Journal of Experimental Biology*, *224*(7), jeb240812. https://doi.org/10.1242/jeb.240812

Terakita, A. (2005). The opsins. *Genome Biology*, *6*(3), 213. https://doi.org/10.1186/gb-2005-6-3-213

Terakita, A., Kawano-Yamashita, E., & Koyanagi, M. (2012). Evolution and diversity of opsins. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling*, *1*(1), 104–111. https://doi.org/10.1002/wmts.6

Thomas, K. N., Robison, B. H., & Johnsen, S. (2017). Two eyes for two purposes: *In situ* evidence for asymmetric vision in the cockeyed squids *Histioteuthis heteropsis* and *Stigmatoteuthis dofleini*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1717), 20160069. https://doi.org/10.1098/rstb.2016.0069

Treves, S., Thurnheer, R., Mosca, B., Vukcevic, M., Bergamelli, L., Voltan, R., Oberhauser, V., Ronjat, M., Csernoch, L., Szentesi, P., & Zorzato, F. (2012). SRP-35, a newly identified protein of the skeletal muscle sarcoplasmic reticulum, is a retinol dehydrogenase. *Biochemical Journal*, *441*(2), 731–741. https://doi.org/10.1042/BJ20111457

Uribe, J. E., & Zardoya, R. (2017). Revisiting the phylogeny of Cephalopoda using complete mitochondrial genomes. *Journal of Molluscan Studies*, *83*(2), 133–144. https://doi.org/10.1093/mollus/eyw052

Vöcking, O., Kourtesis, I., Tumu, S. C., & Hausen, H. (2017). Co-expression of xenopsin and rhabdomeric opsin in photoreceptors bearing microvilli and cilia. *eLife*, *6*, e23435. https://doi.org/10.7554/eLife.23435

Vöcking, O., Leclère, L., & Hausen, H. (2021). The rhodopsin-retinochrome system for retinal re-isomerization predates the origin of cephalopod eyes. *BMC*

Ecology and Evolution, *21*(1), 215. https://doi.org/10.1186/s12862-021-01939-x

Vöcking, O., Macias-Muñoz, A., Jaeger, S. J., & Oakley, T. H. (2022). Deep Diversity: Extensive Variation in the Components of Complex Visual Systems across Animals. *Cells*, *11*(24), 3966. https://doi.org/10.3390/cells11243966

Warrant, E. J. (1999). Seeing better at night: Life style, eye design and the optimum strategy of spatial and temporal summation. *Vision Research*, *39*(9), 1611–1630. https://doi.org/10.1016/S0042-6989(98)00262-4

Warrant, E. J., & Locket, N. A. (2004). Vision in the deep sea. *Biological Reviews*, *79*(3), 671–712. https://doi.org/10.1017/S1464793103006420

Wollesen, T., Rodriguez Monje, S. V., Oel, A. P., & Arendt, D. (2023). Characterization of eyes, photoreceptors, and opsins in developmental stages of the arrow worm *Spadella cephaloptera* (Chaetognatha). *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *340*(5), 342–353. https://doi.org/10.1002/jez.b.23193

Yau, K.-W., & Hardie, R. C. (2009). Phototransduction Motifs and Variations. *Cell*, *139*(2), 246–264. https://doi.org/10.1016/j.cell.2009.09.029

Yoshida, M., Hirota, K., Imoto, J., Okuno, M., Tanaka, H., Kajitani, R., Toyoda, A., Itoh, T., Ikeo, K., Sasaki, T., & Setiamarga, D. H. E. (2022). Gene Recruitments and Dismissals in the Argonaut Genome Provide Insights into Pelagic Lifestyle Adaptation and Shell-like Eggcase Reacquisition. *Genome Biology and Evolution*, *14*(11), evac140. https://doi.org/10.1093/gbe/evac140

Yoshida, M., & Ogura, A. (2011). Genetic mechanisms involved in the evolution of the cephalopod camera eye revealed by transcriptomic and developmental studies. *BMC Evolutionary Biology*, *11*(1), 180. https://doi.org/10.1186/1471-2148-11-180

Yoshida, M., Ogura, A., Ikeo, K., Shigeno, S., Moritaki, T., Winters, G. C., Kohn, A. B., & Moroz, L. L. (2015). Molecular Evidence for Convergence and Parallelism in Evolution of Complex Brains of Cephalopod Molluscs: Insights from Visual Systems. *Integrative and Comparative Biology*, *55*(6), 1070–1083. https://doi.org/10.1093/icb/icv049

Young, J. Z. (1962). The optic lobes of *Octopus vulgaris*. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *245*(718), 19–58. https://doi.org/10.1098/rstb.1962.0005

Young, J. Z. (1991). Light has many meanings for cephalopods. *Visual Neuroscience*, *7*(1–2), 1–12. https://doi.org/10.1017/S0952523800010907

Zhang, H., Futami, K., Horie, N., Okamura, A., Utoh, T., Mikawa, N., Yamada, Y., Tanaka, S., & Okamoto, N. (2000). Molecular cloning of fresh water and deep-sea rod opsin genes from Japanese eel *Anguilla japonica* and expressional analyses during sexual maturation [1]. *FEBS Letters*, *469*(1), 39–43. https://doi.org/10.1016/S0014-5793(00)01233-3

Zhang, T., Cao, L.-H., Kumar, S., Enemchukwu, N. O., Zhang, N., Lambert, A., Zhao, X., Jones, A., Wang, S., Dennis, E. M., Fnu, A., Ham, S., Rainier, J., Yau, K.-W., & Fu, Y. (2016). Dimerization of visual pigments in vivo. *Proceedings of the National Academy of Sciences*, *113*(32), 9093–9098. https://doi.org/10.1073/pnas.1609018113

Zhang, Y., Mao, F., Mu, H., Huang, M., Bao, Y., Wang, L., Wong, N.-K., Xiao, S., Dai, H., Xiang, Z., Ma, M., Xiong, Y., Zhang, Z., Zhang, L., Song, X., Wang, F., Mu, X., Li, J., Ma, H., … Yu, Z. (2021). The genome of Nautilus pompilius illuminates eye evolution and biomineralization. *Nature Ecology & Evolution*, *5*(7), 927–938. https://doi.org/10.1038/s41559-021-01448-6

APPENDIX

I Species under investigation

This section provides a summary of the main characteristics of the species analysed during this investigation. All data and figures, excluding those for *Octopus vulgaris*, are from (Jereb et al., 2005).

***Argonauta argo* (Linnaeus, 1758)**

Also known as the paper nautilus. It is the most studied species among the four in the family Argonautidae, a group of pelagic incirrate octopods. Within the octopods, *Argonauta argo* is quite peculiar.

The species is characterized by large females that possess an egg case resembling the shells of nautiloids, while the males are small and unsheltered. The male's reproductive arm is a modified hectocotylus (the third left arm), which develops a sperm sac beneath the left eye. This arm typically detaches from the male and transfers the sperm sac to the female. Eggs are kept inside the female's shell, and small male individuals have occasionally been found there as well.

Argonauta argo is an epipelagic species with a global distribution between approximately 40°N and 40°S, and it can reach depths of up to 300 meters.

Appendix Figure 1 – *Argonauta argo*. Photo of Julian Finn @Fao.

***Eledone cirrhosa* (Lamarck, 1798)**

Also known as the horned octopus or white octopus, is a species distributed along the coasts of Iceland and the Mediterranean Sea. It can be found in shallow environments down to depths of 500 meters, occupying a wide range of habitats, from sandy areas to rocky reefs.

Like all members of the genus *Eledone*, it is characterized by a single row of suckers on each arm. However, unlike *Eledone moschata*, E. cirrhosa has a small cirrus near the eye and a reddish coloration. It preys on a variety of animals, including small crustaceans, vertebrates, and other molluscs.

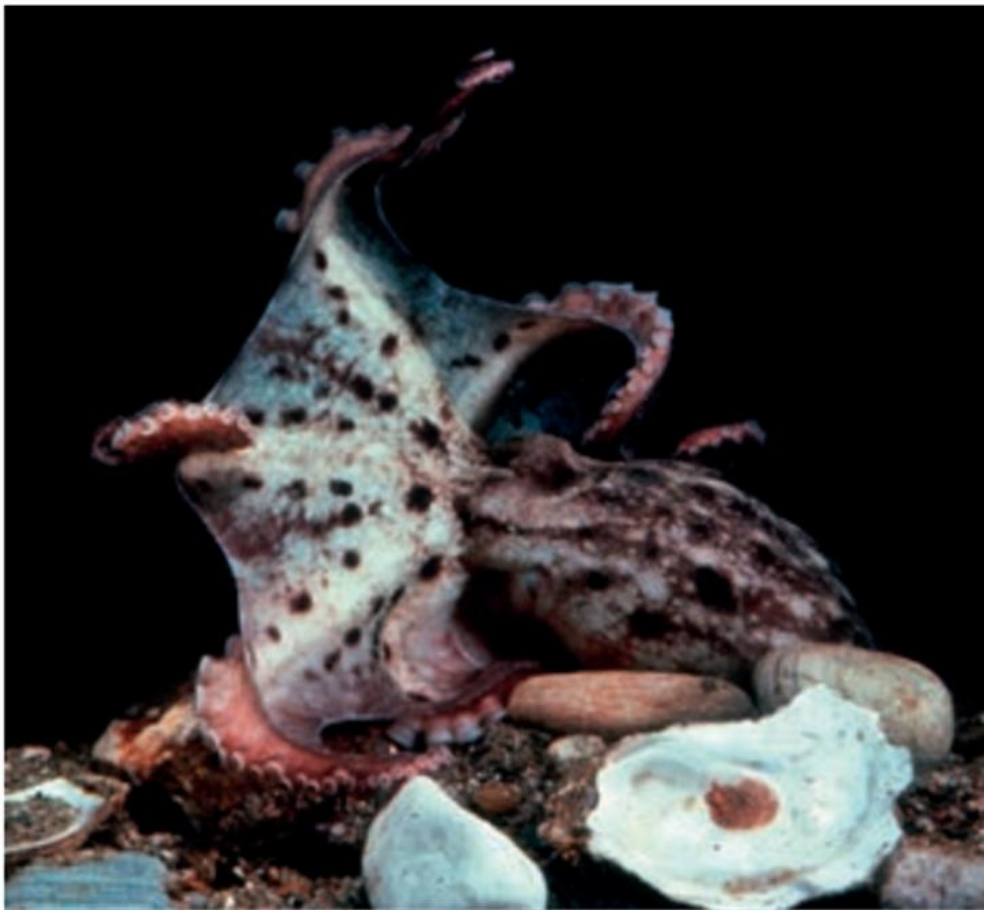
The reproductive season occurs between May and September, peaking in July. The hatchlings are likely benthic, based on the size of the eggs. *Eledone cirrhosa* typically has a lifespan of up to three years. This species is highly commercialized in the Mediterranean Sea.

Appendix Figure 2 – *Eledone cirrhosa*. Photo of Jean Lecomte @Fao.

***Eledone moschata* (Lamarck, 1798)**

Commonly known as the musky octopus, is the second most abundant species of the genus *Eledone* found in the Mediterranean Sea. It can live at depths of up to 300 meters but is also found in shallower environments. *E. moschata* primarily feeds on crustaceans.

Little is known about its spawning season, and the hatchlings are believed to be benthic. Although this species has commercial value, it is generally considered less valuable than *Eledone cirrhosa*.



Appendix Figure 3 – *Eledone moschata*. Photo of Jean Lecomte @Fao.

***Callistoctopus macropus* (Risso, 1826)**

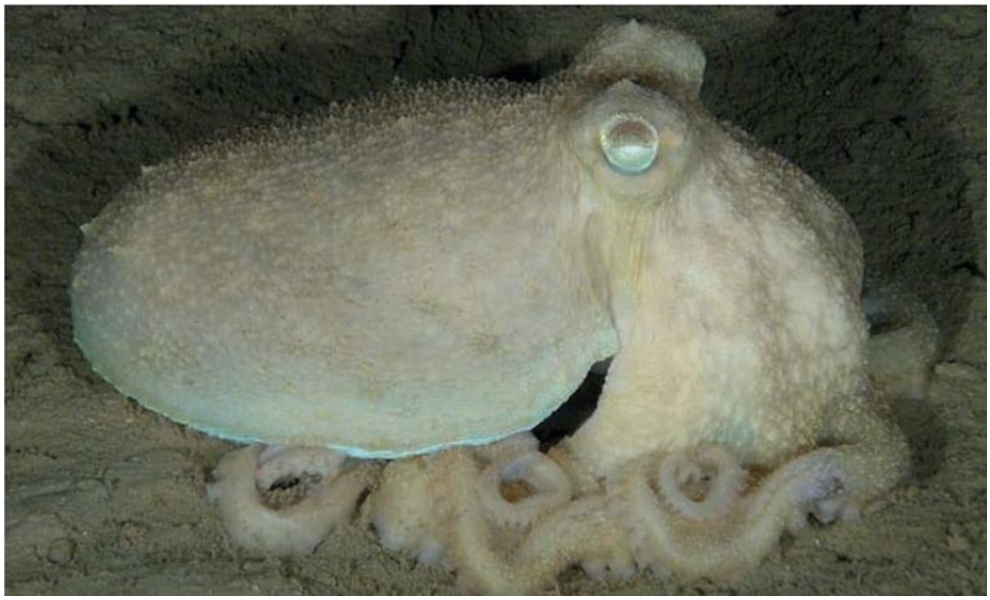
Commonly known as the white-spotted octopus, is a species widely distributed in the eastern North Atlantic Ocean and along Mediterranean coasts. It is characterized by its red mantle with white spots and extremely elongated arms, which can be up to seven times the length of its mantle, with the dorsal arms being more robust.

This species is predominantly nocturnal, inhabiting littoral waters up to a depth of 80 meters. It feeds on small fish, crustaceans, and other cephalopods. Although it has some commercial value, it is relatively rare compared to *Octopus vulgaris*.

Appendix Figure 4 – *Callistoctopus macropus*. Photo of Mark Norman @Fao.

***Scaeurgus unicirrhus* (Troschel, 1857)**

Also known as the unicorn octopus, this species is distributed in the Mediterranean Sea and along the west coast of the Atlantic Ocean. It is recognizable by its unevenly distributed rows of suckers and a mantle that extends up to 1/5 of the total arm length. It has moderate commercial value, often being sold as *Octopus vulgaris*, though it is generally considered less valuable in the market.



Appendix Figure 5 – *Scaeurgus unicirrhus*. Photo of Oceana Ranger Expedition @Fao.

***Pteroctopus tetracirrhus* (Delle Chiaje, 1830)**

Also known as the four-horned octopus, this species is distributed in the eastern Atlantic Ocean and the Mediterranean Sea. It is the only species in this study known to reach depths of up to 700 meters, with occasional reports of depths up to 750 meters. The species is characterized by a pair of sucker rows unevenly distributed on each arm and the presence of cirri around the eyes.

In the water, it displays small horn-like structures covering the mantle. When fished, it is easily recognized by its gelatinous, soft skin and resembles the collapsed "blob" appearance typical of deep-sea animals brought to the surface and subjected to lower pressure. The species primarily feeds on small crustaceans.

Appendix Figure 6 - *Pteroctopus tetracirrhus*. Photo of Oceana Europe @Fao.

***Octopus vulgaris* (Cuvier, 1797)**

Also known as the common octopus, *Octopus vulgaris* is the most studied cephalopod species, renowned for its well-developed intelligence. It has a global distribution and is typically found at depths of up to 100 meters, though it has occasionally been reported at depths of 250 meters.

This species is primarily nocturnal, but it is not uncommon to spot them during the day, particularly in rocky habitats. *O. vulgaris* can be recognized by its large size (some specimens can reach up to 10 kg), the distinctive texture of its skin with regular patches, and the evenly distributed pairs of suckers along each arm.

It feeds on a wide range of prey, including crustaceans, small fish, and primarily shelled molluscs. *O. vulgaris* is commercially valued, especially in the Mediterranean, where it is in high demand due to its abundance and the ease with which it can be fished.

Appendix Figure 7 – *Octopus vulgaris*.

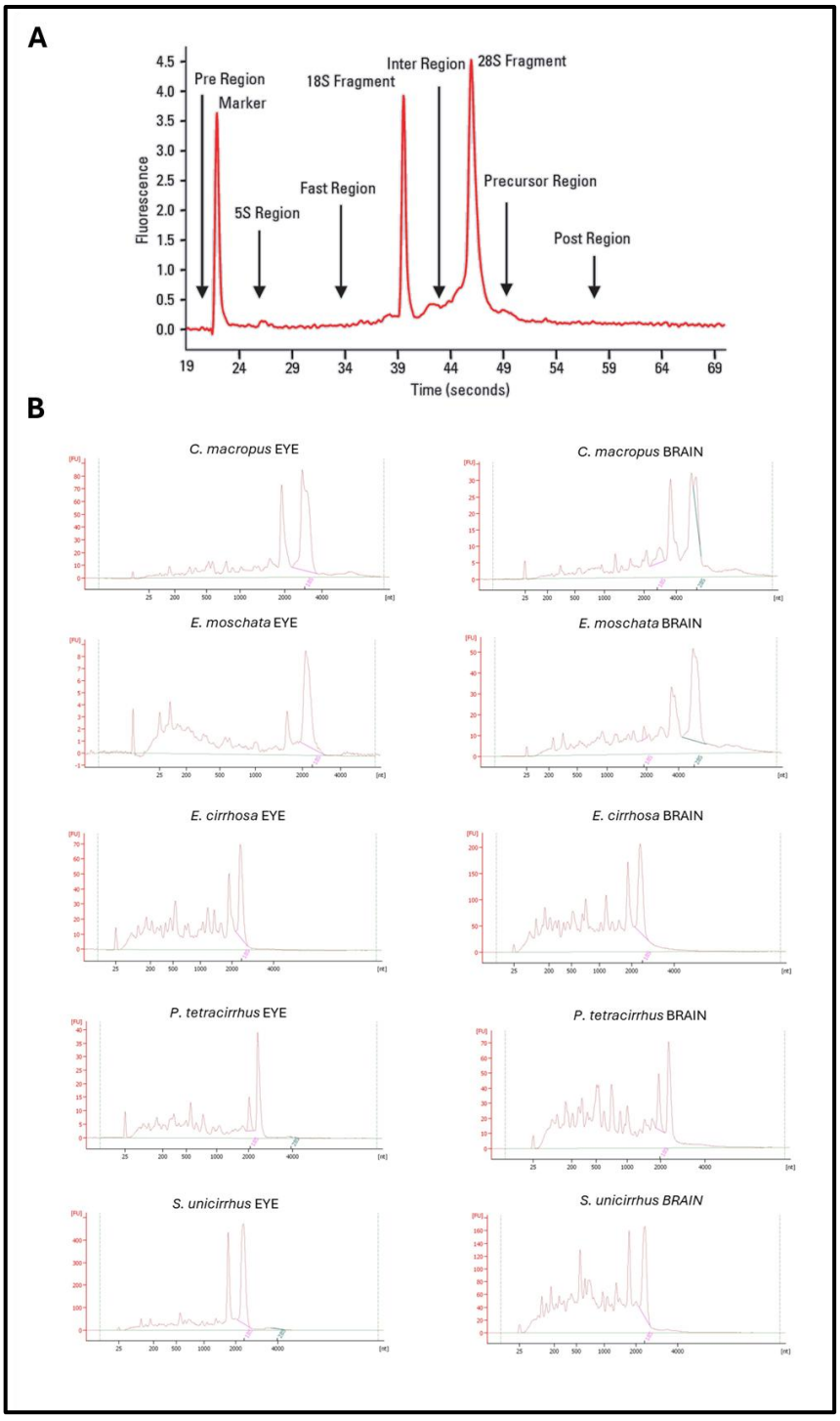
II RT-qPCR primers

Target gene	Primer couple (Forward and Reverse 5'---3')			Primer efficiency
AAR-opsin1	5	CCTCGCTATGTCTGACTTGTCTTT	3'	107,6197502
	5	TGACATGAATCCGAAGATACC	3'	
AAR-opsin2	5	TTACAAGAGAAACGACCGGAT	3'	106,1689403
	5	GCTCTTCTTTTAGTTTTGTCC	3'	
AAXenopsin1	5	CCTGCCTGCTGTACGCCCTA	3'	104,3590295
	5	TGTGTGGTCCCTTTGCCCCT	3'	
AAXenopsin2	5	CTCGTGAAGTTGGGCTGGCG	3'	106,5173791
	5	CCGACACCACAGCGAATGGG	3'	
AARetinochrome	5	TCTCACATCTGGATGCTGTTTGCA	3'	125,6460973
	5	ACATGGAAGCAATCAAGGTGT	3'	
AAPeropsin	5	CTTCTTGGGAATGTTATTTGG	3'	96,83193971
	5	CCAGATTCCAGCAATGATCC	3'	
AAPseudopsin	5	GCAGCGCTGTCTGTGAAGGG	3'	114,9421025
	5	ACGGCTGCGTTGGTTAACGG	3'	
OVR-opsin1	5	GTAGACCTATGGCGGCATCC	3'	106,7419265
	5	ATGCTCCCCAGTTGAAGACG	3'	
OVR-opsin2	5	TTACAAGAGCATCGACCGGCT	3'	126,578004
	5	ACTCGTCTGGCAGCTTTCTCG	3'	
OVXenopsin	5	CCTCGGACATCGCATATATTG	3'	102,7693272
	5	GCATAGGGCATCCATGATGC	3'	
OVRetinochrome	5	ACAGCGAGAAGGAAAAGGACC	3'	101,5538681
	5	GCAAATCCAGACCACGAGAC	3'	

OVPeropsin	5	CTTCACAGGGATGTTGTTTGG	3’	114,8591002
	5	ATCCATACACAGGCAGCAATTAAG	3’	
OVPseudopsin	5	GGACTATAACGGCGTTGAGTT	3’	112,1669522
	5	TGTCAAAGATGAATCGGCTATAACC	3’	
ElongationFactor1alpha	5	AGGCCGAGAGAGAACGTGGT	3’	111,4918193
	5	GCAGCAACAACCAGCACAGC	3’	

Appendix Table 1 – List for real time quantitative PCR (RT-qPCR) primers. The table shows the selected RT-qPCR primers designed, when possible, in exon boundaries and with an annealing temperature of 62°C. In the table are shown the target genes, their respective primer pair and the primer pair efficiency. AA indicates primers designed on *A. argo* genome and OV indicates primers designed on *O. vulgaris* genome.

III Bioanalyzer Quality check of total RNA for transcriptome sequencing



Appendix Figure 8 – Agilent bioanalyzer results. A) Example of a hight quality eukaryote total RNA on AgilentTM 2000 Bioanalyzer™ showing the different regions. B) on AgilentTM 2000 Bioanalyzer™ results of total RNA extracted from the samples. EYE refers to RNA extracted from the retina, BRAIN to RNA extracted from the Optic lobe. Notice the lack of the 28S peak and the presence of two peaks nearby the 18S fragment. Picture in A) From Mueller et al. (2004).

IV Primer list for gene cloning

In this section there is the list of all the primers that led to successful amplifications.

R-opsin1 – 1354 bp

UP	ATGGTCGAATCAAGCACGTTAGT	TM 65.6 °C
DW	CCTGATTGTCGACTCCTTGGGG	TM 62.2 °C

R-opsin1 – 1146 bp (more efficient but shorter)

UP	CGTAGACATCCATCCTCATTG	TM 61.9 °C
DW	GTGGTGGGTAGCCTTGAG	TM 60.8 °C

R-opsin2 - 1008 bp (works only in *O. vulgaris*, full lenght)

UP	ATGTCTTCTATCGATATGTGTTTTG	TM 59.9°C
DW	TTAGACTGTCTCCGATTCCTT	TM 59.4°C

Xenopsin – 1347 bp (works only in *O. vulgaris*, full length)

UP	ATGGCATTTTCGACAGCTCTCAGC	TM 71.3°C
DW	TCAAACACACGTCATATCTTCATCGC	TM 69.6°C

Xenopsin1 – 670 bp (for A. *argo*)

UP	CCCTCACTGTCATGCCTACC	TM 64.6°C
DW	ATGACGGGCATAACAAGGCA	TM 67.4°C

Xenopsin2 – 762 bp (for *A. argo*)

UP	ATGACGTCACAGTCACCACC	TM 63.8°C
DW	GATGACGAGCAGCGGTAGAA	TM 65.4°C

Retinochrome – 915 bp (full length)

UP	ATGTTTGGTGTTCATCATAGC	TM 58.9°C
DW	TTAAGGCTTCTTGGATTCTGC	TM 62.2°C

Retinochrome - 840 bp (designed for *A.argo*)

UP	GCTCTTTACTTGATAATTGGTG	TM 58.9°C
DW	TTAAGGCTTCTTGGATTCTGC	TM 62.2°C

Peropsin - 978 bp (works only in *O. vulgaris*, full length)

UP	ATGGAAAATGTCAGTGAAATATCG	TM 62.6°C
DW	TTAGCACATCAAATTTTTCATTC	TM 59.6°C

Pseudopsin -1359 bp

UP	CAGCAGCTACAGCAGCGGTTG	TM 70.4°C
DW	CCGTAATAGTCCGTGACATGCC	TM 67.3°C

Pseudopsin - 945 bp (shorter but more performing)

UP	CAGCTTCCGTAAAATGGC	TM 60.8°C
DW	TGTTGGTAAGGTCCCGTAA	TM 60.4°C

V RT-qPCR results

V.I Raw Data

The tables of these sections show the output data of RT-qPCR analysis

Octopus vulgaris		EXPERIMENT 1			
	Sample	**Target**	**Cq Mean**	**ΔCq**	**$2^{-\Delta Cq}$**
				2	
	EYE	RH1	12,20362384	-3,37955	10,40749158
	EYE	RH2	26,80728356	11,22411	0,000418029
	EYE	XEN	25,54556797	9,962394	0,001002353
	EYE	RET	15,57431211	-0,00886	1,006161724
	EYE	PER	29,56159164	13,97842	6,19551E-05
	EYE	PSE	25,81550777	10,23233	0,000831306
	EYE	2	15,58317432		
	BRAIN	RH1	19,05238656	4,501796	0,044139183
	BRAIN	RH2	25,03796913	10,48738	0,000696601
	BRAIN	XEN	30,52722815	15,97664	1,55079E-05
	BRAIN	RET	22,59674961	8,046159	0,003783247
	BRAIN	PER	24,74083447	10,19024	0,000855915
	BRAIN	PSE	23,68516747	9,134577	0,001779173
	BRAIN	2	14,55059029		
	SUCKERS	RH1	24,70280156	9,621707	0,001269339
	SUCKERS	RH2	27,62210878	12,54101	0,000167795
	SUCKERS	XEN	NG	-15,0811	34662,64856
	SUCKERS	RET	26,63005044	11,54896	0,000333747
	SUCKERS	PER	29,7947317	14,71364	3,72181E-05
	SUCKERS	PSE	25,16871066	10,08762	0,00091902
	SUCKERS	2	15,08109427		
	MUSCLE	RH1	27,73263555	13,42757	9,07607E-05
	MUSCLE	RH2	26,90705343	12,60199	0,000160851
	MUSCLE	XEN	NG		
	MUSCLE	RET	21,64586955	7,340807	0,006168748
	MUSCLE	PER	NG		
	MUSCLE	PSE	23,82498842	9,519926	0,001362125
	MUSCLE	2	14,30506289		

Appendix Table 2 – Results of the first quantitative PCR (RT-qPCR) experiment conducted on *O. vulgaris*. EYE indicates the retina, BRAIN the optic lobe, SUCKERS the proximal suckers of the right dorsal tentacle, MUSCLE the retractor muscle of the mantle, RH r-opsin, XEN xenopsin, RET retinochrome, PER peropsin, PSE pseudopsin and 2 the control (Elongation Factor 1-Alpha-like).

Argonauta argo **EXPERIMENT 1**

Sample	Target	Cq Mean	ΔCq	$2^{-\Delta Cq}$
			2	
EYE	RH1	17,4845726	-2,07082	4,201254415
EYE	RH2	31,5561592	12,00077	0,000244011
EYE	XEN1	30,3147493	10,75936	0,000576914
EYE	XEN2	29,9515961	10,3962	0,000742046
EYE	RET	22,094127	2,538734	0,172093648
EYE	PER	31,9607004	12,40531	0,000184345
EYE	PSE	29,6793945	10,124	0,000896132
EYE	2	19,5553927		
BRAIN	RH1	22,906	6,41	0,01175974
BRAIN	RH2	NG		
BRAIN	XEN1	27,7360405	11,24004	0,000413438
BRAIN	XEN2	30,215	13,719	7,416E-05
BRAIN	RET	23,5328208	7,036821	0,007615631
BRAIN	PER	30,8985948	14,40259	4,61729E-05
BRAIN	PSE	22,417	5,921	0,016504463
BRAIN	2	16,496		
SUCKERS	RH1	29,5012172	12,83057	0,000137282
SUCKERS	RH2	27,5498864	10,87924	0,000530913
SUCKERS	XEN1	34,481758	17,81111	4,34832E-06
SUCKERS	XEN2	NG		NG
SUCKERS	RET	26,7074015	10,03675	0,000951999
SUCKERS	PER	31,8629834	15,19234	2,67086E-05
SUCKERS	PSE	27,5590275	10,88838	0,000527559
SUCJERS	2	16,6706484		
MUSCLE	RH1	27,2510059	10,77098	0,000572285
MUSCLE	RH2	28,499513	12,01949	0,000240865
MUSCLE	XEN1	30,7260781	14,24605	5,14649E-05
MUSCLE	XEN2	32,6252723	16,14525	1,37974E-05
MUSCLE	RET	26,4510522	9,971026	0,000996374
MUSCLE	PER	NG		NG
MUSCLE	PSE	26,4865396	10,00651	0,000972164
MUSCLE	2	16,4800265		

Appendix Table 3 – Results of the first quantitative PCR (RT-qPCR) experiment conducted on *A. argo*. EYE indicates the retina, BRAIN the optic lobe, SUCKERS the proximal suckers of the right dorsal tentacle, MUSCLE the retractor muscle of the mantle, RH r-opsin, XEN xenopsin, RET retinochrome, PER peropsin, PSE pseudopsin and 2 the control (Elongation Factor 1-Alpha-like).

Octopus vulgaris **EXPERIMENT 2**

Sample	Target	Cq Mean	ΔCq	$2^{-\Delta Cq}$
			2	
EYE	RH1	16,193	-3,22534	9,352410686
EYE	RH2	40	20,58166	6,37239E-07
EYE	XEN	25,503	6,084662	0,014734463
EYE	RET	16,5781056	-2,84023	7,161355714
EYE	PER	32,1349089	12,71657	0,00014857
EYE	PSE	25,775	6,356662	0,01220265
EYE	2	19,4183383		
BRAIN	RH1	19,698	4,412476	0,046958286
BRAIN	RH2	27,5855797	12,30006	0,000198296
BRAIN	XEN	30,945	15,65948	1,93209E-05
BRAIN	RET	23,7799433	8,49442	0,00277284
BRAIN	PER	25,9334841	10,64796	0,000623224
BRAIN	PSE	25,6141735	10,32865	0,000777618
BRAIN	2	15,2855236		
SUCKERS	RH1	25,8663988	9,977301	0,000992049
SUCKERS	RH2	35,4041077	19,51501	1,33474E-06
SUCKERS	XEN	NG		1
SUCKERS	RET	26,8220082	10,93291	0,000511524
SUCKERS	PER	33,7263276	17,83723	4,27031E-06
SUCKERS	PSE	28,0741751	12,18508	0,000214747
SUCKERS	2	15,8890981		
MUSCLE	RH1	28,7303459	13,38577	9,34291E-05
MUSCLE	RH2	34,0964135	18,75184	2,26535E-06
MUSCLE	XEN	NG		1
MUSCLE	RET	22,0186767	6,6741	0,00979255
MUSCLE	PER	34,6350394	19,29046	1,55952E-06
MUSCLE	PSE	30,1456084	14,80103	3,50304E-05
MUSCLE	2	15,344577		

Appendix Table 4 – Results of the second quantitative PCR (RT-qPCR) experiment conducted on *O. vulgaris*. EYE indicates the retina, BRAIN the optic lobe, SUCKERS the proximal suckers of the right dorsal tentacle, MUSCLE the retractor muscle of the mantle, RH r-opsin, XEN xenopsin, RET retinochrome, PER peropsin, PSE pseudopsin and 2 the control (Elongation Factor 1-Alpha-like).

Argonauta argo **EXPERIMENT 2**

Sample	Target	Cq Mean	ΔCq 2	$2^{-\Delta Cq}$
EYE	RH1	16,025	-3,18254	9,079043085
EYE	RH2	NG		
EYE	XEN1	29,438	10,23046	0,000832386
EYE	XEN2	30,215	11,00746	0,000485763
EYE	RET	21,8512971	2,643757	0,160011019
EYE	PER	30,874	11,66646	0,000307642
EYE	PSE	29,671	10,46346	0,000708247
EYE	2	19,2075402		
BRAIN	RH1	22,906	6,41	0,01175974
BRAIN	RH2	NG		
BRAIN	XEN1	27,7360405	11,24004	0,000413438
BRAIN	XEN2	30.9999	14,404	4,61279E-05
BRAIN	RET	23,5328208	7,036821	0,007615631
BRAIN	PER	30,8985948	14,40259	4,61729E-05
BRAIN	PSE	22,417	5,921	0,016504463
BRAIN	2	16,496		
SUCKERS	RH1	26,6735025	9,556408	0,001328112
SUCKERS	RH2	32,8037103	15,68662	1,89609E-05
SUCKERS	XEN1	33,6866982	16,5696	1,02814E-05
SUCKERS	XEN2	NG		
SUCKERS	RET	27,21407 54	10,09698	0,000913074
SUCKERS	PER	34,9006844	17,78359	4,43206E-06
SUCKERS	PSE	27,409	10,29191	0,000797678
SUCJERS	2	17,1170945		
MUSCLE	RH1	28,5318128	11,86116	0,000268803
MUSCLE	RH2	NG		
MUSCLE	XEN1	31,1813682	14,51072	4,28389E-05
MUSCLE	XEN2	35,156615	18,48597	2,72376E-06
MUSCLE	RET	25,5040601	8,833412	0,00219219
MUSCLE	PER	NG		
MUSCLE	PSE	27,8765459	11,2059	0,000423339
MUSCLE	2	16,6706484		

Appendix Table 5 – Results of the second quantitative PCR (RT-qPCR) experiment conducted on *A. argo*. EYE indicates the retina, BRAIN the optic lobe, SUCKERS the proximal suckers of the right dorsal tentacle, MUSCLE the retractor muscle of the mantle, RH r-opsin, XEN xenopsin, RET retinochrome, PER peropsin, PSE pseudopsin and 2 the control (Elongation Factor 1-Alpha-Like).

V.II Normalization

The table below show the results normalization of RT-qPCR data

A.argo		**Experiment 1 2^(-DeltaCq)**	**Experiment 2 2^(-DeltaCq)**	**AVG**	**ST DEV**	**LN(AVG)**	**LN(AVG)- MIN** MIN=RH1 Muscle
EYE	RH1	4,201254415	9,079043085	6,640149	3,449117445	1,893134365	9,667096468
BRAIN	RH1	0,144912889	0,01175974	0,078336	0,094153494	-2,546743996	5,227218107
SUCKERS	RH1	0,000137282	0,001328112	0,000733	0,000842044	-7,218778547	0,555183556
MUSCLE	RH1	0,000572285	0,000268803	0,000421	0,000214594	-7,773962103	
EYE	RH2	0,000244011	NG				
BRAIN	RH2	0,00061194	NG				
SUCKERS	RH2	0,000530913	1,89609E-05	0,000275	0,000362005	-8,198969537	
MUSCLE	RH2	0,000240865	NG	0,000241		-8,331273205	
EYE	XEN1	0,000576914	0,000832386	0,000705	0,000180646	-7,257809504	0,516152599
BRAIN	XEN1	0,000546512	0,000413438	0,00048	9,40972E-05	-7,641776436	0,132185667
SUCKERS	XEN1	4,34832E-06	1,02814E-05	7,31E-06	4,19532E-06	-11,82560248	
MUSCLE	XEN1	5,14649E-05	4,28389E-05	4,72E-05	6,09952E-06	-9,962136488	
EYE	XEN2	0,000742046	0,000485763	0,000614	0,00018122	-7,395671005	0,378291098
BRAIN	XEN2	0,000158204	7,416E-05	0,000116	5,94282E-05	-9,060352365	
SUCKERS	XEN2	NG	NG				
MUSCLE	XEN2	1,37974E-05	2,72376E-06	8,26E-06	7,83023E-06	-11,70401649	
EYE	RET	0,172093648	0,160011019	0,166052	0,008543709	-1,79545228	5,978509823
BRAIN	RET	0,00524794	0,007615631	0,006432	0,00167421	-5,046503133	2,727458969
SUCKERS	RET	0,000951999	0,000913074	0,000933	2,75239E-05	-6,977602622	0,796359481
MUSCLE	RET	0,000996374	0,00219219	0,001594	0,00084557	-6,441332081	1,332630022
EYE	PER	0,000184345	0,000307642	0,000246	8,71845E-05	-8,310206315	
BRAIN	PER	0,000607187	4,61729E-05	0,000327	0,000396697	-8,026529067	
SUCKERS	PER	2,67086E-05	4,43206E-06	1,56E-05	1,57519E-05	-11,07014259	
MUSCLE	PER	NG	NG				
EYE	PSE	0,000896132	0,000708247	0,000802	0,000132854	-7,128165938	0,645796165
BRAIN	PSE	0,005657206	0,016504463	0,011081	0,007670169	-4,502538275	3,271423828
SUCKERS	PSE	0,000527559	0,000797678	0,000663	0,000191003	-7,319310736	0,454651366
MUSCLE	PSE	0,000972164	0,000423339	0,000698	0,000388077	-7,267647409	0,506314694
REFERENCE				1		0	7,773962103

O. vulgaris		**Experiment 1 2^(-DeltaCq)**	**Experiment 2 2^(-DeltaCq)**	**AVG**	**ST DEV**	**LN(AVG)**	**LN(AVG)- MIN** MIN=RH1 Muscle
EYE	RH1	10,40749158	9,352410686	9,879951	0,746054858	2,290507566	11,58319867
BRAIN	RH1	0,044139183	0,046958286	0,045549	0,001993407	-3,088972433	6,203718672
SUCKERS	RH1	0,001269339	0,000992049	0,001131	0,000196073	-6,78492374	2,507767366
MUSCLE	RH1	9,07607E-05	9,34291E-05	9,21E-05	1,88682E-06	-9,292691106	
EYE	RH2	0,000418029	6,37239E-07	0,000209	0,000295141	-8,471583621	0,821107484
BRAIN	RH2	0,000696601	0,000198296	0,000447	0,000352355	-7,711948463	1,580742642
SUCKERS	RH2	0,000167795	1,33474E-06	8,46E-05	0,000117705	-9,377992899	
MUSCLE	RH2	0,000160851	2,26535E-06	8,16E-05	0,000112137	-9,4141961	
EYE	XEN	0,001002353	0,014734463	0,007868	0,009710068	-4,844899532	4,447791574
BRAIN	XEN	1,55079E-05	1,93209E-05	1,74E-05	2,69622E-06	-10,95821266	
SUCKERS	XEN	34662,64856	NG				
MUSCLE	XEN	NG	NG				
EYE	RET	1,006161724	7,161355714	4,083759	4,35237941	1,407017819	10,69970892
BRAIN	RET	0,003783247	0,00277284	0,003278	0,000714465	-5,720508477	3,572182629
SUCKERS	RET	0,000333747	0,000511524	0,000423	0,000125707	-7,768999941	1,523691165
MUSCLE	RET	0,006168748	0,00979255	0,007981	0,002562415	-4,830735596	4,46195551
EYE	PER	6,19551E-05	0,00014857	0,000105	6,12459E-05	-9,159053329	0,133637777
BRAIN	PER	0,000855915	0,000623224	0,00074	0,000164537	-7,209441711	2,083249395
SUCKERS	PER	3,72181E-05	4,27031E-06	2,07E-05	2,32976E-05	-10,78324459	
MUSCLE	PER	NG	1,55952E-06				
EYE	PSE	0,000831306	0,01220265	0,006517	0,008040755	-5,033344546	4,259346559
BRAIN	PSE	0,001779173	0,000777618	0,001278	0,000708206	-6,662149178	2,630541928
SUCKERS	PSE	0,00091902	0,000214747	0,000567	0,000497996	-7,475357224	1,817333881
MUSCLE	PSE	0,001362125	3,50304E-05	0,000699	0,000938397	-7,266464356	2,026226749
REFERENCE						0	9,292691106

Appendix Table 6 – RT-qPCR) normalization results. EYE indicates the retina, BRAIN the optic lobe, SUCKERS the proximal suckers of the right dorsal tentacle, MUSCLE the retractor muscle of the mantle, RH r-opsin, XEN xenopsin, RET retinochrome, PER peropsin, PSE pseudopsin and 2 the control (Elongation Factor 1-Alpha-like).

VI Looking for opsins in optic lobe single cell data

With the help of Lorenza Rusciano, the level expression of opsin in the optic lobe single-cell RNA of *Octopus bimaculoides* (Songco-Casey et al., 2022) was investigated. In this study, the experiment was conducted on juvenile specimens (1.5 months from hatching). First, homologous opsin sequences were searched in a sequence file used for annotation via a simple BLASTn search and annotated according to the table below (Appendix Table 7). The *O. vulgaris* peropsin ortholog was absent from the annotation.

Gene	Sequence name
R-opsin1	obimac0017169
R-opsin2	obimac0015293
Xenopsin	obimac0025385
Retinochrome	obimac0009479
Peropsin	absent
Pseudopsin	obimac0006929

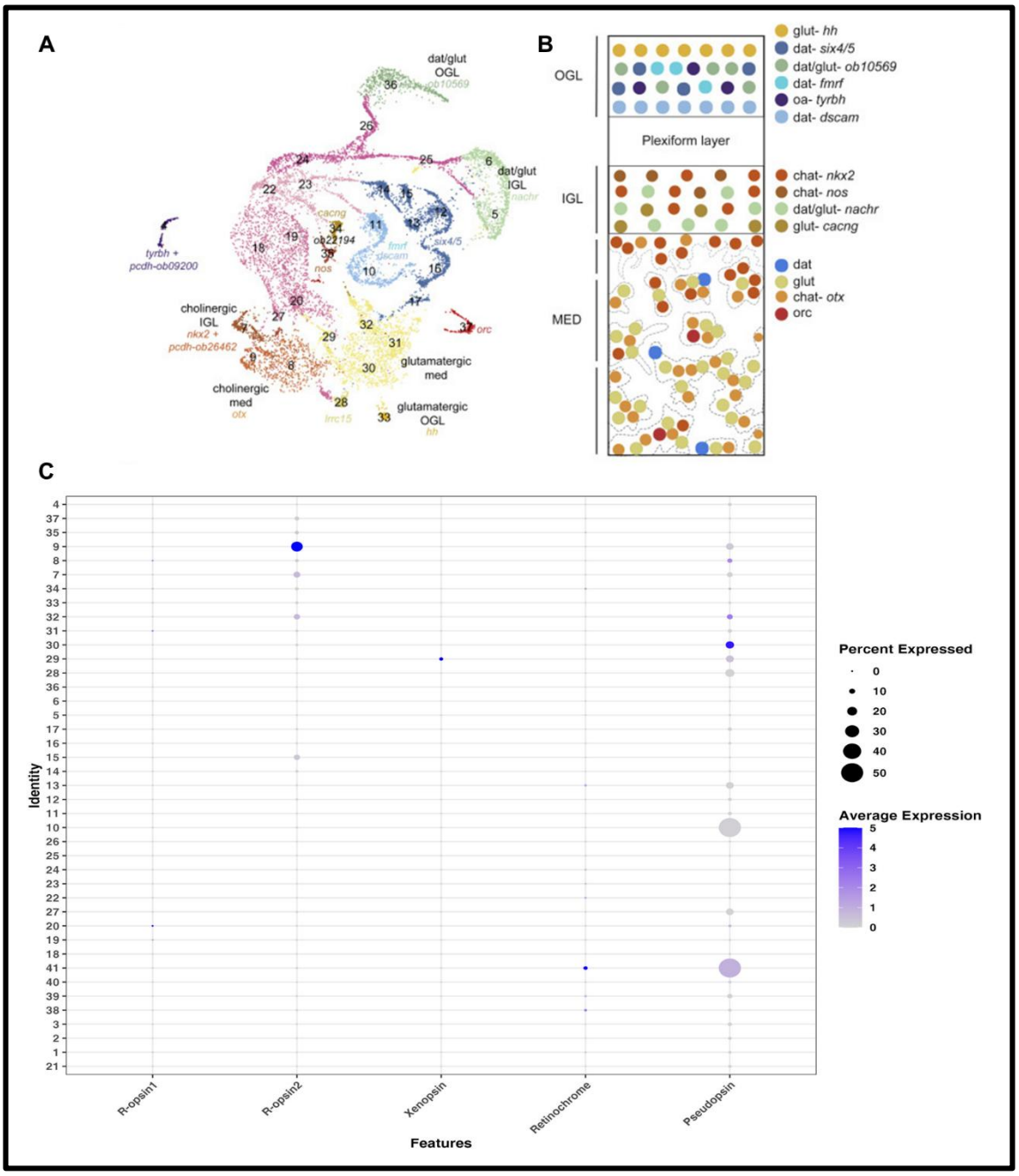
Appendix table 7– Opsin sequences. The table shows the opsin sequences name as available in Songco-Casey et al. (2022) single cell transcriptome.

For convenience, the optic lobe medulla will be referred to as MED, the inner granular layer as IGL, and the outer granular layer as OGL. Our results show that r-opsin1 is highly expressed in a few cell types classified as immature neurons (19-20) and moderately expressed in cholinergic cells in MED and IGL (8) and glutamatergic cells in MED (31). On the other hand, r-opsin2 is well expressed in cholinergic cells in MED (7-8-9) and moderately expressed in a few dopaminergic

cells in OGL (15) and glutamatergic cells in MED (32). Xenopsin is well expressed in glutamatergic cells in MED (29). Regarding the photoisomerases, retinochrome shows a proficient level of expression in glial (38) and endothelial (39-41) cells (number reported in supplementary data, Songco-Casey et al. 2022). The enigmatic pseudopsin is expressed in many cell types, including dopaminergic cells in OGL (10), cholinergic cells in MED (8), immature neurons (20, 27), and endothelial cells (41), with peak expression in glutamatergic cells in MED (30, 32). It was not possible via BLASTn search to find peropsin in the sequence files; therefore, data are missing.

Overall, in *O. bimaculoides* juveniles, all the opsins are expressed in the MED region and might play a role in integrating signals coming from the outer layer of the optic lobe and the eye, integrating the visual information with the light reaching these areas.

The low volume of the juvenile optic lobe may allow more photons to reach deeper layers and, therefore, the MED, justifying this unexpected pattern of expression in the most central areas. Conversely, as the optic lobe increases in size, opsin expression may change, becoming more pronounced in the outer layers, as observed in *Sepia officinalis*, where r-opsin1 expression is more evident in the OGL and IGL. Retinochrome appears to be highly expressed in cells that are less likely to be photoreceptive but contribute to supporting functions. Pseudopsin is extremely expressed in many cell types, indicating its involvement in important intercellular communication and regulation mechanisms, as suggested by its high expression in the medulla.



Appendix Figure 9 – Opsin expression in single-cell data of *Octopus bimaculoides* juvenile optic lobe. Distinct cell clusters and cell types are identified in the *O. bimaculoides* optic lobe (A), distributed across separate layers (B). Opsins are expressed within various cell clusters in the optic lobe (see main text), with cluster numbers corresponding to those in A. Images in A and B are from Songco-Casey et al. (2022).