Abstract

Eye evolution is driven by the evolution of visually guided behavior. Accumulation of gradually more demanding behaviors have continuously increased the performance requirements on the photoreceptor organs. Starting with nondirectional photoreception, I argue for an evolutionary sequence continuing with directional photoreception, low-resolution vision, and finally, high-resolution vision. Calculations of the physical requirements for these four sensory tasks show that they correlate with major innovations in eye evolution and thus work as a relevant classification for a functional analysis of eye evolution. Together with existing molecular and morphological data, the functional analysis suggests that urbilateria had a simple set of rhabdomeric and ciliary receptors used for directional photoreception, and that organ duplications, positional shifts and functional shifts account for the diverse patterns of eyes and photoreceptors seen in extant animals. The analysis also suggests that directional photoreception evolved independently at least twice before the last common ancestor of bilateria and proceeded several times independently to true vision in different bilaterian and cnidarian groups. This scenario is compatible with Pax-gene expression in eye development in the different animal groups. The whole process from the first opsin to high-resolution vision took about 170 million years and was largely completed by the onset of the Cambrian, about 530 million years ago. Evolution from shadow detectors to multiple directional photoreceptors has further led to secondary cases of eye evolution in bivalves, fan worms, and chitons.

Keywords: Eye evolution, Photoreceptor, Vision, Visually guided behavior

Introduction

One of the first steps in eye evolution must have been the appearance of a light-dependent chemical reaction coupled to a signaling system. All animal phyla, with the exceptions of placozoans and sponges, employ the opsin-class of 7-transmembrane proteins for this purpose (Plachetzki et al., 2007; Porter et al., 2011; Feuda et al., 2012; Schnitzler et al., 2012). Other crucial and early steps in eye evolution must have been the development of specialized photoreceptor neurons capable of holding substantial quantities of opsin and the association with screening pigment to generate directionality (Arendt et al., 2009; Nilsson, 2009). Arrays of photoreceptors forming retinas and the development of optics to focus light must later on have been essential steps in the evolution of eyes and true vision. But what were the driving forces that made evolution follow the route from simple light sensitivity to acute spatial vision? Eyes or any other sensory structure would only evolve if they cause an increase in fitness, and this increase is generated through behaviors that rely on sensory information. In fact, one of the distinguishing features of animals, as compared to plants, fungi, and unicellular organisms, is their ability to process vast quantities of external information and generate complex behaviors in response to the information (Atick, 1992).

Obviously, selection acts both on the eyes and the visually guided behavior that the nervous system is able to generate (Endler, 1992). But it can be argued that selection acts more directly on the behavior, which may be more or less successful depending on the quality of the information provided by the sensory organs (Nilsson, 2009). The whole-evolutionary process would then start by genetic variation that causes variation in sensory structure and function, and this in turn allows for variation in the behavior that relies on the sensory information, which ultimately causes the change in fitness that selection can act on (Fig. 1). From this point of view, eye evolution is a consequence of the evolution of visually guided behaviors and any attempt to understand the evolution of eyes would require that we first understand the evolution of visually guided behaviors. Here, I will take that approach by first classifying behaviors that rely on information from photoreceptors and then relate that to the structures and functions needed to quantitatively provide the necessary sensory discrimination. The result reveals a surprising correlation between major classes of photoreceptor-controlled behavior and key steps in eye evolution. To get the whole picture, I proceed to map the behavioral classes and major functional innovations onto current phylogenetic trees, and finally, I focus on specific events of particular importance for reconstructing the evolution of photoreception and vision.
Classifying photoreceptor-controlled behavior

Apart from true vision, simpler types of light sensing are well known for controlling a range of different behaviors. In a classification of behaviors aimed at understanding the origin of vision, these behaviors are of course of crucial importance because they are likely to have preceded the evolution of “visually guided” behaviors. The behaviors of interest must thus cover the broader context of photoreceptor-controlled behaviors. I here elaborate on the classification into major classes of sensory tasks and related behaviors that were recently introduced (Nilsson, 2009; Land & Nilsson, 2012). The classification is a sequence of four classes of behaviors defined by the sensory tasks, starting with those that require the smallest amount of information, and proceeding through behaviors that require gradually larger amounts and more complex information from photoreceptors. As the amount of information grows, the behaviors gradually become more complex and gradually require more sophisticated sensory organs and neural processing.

Class I: Nondirectional photoreception

By just monitoring the ambient light intensity, it is possible to get information for a number of important behaviors such as directly triggering or turning off behaviors in relation to the time of day or indirectly by providing input to a circadian clock (Bennett, 1979). Aquatic animals can also use the ambient intensity as a depth gauge to control their vertical position in the water column (Lythgoe, 1979). Burrowing animals may use nondirectional photoreception to trigger appropriate behaviors when the animal breaks the substrate surface and enters the lit world. Shadow detection can inform animals when they move into or out of shadows or when a shadow moves in on the animal. Warning for harmful levels of UV radiation is another task that is essential for many animals (Paul & Gwynn-Jones, 2003; Leech et al., 2005).

Class II: Directional photoreception

If the photoreceptor is shielded such that it detects light in some directions but not in others, it is well suited to determine the direction toward or away from the light in a moving animal. The obvious behaviors guided by directional photoreception are phototaxis based on scanning body movements (Jékely et al., 2008). The photoreceptor’s directionality in combination with changes in body orientation will provide a direct feedback to steering such that the animal can move toward the light or seek out a dark refuge. A directional photoreceptor can also be used as an optical “statocyst” to orient the body posture in relation to the light (Nilsson, 2009). Shadow detection, which was listed under class I, can also be performed by a directional photoreceptor. If the directionality becomes sufficiently narrow, a photoreceptor can warn about movements of objects, and this can lead to higher behaviors classes.

Class III: Low-resolution vision

An array of directional photoreceptors, such as those in pit or cup eyes (as well as in more advanced eyes), will make simultaneous readings of the luminance in different directions. Even with very crude resolution such information can serve a large number of important tasks. Monitoring of self-motion in relation to the stationary world is a task that is ideally performed using low spatial frequencies, which are also the most economical to process. Knowledge about the flow field caused by self-motion can be used to control speed and direction of locomotion as well as the trajectory in relation to surrounding structures (Theobald et al., 2009). Object avoidance (anticollision) responses also belong to this class of behaviors. Finding a suitable habitat and maintaining a position within it are further tasks that require very modest spatial resolution. In fact, most navigation and visual feedback from the inanimate surrounding can be classified as low-resolution vision. This class introduces imaging and spatial resolution (which would require a minimum of two resolved pixels). Thus, it represents the first true eyes and the origin of vision (rather than just photoreception, as in classes I and II).

Class IV: High-resolution vision

By just increasing the spatial resolution, it is possible to engage in detection, pursuit, and communication with other animals. At distances exceeding a few body lengths, predators, prey, and conspecifics will occupy only small visual angles. But information about other animals within movement range allows for a huge range of advanced visually guided behaviors. Detection and pursuit of prey, predator detection and escape, mate recognition and visual communication are perhaps the most basic and common behaviors in this class. Flower detection in floral visitors and seed detection in seed eaters are also tasks that clearly rely on high spatial resolution. For a few advanced species, making and using tools add to this behavioral class of visually guided behaviors. Reading this text is one of many specifically human behaviors that require high spatial resolution.

It may seem self-evident that class IV tasks must have evolved from class III tasks that they in turn evolved from class II and that class I tasks were the starting point. The four classes are thus different stages in the evolution of eyes and vision. Once evolution has reached a particular class, it is of course possible for the process to go either way, such that simpler behaviors evolve from more complex ones. But here we are mainly concerned with the first appearance of each class. Good support for the assumption that the behavioral classes originally evolved in ascending order is offered by the fact that the amount and rate of information fed to the
nervous system increases steeply for each higher class of behaviors [for actual numbers, see eqn. (4) and Table 1 below].

**Fundamental performance requirements for the four classes**

To further corroborate the hypothesis that each of the behavioral classes originally evolved from the next lower class, we go on to compare the physical parameters that have to be met by the organs of light detection. More specifically, we consider the demands on integration time (response-speed), detection angle, and precision of detection. With knowledge about these functional requirements, at the organ level, we can correlate the evolution of the behavioral classes with the evolution of photoreceptor cells and visual organs.

**Integration time**

Quite clearly, the integration time needed for avoiding motion blur in high-resolution vision is much shorter than that needed for monitoring the daily light cycle. Also, for a small planktonic organism to measure depth in the sea, changes will be slow, and a long integration time is desired to extract the relevant information. The nondirectional light sensitivity needed for class I tasks relies on the temporal variation in the radiance from the main light source, such as the sun or the moon, and this variation is typically slow. Integration times of about 10 min are long enough to filter out faster signal variations that have other causes and short enough to provide good temporal resolution for tracking diel radiance-variations or monitor water depth in planktonic organisms. Experimental measurements of the integration time in nondirectional photoreceptors are in the range of 10–30 min (Gotow & Nishi, 2008). For detection of shadows that the animal moves into or that moves over the animal, much shorter integration times are useful. For directional photoreception and true vision, the animals’ self-motion will also cause much faster signal changes, calling for integration times of 1 s or shorter (Nilsson, 2009).

**Detection angle**

The required detection angle becomes narrower for each higher class. By definition, nondirectional photoreception (class I) should not be restricted in angle, and with even slight directionality, changes in body posture would contaminate the information on the general ambient radiance. Unshielded photoreceptors in small transparent animals can achieve truly nondirectional detection, but in larger and pigmented animals, shielding by the body inevitably causes directionality, although it is in principle not desirable. In contrast, class II tasks require directionality, but for phototaxis and control of body posture, the detection angle should not be too small to the point where the photoreceptor responds to image details rather than guiding toward brighter or dimmer parts of the habitat. Photoreceptors for phototaxis typically collect light over 100–180° (see Jékely et al., 2008). Classes III and IV require much smaller detection angles because these will directly determine the visual acuity, which range from 30° down to fractions of a degree (Nilsson, 2009; Land & Nilsson, 2012).

**Table 1. Values used for calculations of minimum intensity for the four classes of photoreceptive tasks**

<table>
<thead>
<tr>
<th></th>
<th>Class I</th>
<th>Class II</th>
<th>Class III</th>
<th>Class IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration time, Δt (s)</td>
<td>600 (1800)</td>
<td>1</td>
<td>0.1 (0.2)</td>
<td>0.05 (0.1)</td>
</tr>
<tr>
<td>Detection angle (°)</td>
<td>360</td>
<td>180</td>
<td>25 (40)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Contrast, ∆I (%)</td>
<td>30 (60)</td>
<td>10 (20)</td>
<td>3 (3–10)</td>
<td>3 (3–10)</td>
</tr>
<tr>
<td>Photon sample, N (number of quanta)</td>
<td>50 (10)</td>
<td>500 (100)</td>
<td>5000 (500)</td>
<td>5000 (500)</td>
</tr>
<tr>
<td>Number of membrane layers, n</td>
<td>2</td>
<td>2 stacking: 50 (500)</td>
<td>2 stacking: 300 (2000)</td>
<td>2 stacking: 1500 (4000)</td>
</tr>
<tr>
<td>Membrane absorption, k_{u} (10^{-7})</td>
<td>0.5 (1)</td>
<td>0.5 (1)</td>
<td>0.5 (1)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Diameter of detection area or aperture (mm)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.005 lens: 0.1 (1)</td>
<td>0.005 lens: 3 (15)</td>
</tr>
<tr>
<td>Minimum intensity, Log_{min} (Log quanta m^{-2} sr^{-1} s^{-1})</td>
<td>11.7 (10.2)</td>
<td>15.8 (14.8)</td>
<td>20.0 (18.0)</td>
<td>23.1 (20.0)</td>
</tr>
<tr>
<td>Cell with no membrane stacking</td>
<td>14.4 (12.5)</td>
<td>17.8 (15.3)</td>
<td>20.3 (17.4)</td>
<td>23.1 (20.0)</td>
</tr>
<tr>
<td>Cell with membrane stacking</td>
<td>15.2 (10.7)</td>
<td>14.8 (10.4)</td>
<td>20.3 (17.4)</td>
<td>23.1 (20.0)</td>
</tr>
<tr>
<td>Cell with membrane stacking and focusing optics</td>
<td>11 × 10^{-3}</td>
<td>6.66</td>
<td>2.8 × 10^{6}</td>
<td>3.5 × 10^{6}</td>
</tr>
</tbody>
</table>

Values in brackets are chosen as the most relaxed end of the range beyond which the task would be too compromised to provide reliable information for the respective class of behaviors. For the number of membrane layers, and the detection area, values are given for the absence and presence of membrane stacking and focusing optics, respectively. The bottom row gives the information capacity in bits/s, calculated from the nonbracketed values for the entire organ, assuming a signal-to-noise ratio of 100 for all photoreceptors and 180° visual fields for classes III and IV.
typically a matter of detecting objects against a background. An accuracy (contrast sensitivity) of about 3% would be sufficient to discriminate most natural objects and their backgrounds (Srinivasan et al., 1982).

The accuracy of intensity detection is directly correlated with the minimum number of photons that are detected per integration time. The reason for this correlation is that photons arrive randomly, and if different intensities are to be discriminated, a statistically significant difference in the number of photons has to be detected. The minimum number of photons, \( N \), needed to discriminate an intensity, \( I \), from another intensity differing by the amount \( \Delta I \), is

\[
N = 2 \left( \frac{R}{\Delta I/I} \right)^2,
\]

where \( R \) is a confidence factor, which is 1.28 for 80% confidence and 1.96 for 95% confidence (Land, 1981). Discrimination of 30, 10, and 3% intensity differences (as assumed above) with a moderate confidence of 80% then requires detection of 36, 330, and 3600 photons per integration time, respectively. If a higher confidence is desirable, say 95%, then the corresponding photon samples must be at least 85, 770, and 8500. Obviously, the required number of photons detected in each integration time increases steeply with higher contrast sensitivity. Assumed values for the parameters discussed above, for each of the four behavioral classes, are summarized in Table 1.

Mathematical expression of the minimum intensity needed for the sensory tasks

It is clear that progress through classes I–IV involves monotonic reductions in both integration time and detection angle, leading to dramatically shrinking numbers of photons detected per integration time. But the required intensity discrimination instead calls for gradually larger photon samples for each higher class of behavioral tasks. That leads to a critical question: are there enough photons in the environment to allow evolution from class I through to class IV? We can easily calculate the minimum intensity \( I_{\text{min}} \) that

![Fig. 2. Minimum intensities from Table 1, for the four classes of sensory tasks, plotted together with the daily variation of natural luminances and daylight intensities at different depths in clear water. Blue indicates calculations for a 10 μm diameter cell with no membrane stacking and no focusing optics (just screening to obtain the desired detection angle). Calculations for membrane stacking are indicated by green and for focusing optics (and membrane stacking) by red. The color gradients at the lower end of the bars show the range of gradually decreasing function between the primary values from Table 1 (dotted lines), and the bracketed values where sensory information is assumed too poor for the task.](https://doi.org/10.1017/S0952523813000035)
a sensory task will need in order to work (see Nilsson, 2009 for a derivation):

\[ I_{mn} = \frac{N}{\Omega Z \Delta t a}, \]

(2)

where \( N \) is the minimum photon sample from eqn. (1), \( \Omega \) is the solid angle of sensitivity, \( Z \) is the sensitive area, \( \Delta t \) is the integration time, and \( a \) is the quantum efficiency of received photons.

The quantum efficiency, \( a \), can be further developed as a function of the number of membrane layers that contain opsin in the photoreceptor cell. For absorption of white light (Warrant & Nilsson, 1998), this becomes (Nilsson, 2009):

\[ a = 0.4 \left( \frac{k_a u}{2.3 + k_a u} \right), \]

(3)

where \( k_a \) is the absorption at peak wavelength of a single membrane layer, \( u \) is the number of membrane layers, and the factor 0.4 is the product of assumed values of 0.5 for the quantum efficiency of transduction and 0.8 for light transmission through the ocular media (Warrant, 1999). The fraction of absorbed light per membrane layer, \( k_a \), was estimated to be \( 1 \times 10^{-3} \) for ciliary photoreceptors and \( 0.5 \times 10^{-3} \) for rhabdomeric receptors (Nilsson, 2009).

In addition to calculating the minimum workable intensity, I also use the same performance values to calculate the information capacity of the four classes (Table 1). Applying the Shannon–Hartley theorem to an eye (see Snyder, 1979), the information capacity, \( H \), can be expressed in bits/s if we use a base 2 logarithm:

\[ H = \frac{n}{\Delta t} \log_2 (1 + snr), \]

(4)

where \( n \) is the number of pixels and \( snr \) is the signal-to-noise ratio of a photoreceptor, here set to 100. The intriguing result of performing this calculation on the assumed performance values for the four different classes of tasks is that the information rate increases roughly 1000-fold for each higher class. Because the information rate depends linearly on the number of pixels, it can vary quite substantially within classes II and III. The calculations nevertheless demonstrate that eye evolution implies an extreme and persistent increase of the amount of information passed to the nervous system.

Using the theory for identifying major innovations in eye evolution

Applying eqn. (2) to the values listed in Table 1 for a cell of typical eukaryotic dimensions (10 μm), with densely packed photopigment in its membrane but no morphological specializations, reveals that it will work down to well below starlight levels (Fig. 2). This is an important conclusion because it demonstrates that nondirectional monitoring of ambient intensities can be performed by cells without any morphological specializations. Such cells can thus be very inconspicuous and only be identified by labeling of the photopigment. Consequently, it is likely that the majority of nondirectional photoreceptors across the animal kingdom are yet to be discovered.

Applying the same calculation to sensory tasks of class II reveals that they will work into mid-dusk intensities or about 4 log units dimmer than bright sunlight. By accepting less reliable detection, the sensory task can be pushed to work another log unit into even dimmer conditions. But there is no way that directional photoreception can work at night-time intensities with morphologically unspecialized receptor cells, unless the integration time or detection angle are extended beyond workable values. Even though there is an obvious uncertainty as to what might be realistic input values for eqn. (2), the general conclusions are robust because the calculated values are mapped on a log scale stretching over almost 10 log units of natural intensities.

Class III sensory tasks (low resolution vision) require even more light, and with the assumed input values, these tasks in fact require light intensities exceeding bright sunlight. By tweaking the input values, it is possible to make class III tasks work under sunlit conditions but not at lower intensities. The results for high-resolution vision (class IV) are even worse: these tasks need intensities about 1000 times brighter than direct sunlight. Obviously, on the planet Earth, it is not a viable solution for vision to have receptor cells that are morphologically unspecialized and obtain directionality by screening only.

One of the reasons for the poor outcome of the calculations is that the cell was assumed to lack morphological specializations, and thus, light passes only two layers of membrane. This means that only 0.02% of the light reaching the cell will be captured and detected. An obvious remedy to this problem is to fold the membrane to form stacks that light has to pass. The more layers in the stack, the larger is the fraction of incident light that will be absorbed and detected. The maximal gain is about 3.4 log units, but with a realistic number of membrane layers, say 2000, the sensitivity gain is 2.5 log units. The rods of rhesus monkeys have about 1000 discs, which make 2000 membrane layers (Young, 1971). Vertebrates with longer rods or invertebrates with long rhabdoms may have significantly larger numbers of membrane layers in their photoreceptors, and they may approach a gain of about 3 log units.

When the sensitivity gain brought by membrane stacking is introduced in the calculations, it is clear that it can rescue class III tasks such that they can be performed into mid-dusk intensities (Fig. 2). For class II tasks, even a small or moderate number of membrane layers would be sufficient to extend operation to intensities significantly below mid dusk. But for high-resolution vision (class IV) membrane, stacking is not enough to allow vision even in bright sunlight.

The only remaining variable of eqn. (2) that can be changed in order to capture more photons is the sensitive area, \( Z \), which has been assumed to be the area of the cell. The situation can be greatly improved by placing a lens in front of the cell. The sensitive area then becomes the lens area rather than the cell area, and with a lens of 3 mm diameter, the gain in sensitivity is an impressive 5.5 log units. With a lens of 15 mm and sacrifices of speed, resolution, and contrast sensitivity, it is possible to push class IV tasks across more than 9 log units into the lower end of night-time intensities. Compound eyes of the apposition type are constrained to have much smaller lenses and will consequently work only in daylight unless the spatial and temporal resolution is sacrificed by summation (see Warrant, 1999). Compound eyes of the superposition type have large effective apertures and perform similar to camera-type eyes of equal size (see Land & Nilsson, 2012).

Focusing by curved mirrors is an alternative and equally efficient solution for increasing the receptive area because the reflector can be just as large as a lens. The reflecting tapeta, which are backing the retina in many lens eyes and compound eyes, have much less of an impact on sensitivity. The reason for the limited sensitivity gain of tapeta is that much of the light has already been absorbed during the first passage through the retina, and only a small fraction remains for the second passage. If 50% is absorbed on the way in, and the tapetum reflects all the remaining 50%, then half of that, i.e., 25% extra, will be absorbed on the way out. On the
log scale of Fig. 2, this would be a hardly noticeable gain of 0.13 log units. Tapeta can thus be safely ignored in this discussion.

In summary, class I tasks can be performed at below starlight intensities with unaided and morphologically unspecialized cells. Class II tasks can also be performed by such plain receptor cells but only down to mid-dusk intensities. For operation at lower intensities, it is necessary to introduce membrane stacking, which is a prerequisite for class III tasks to function even at daylight intensities. To extend class III tasks below mid-dusk intensities, it is necessary to add focusing optics. High-resolution vision requires both membrane stacking and focusing optics for all intensities, and sufficiently large apertures allow operation down to night-time intensities. The classification of sensory tasks for different classes of behaviors can thus be translated to corresponding types of sensory structure or function, and the resulting sequence of functional innovations define the key steps in eye evolution (Fig. 3). These four fundamental innovations in eye evolution (light sensitivity, screening structures, membrane stacking, and focusing optics) may very well qualify as “key innovations” in the sense of Simpson (1944).

Relating sensory tasks to sensory structure across the animal kingdom

From the exercise of analyzing functional requirements for the four behavioral classes, it is obvious that screening pigment, membrane stacking, and focusing optics must have been essential evolutionary innovations, each enabling transition to a higher class of sensory task. This means that the sensory organs required for the different classes of sensory tasks can be identified by their structure, and it is possible to transfer the classification of tasks to a corresponding evolutionary classification of photoreceptor organs.

Class I tasks would then be associated with cells lacking any kind of membrane stacking and any proximity to screening pigment or other light-shielding structures. Such light sensitive cells are indeed known from many corners of the animal kingdom (e.g., Provencio et al., 1998, Gotow & Nishi, 2007, 2008; Peirson et al., 2009; Ramirez et al., 2011; Backfisch et al., 2013). Fineal photoreceptors and polychaete brain photoreceptors are clearly serving class I tasks but are unusual in that they display some membrane stacking. This is discussed later under the heading “The evolution and fate of ciliary photoreceptors in bilateria.”

The range of functions involving class I photoreception is impressive, with many surprising examples. Interneurons can serve as class I photoreceptors by using the ambient light intensity to modulate the signal they pass. The photosensitive retinal ganglion cells in vertebrate eyes are irradiance-detecting interneurons that are thought to have such a modulating function (Peirson et al., 2009), but they differ from typical class I detectors because their position in the eye makes them highly directional. Nonneural cells may also express opsin to control various processes according to the ambient luminance. Possible examples are the opsins expressed in jellyfish gonads (Suga et al., 2008), and several different tissues in sea urchins (Raible et al., 2006), but it cannot be excluded that these opsins have roles other than light detection. The firing of cnidianian nematocysts is controlled by nearby neurons expressing opsins (Plachetzki et al., 2012), and squid photophores contain opsin-expressing cells that monitor the animal’s own bioluminescence (Tong et al., 2009).

A pigment cell and one or two photoreceptor cells typically form sensory organs serving class II tasks (Fig. 4B). Such organs are common in planktonic larvae of many different phyla, and the photoreceptor cells typically form membrane stacks of a few 10s to several 100 layers (Eakin, 1972; Salvini-Plawen & Mayr, 1977; Jékely et al., 2008). Some median eyes, such as the nauplius eyes (frontal eyes) of crustaceans are directional photoreceptors that possibly aid in controlling body orientation (Elofsson, 2006). The paired eyespots of Acoels are unique in having a rather unspecialized membrane without microvillar or ciliary extensions (Yamasu, 1991). The sensitivity calculations also indicate that membrane stacking is not necessary for phototaxis in well-lit habitats. But
The sensory structures related to class III tasks (low resolution vision) are pit and cup eyes that are widely spread across many different animal phyla but often poorly studied (Eakin, 1972; Vanleeuwen & Coomans, 1976; Salvini-Plawen & Mayr 1977; Vanleeuwen, 1982; Arendt & Wittbrodt, 2001). The eyes of flatworms (Platyhelminthes) frequently serve as examples for lens-less cup eyes and these consist of a pigment cup, often formed by a single cell, and a varying number of photoreceptor cells (Fig. 4C). Some flatworm species have a single pair of large cup eyes with numerous receptors in each, whereas others have many smaller eyecups with few or even just a single receptor in each. This difference illustrates the origin of the two fundamental solutions to spatial vision: simple and compound eyes.

Eyes built for low-resolution vision often have lens-like structures. The eyes of box jellyfish (Fig. 4D), polychaete worms, and gastropods are excellent examples. At least in box jellyfish, it is known that the lenses are too weak to focus a sharp image on the retina (Nilsson et al., 2005; O’Connor et al., 2009), and it is likely that the same is true for polychaete eyes, most gastropod eyes and other small cup-eyes with a lens-like structure filling the cup. This type of under-focused lens eye (Land & Nilsson, 2012) will provide a better photon catch than lens-less eyes, and for low spatial resolution, a perfect focus would not further improve the situation. Expressed another way, a lens-less cup eye can generate small detection angles by reducing the size of the aperture, but this will be at the expense of the photon catch. The under-focused lens will allow smaller detection angles without corresponding losses in photon catch. For even smaller detection angles (class IV), the lenses will have to place a sharp image on the retina to avoid unnecessary losses in photon catch. The lenses of a few gastropods such as Littorina and Strombus form sharp images on the retina, but they are not known to use the high performance for more than class III tasks (Seyer, 1992, 1994).

Low-resolution vision can also be performed by dispersed photoreceptors. Sea urchins are able to detect and move toward dark objects (Blevins & Johnsen, 2004; Yerramilli & Johnsen, 2009). This behavior is based on opsin expressed in their tube feet, and screening by the spines or the dark calcite skeleton provides directionality (Lesser et al., 2011; Ullrich-Lüter et al., 2011). Because tube feet are distributed across the entire body, the arrangement acts like a dispersed compound eye. The dispersed lenses on brittle stars may serve a similar function (Aizenberg et al. 2001), whereas starfish have concentrated their organs for class III tasks to compound eyes, called the optic cushions, on the tip of each arm (Yoshida & Ohtsuki, 1968; Penn & Alexander, 1980).

Eyes for class IV, high-resolution vision, have only evolved in a few animal groups: vertebrates, cephalopods, and arthropods. They can be of both camera type and compound type, but common to all is that they rely on focusing optics. The compound eyes of insects and crustaceans give room only for rather small lenses, and this is reflected in a lower resolution compared to camera-type eyes of the same size. Space constraints in the compound eyes of the smallest insects and crustaceans are incompatible with high-resolution vision, and in these cases, the animals can only use their eyes for class III
tasks. Also, for all those vertebrates, cephalopods and arthropods that rely on class IV vision, low-resolution vision remains essential and is performed by the same eyes. Within the arthropods, it seems that high-resolution vision may have originated independently in insects and spiders (mainly jumping spiders and wolf spiders).

Just like with the other classes of tasks, there are a couple of isolated and unusual cases that deserve mentioning. Heteropod snails and Alciopid polychaetes are visually guided predators outside the three groups of animals that otherwise are unique in mastering class IV behaviors. But these cases are not known to use their eyes for other class IV tasks, and they may represent rather recent attempts to advance from class III to class IV. Another unusual group is copepod crustaceans. They lack compound eyes and in some cases, such as Pontella, Copilia, and Sapphirina, their normally simple frontal eyes have evolved into special purpose eyes with narrow visual fields, and target-finding retinas used for spotting mates or prey (see Land & Nilsson, 2006).

Focusing by curved mirrors typically produce very compressed images, and ray paths that make high spatial resolution (class IV) difficult. Eyes focusing by concave mirrors are consequently rare and only found in aquatic animals from dim habitats, such as scallop, some ostracod and hyperiid crustaceans, and some deep-sea fish (for details and references see Land & Nilsson, 2012). Some types of compound eyes, mainly in crustaceans, also use mirrors for focusing, but these solutions are more successful and represented in a larger number of species (Nilsson, 1990). Some of the eyes using mirrors for imaging are limited to serve class III tasks, but others are probably involved in class IV.

**Visual systems evolve through acquisition of increasing numbers of behaviors**

Most vertebrates display a huge number of different behaviors that are aided by vision in one way or another. But all those behaviors were of course not acquired at the same time. Each new behavior requires modification and expansion of the sensory processing circuits and motor circuits, and the natural evolutionary process must be that new behaviors are acquired sequentially. Some of the new behaviors will have slightly higher sensory performance requirements, and that will drive evolution of the sensory organs, and also allow even more demanding behaviors to enter the scene. Eye evolution is thus driven by sequential additions of new and more demanding behaviors. The numbers of behaviors that rely on classes I and II sensory tasks are small. They increase to moderate for class III and then rise sharply for class IV. This is also reflected in the brain size of the animals (Land & Nilsson, 2006).

A common conception is that eyes evolved from poor to perfect. But in the light of the reasoning here, it is obvious that eyes, as well as simpler photoreceptive organs, all can be expected to be optimized for the behavioral tasks that they support. A more correct view is thus that eyes evolved from performing few and simple tasks excellently to performing many complex tasks excellently.

This said, there are many cases where photoreception have reverted back to lower functional classes when changes in life style have altered or removed selection for good visual performance.

**Preadaptations and transitions between classes of sensory tasks**

The introduction of four distinct classes of photoreceptive sensory tasks and corresponding classes of sensory morphology are associated with the major innovations identified above. But none of these innovations are likely to suddenly have appeared fully functional. They each involve complex morphology/physiology that must have evolved in numerous steps, driven by selection for improved function. And the final function may not have been the same as the function that originally started the evolution of the innovation. There are thus good reasons to search for mechanisms that might have facilitated the introduction of the different innovations.

The transition from class I to class II would require a substantial reduction of the integration time (improvement of response speed) and the introduction of light shielding material close to the receptor. Both these modifications would compromise the slow and nondirectional irradiation monitoring of a class I detector. But if a duplication of the receptive system frees one of the copies from the original constraints, it can gradually become faster and start to act as a shadow detector rapidly informing the animal when it drifts into a shaded area of the habitat. Direct shadows also have a directional quality, which would favor the introduction of screening pigment to determine the direction to the shading object. Screening pigment itself may already have been around in low concentrations to repair or reduce photic damage. Melanin cannot only sustain photon absorption, but it also acts as an anti-oxidant that can prevent damage by photon-generated free radicals (Meredith & Riesz, 2004; McGraw, 2005). Screening could also evolve as a by-product of increased body size, where the body itself would shield light from some directions.

The membrane stacking that must be present for class III tasks to evolve is already an advantage for class II tasks. And the membrane stacking structures, cilia and microvilli are present also in sponges, suggesting that the ability to form these organelles evolved for other reasons long before they became useful in photoreception. Just like screening pigment, membrane stacking may thus have been rather easily recruited when photoreception gradually became involved in more demanding tasks.

The introduction of a lens is an obvious way for allowing the transition from class III to class IV. But focusing can be introduced gradually and even weak focusing provides an improved photon catch for class III tasks. There are also several additional reasons that may have motivated the evolutionary introduction of a lens. In some of the box jellyfish (O’Connor et al., 2009), and also in polychaete worms, the cup eyes are filled with a lens-like transparent material. But it has practically no focusing power. Instead, it may prevent foreign material to enter the eyecup, it may act as a scaffold around which a smoothly curved retina can be formed, and it may also act as a filter removing short wavelengths that are not used for vision but could cause damage to the retina. There are thus several reasons to develop the material that lenses can later evolve from.

Small lens-like structures are known from some larval eyespots for directional photoreception (Blumer, 1994; Passamaneck et al., 2011). These are too small to improve photon catch, but they could aid in obtaining directionality, or they may have a role as spectral filters. The origin of these lens-like structures in class II photoreceptor organs could also result from larval eyes being derived from more advanced adult eyes in accordance with the larvae being derived from more ancient body plans of adults (Hejnol & Martindale, 2008; Raff, 2008).

**The need for an extra class of photoreceptive tasks**

Scanning the animal kingdom, it is clear that the four classes of behaviors and corresponding morphologies, normally followed in
eye evolution, cannot comfortably accommodate all known photoreceptor organs. Specifically, the classification does not have room for some prominent photoreceptor organs that also happen to be of evolutionary importance. This type is the alarm photoreceptors on the mantle edge of bivalves and the tentacles of sabellid polychaetes (Nilsson, 1994). These clams and fan-worms have multiple directional photoreceptors that are assembled such that they form eye-like structures. Clams may have cup or pit eyes as well as compound eyes on the mantle edge (Fig. 5) or on tentacles around the siphons (Barber & Land, 1967; Barber et al., 1967; Adal & Morton, 1973; Nilsson, 1994; Land, 2003). These bivalves use the eyes to detect potential predators and respond by closing the shell or retracting the siphon to prevent predators biting pieces from the clam’s tissue. Sabellid polychaetes have varying numbers of compound eyes on their feeding tentacles (Fig. 5), and they are analogously used to trigger protective withdrawal into the worm’s clay tube.

The reason that these photoreceptor organs do not fit in class III is that the animals are not likely to obtain any image information from them. To the animal, it is of no consequence which of the many photoreceptors that signal a threat because the animal can only respond to it in one way that has no direction: closing the shell if it is a clam and retracting into the clay tube if it is a fan-worm. These organs do not deliver any image information, and strictly speaking, they are not true eyes. They are multiple directional photoreceptors and in that sense, they are closer to class II than to class III. But because the function is so specific and the physical characteristics differ substantially from other class II receptors, it seems motivated to form a class IIb for multiple directional photoreceptors serving an alarm function.

The concave mirror eyes of scallops (Land, 1965) contain two retinas, and only the distal one is believed to warn for predators, whereas the proximal retina may be used for habitat selection (Speiser & Johnsen, 2008) and may properly be classified as serving a class III task. Scallops differ from the other clams because they can swim by clapping their shells (Caddy, 1968), and image information may thus be extracted from the eyes.

The dispersed eyes on the dorsal side of chitons (Speiser et al., 2011) may be another example of receptors for class IIb tasks. These receptors use aragonite lenses to obtain a narrow detection angle, and they are known to elicit a protective suction response to small stimuli above the animal. Because this suction response is nondirectional, it would qualify for class IIb, but because chitons are mobile, it cannot be excluded that they also use the dispersed receptors as a large compound eye for class III tasks, in the same way as sea urchins use their dispersed photoreceptors (Blevins & Johnsen, 2004).

Class IIb photoreception does not fit into the linear succession of classes I–IV. Clearly, class IIb must have been preceded by classes I and II (Fig. 6), but continuation to class IIb would only make sense in a world with mobile predators. Because the first macroscopic mobile predators probably already had reached at least class III, it is likely that class IIb represents a later and secondary line of evolution of photoreception. It is also clear that strictly, class IIb does not fulfill the requirements for being termed spatial vision because an image need not be used. But it would be to take the definitions too far to not use the term “eyes” for the elaborate photoreceptive organs on clams and fan-worms.

Sensory tasks and animal phylogeny

The functional arguments developed in this paper suggest that class I tasks originated first, and that the higher classes evolved in succession. Different animal groups have reached different levels, and some have reverted back to lower classes of sensory tasks. With the suggested evolutionary sequence (Fig. 6), it is now possible to map out the evolutionary events on a phylogenetic tree (Fig. 7). Even though the exact outcome depends on the phylogeny, which is notoriously hard to establish, the tree reveals a pattern: the first

Fig. 5. Eyes serving class IIb tasks (optical predator alarms). The reflector-cup eyes of scallop, *Pecten maximus* (A) and the lens-less compound eyes of ark clams (B) are found in large numbers on the mantle edge. In fan worms (*Sabella melanostigma*), pairs of eyes are found at regular intervals along the feeding tentacles (C). A section through the fan-worm eyes (D) reveals ommatidia shielded by pigment tubes, and lenses that improve both photon catch and directionality (Nilsson, 1994).
classes may have evolved once or just a few times (homology based on similarities in membrane stacking, opsin family, transduction cascade and developmental genes), but higher classes originated multiple times. Because there are different families of opsin proteins with distinct transduction cascades, the full complexity of eye evolution will only be visible with the tree of Fig. 7 plotted independently for c-opsins, r-opsins, and RGR/Go -opsins (Plachetzki et al., 2007; Porter et al., 2011), but currently, the expression patterns of the different opsins are insufficiently known for most of the animal kingdom.

If the tree of Fig. 7 is compared with timing data for the branching points (Peterson et al., 2008; Erwin et al., 2011), the baffling conclusion is that nearly all fall in a period from 700 to 530 million years ago. This era, spanning 170 million years, was completed just before the first macroscopic animals started to generate the fossils we can see today. High-resolution vision may have originated during the early phase of the Cambrian explosion about 540 million years ago (see, e.g., Peterson et al., 2011), but classes I, II, and III must have gradually appeared well before the Cambrian period when animals were small, simple, and without shells or skeletons that would have left fossil remains (Hejnol & Martindale, 2008; Raff, 2008). In the rest of this paper, I will develop functional arguments as attempts to understand important events in the evolution of photoreception and eyes.

The origin and diversification of photoreception

Several studies have indicated a common origin for opsins and melatonin receptors (Fredriksson et al., 2003; Plachetzki et al., 2010; Feuda et al., 2012) and this may offer a functional explanation to the evolution of opsins. Melatonin is involved in circadian control not only in animals but also in unicellular eukaryotes (Hardeland et al., 1995). The original involvement of melatonin in mediation of photoperiodic signals is likely to be based on the fact that melatonin is irreversibly oxidized and degraded by light. With a steady synthesis, its concentration can thus be used to signal the onset of dawn and dusk. The evolution of a G-protein-coupled receptor for melatonin detection might have led to a much more efficient signaling system for melatonin levels. But because melatonin is irreversibly degraded by light, a far more efficient light receptor would be generated if melatonin were replaced by retinal with its flip-flop conformational properties. The shared origin with the melatonin receptor thus suggests that the original role of the first opsin was to monitor the daily light cycle and provide input to a circadian control, in agreement with the proposed evolutionary sequence of sensory tasks.

The phylogenetic relationship between the basic opsin-classes appears difficult to resolve (Plachetzki et al., 2007, 2010; Kozmik et al., 2008; Suga et al., 2008; Porter et al., 2011; Feuda et al., 2012; Schnitzler et al., 2012) and this may indicate that there was a rapid initial divergence, possibly reflecting a functional problem: when an opsin responds to light, its bound chromophore (retinal) changes conformation from 11-cis to all-trans. For opsins to work continuously, there must be an efficient mechanism for converting the all-trans state of retinal to its original 11-cis form. In the c-opsins of vertebrates, all-trans retinal dissociates from the opsin molecule and is reconverted to 11-cis through a complex process of transport and enzyme conversion (Kusakabe et al., 2009; Goldsmith, 2013). Insect r-opsins are instead bi-stable, such that the opsin can hold both isomers of retinal, and reconversion is accomplished by the absorption of yet another photon. Some of the RGR/Go family of opsins are reconversion enzymes that do exactly the opposite to vertebrate c-opsins. If the first opsin was similar to a c-opsin, which some opsin phylogenies suggest, then there would have been an immediate need to evolve a reconversion mechanism, and the evolution of RGR/Go opsins and r-opsins may have been evolutionary responses to this need (Nilsson, 2009). Compared to bilateria, the ctenophores and cnidarians seem to have different sets of opsin families (see references above and Mason et al., 2012; Schnitzler et al., 2012), and it might be very rewarding to investigate the photopigment regeneration mechanisms in these groups.

An important fact is that the different opsin families couple to different G-protein transduction cascades. The functional reason for this diversity is not understood. One possibility is that a switching to different transduction mechanisms happened after a gene duplication leading to opsin siblings in separate cells. Another possibility is that two sister opsins shared the same cell, and that different transduction pathways allowed for a chromatic antagonistic system similar to that discovered in a lizard parietal eye (Solessio & Engbretson, 1993; Su et al., 2006). Such a system has obvious advantages for class I tasks, such as a depth gauge because the spectral composition varies predictably with depth, and is almost independent of weather or time of day.

The association of microvilli and cilia with r- and c-opsins

Ever since Eakin’s theories on photoreceptor evolution (Eakin, 1963), the distinction between photoreceptors with microvilli or cilia has been considered of great phylogenetic significance. The calculations performed in this paper suggest that the cilia and microvilli were recruited for photoreception only after class II tasks had evolved, and photoreceptor cells had been associated with screening pigment. The presence of the two different solutions indicates that
class II tasks, and thus the first step from irradiance detection to an eye, have evolved independently at least twice.

This view is further supported by the fact that vertebrate and cnidarian eyes with ciliary photoreceptors are associated with the screening pigment melanin (Kozmik et al., 2008), whereas arthropod and other invertebrate groups with eyes based on rhabdomeric photoreceptors are associated with ommochrome pigments (Needham, 1974). It thus seems that directional photoreception, with screening pigment association and membrane stacking, evolved independently from different class I receptors with c-opsin and r-opsin, respectively.

Because there are only two options for efficient membrane stacking available, it may be wise to leave room for alternative explanations for the presence of ciliary or microvillar membrane extensions in specific cases: if it has evolved more than twice, we will still only see two solutions! There are also a handful of cases where animals inexplicably seem to have the “wrong” type of membrane stacking or both types in the same receptor, or change from one type to the other during ontogeny (Eakin & Westfall, 1965; Eakin, 1972; Eakin & Brandenberger, 1981; Blumer, 1995, 1996; Passamaneck et al., 2011; Porter et al., 2011). Importantly, eye development in bilateria and cnidaria has been demonstrated to be specified by nonhomologous control genes (Pax-6 in Bilateria, and Pax-A in Cnidaria; Suga et al., 2010), which suggests independent transition from class I to class II in these two major branches of the animal kingdom.

The evolution and fate of ciliary photoreceptors in Bilateria

As lateral eyes, the vertebrate eye is unique among bilaterians in that it has ciliated photoreceptors with c-opsin and a transduction cascade based on phosphodiesterase. Vertebrate pineal and parietal photoreceptors are also of the same type and it seems clear that these median photoreceptors share a common past with our lateral eyes (Lamb et al., 2007). Both cephalochordates and tunicate larvae have small median eyespots (class II) with ciliary photoreceptors (Kusakabe et al., 2001; Lacalli, 2004) and it seems possible that these resemble the early deuterostome condition.

Also protostomes have median ciliary photoreceptors (Platynereis: Arendt et al., 2004; Apis: Velarde et al., 2005), but these are not associated with any pigment and are indicated to be involved in control of the photoperiod (class I task). The presence of some, but very limited membrane stacking both in the pineal photoreceptors of nonmammalian vertebrates and in the homologous circadian photoreceptors of polychaetes (Arendt et al., 2004), may indicate that both share a history as median class II receptors in an urbilaterian ancestor (Fig. 8). According to the calculations presented above, membrane stacking would be unnecessary in nondirectional photoreceptors, but if these receptors evolved from a pigmented class II eyespot, similar to the homologous structures in tunicate larvae and cephalochordates, it would imply that the tendency to fold the ciliary membrane in vertebrate pineal, and Platynereis median receptors are rudiments from earlier directional tasks.
If urbilateria had a median ciliary ocellus for class II tasks, it clearly had different fates in the deuterostome and protostome lineages. In polychaetes, it has reverted from class II to class I tasks, and although it is present in some insects, (Velarde et al., 2005) it appears to be gone in Drosophila or at least is no longer photosensitive (Arendt et al., 2004). The distribution of ciliary photoreceptors in deuterostomes suggest that in the early evolution of vertebrates, the median eye separated into median and lateral eyes after which the median part lost its screening pigment and reverted to class I tasks, and the lateral parts developed spatial vision and went from class II via class III to class IV (Fig. 8).

The evolution and fate of rhabdomeric photoreceptors in bilateria

With very few exceptions, the lateral cephalic eyes or eyespots of protostome phyla have rhabdomeric photoreceptors with r-opsin and a transduction pathway based on phospholipase C. This strongly suggests that the last common protostome ancestor had paired cephalic photoreceptors of this kind. Because a significant number of protostome phyla only have directional photoreceptors (largely omitted from Fig. 7), it is most parsimonious to assume that this was the ancestral condition, although diverse screening-pigment identities suggest a common ancestor without pigment-associated photoreceptors (Vopalensky & Kozmik, 2009). Development into class III is likely to have happened several times. The different eyecup solutions, e.g., eversible cups in molluscs and inverse in flatworms, also support the view that the transition from class II to class III happened independently in different protostome groups.

A pigmented rhabdomeric eyespot is present also in cephalochordates, suggesting that the conditions of the protostome ancestor may have been present already in urbilateria. Even stronger support for that view comes from the light sensitive subset of vertebrate retinal ganglion cells that display the molecular identity of rhabdomeric photoreceptors and have neural connections homologous to those of lateral eyes in protostomes (Arendt et al., 2004). Taken together, there are strong indications that urbilateria had pigmented lateral eyespots. If these were directional class II receptors in the first protostome descendant, as argued above, it is fair to suggest that they were no more advanced in urbilateria.

Among the protostomes, the lateral rhabdomeric receptors have been retained and elaborated in different ways in different phyla. Polychaetes have two pairs of cup eyes and a third pair of pigmented lateral eyespots (Backfisch et al., 2013), and arthropods have both lateral and median eyes with rhabdomeric receptors. Duplication and subsequent functional segregation have probably been a recurring theme in eye evolution (Oakley, 2003; Arendt et al., 2004). In crustaceans and insects, the lateral eyes have advanced to serve class IV tasks, and the median photoreceptors have other roles (Fig. 8). In crustaceans, the median eyes (nauplius eyes or frontal eyes) are in most cases directional receptors, possibly working as optical statocysts, and in insects, the same eyes have become the dorsal ocelli, which are essential for flight stabilization (Srinivasan & Zhang, 2004).

Some spiders, especially salticids and lycosids (jumping spiders and wolf spiders) have developed excellent eyes for high-resolution vision in the forward direction (see Land 1985). Spiders typically have four pairs of lens eyes with different evolutionary past. Three of the pairs are termed secondary eyes and may have originated from lateral compound eyes in their aquatic ancestors. The fourth pair is termed principle eyes. It develops separately from the secondary eyes and has neural connections indicating that it is not homologous with the compound eyes of their ancestors (Blest, 1985; Land, 1985). Interestingly, salticids have recruited the principal eyes for class IV vision, whereas lycosids have recruited the frontal-most pair of the secondary eyes for the same purpose. Although both the principal and the secondary eyes of spiders are rhabdomeric, their ontogeny implies that class IV vision has evolved independently in salticids and lycosids.

Throughout the animal kingdom, except in vertebrates, it seems that rhabdomeric receptors have all the advanced roles, from class II to class IV, leaving only class I to ciliary receptors. The

---

**Fig. 8.** Schematic view of the head region of bilateria indicating the evolutionary shift of position and function of photoreceptor organs argued for in the text. The photoreceptor/opsin types are denoted by (r) for rhabdomeric/r-opsin and (c) for ciliary/c-opsin. Note that the vertebrate lateral eyes are composite structures formed by fusion of components from different photoreceptor organs in Ur bilateria. It is also possible that vertebrate ancestors went through a stage with only median eyes (Vopalensky et al., 2012). Polychaetes have only made functional shifts and duplicated their lateral rhabdomeric eyes, whereas crustaceans/insects have formed a new median eye from rhabdomeric photoreceptors.
rhabdomeric receptors seem to have been novelties that arose in early bilaterians or in a bilaterian–cnidian ancestor (Feuda et al., 2012; Mason et al., 2012), and it may be that they took over functions from older ciliary receptor because rhabdomeric receptors have a better signal-to-noise ratio, faster signaling, and a larger dynamic range (Fain et al., 2010). But r-opsins are not only expressed in eyes. In polychaetes, r-opsins are expressed both in the ventral nerve cord and in the parapodia (Backfisch et al., 2013).

The fate of lateral rhabdomeric eyes in deuterostomes is very different. Cephalochordates and urochordate larvae, only have small median eyespots (Kusakabe et al., 2001; Lacalli, 2004; Vopalsensky et al., 2012), and in echinoderms, rhabdomeric receptors are found in the optic cushions of starfish (Penn & Alexander, 1980), and in the tube feet of sea urchins (Lesser et al., 2011; Ullrich-Lüter et al., 2011). The most remarkable fate of rhabdomeric receptors is clearly that in vertebrates (Fig 4), where the rhabdomeric photoreceptors appear to have merged with ciliary ocelli to form a composite organ performing class IV tasks with the ciliary receptors and class I tasks with the rhabdomeric receptors (Arendt et al., 2004).

The employment of Pax-6 genes for controlling eye development in both protostome and vertebrate lateral eyes has been taken as evidence that eyes only evolved once (Gehring & Ikeo, 1999; Gehring, 2005, 2011). The more complex scenario proposed here and by Arendt et al. (2004) and Arendt (2008) is also compatible with shared Pax-6 expression (Vopalsensky & Kozmik, 2009). The origin of this unique arrangement, where a ciliary eyespot got connected to the brain through interneurons that once were receptors in a rhabdomeric eyespot, may seem remarkable (Vopalsenský et al., 2012). But in a small ancestor, with only few cells in each type of eyespot, and short distances in a simple nervous system, the novel association may have been rather straightforward.

Even if Pax-6 may have been responsible for the development of lateral rhabdomeric eyespots in urbilateria, it does not mean that r-opsins must have originated in the same process. Non-cephalic and Pax-6-independent expression of r-opsins in Platynereis (Backfisch et al., 2013) suggests that r-opsins existed before Pax-6-specified photoreceptors originated in our common bilaterian ancestor. Alternatively, Pax-6 dependence predates the split between r-opsins and c-opsins, but after organ duplication and diversification, it was retained only for some photoreceptors (Vopalsensky & Kozmik, 2009).

### Relationship between cnidarian and bilaterian photoreceptors

The majority of recent opsin phylogenies indicate that the cnidarians split from bilaterians before the three opsin-families c-opsin, r-opsin, and RGR/G, diverged, but two recent studies suggest that also cnidarians have similar opsin families (Feuda et al., 2012; Mason et al., 2012). Cnidarian visual opsins are most similar to bilaterian c-opsins, and in the eyes of jellyfish, the photoreceptors carry ciliary membrane extensions. This of course raises the question if the ciliary eyespots assumed to be present in urbilateria share a common ancestry with cnidian-pigmented photoreceptors. To approach this question, it is first necessary to consider the possible body plan of the last common ancestor of bilateria and cnidaria. This is likely to have been an organism simpler than the urbilateria and probably at a planuloid level of organization (Raff, 2008; Hejnol & Martindale, 2008). The class III eyes of some modern cnidarian medusae may be of little significance here if the medusoid mode of life developed in cnidaria after the split from bilateria. Cnidarian polyps express opsin (Plachetzki et al., 2007) but not in association with screening pigment. Ctenophora also express opsins without any association with screening pigment, so the ancestor to cnidarians, ctenophorans, and bilaterians must clearly have had class I photoreception.

Of more evolutionary interest are the single-cell eyespots of cubozoan larvae. These planula larvae carry some 10–20 eyespots randomly spread around the mid or rear part of the body. These cells have a single motile cilium surrounded by microvilli protruding from the inside of a pigmented cup (Fig. 4A; Nordström et al., 2003). The arrangement of this directional photoreceptor is unique in the animal kingdom. It is not at all unlikely that the urbilaterian ciliary ocellus evolved by modification from a shared ancestor with eyespots like those of cubozoan larvae (Arendt et al., 2009), but there is yet no convincing evidence for such a hypothesis. The fact that development of cnidarian and bilaterian eyes is controlled by nonhomologous Pax-genes (Suga et al., 2010), instead lends support for independent transitions from class I to class II photoreception in cnidarians and bilaterians. Until we know more about the molecular components of cubozoan larval eyespots, it will remain an open question whether the transition from class I to class II tasks occurred independently in cnidaria and bilateria, or if directional photoreception was a feature of our common ancestor.

### The origin of class IIb photoreceptors and eyes

The alarm photoreceptors of clams, fan worms, and chitons detect moving predators. They can be interpreted as high-performance versions of shadow detectors and are able to respond to distant movements even if no direct shadow is cast on the receptors (Nilsson, 1994; Land, 2003; Speiser et al., 2011). It is also very likely that class IIb tasks evolved to improve shadow responses. Because the directionality of individual receptors may be as narrow as 10°, multiple receptors are necessary to cover all possible directions. The evolutionary origin would thus be from fast shadow detectors (class I) via directional shadow receptors (class II) to multiple directional receptors (class IIb).

A continuation to low-resolution vision is a possibility, but it would require substantial expansion of the neural processing. The mantle eyes of stationary clams (ark clams and giant clams) and the tentacular eyes of fan worms are not likely to evolve into true (imaging) eyes because the animals are not able to generate any directional responses. But scallops can swim by clapping the shells, and their reflector-cup eyes are thought to be involved in habitat selection using true image vision (a class III task), in addition to acting as optical alarm systems (Land, 1966). The dispersed eyes on the dorsal plates of chitons may also have a dual role for habitat selection and predator alarm and would thus have made the transition from class IIb to class III.

Both from the general design and from the photoreceptor-cell morphology, these clam, fan worm, and chiton eyes are unconventional. Their positions, on the mantle of clams, the tentacles of fan worms, or the dorsal plates of chitons, are not homologous to the placement of true eyes or directional eyespots. Some of these eyes have rhabdomeric receptors, i.e., the cup eyes of ark clams, one of two retinal layers in the reflector-cup eyes of scallops and the dispersed ocelli of chitons, but the majority have ciliated receptors, i.e., the compound eyes of ark clams and fan worms (Fig. 5), the cup eyes of giant clams, cockles, and the other retinal layer in scallop eyes. The opsin identity or transduction pathway is known only for the mantle eyes of scallops. Here, the ciliary receptors use the alarm photoreceptors of clams, fan worms, and chitons detect moving predators. They can be interpreted as high-performance versions of shadow detectors and are able to respond to distant movements even if no direct shadow is cast on the receptors (Nilsson, 1994; Land, 2003; Speiser et al., 2011). It is also very likely that class IIb tasks evolved to improve shadow responses. Because the directionality of individual receptors may be as narrow as 10°, multiple receptors are necessary to cover all possible directions. The evolutionary origin would thus be from fast shadow detectors (class I) via directional shadow receptors (class II) to multiple directional receptors (class IIb).

A continuation to low-resolution vision is a possibility, but it would require substantial expansion of the neural processing. The mantle eyes of stationary clams (ark clams and giant clams) and the tentacular eyes of fan worms are not likely to evolve into true (imaging) eyes because the animals are not able to generate any directional responses. But scallops can swim by clapping the shells, and their reflector-cup eyes are thought to be involved in habitat selection using true image vision (a class III task), in addition to acting as optical alarm systems (Land, 1966). The dispersed eyes on the dorsal plates of chitons may also have a dual role for habitat selection and predator alarm and would thus have made the transition from class IIb to class III.

Both from the general design and from the photoreceptor-cell morphology, these clam, fan worm, and chiton eyes are unconventional. Their positions, on the mantle of clams, the tentacles of fan worms, or the dorsal plates of chitons, are not homologous to the placement of true eyes or directional eyespots. Some of these eyes have rhabdomeric receptors, i.e., the cup eyes of ark clams, one of two retinal layers in the reflector-cup eyes of scallops and the dispersed ocelli of chitons, but the majority have ciliated receptors, i.e., the compound eyes of ark clams and fan worms (Fig. 5), the cup eyes of giant clams, cockles, and the other retinal layer in scallop eyes. The opsin identity or transduction pathway is known only for the mantle eyes of scallops. Here, the ciliary receptors use
an opsin of the RGR/Go family (Kojima et al., 1997), which is in notable contrast to ciliary cephalic eyes or photoreceptors. Both the placement and the receptor types strongly indicate that these eyes are not standard cephalic eyes moved to a new location. It is more likely that they have evolved by co-option of a mosaic of different components. If this is correct, the alarm photoreceptors alone will account for several cases of independently evolved eyes, especially, if multiple directional photoreceptors are classified as eyes.

Conclusions

In this paper, I have tried to demonstrate that a functional analysis is essential for understanding eye evolution. Together with existing molecular and morphological data, functional arguments has yielded the following main conclusions:

- A classification into four classes of photoreceptive tasks (nondirectional photoreception, directional photoreception, multiple directional photoreception, low resolution vision, and high resolution vision) condenses the otherwise nebulous plethora of behaviors and photoreceptor organs into manageable units that facilitate an evolutionary analysis.
- The four classes correlate with major innovations in eye evolution: light detection, light screening, membrane stacking, and focusing optics.
- For each evolutionary transition to a higher class, the rate of information delivered to the nervous system increases by several orders of magnitude. This provides opportunities for more advanced behaviors and requires larger brains.
- Visual systems evolve through sequential acquisition of new and gradually more demanding tasks for new behaviors.
- A synthesis of available data suggests that urbilateria had a simple set of rhabdomeric and ciliary eyespots for directional photoreception, and that this has developed and diversified by organ duplication, positional shifts, task shifts, and losses in different ways in different branches of the animal tree.
- Opsin photoreception evolved once, and its original task was nondirectional monitoring of irradiance. This was followed by a few independent instances of transition to directional photoreception, and from there, numerous transitions to low-resolution vision, and finally, a smaller number of transitions to high-resolution vision. The scenario is compatible with observed Pax-gene expression in eyes of cnidarians, protostomes, and vertebrates.
- Secondarily, low-resolution vision has evolved as dispersed photoreceptors in echnodermns and possibly also in chitonns.
- Eye-like structures for multiple directional photoreception for predator alarm has evolved on the mantle of bivalves and on the tentacles of sabellid polychaetes. In the case of scallops, this predator alarm has evolved on the mantle of bivalves and on the tentacles of sabellid polychaetes. In the case of scallops, this has evolved into low-resolution vision.

A major difficulty in reconstructing evolutionary events, especially those dating very far back in time, is that they must be based on correct phylogenetic relationships between animal taxa. Unfortunately, this is a persistent problem, especially concerning the position of key taxa such as aplacozoans, cnidarians, ctenophores, and accoel worms. Another limitation is that most of eye evolution was completed just before the first animal fossils were formed. The latter problem, however, also carries important information. The lack of fossils implies that the animals in which much of eye evolution took place were small and soft bodied. For this reason, it can be rewarding to study eyes, photoreceptors, nervous systems, and behavior in the extant groups that are also small and soft bodied, such as flatworms, accoels, and various larval forms.

Eye evolution is an integral part of animal evolution, and the development of vision together with locomotion may together have been the single most important process in the evolution of macroscopic animals and higher trophic levels in ecological systems.

Acknowledgments

This work was supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. I am grateful to Eric Warrant for valuable discussions, and to John Endler and Todd Oakley and for helpful comments on the manuscript.

References


Eye evolution and its functional basis


