Adaptation to nocturnality – learning from avian genomes

Diana Le Duc[1)2)*] *and Torsten Schöneberg*[1)]

The recent availability of multiple avian genomes has laid the foundation for a huge variety of comparative genomics analyses including scans for changes and signatures of selection that arose from adaptions to new ecological niches. Nocturnal adaptation in birds, unlike in mammals, is comparatively recent, a fact that makes birds good candidates for identifying early genetic changes that support adaptation to dim-light environments. In this review, we give examples of comparative genomics analyses that could shed light on mechanisms of adaptation to nocturnality. We present advantages and disadvantages of both ''data-driven'' and ''hypothesis-driven'' approaches that lead to the discovery of candidate genes and genetic changes promoting nocturnality. We anticipate that the accessibility of multiple genomes from the Genome 10K Project will allow a better understanding of evolutionary mechanisms and adaptation in general.

Keywords:
adaptation; genome sequencing; nocturnality

Introduction

Since their first availability about 10 years ago, the next-generation sequencing (NGS) technologies have been shown to be powerful tools in many fields of genetics and evolutionary biology. A central question in biology concerns the genetic basis of traits that evolve in response to environmental factors. There are a number of traits that evolved independently in different species, and these may offer a unique opportunity to identify genes underlying convergent phenotypes. Nocturnality is an obvious trait, that, according to current phylogenies, is thought to have evolved multiple times in birds, and species from a number of distinct taxa are either fully or partially nocturnal [1, 2].

Modifications of the sensory system are among the most common changes that occur when shifting from a diurnal to a nocturnal lifestyle. Visual adaptation may lead to either enhanced visual sensitivity, or to a decreased reliance on vision, depending on how adaptation to the niche occurs. The reduction of the visual system is often accompanied by the enhancement of other sensory systems such as olfaction, tactile senses, and hearing [3–5]. Understanding the molecular basis underlying the biochronology and the unusual state of nocturnal adaptation of birds will allow for a better overview of the mechanisms involved. Moreover, this can shed light on the reverse mechanism, namely mammalian adaptation to the diurnal niche after escaping the "nocturnal bottleneck". According to the "nocturnal bottleneck" hypothesis during the Mesozoic era eutherian mammals competed with diurnal reptiles (e.g. dinosaurs); once the dinosaurs went extinct, mammals diversified and started occupying the diurnal niche, which lead to sensorial adaptations, yet without erasing completely hallmarks of the nocturnal evolution time [6].

Another interesting aspect of nocturnality is that it might be an example to address the question whether "genes follow behavior" or "behavior follows genes" in evolution. One can contrast the paradigm of the modern evolutionary synthesis (neo-Darwinian synthesis), which puts the emphasis entirely on genes driving the organisms' evolution, and speculate that a shift to activity in the dark may have come first, and then random genetic changes occurred, some of which provided physiological/morphological support for adaptation to the niche.

The advent of NGS technologies with their reduced sequencing cost and high throughput made possible whole-genome sequencing and de novo assembly. In 2014, an initial set of 48 bird genomes were released with the main aim of building the Neoaves phylogeny [7]. In 2015, the "Bird 10,000 Genomes Project" announced the intention to generate

DOI 10.1002/bies.201600006

[1)] Medical Faculty, Rudolf Schönheimer Institute of Biochemistry, University of Leipzig, Leipzig, Germany
[2)] Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

***Corresponding author:**
Diana Le Duc
E-mail: diana_leduc@eva.mpg.de

representative draft genome sequences from all extant bird species within the next 5 years [8].

In this essay, we will discuss the scientific potential of such genomic adventures and the current opportunities and limitations of comparative genome analyses. We will use nocturnal adaptation in birds as an example to illustrate the standard comparative genome analyses that aim to identify the selective pressures and targets of selection in genomes.

Genomic changes accompany adaptation to nocturnality

With the advent of NGS and genome assembly of non-model organisms, we now have an unprecedented opportunity to supplement the "hypothesis-driven" approach with the "data-driven" inquiry of genomic changes, which are responsible for adaptation. While NGS does not require a prior known paradigm to deliver gene candidates involved in shaping a trait, this cannot completely replace the traditional routes of discovery. The main reason for this is that in the absence of – for example – a strong genomic signal of selection, one can miss important genes. The "hypothesis-driven" inquiry of genes is based on their functional involvement in senses/biological systems shown to be affected by nocturnality, e.g. vision, olfaction, tactile sense, hearing, and biorhythm. Nevertheless, the "hypothesis-driven approach" is limited only to previously described changes, and thus lacks the ability to uncover new mechanisms and genes potentially related to nocturnal adaptation. Hence, the use of both approaches can lead to a more comprehensive description of genetic changes underlying the phenotype.

Comparative genomic analyses have benefits and challenges

Since the genome determines a multitude of phenotypes that an organism possesses, the availability of genome sequences from multiple species provides an unprecedented opportunity to discover which genomic changes

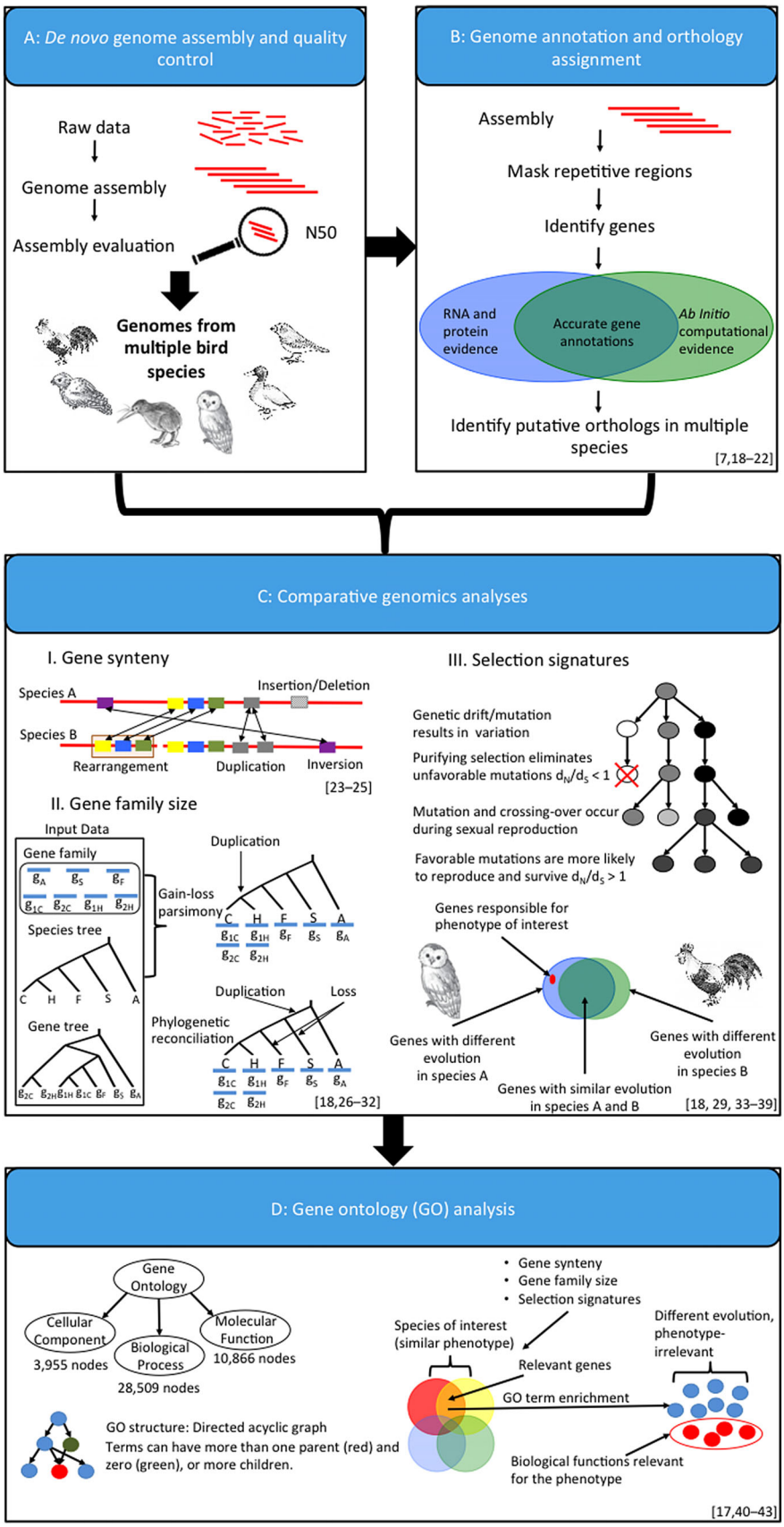


Figure 1. Continued.

Figure 1. A typical comparative genomic analyses workflow. **A:** After generating the raw data, usually by NGS, the first goal is to produce a reasonably contiguous genome assembly. The most common metrics for assembly evaluation is N50 (N50 is defined as the scaffold length such that using equal or longer scaffolds produces half the bases of the assembly. That is the N50 value denotes that 50% of the entire assembly length is contained in scaffolds that have an equal or bigger length than N50), which measures assembly contiguity [13]. Most currently assembled genomes are draft assemblies, and "high-quality assemblies" are considered to be at least 90% complete [14]. The "Bird 10,000 Genomes Project" intends to assemble representative draft genomes from all extant bird species by 2020 [8] – available bird genomes (Fig. 2). **B:** Once a reasonably contiguous genome assembly has been produced the next step is to identify important elements such as coding and non-coding sequence regions, and to determine their functions. The completeness and contiguity of a genome assembly, as well as the genome annotation availability of closely related species, are important features that influence the extent to which the complete annotation of a new genome is possible. For a detailed assessment of genome contiguity and annotation quality, we refer the reader to the review of Yandell and Ence [15]. The figure briefly shows common annotation pipeline steps. See Fig. 3 for a hypothetical evolutionary history of a gene. **C:** I. Synteny and chromosome assignment of bird genes are usually identified by whole genome alignment to the chicken and zebra finch genomes. Unpreserved synteny could be functionally relevant for the different phenotypes in different species. II. Image after Capra et al. [16]: in a simple model of gain-loss method, only one duplication event would be inferred. However, reconciliation with the gene tree assumes one duplication event and two loss events. The branch where duplication/loss occurs is very important for comparative genomics analyses. III. Natural selection allows organisms better adapted to the environment to survive and reproduce more effectively. Variation occurs by random drift and unfavorable mutations are eliminated over time, while advantageous ones will finally prevail in the population. Codon-based methods for detecting selective forces are prevalent in comparative genomics. If on one lineage, a gene possesses more non-synonymous substitutions (changes the encoded amino acid) per non-synonymous site (d_N) than synonymous substitutions (preserved amino acid) per synonymous site (d_S), then d_N/d_S ratio is >1, suggestive of advantageous mutations/positive selection. Conversely, if d_N/d_S ratio is <1, mutations are most probably unfavorable and will be eliminated by purifying selection. When comparing multiple species, genes with different selection signatures may be responsible for the different phenotypes. However, only a very small fraction of those different genes contributes to the phenotype of interest. **D:** The gene ontology (GO) classifies gene products in three parent domains: cellular component, the parts of a cell or its extracellular environment; molecular function, the elementary activities of a gene product at the molecular level, such as binding or catalysis; and biological process, operations or sets of molecular events with a defined beginning and end, pertinent to the functioning of integrated living units: cells, tissues, organs, and organisms [17]. GO is structured as a directed acyclic graph such that each term is related to one or more other parent nodes, and sometimes to more children nodes. If multiple species possess a phenotype of interest (e.g. nocturnality), while others do not, genes that show signals of different evolution will cluster in similar GO categories. However, while the analysis outputs multiple candidate pathways, only a part of these GO categories will be truly responsible for the evolution of the phenotype under study. Additional readings are given in brackets in each panel.

underlie particular phenotypic changes between species. Here, we will not go into details of the different NGS technologies or genome assembly tools, which represent the data basis for such approaches, but rather refer to recent reviews [9–12]. To answer questions related to genotype-phenotype association, one has to correlate genetic changes with the phenotype of interest (e.g. nocturnality) across a phylogeny. The major advantage of this method is that it can uncover genes that are candidates for influencing the inquired phenotype, without any prior knowledge about their functionality. This implies discovery of novel players involved in functional shaping of the organism.

The typical workflow for comparative genomics is shown in Fig. 1. This workflow only focuses on a subset of the analyses that may contribute to understanding the genomic changes related to nocturnal adaptation.

Once a genome has been assembled and orthologous sequences are assigned, comparative genomic analyses are used to identify gene and gene family expansion (gain) and contraction (loss) and signatures of selections (Fig. 1). To detect changes that are relevant for adaptation to nocturnality, one can make use of already assembled bird genomes [7, 18, 19, 29, 44–48] and compare nocturnal versus diurnal species (Fig. 2). A more comprehensive bird phylogeny [49] could be very useful to decide which bird genomes should be

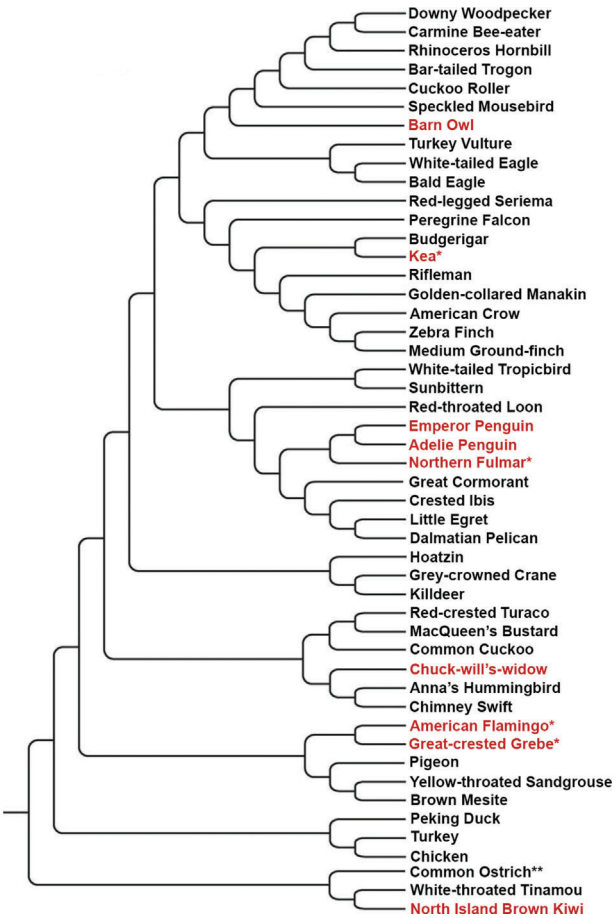


Figure 2. Phylogenetic tree of the avian genomes sequenced to date. The tree is adapted from studies by Jarvis et al. [7] and Zhang et al. [48] and additionally includes the North Island Brown Kiwi [18]. Species in red are nocturnal; * denotes semi-nocturnality; ** the position of the ostrich in the *Palaeognathae* clade varies according to which genome loci are considered [18].

additionally sequenced to reduce false positive results. The result of such analyses (comparative genomics and gene ontology analyses) is typically a list of processes or functions (for example metabolic pathways) that are significantly associated with, for example, nocturnality. For example, many comparative genomics studies have revealed a higher diversification of the olfactory receptors repertoire [18, 32, 50]. Although this is an expected biological trait, the unbiased result of comparative genomics offers further support to the importance of olfaction in low-light environments. Nevertheless, we caution, that in the case when such testing was realized without an a priori hypothesis, the different evolution of genes cannot be validated because results make sense a posteriori. Such an approach can easily lead to over-interpretation of the results, and storytelling [43]. The outcome of such analysis should be rather that any enriched pathway, whether it makes biological sense or not, ought to be considered as a potential candidate.

However, the statistical power of comparative genomics is limited mostly by the low number of available species. Also, the lineage-specific biology is lost if the time-scale of the relationship of species being studied is widened [51]. Intuitively, the complexity and difficulty of the between-species phenotype-genotype associations resides in the fact that genomes differ by millions of changes. Moreover, differences between species are not at the level of just one trait, such that the presumably complex genetics of the inquired phenotype (nocturnality) and the very large search space of bird genomes impose great limitations for comparative genetics. Hence, to avoid over-interpretation of comparative genomics output [43], it is crucial to experimentally follow up on the generated hypotheses. Another approach would be limiting the search space to genes known to play a role in the development of the trait of interest – nocturnality. This is the "hypothesis-driven" approach and will be discussed in the following section.

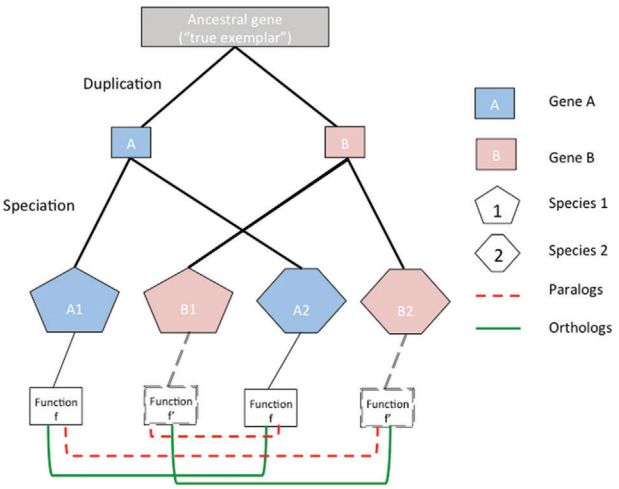


Figure 3. Hypothetical evolutionary history of a gene. Orthologs and paralogs are two different types of homologous genes that differ in the way they arose. Orthologs are genes in different species that evolved from a common ancestral gene given a speciation event, while paralogs evolved by duplication in the same genome. The orthology among species can result in a many-to-many relationship where an ortholog group is formed by groups of paralogs in each genome. However, for the set of paralogs in a genome each copy is the direct descendant of an ancestral gene. The ancestral gene ("true exemplar") experienced a duplication event, which resulted in genes A and B (paralogs). After speciation, the resulting species 1 contained genes A1 and B1, while species 2 contained genes A2 and B2. All genes are homologous to one another, and A genes are paralogs of B genes, while 1 genes are orthologs of 2 genes.

Adaptation to nocturnality shapes senses

The "hypothesis-driven" approach, unlike comparative genomics – "data-driven" approach, relies on prior knowledge about the known or presumed physiology of the phenotype under investigation. The rationale of the "hypothesis-driven" approach considers that the quantitative genetic variation of the phenotype under investigation is determined by functional mutation of putative genes (Fig. 3). The main advantage of such a method is the reduction of the search space for genetic changes responsible for the phenotype of interest. This approach can also prove to be a more cost-effective method than comparative genomics. However, the greatest limit of the method is the reliance on a priori knowledge about the molecular aspects of the phenotype. Unfortunately, the detailed genetic architecture of most biological traits is very limited [52]. Moreover, the candidate-gene approach has been criticized for non-replication of results in follow-up subsequent association studies [53].

The limitations of the "hypothesis-driven" approach can at least partially be overcome by the combined use with comparative genomics. This strategy makes use of cross-species comparisons to characterize the effect of multiple putative candidate genes, for which no a priori information is available [54].

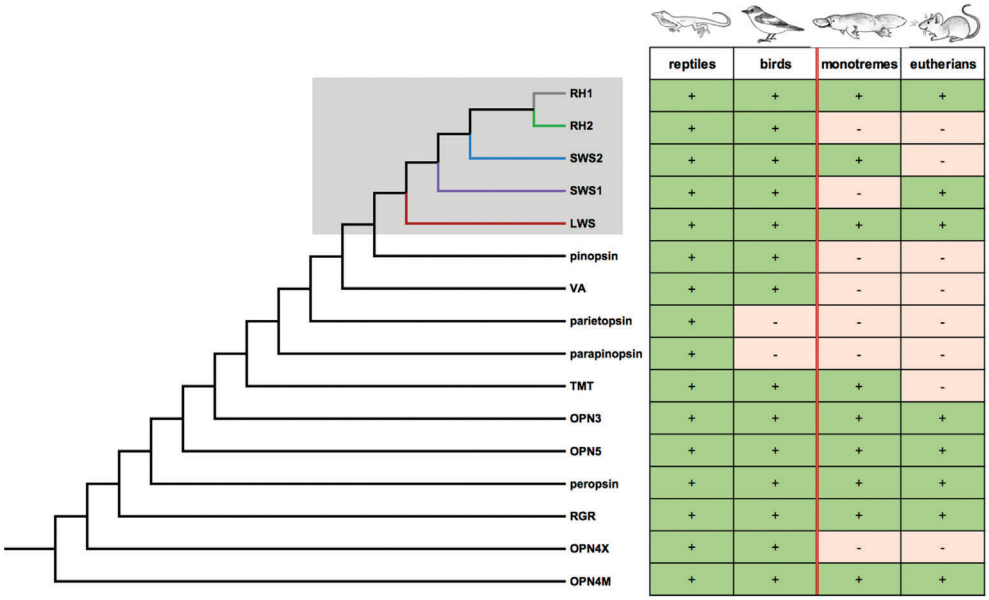
In a classical view, the "hypothesis-driven" inquiry of genes is based on their functional involvement in senses/biological systems shown to be affected by nocturnality, e.g. vision, olfaction, tactile sense, hearing, and biorhythm.

Changes in vision-related genes can be markers for nocturnality

Nocturnal animals generally have well-developed senses that allow adaptation to low-light (scotopic) conditions. According to which environmental niche the animal belongs, to support behaviors like e.g. feeding, mating, predation, some of the essential senses might be better developed while the superfluous ones tend to be regressed.

The intensity and wavelength of light available for vision act as selective forces on the evolution of the vertebrate visual system such that the size and shape of the eyes exhibit high variability according to the activity pattern [4, 6, 55, 56]. Although eye anatomy seems to be an excellent indicator of diurnality versus nocturnailty [4, 6], it is very difficult to depict the genetic cause that leads to these morphological differences of the eye.

In contrast, genes responsible for color vision have been very well characterized. Cone and rod photoreceptors contain photopigments, which are represented by an opsin linked to a



reptiles	birds	monotremes	eutherians
+	+	+	+
+	+	-	-
+	+	+	-
+	+	-	+
+	+	+	+
+	+	-	-
+	+	-	-
+	-	-	-
+	-	-	-
+	+	+	-
+	+	+	+
+	+	+	+
+	+	+	+
+	+	+	+
+	+	-	-
+	+	+	+

Figure 4. Phylogenetic tree of visual and non-visual opsins (adapted from Ref. [6]). Branches to visual opsins (highlighted in gray) are colored according to the mean spectral sensitivity of the opsin. The summary table shows the presence of orthologs in non-mammals (reptiles and birds) and mammals (monotremes and eutherians). LWS, long-wavelength-sensitive opsin; OPN3, panopsin/encephalopsin; OPN5, neuropsin; OPN4M, mammalian-like melanopsin; OPN4X, xenopus-like melanopsin; RGR, retinal G-protein-coupled receptor; RH1, middle-wavelength-sensitive rhodopsin 1 (rod); RH2, middle-wavelength-sensitive rhodopsin 2 (cone); SWS1, short-wavelength-sensitive opsin 1; SWS2, short-wavelength-sensitive opsin 2; TMT, teleost multiple tissue opsin; VA, vertebrate ancient opsin.

vitamin A-derived light-sensitive retinal chromophore. Rod photoreceptors are specialized for dim light vision, whereas cones function in daylight and are responsible for color vision. To see colors, at least two spectrally distinct classes of cone photoreceptors must be present. There are five subtypes of visual pigments: a single rod opsin (RH1) and four cone subtypes (SWS1, SWS2, RH2, and LWS) that detect wavelengths in the range from UV to near-infrared [6, 57] (Fig. 4). Phylogenetic analysis of the gene sequence identity revealed that cone pigments preceded the rod pigment evolution [58]. The sequence identity of cone opsins is around 40%. Conversely, the RH2 (cone) and RH1 (rod) opsins show an identity of around 80%, which supports a more recent split of the *Rh1* and *Rh2* genes, with *Rh1* arising from an ancestral duplication of the *Rh2* cone opsin gene [58, 59] (Fig. 4).

Mammals display only two classes of cone pigment genes [59]. Thus, first the *Rh2* gene was lost, such that the common ancestor of all mammals had only three other types of cone pigments (trichromacy) [60]. Next, *Sws1* and *Sws2* genes have become inactivated in monotremes and eutherians/marsupials, respectively [6, 58, 59] (Fig. 4). This is believed to be related to their early evolution about 150–200 million years ago, when they were undergoing the nocturnal bottleneck [6, 56, 59]. It seems thus very plausible, that in the absence of specific light impulses, given a relaxed selection constraint cone pigments-coding genes tend to be lost [61]. This would be a true example where the "fate of a gene follows behavior" (examples see below).

On the other hand, birds, which have always been evolving in diurnal environments, have generally retained all four classes of cone visual pigments [62], which provide for tetrachromacy. However, little is known about the cone pigments in nocturnal bird species [59]. In the few studied species, the retina is dominated by rods (80–90%), compared with only 20–30% in diurnal species [63, 64]. In the nocturnal tawny owl, *Strix aluco*, three classes of cones have been found (*Lws*, *Rh2*, and *Sws2*), suggesting a nocturnal-determined loss of genes [63]. This is probably not common to nocturnal bird species, because in another nocturnal species, the kiwi (*Apteryx mantelli*) deleterious mutations were observed in both, *Rh2* and *Sws1*.

Features of nocturnality do not always depend on a classical day/night cycle. The aquatic lifestyle and marked seasonal changes in the length of daylight – e.g. at the Arctic and Antarctica circle – could affect the visual abilities as well as the non-visual phototransduction. Pseudogenizations of *Rh2* and the pinopsin gene were found in the Adélie and Emperor penguins [65].

We lack the means to test whether genomic changes have occurred as a consequence of the nocturnal bottleneck in the evolution of mammals. However, by understanding how nocturnal adaptation of birds shapes genes we can have a more informed opinion on the reverse mechanism, namely the mammalian adaptation to the diurnal niche after escaping the bottleneck.

To this end, it is interesting to detect selective pressures on opsins using methods as described above. It is also interesting to detect mutations that led to loss-of-function of these pigments. This approach can be facilitated by multiple sequence alignment of the genes from nocturnal and diurnal birds. It was previously shown that comparative sequence data from orthologs are suitable to predict the functional relevance of mutations in a model protein [66], such that

amino acid positions, which are conserved across diurnal species, but vary in the nocturnal birds, might be functionally relevant.

Nocturnality is accompanied by high diversity in olfaction-related genes

Because visual input is rather limited under nocturnality increased sensitivity of other senses, such as smell, may become more relevant [67] for finding food or for locating danger or a mate. Olfactory receptors (ORs) mediate the detection of odorants by vertebrates [68]. It is believed that the total number of OR genes and the proportion of intact OR genes in a genome are indicative for how developed the sense of smell is in that particular organism [69]. OR genes have evolved by multiple tandem duplications [70] and display a large number of pseudogenes [71], which complicate the annotation process of these genes [18]. Given the large number of similar genes, *de Bruijn* graph assemblers [72–74] tend to overcollapse these regions and the final result may be a lower number of assembled genes than that which truly exist. Moreover, for an NGS de novo assembled genome, the publically deposited data are usually the assembly, and not the raw sequences, such that a posteriori estimation of the collapsed regions [18] is a tedious task. We have shown that even for well-annotated, Sanger-sequenced genomes, like the chicken, revisions of the annotation can lead to changes in the total number of ORs and hence in the proportion of intact genes [18].

A wider range of odors is thought to be detected in the presence of higher genetic variance of the OR family in a species [75]. Thus, the sequence variation of the OR genes could be a good indicator for the olfactory abilities in an organism. We proposed using Shannon entropy (H) to estimate the diversity of ORs [18] because H is considered a sensitive tool for estimating the diversity in a system [76, 77].

To infer the impact of nocturnality on the evolution of OR genes in birds, one would first need to annotate the receptors. However, this is a more demanding task than just considering the de novo annotations. To ensure a better annotation, one would need to perform additional steps including: (i) identifying sequences that contain the Pfam [78] 7tm_4 domain (olfactory domain); (ii) ideally, using sequences of ORs from well-annotated bird genomes like chicken and zebra finch, one can build hidden Markov models profiles to further search the proteome and retrieve non-redundant hits in the organisms of interest. After performing the annotation step, a multiple alignment of the translated OR sequences can be used to infer the variation at each amino acid position and to detect the diversity in the OR system in nocturnal versus diurnal birds.

Previous studies have shown that nocturnality in birds is well-correlated with the size of the olfactory bulbs [5]. Interestingly, the total number of ORs is correlated with the olfactory bulb size, but not the proportion of intact genes [79]. Since we have shown that technically the annotation of ORs with the current state of the genome assemblies is prone to errors [18], we believe that considering the diversity of the system, rather than the absolute number of genes, is a more reliable measure for olfactory acuity. This was shown in one nocturnal species, *Apteryx mantelli* [18], but we expect that the availability of more nocturnal bird genomes will allow a better overview on the impact that nocturnality has on the diversity of ORs. The ORs are ideal candidates for speculation that "behavior follows genes." Given the genetic diversity and dynamics of ORs, a bird may find prey just by chance, which will influence its behavior and the environmental niche to which the bird adapts.

Nocturnal animals may have a developed tactile sense

Nocturnal animals also frequently develop an enhanced tactile sense, which may prove a useful tool for foraging. The diving ducks (*Aythya*, mainly *A. nyroca* [80] and *A. fuligula* [81]) are predominantly nocturnal in the foraging activity and find their food by touch at the bottom of lakes [82]. Probe-foraging birds often rely on other senses than vision for prey-detection. The neognathous shorebirds *Scolopacidae*, a family of probing birds, detect their buried prey using specialized vibration and pressure-sensitive mechanoreceptors which form a honeycomb of sensory pits in the bill-tip. Their feeding activity follows the tides as much as the day/night cycle resulting in activity at night if the tides are right. While these birds are diurnal, a scolopacid-type bill-tip organ was described in the nocturnal *Apteryx mantelli*, suggesting their reliance not only on olfaction, but also on touch for their prey-detection [3].

While intuitively the tactile sense should be more developed in nocturnal animals, studies showing a clear correlation between nocturnality and touch are currently lacking, to our knowledge. In the reverse hypothesis – that of the mammal "nocturnal bottleneck" – tactile vibrissae (whiskers) are supposed to have evolved as a result of nocturnality [6, 83]. Whereas morphological description of the mechanoreceptors seems easier to be performed, analyzing the sense of touch from a genetic perspective proves to be more difficult. The reason for this is that even in simpler model organisms, like *Drosophila melanogaster*, mechanoreceptors' development is regulated by a complicated gene network [84]. Thus, understanding the molecular mechanism of touch in nocturnal and diurnal birds may be unrealistic at the current state of our knowledge.

We propose making use of the GO annotations to investigate the nocturnal-driven adaptation of touch. To this end, there are two GO terms associated with touch: "sensory perception of touch" (GO:0050975) and "detection of mechanical stimulus involved in sensory perception of touch" (GO:0050976) (Table 1). While experimental studies are still limited for the genes of these GOs, it would be interesting to check signals of selection acting on the genes from nocturnal versus diurnal birds. Also, mutations that occur at certain amino acid positions with higher frequency under nocturnality may be informative for developing future functional studies.

Nocturnality can lead to better hearing in birds

Another trait shaped by nocturnality in mammals is high-frequency hearing [6, 85]. In the presence of other less-developed traits, hearing can facilitate locating the prey or

Table 1. GO categories relevant for testing genes involved in touch and hearing

GO annotation	GO term	No. of genes in node
GO categories related to touch		
Sensory perception of touch	GO:0050975	14
Detection of mechanical stimulus involved in sensory perception of touch	GO:0050976	9
GO categories related to hearing		
Auditory receptor cell morphogenesis	GO:0002093	5
Sensory perception of sound	GO:0007605	140
Auditory receptor cell fate commitment	GO:0009912	3
Response to auditory stimulus	GO:0010996	13
Auditory behavior	GO:0031223	3
Inner ear morphogenesis	GO:0042472	69
Middle ear morphogenesis	GO:0042474	22
Auditory receptor cell differentiation	GO:0042491	17
Auditory receptor cell fate specification	GO:0042667	1
Auditory receptor cell fate determination	GO:0042668	3
Regulation of auditory receptor cell differentiation	GO:0045607	3
Negative regulation of auditory receptor cell differentiation	GO:0045608	3
Positive regulation of auditory receptor cell differentiation	GO:0045609	1
Inner ear development	GO:0048839	56
Auditory receptor cell stereocilium organization	GO:0060088	13
Inner ear receptor cell differentiation	GO:0060113	9
Auditory receptor cell development	GO:0060117	7
Inner ear receptor cell development	GO:0060119	10
Inner ear receptor stereocilium organization	GO:0060122	16
Negative regulation of inner ear receptor cell differentiation	GO:2000981	2
Positive regulation of inner ear receptor cell differentiation	GO:2000982	1

The genes annotated for each GO category are not unique to the category (node), but may be shared among multiple nodes. The genes can be retrieved from GO (http://geneontology.org/) using the provided identifier.

hiding from a predator, as well as communicating and mating. Gregarious nocturnal animals rely more on sound than diurnal counterparts, in which case communication or territorial boundary delimitations can be intermediated through vision as well [86]. This implies that nocturnal animals should have a very fine sense of hearing. Indeed, nocturnal species that rely on hearing, such as bush babies, fennecs, and insectivorous bats, have well-developed ears with enlarged pinnae to collect sounds into the ear canal and locate its source. Unlike mammals, birds lack outer ears, but they do possess internal ears, into which sounds are funneled by specialized auricular feathers.

Again, the morphological characteristics can be relatively easily assessed, but a genomic comparison between nocturnal and diurnal bird species has not yet been performed. Because there is no genomic hallmark for fine hearing, the proposed way to test the hypothesis as to whether hearing evolves differently under nocturnality involves the GO annotation. Unlike, the case of touch, there are more than 200 genes in the human genome that are involved in the morphogenesis of the inner and middle ear, as well as genes coding for receptor organizations or neurological response to an auditory stimulus. It is thus understandable why these genes cannot be all manually screened. However, to make this process easier, we provide a list with 21 annotated GO categories related to development of auditory sense (Table 1). Testing whether genes annotated in these categories show faster/slower evolution in nocturnal birds would give a better insight into the genetic basis of hearing adaptation.

Nocturnality leads to changes in biorhythm

Non-visual photoreception plays an important role in nocturnality [6]. This can influence the endocrine biorhythms with different time period, i.e. circadian and circannual rhythms for example.

There is increasing evidence that the pineal gland plays a crucial role in the regulation of the avian biorhythms. The pineal gland influences seasonal breeding in birds [87], as well as the migratory instinct ("Zugunruhe") and orientation [88]. The regulation of pineal melatonin secretion is essentially different in mammals and birds [89]. As in the case of any regulated system, there are three components that contribute to the regulation: input, central part, and output. In the case of the mammals, the input is located in the visual pathways, the central clock is in the suprachiasmatic nucleus, and the output is in the pineal gland. For birds, all three elements are located in the pineal gland [89]. The input channel in birds is represented by light receptors, which are the rod- and cone-like pineal cells: these are coupled through a biochemical pacemaker to the melatonin-producing pinealocytes (output channel). Because in mammals, the pineal melatonin secretion regulatory system has components situated in three different locations, while preserving the mammalian pineal gland in vitro, only a low basal secretion of melatonin can be recorded. Conversely, the avian pineal maintains its reactivity in vitro. This autonomy of the rhythm is not maintained by rhythmic changes of environmental light levels, but it is believed to be genetically coded in clock-related genes [90–92], including pinopsin, a non-visual opsin present in birds; however, this non-visual opsin is missing in mammals [89, 93] (Fig. 4). Hence, comparing selection pressures acting on genes related to pineal development and function in nocturnal versus diurnal birds could shed more light on this system's evolution. Moreover, by comparing evolution of

these genes in mammals and nocturnal birds, one could probably infer whether the essential difference in the pineal secretion regulation is related to mammals passing through the nocturnal bottleneck.

Non-visual photoreception can also be mediated through non-visual pigments, which consist of an opsin linked to a retinal chromophore. Figure 4 shows the presence of non-visual opsins in reptiles and birds versus mammals. The three opsins missing from mammals, but present in birds and reptiles, are pinopsin, OPN4X, and VA. OPN4X and OPN4M arose from the duplication of the ancestral melanopsin early in vertebrate evolution [94]. Although both orthologs were retained in non-mammalian vertebrates, OPN4X was lost in mammals [94, 95]. To date, there is no clear explanation why OPN4M, unlike OPN4X, survived in all studied extant vertebrates. It would thus be interesting to closely inspect these genes for inactivating mutations and selective pressures (see above) in nocturnal versus diurnal birds. If the OPN4X loss is related to mammal evolution in a low-light environment, one would expect different selective pressures under nocturnality compared to diurnality in birds.

Similarly, the VA opsin was lost in mammals, while in birds it is involved in regulating seasonal breeding via direct hypothalamic photosensitive cells located deep in the brain [96, 97]. It has been suggested that the concomitant loss of VA [95] and deep-brain photoreception in mammals [98] supports the evolutionary significance of a nocturnal bottleneck [6]. It would thus be worth testing whether the VA opsin is evolving under a relaxed selective constraint in nocturnal birds.

Conclusions and outlook

Animals active at night depend on more than their vision to integrate their surroundings. Nocturnal animals tend to have at least one highly developed sense. While morphological information has been used extensively to identify adaptations to nocturnality, our genetic understanding of these changes is not as advanced. This adaptation is most probably very complex, and did not include changes only in components of the senses and biorhythm, but also in components of temperature regulation, energy expenditure, coat coloring, and UV-light-caused DNA repair. It is also reasonable to assume that the changes in components or pathways promoting nocturnality are not always the same in different species, and that the sum of some subset of those changes may result in a convergent phenotype. Strategically, it would be straightforward to compare several genomes from diurnal and nocturnal species within an order, a family or even better within a genus. Further, some convergent sub-phenotypes found also in diurnal species, such as inactivation of cone opsins and reliance on tactile sense (see above), can be recruited for comparative purposes as well. Unfortunately, most signals that will be detected in candidate genes in comparative genomic ventures are not as easy to interpret as the inactivation of e.g. a cone opsin. Most such studies will end up with lists of GO categories, candidate genes or gene variants which then need to be tested experimentally for their relevance in the respective trait. Nevertheless, making use of the rich genomic information, candidate genes can be explored in other in vitro or in vivo model systems in the view of sensorial or biorhythm tuning. Using comparative genomics to understand the genetic basis of nocturnality related changes in birds will shed light on gene relations to various biological systems, and in the end it will facilitate understanding the genome itself.

Acknowledgements

The authors would like to thank Andrew Fidler, Janet Kelso, and Andrew Moore for helpful suggestions and input on the paper.

The authors have declared no conflicts of interest.

References

1. **Braun MJ, Huddleston CJ**. 2009. A molecular phylogenetic survey of caprimulgiform nightbirds illustrates the utility of non-coding sequences. *Mol Phylogenet Evol* **53**: 948–60.
2. **Ericson PG, Anderson CL, Britton T, Elzanowski A**, et al. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. *Biol Lett* **2**: 543–7.
3. **Cunningham S, Castro I, Alley M**. 2007. A new prey-detection mechanism for kiwi (Apteryx spp.) suggests convergent evolution between paleognathous and neognathous birds. *J Anat* **211**: 493–502.
4. **Hall MI, Kamilar JM, Kirk EC**. 2012. Eye shape and the nocturnal bottleneck of mammals. *Proc Biol Sci/Roy Soc* **279**: 4962–8.
5. **Healy S, Guilford T**. 1990. Olfactory-bulb size and nocturnality in birds. *Evolution* **44**: 339–46.
6. **Gerkema MP, Davies WI, Foster RG, Menaker M**, et al. 2013. The nocturnal bottleneck and the evolution of activity patterns in mammals. *Proc Biol Sci/Roy Soc* **280**: 20130508.
7. **Jarvis ED, Mirarab S, Aberer AJ, Li B**, et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**: 1320–31.
8. **Zhang G, Rahbek C, Graves GR, Lei F**, et al. 2015. Genomics: bird sequencing project takes off. *Nature* **522**: 34.
9. **Ekblom R, Wolf JB**. 2014. A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl* **7**: 1026–42.
10. **Hodkinson BP, Grice EA**. 2015. Next-generation sequencing: a review of technologies and tools for wound microbiome research. *Adv Wound Care* **4**: 50–8.
11. **Nagarajan N, Pop M**. 2013. Sequence assembly demystified. *Nat Rev Genet* **14**: 157–67.
12. **van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C**. 2014. Ten years of next-generation sequencing technology. *Trends Genet: TIG* **30**: 418–26.
13. **Miller JR, Koren S, Sutton G**. 2010. Assembly algorithms for next-generation sequencing data. *Genomics* **95**: 315–27.
14. **Chain PS, Grafham DV, Fulton RS, Fitzgerald MG**, et al. 2009. Genomics. Genome project standards in a new era of sequencing. *Science* **326**: 236–7.
15. **Yandell M, Ence D**. 2012. A beginner's guide to eukaryotic genome annotation. *Nat Rev Genet* **13**: 329–42.
16. **Capra JA, Stolzer M, Durand D, Pollard KS**. 2013. How old is my gene?. *Trends Genet* **29**: 659–68.
17. **Ashburner M, Ball CA, Blake JA, Botstein D**, et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25–9.
18. **Le Duc D, Renaud G, Krishnan A, Almen MS**, et al. 2015. Kiwi genome provides insights into evolution of a nocturnal lifestyle. *Genome Biol* **16**: 147.
19. International Chicken Genome Sequencing C. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**: 695–716.
20. **Warren WC, Clayton DF, Ellegren H, Arnold AP**, et al. 2010. The genome of a songbird. *Nature* **464**: 757–62.

21. **Curwen V, Eyras E, Andrews TD, Clarke L**, et al. 2004. The Ensembl automatic gene annotation system. *Genome Res* **14**: 942–50.
22. **Cunningham F, Amode MR, Barrell D, Beal K**, et al. 2015. Ensembl 2015. *Nucleic Acids Res* **43**: D662–9.
23. **Lovell PV, Wirthlin M, Wilhelm L, Minx P**, et al. 2014. Conserved syntenic clusters of protein coding genes are missing in birds. *Genome Biol* **15**: 565.
24. **Li W, Cowley A, Uludag M, Gur T**, et al. 2015. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res* **43**: W580–4.
25. **Kuhn RM, Karolchik D, Zweig AS, Wang T**, et al. 2009. The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res* **37**: D755–61.
26. **Olson MV**. 1999. When less is more: gene loss as an engine of evolutionary change. *Am J Hum Genet* **64**: 18–23.
27. **Lespinet O, Wolf YI, Koonin EV, Aravind L**. 2002. The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res* **12**: 1048–59.
28. **Huang Y, Li Y, Burt DW, Chen H**, et al. 2013. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat Genet* **45**: 776–83.
29. **Zhan X, Pan S, Wang J, Dixon A**, et al. 2013. Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. *Nat Genet* **45**: 563–6.
30. **Li H, Coghlan A, Ruan J, Coin LJ**, et al. 2006. TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res* **34**: D572–80.
31. **De Bie T, Cristianini N, Demuth JP, Hahn MW**. 2006. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**: 1269–71.
32. **Wang G, Shi P, Zhu Z, Zhang YP**. 2010. More functional V1R genes occur in nest-living and nocturnal terricolous mammals. *Genome Biol Evol* **2**: 277–83.
33. **Kimura M**. 1984. *The Neutral Theory of Molecular Evolution*. New York, USA: Cambridge University Press.
34. **Yang Z, Nielsen R**. 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* **17**: 32–43.
35. **Messier W, Stewart CB**. 1997. Episodic adaptive evolution of primate lysozymes. *Nature* **385**: 151–4.
36. **Crandall KA, Hillis DM**. 1997. Rhodopsin evolution in the dark. *Nature* **387**: 667–8.
37. **Nielsen R, Yang Z**. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **148**: 929–36.
38. **Yang Z**. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**: 1586–91.
39. **Yang Z, Nielsen R, Goldman N, Pedersen AM**. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**: 431–49.
40. Gene Ontology C. 2015. Gene Ontology Consortium: going forward. *Nucleic Acids Res* **43**: D1049–56.
41. **Bonaccorsi ED, Ferreira AJ, Chambergo FS, Ramos AS**, et al. 2006. Transcriptional response of the obligatory aerobe Trichoderma reesei to hypoxia and transient anoxia: implications for energy production and survival in the absence of oxygen. *Biochemistry* **45**: 3912–24.
42. **Lin Y, Wu H**. 2009. Comparative analysis of aerobic and anaerobic prokaryotes to identify correlation between oxygen requirement and gene-gene functional association patterns. *Genome Inform* **23**: 72–84.
43. **Pavlidis P, Jensen JD, Stephan W, Stamatakis A**. 2012. A critical assessment of storytelling: gene ontology categories and the importance of validating genomic scans. *Mol Biol Evol* **29**: 3237–48.
44. **Dalloul RA, Long JA, Zimin AV, Aslam L**, et al. 2010. Multi-platform next-generation sequencing of the domestic turkey (Meleagris gallopavo): genome assembly and analysis. *PLoS Biol* **8**: e1000475.
45. **Mueller JC, Kuhl H, Timmermann B, Kempenaers B**. 2016. Characterization of the genome and transcriptome of the blue tit Cyanistes caeruleus: polymorphisms, sex-biased expression and selection signals. *Mol Ecol Resour* **16**: 549–61.
46. **Poelstra JW, Vijay N, Bossu CM, Lantz H**, et al. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **344**: 1410–4.
47. **Shapiro MD, Kronenberg Z, Li C, Domyan ET**, et al. 2013. Genomic diversity and evolution of the head crest in the rock pigeon. *Science* **339**: 1063–7.
48. **Zhang G, Li C, Li Q, Li B**, et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**: 1311–20.
49. **Prum RO, Berv JS, Dornburg A, Field DJ**, et al. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**: 569–73.
50. **Liu Y, Zhou Q, Wang Y, Luo L**, et al. 2015. Gekko japonicus genome reveals evolution of adhesive toe pads and tail regeneration. *Nat Commun* **6**: 10033.
51. **Cooper GM, Shendure J**. 2011. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat Rev Genet* **12**: 628–40.
52. **Zhu M, Zhao S**. 2007. Candidate gene identification approach: progress and challenges. *Int J Biol Sci* **3**: 420–7.
53. **Tabor HK, Risch NJ, Myers RM**. 2002. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* **3**: 391–7.
54. **Kim J, Lee T, Kim TH, Lee KT**, et al. 2012. An integrated approach of comparative genomics and heritability analysis of pig and human on obesity trait: evidence for candidate genes on human chromosome 2. *BMC Genomics* **13**: 711.
55. **Vega-Zuniga T, Medina FS, Fredes F, Zuniga C**, et al. 2013. Does nocturnality drive binocular vision? Octodontine rodents as a case study. *PloS ONE* **8**: e84199.
56. **Walls GL**. 1942. *The Vertebrate Eye and its Adaptive Radiation*. Bloomfield Hills: Cranbrook Institute of Science.
57. **Yokoyama S**. 2000. Molecular evolution of vertebrate visual pigments. *Prog Retin Eye Res* **19**: 385–419.
58. **Okano T, Kojima D, Fukada Y, Shichida Y**, et al. 1992. Primary structures of chicken cone visual pigments: vertebrate rhodopsins have evolved out of cone visual pigments. *Proc Natl Acad Sci U S A* **89**: 5932–6.
59. **Hunt DM, Carvalho LS, Cowing JA, Davies WL**. 2009. Evolution and spectral tuning of visual pigments in birds and mammals. *Philos Trans R Soc Lond B Biol Sci* **364**: 2941–55.
60. **Davies WL, Carvalho LS, Cowing JA, Beazley LD**, et al. 2007. Visual pigments of the platypus: a novel route to mammalian colour vision. *Curr Biol: CB* **17**: R161–3.
61. **Zhao H, Rossiter SJ, Teeling EC, Li C**, et al. 2009. The evolution of color vision in nocturnal mammals. *Proc Natl Acad Sci U S A* **106**: 8980–5.
62. **Hart NS, Hunt DM**. 2007. Avian visual pigments: characteristics, spectral tuning, and evolution. *Am Nat* **169**: S7–26.
63. **Bowmaker JK, Martin GR**. 1978. Visual pigments and colour vision in a nocturnal bird, Strix aluco (tawny owl). *Vision Res* **18**: 1125–30.
64. **Rojas LM, Ramirez Y, McNeil R, Mitchell M**, et al. 2004. Retinal morphology and electrophysiology of two caprimulgiformes birds: the cave-living and nocturnal oilbird (Steatornis caripensis), and the crepuscularly and nocturnally foraging common pauraque (Nyctidromus albicollis). *Brain Behav Evol* **64**: 19–33.
65. **Li C, Zhang Y, Li J, Kong L**, et al. 2014. Two Antarctic penguin genomes reveal insights into their evolutionary history and molecular changes related to the Antarctic environment. *GigaScience* **3**: 27.
66. **Coster M, Wittkopf D, Kreuchwig A, Kleinau G**, et al. 2012. Using ortholog sequence data to predict the functional relevance of mutations in G-protein-coupled receptors. *FASEB J* **26**: 3273–81.
67. **Streidter GF**. 2005. *Principles of Brain Evolution*. Sunderland, MA: Sinauer Associates Inc.
68. **Buck L, Axel R**. 1991. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**: 175–87.
69. **Steiger SS, Fidler AE, Kempenaers B**. 2009. Evidence for increased olfactory receptor gene repertoire size in two nocturnal bird species with well-developed olfactory ability. *BMC Evol Biol* **9**: 117.
70. **Niimura Y, Nei M**. 2003. Evolution of olfactory receptor genes in the human genome. *Proc Natl Acad Sci U S A* **100**: 12235–40.
71. **Newman T, Trask BJ**. 2003. Complex evolution of 7E olfactory receptor genes in segmental duplications. *Genome Res* **13**: 781–93.
72. **Compeau PE, Pevzner PA, Tesler G**. 2011. How to apply de Bruijn graphs to genome assembly. *Nat Biotechnol* **29**: 987–91.
73. **Luo R, Liu B, Xie Y, Li Z**, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* **1**: 18.
74. **Zerbino DR, Birney E**. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**: 821–9.
75. **Hasin-Brumshtein Y, Lancet D, Olender T**. 2009. Human olfaction: from genomic variation to phenotypic diversity. *Trends Genet: TIG* **25**: 178–84.
76. **Margulies DH, Natarajan K, Rossjohn J, McCluskey J**. 2012. *Fundamental Immunology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.
77. **Shannon CE.** 1948. A mathematical theory of communication. *Bell System Tech J* **27**: 379; 623.
78. **Finn RD, Bateman A, Clements J, Coggill P**, et al. 2014. Pfam: the protein families database. *Nucleic Acids Res* **42**: D222–30.

79. **Steiger SS, Fidler AE, Valcu M, Kempenaers B**. 2008. Avian olfactory receptor gene repertoires: evidence for a well-developed sense of smell in birds?. *Proc Biol Sci/Roy Soc* **275**: 2309–17.
80. **Houhamdi M, Samraoui B**. 2008. Diurnal and nocturnal behaviour of Ferruginous Duck Aythya nyroca at Lac des Oiseaux, northeast Algeria. *Ardeola* **55**: 59–69.
81. **Pedroli J-C**. 1982. Activity and time budget of tufted ducks on Swiss lakes during winter. *Wildfowl* **33**: 105–12.
82. **Daan S**. 1981. *Handbook of Behavioral Neurobiology*. New York: Plenum Press.
83. **Muchlinski MN**. 2010. A comparative analysis of vibrissa count and infraorbital foramen area in primates and other mammals. *J Hum Evol* **58**: 447–73.
84. **Furman DP, Bukharina TA**. 2009. The gene network determining development of Drosophila melanogaster mechanoreceptors. *Comput Biol Chem* **33**: 231–4.
85. **Coleman MN, Boyer DM**. 2012. Inner ear evolution in primates through the Cenozoic: implications for the evolution of hearing. *Anat Rec* **295**: 615–31.
86. **Roots C**. 2006. *Nocturnal Animals*. Conneticut: Greenwood Publishing Group.
87. **Wolfson A**. 1970. Light and darkness and circadian rhythms in the regulation of annual reproductive cycles in birds. *Colloq Int Cent Nat Rech Sci* **172**: 93–119.
88. **Schneider T, Thalau HP, Semm P, Wiltschko W**. 1994. Melatonin is crucial for the migratory orientation of pied flycatchers (Ficedula Hypoleuca Pallas). *J Exp Biol* **194**: 255–62.
89. **Csernus V, Mess B**. 2003. Biorhythms and pineal gland. *Neuro Endocrinol Lett* **24**: 404–11.
90. **Mess B, Csernus V**. 2002. Rhythmic pineal melatonin productioninbirds. In: Haldar C, Singaravel M, Maitra SK, eds; *Treatise on Pineal Gland and Melatonin*. Plymouth, UK: Science Publishers Inc. pp 177–98.
91. **Robertson LM, Takahashi JS**. 1988. Circadian clock in cell culture: I. Oscillation of melatonin release from dissociated chick pineal cells in flow-through microcarrier culture. *J Neurosci* **8**: 12–21.
92. **Yoshimura T, Suzuki Y, Makino E, Suzuki T**, et al. 2000. Molecular analysis of avian circadian clock genes. *Brain Res Mol Brain Res* **78**: 207–15.
93. **Max M, McKinnon PJ, Seidenman KJ, Barrett RK**, et al. 1995. Pineal opsin: a nonvisual opsin expressed in chick pineal. *Science* **267**: 1502–6.
94. **Bellingham J, Chaurasia SS, Melyan Z, Liu C**, et al. 2006. Evolution of melanopsin photoreceptors: discovery and characterization of a new melanopsin in nonmammalian vertebrates. *PLoS Biol* **4**: e254.
95. **Davies WL, Hankins MW, Foster RG**. 2010. Vertebrate ancient opsin and melanopsin: divergent irradiance detectors. *Photochem Photobiol Sci* **9**: 1444–57.
96. **Davies WI, Turton M, Peirson SN, Follett BK**, et al. 2012. Vertebrate ancient opsin photopigment spectra and the avian photoperiodic response. *Biol Lett* **8**: 291–4.
97. **Halford S, Pires SS, Turton M, Zheng L**, et al. 2009. VA opsin-based photoreceptors in the hypothalamus of birds. *Curr Biol: CB* **19**: 1396–402.
98. **Hazlerigg D**. 2012. The evolutionary physiology of photoperiodism in vertebrates. *Prog Brain Res* **199**: 413–22.