Genomic signals of selection in the adaptative history of owls (Strigiformes)

Dissertation

Fakultät für Biologie
Ludwig-Maximilians-Universität München

durchgeführt am

Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany

vorgelegt von

Pamela Espíndola-Hernández

2022

Genomic signals of selection in the adaptative history of owls (Strigiformes)

Dissertation

Fakultät für Biologie
Ludwig-Maximilians-Universität München

durchgeführt am

Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany

vorgelegt von

Pamela Espíndola-Hernández

2022

Erstgutachter: **Prof. Dr. Bart Kempenaers**

Zweitgutachter: **Prof. Dr. John Parch**

Tag der Abgabe: 02.11.2022

Tag der mündlichen Prüfung: 26.01.2023

Diese Dissertation wurde unter der Leitung von **Dr. Jakob Mueller** angefertigt.

Contents

Abstract ……..……..... iv

General introduction ……..……..... 3

Chapter 1 – *Genomic Evidence for Sensorial Adaptations to a Nocturnal Predatory Lifestyle in Owls* …….……..…. 15

Chapter 2 – *Genomic signatures of the evolution of a diurnal lifestyle in Strigiformes* …….……..…….. 47

Chapter 3 – *Comparing two methods for the identification of the W chromosomes of owls* ……..... 69

General discussion …….... 91

Acknowledgments …….….... 101

Curriculum vitae …….……..….... 105

Statutory declaration and statement …….…….…. 108

Abstract

A *genomic signal of selection* can be inferred by the identification of an accumulation of genetic changes at a certain branch of a phylogenetic tree. Considering the whole genome of any vertebrate species, most of the genetic changes have no functional effect on the organisms and they can get fixed in a population under *neutral evolution*. In genomic regions that play a relevant role for the organism, for example, coding or regulatory regions, most genetic changes have negative effects on the organisms and, most of the time, *purifying selection* prevents their accumulation in natural populations. In protein-coding genes, for example, a change in the second nucleotide of a codon usually causes structural and functional consequences for the encoded protein. These genetic changes are known as *non-synonymous substitutions*. For that reason, it is very unexpected *to see* an accelerated rate of changes, such as *non-synonymous substitutions*, in a specific natural population or species. An accelerated accumulation of genetic changes in a population or species – that is adapted to a particular environment – has probably advantageous effects on the organisms, which fixation is promoted by *positive selection*. I choose the owls (Strigiformes) as the study system, and I looked for the *genomic signals of selection* in relation to their adaptations as nocturnal raptors, using a genome-wide comparative approach. Among the several Owls´ adaptations to the *nocturnal niche* are their large eyes with a duplex retina dominated by rods, excellent sense of hearing, cryptic plumage coloration, and silent flying.

Here, I present my dissertation with the aim of contributing to the understanding of the genomic fingerprints of the evolution of owls and their adaptations as nocturnal raptors. This dissertation is comprised of three chapters. Chapter 1 supports the hypothesis of the diurnal ancestor of all living birds and an independent adaptive history of owls as nocturnal birds of prey, through the identification of accumulated genetic changes at the ancestral branch of the owls in genes that are functionally associated with nocturnal hunting. Chapter 2 suggests that the evolution of gene regulation might have played a predominant role in the shift to a diurnal lifestyle in owls. This conclusion is based on the higher acceleration of substitution rates at noncoding elements, and on the functional association of the accelerated protein-coding genes in diurnal owls with the regulation of gene expression. Finally, Chapter 3 shows the comparison of two methods for the identification of the sex-restricted W chromosome of owls. The results suggest that the method based on the *de-novo* assembly of a female owl´s genome and remapping male and female genomes from the same species can detect more sex-specific genome regions than methods based on mapping reads to a closely related species.

General Introduction

A good part of the adaptations of owls (Strigiformes) to their nocturnal lifestyle is based on their crypticity. The plumage colors of the owls have combinations of brown, gray, and white, producing a convenient blend with the bark of the trees. In addition, their feathers have a velvet-like texture that allows owls to fly very silently. They take advantage of the darkness by being almost undetectable to their prey. However, their crypticity and nocturnality have not been an obstacle to our fascination for them throughout human history. The owl's image is part of the mythology and literature of many cultures (Weinstein 1990; Eason 2008; Pasierowska 2017). The symbolism associated with owls has ranged from the more benevolent to the rather evil connotations. Nowadays, owls enjoy strong popularity in our society, *e.g.*, as a symbol of wisdom, as iconic nocturnal animals, and as a charismatic gaze. This popularity is also reflected in scientific research, through the publication of hundreds of articles about owls[1]. Even so, no owl species can be considered a research model species. There are many unknown aspects of its natural history, physiology, behavior, and, particularly, owl genetics.

The genetic studies on owls mostly focused on understanding owl population structures with ecological and conservational approaches (*e.g.*, Hanna *et al.* 2018; Macías-Duarte *et al.* 2019; Mendelsohn *et al.* 2020; Fujito *et al.* 2021; Cumer *et al.* 2022). Some other genetic studies on owls have focused on the genes that encode visual pigments in these species (*e.g.*, Borges *et al.* 2015, 2019; Wu *et al.* 2016; Emerling 2018; Höglund *et al.* 2019). However, the possibility of understanding the evolution of owls using large datasets of genetic information has emerged very recently. Nowadays, evolutionary

[1] The Web of Science Core Collection finds 911 results for [owls (Title) OR strigi* (Title) NOT monkey* (All Fields)] in the last five years. https://www.webofscience.com/ search on 24.10.2022.

biology studies based on the comparison of whole genome sequences from several species of vertebrates – such as owls – are possible due to the development of high-throughput sequencing technologies, the reduction of sequencing costs, and the rapid growth of bioinformatics as a well-established discipline (Nielsen 2005; Vitti *et al.* 2013; Booker *et al.* 2017; Jones *et al.* 2019).

In the following paragraphs, I will introduce essential concepts that form the theoretical background of my dissertation, such as the study system, the nocturnal niche, and some models used for the study of adaptation. At the end of this general introduction, I will present the aims and chapters that form my dissertation as a cumulative thesis.

The study system: the owls

Owls are taxonomically designated in the Strigiformes order, which is divided into two families, Strigidae and Tytonidae (del Hoyo *et al.* 1999; König and Weick 2008; Wink *et al.* 2009; Ponder and Willette 2015). Owls are raptors, also called "birds of prey", with several adaptations to the nocturnal niche. Similar to other raptors, such as hawks and eagles, owls have forward-looking eyes, claws, curved beaks, cryptic plumage coloration, and reversed sexual size dimorphism (del Hoyo *et al.* 1999; König and Weick 2008). Many owl species can localize their prey in darkness guided by their sense of hearing, which is supported by morphological features such as the asymmetrical position of the ears and the presence of a facial disk of feathers (Payne 1971). The feathers of most owl species have a particular structure that makes them softer compared to feathers from other birds, i.e., a serrated leading edge, a fringe trailing edge, and very fine barbules (Kopania 2016; Sagar *et al.* 2017). The soft feathers of owls minimize their noise at flying, which

might improve the owls´ hunting success by preventing their hearing abilities from being hampered and avoiding being detected by their prey (Kopania 2016; Sagar *et al.* 2017). In several aspects, the eyes of owls are similar to the eyes of other nocturnal animals, which are typically large and with a duplex retina dominated by rods that bring an enhanced visual sensitivity but poor color discrimination (Walls 1942; Fite 1973; Martin *et al.* 2004; Corfield *et al.* 2015; Emerling 2018).

The nocturnal niche

The niche that certain species occupy might refer to the environmental conditions, *e.g.*, the temperature, pH, and light conditions, or the set of resources the organism needs to live there. The first is known as *Grinnellian* and the latter as *Eltonian niche* (Grinnell 1917; Elton 1927; Soberón 2007). Every species has their niche borders and optimal conditions, and, normally, several species compete for the same or similar ecological niche. The differential use of the space in time is known as the *diel activity niche* and allows to expand the *Grinnellian* and *Eltonian* niche of the species, reducing negative interaction among them (Carothers and Jaksić 1984; Kronfeld-Schor and Dayan 2003). Nocturnal organisms are active at night, they start activity at night, or increase activity at night, as defined by Orlando Park (1940). Park intensively described the ecological importance of what he called "*Nocturnal Problem*" (Park 1940).

From the evolutionary biology perspective, the niche should be such that it keeps conserved in every lineage. This is niche conservatism, which refers to the phylogenetic constraints on the evolution of certain traits. In general, there are phylogenetic constraints on the evolution of diel activity patterns (Roll *et al.* 2006; Anderson and Wiens 2017), mostly due to morphological and physiological limitations such as thermal tolerances,

visual systems, and biological rhythms (Kronfeld-Schor and Dayan 2003). In other words, it is expected that species that are phylogenetically closely related have similar niche, including the diel niche. However, some species might shift their expected diel activity pattern in response to competition. These species are named 'time-shifter' (Jaksić 1982; Carothers and Jaksić 1984; Pei *et al.* 2018). As most living birds are diurnal, the nocturnal birds can be considered as 'time-shifter', such as kiwis and owls.

Le Duc and Schöneberg (2016) revised the nocturnality in birds – which evolved independently in different bird taxa – and presented it as a convinient model to study recent genetic changes and convergent evolution, through the comparison of genetic sequences related with adaptations to nocturnality (Le Duc *et al.* 2015; Le Duc and Schöneberg 2016).

In birds, adaptation to nocturnality involves many phenotypical changes in sensorial systems, such as accurate vision and hearing, as well as modifications in feathers structure, biorhythm, coloration patterns, etc. The diel activity pattern is associated with eye shape and size (Hall and Ross 2007; Lisney *et al.* 2012), size of olfactory bulbs (Healy and Guilford 1990), neural visual pathways (Gutiérrez-Ibáñez *et al.* 2013), and iris coloration (Passarotto *et al.* 2018b, 2018a). However, the genetic basis for nocturnal adaptations in birds has mostly been studied in the visual system of kiwi (*Apteryx mantelli*) (Le Duc *et al.* 2015) and the Barn owl (*Tyto alba*) (Le Duc *et al.* 2015; Borges *et al.* 2015; Emerling 2018).

The substitution rates as genomic signals of selection

Depending on the environment, mutations can be either beneficial or detrimental for the organism that carries them. Beneficial mutations increase life expectancy and reproductive success, as a consequence, these mutations are maintained from generation to generation, fixing the trait on the populations. These mutations are selected in the evolution of the species, and the process can be traced by comparing among genomes of different species in a phylogeny.

There are two categories of mutations in the coding part of the genome, *Synonymous* and *non-synonymous mutations*, depending on the codon position and whether they change the amino acid sequence of the encoded protein, or not. *Synonymous substitutions* are mostly neutral because they do not change the amino acid sequence of the encoded protein, whereas *non-synonymous substitutions* do modify the structure and might, as a consequence, modify the function of the protein. Therefore, *non-synonymous substitutions* are often deleterious and are negatively affected by natural selection (Kreitman and Akashi 1995). The ratio between *non-synonymous rate* to *synonymous substitutions rate* can indicate how natural selection has acted on sequence changes. This is called "ω ratio" or ω = dN / dS, where dN is the non-synonymous substitutions rate and dS is the synonymous substitutions rate (Yang and Nielsen 2000; Nielsen 2005). Thus, from the values of this ω ratio we can infer positive selection (when $\omega > 1$ the mutation was adaptive), neutral evolution (when $\omega = 1$), or negative selection (when $\omega < 1$) (Kreitman and Akashi 1995; Anisimova and Yang 2007; Yang 2007).

My dissertation is composed of three chapters that tackle different aspects of the study of owls' evolution by genome-wide comparative approach, with the aim of contributing to the understanding of the genomic fingerprints of the evolution of owls and their adaptation as nocturnal raptors. In Chapter 1, I assessed positive/negative selection in genes associated with adaptation to nocturnal conditions in owls, by nonsynonymous/synonymous substitution rate (ω=dN/dS) estimation at the ancestral branch of the owls to identify the selection signals of their adaptations as nocturnal raptors. In Chapter 2, I searched for signals of accelerated substitution rates in the genome of three diurnal owl species, with the aim of understanding the respective role of coding and non-coding genomic elements in the emergence of a diurnal lifestyle in the owls' clade. Finally, in Chapter 3, I contrasted the contribution of *de-novo* and reference-guided assembly strategies, in particular to the study of highly repetitive sex-restricted genomic regions, with the aim of identifying regions of the sex chromosomes of owls.

References

Anderson, S. R., and J. J. Wiens, 2017 Out of the dark: 350 million years of conservatism and evolution in diel activity patterns in vertebrates. Evolution (N. Y). 71: 1944–1959.

Anisimova, M., and Z. Yang, 2007 Multiple hypothesis testing to detect lineages under positive selection that affects only a few sites. Mol. Biol. Evol. 24: 1219–1228.

Booker, T. R., B. C. Jackson, and P. D. Keightley, 2017 Detecting positive selection in the genome. BMC Biol. 15: 1–10.

Borges, R., J. Fonseca, C. Gomes, W. E. Johnson, S. J. O'Brien *et al.*, 2019 Avian Binocularity and Adaptation to Nocturnal Environments: Genomic Insights from a Highly Derived Visual Phenotype. Genome Biol. Evol. 11: 2244–2255.

Borges, R., I. Khan, W. E. Johnson, M. T. P. Gilbert, G. Zhang *et al.*, 2015 Gene loss, adaptive evolution and the co-evolution of plumage coloration genes with opsins in birds. BMC Genomics 16: 751.

Carothers, J. H., and F. M. Jaksić, 1984 Time as a Niche Difference: The Role of Interference Competition. Oikos 42: 403–406.

Corfield, J. R., S. Parsons, Y. Harimoto, and M. L. Acosta, 2015 Retinal Anatomy of the New Zealand Kiwi: Structural Traits Consistent With Their Nocturnal Behavior. Anat. Rec. 298: 771–779.

Cumer, T., A. P. Machado, F. Siverio, S. I. Cherkaoui, I. Roque *et al.*, 2022 Genomic basis of insularity and ecological divergence in barn owls (Tyto alba) of the Canary Islands. Hered. 2022 1295 129: 281–294.

Le Duc, D., G. Renaud, A. Krishnan, M. S. Almén, L. Huynen *et al.*, 2015 Kiwi genome provides insights into evolution of a nocturnal lifestyle. Genome Biol. 16: 1–15.

Le Duc, D., and T. Schöneberg, 2016 Adaptation to nocturnality – learning from avian

genomes. BioEssays 38: 694–703.

Eason, C., 2008 *Fabulous creatures, mythical monsters, and animal power symbols: a handbook*. Greenwood Publishing Group.

Elton, C. (1927), 1927 *Animal Ecology*. Sedgwick and Jackson, London.

Emerling, C. A., 2018 Independent pseudogenization of CYP2J19 in penguins, owls and kiwis implicates gene in red carotenoid synthesis. Mol. Phylogenet. Evol. 118: 47–53.

Fite, K. V., 1973 Anatomical and behavioral correlates of visual acuity in the great horned owl. Vision Res. 13: 219-IN2.

Fujito, N. T., Z. R. Hanna, M. Levy-Sakin, R. C. K. Bowie, P. Y. Kwok *et al.*, 2021 Genomic Variation and Recent Population Histories of Spotted (Strix occidentalis) and Barred (Strix varia) Owls. Genome Biol. Evol. 13: 1–12.

Grinnell, J., 1917 The Niche-Relationships of the California Thrasher. Auk 34: 427–433.

Gutiérrez-Ibáñez, C., A. N. Iwaniuk, T. J. Lisney, and D. R. Wylie, 2013 Comparative study of visual pathways in Owls (Aves: Strigiformes). Brain. Behav. Evol. 81: 27–39.

Hall, M. I., and C. F. Ross, 2007 Eye shape and activity pattern in birds. J. Zool. 271: 437–444.

Hanna, Z. R., J. P. Dumbacher, R. C. K. Bowie, J. B. Henderson, and J. D. Wall, 2018 Whole-Genome Analysis of Introgression Between the Spotted Owl and Barred Owl (Strix occidentalis and Strix varia, Respectively; Aves: Strigidae) in Western North America. G3 (Bethesda). g3.200754.2018.

Healy, S., and T. Guilford, 1990 Olfactory-Bulb Size and Nocturnality in Birds. Evolution (N. Y). 44: 339–346.

Höglund, J., M. Mitkus, P. Olsson, O. Lind, A. Drews *et al.*, 2019 Owls lack UV-

sensitive cone opsin and red oil droplets, but see UV light at night: Retinal transcriptomes and ocular media transmittance. Vision Res. 158: 109–119.

del Hoyo, J., A. Elliott, J. Sargatal, and J. Cabot, 1999 *Handbook of the birds of the world*. Lynx Edicions, Barcelona.

Jaksić, F. M., 1982 Inadequacy of activity time as a niche difference: the case of diurnal and nocturnal raptors. Oecologia 52: 171–175.

Jones, C. T., E. Susko, and J. P. Bielawski, 2019 Looking for darwin in genomic sequences: Validity and success depends on the relationship between model and data, pp. 399–426 in *Methods in Molecular Biology*, Humana Press Inc.

König, C., and F. Weick, 2008 *Owls of the World*. Christopher Helm. A & C Black, London.

Kopania, J., 2016 Acoustics Parameters theWings of Various Species of Owls. Inter-Noise 2016 2868–2876.

Kreitman, M., and H. Akashi, 1995 Molecular evidence for natural selection. Annu. Rev. Ecol. Syst. 26: 403–422.

Kronfeld-Schor, N., and T. Dayan, 2003 Partitioning of Time as an Ecological Resource. Annu. Rev. Ecol. Evol. Syst. 34: 153–181.

Lisney, T. J., A. N. Iwaniuk, M. V. Bandet, and D. R. Wylie, 2012 Eye Shape and Retinal Topography in Owls (Aves: Strigiformes). Brain. Behav. Evol. 79: 218–236.

Macías-Duarte, A., C. J. Conway, G. L. Holroyd, H. E. Valdez-Gómez, and M. Culver, 2019 Genetic Variation among Island and Continental Populations of Burrowing Owl (Athene cunicularia) Subspecies in North America. https://doi.org/10.3356/JRR-18-00002 53: 127–133.

Martin, G., L. M. Rojas, Y. Ramírez, and R. McNeil, 2004 The eyes of oilbirds (Steatornis caripensis): Pushing at the limits of sensitivity. Naturwissenschaften

91: 26–29.

Mendelsohn, B., B. Bedrosian, S. M. Love Stowell, R. B. Gagne, M. E. F. LaCava *et al.*, 2020 Population genomic diversity and structure at the discontinuous southern range of the Great Gray Owl in North America. Conserv. Genet. 21: 693–706.

Nielsen, R., 2005 Molecular Signatures of Natural Selection. Annu. Rev. Genet. 39: 197–218.

Park, O., 1940 Nocturnalism--The Development of a Problem. Ecol. Monogr. 10: 485–536.

Pasierowska, R. L., 2017 “Screech owls allus holler 'round the house before death”: Birds and the souls of black folk in the 1930s American South. J. Soc. Hist. 51: 27–46.

Passarotto, A., D. Parejo, A. Cruz-Miralles, and J. M. Avilés, 2018a The evolution of iris colour in relation to nocturnality in owls. J. Avian Biol. 49:.

Passarotto, A., D. Parejo, V. Penteriani, and J. M. Avilés, 2018b Colour polymorphism in owls is linked to light variability. Oecologia 187: 61–73.

Payne, R. S., 1971 Acoustic location of prey by barn owls (Tyto alba). J. Exp. Biol. 54: 535–573.

Pei, Y., M. Valcu, and B. Kempenaers, 2018 Interference competition pressure predicts the number of avian predators that shifted their timing of activity. Proc. R. Soc. B Biol. Sci. 285: 20180744.

Ponder, J. B., and M. M. Willette, 2015 Strigiformes, pp. 189–198 in *Fowler’s Zoo and Wild Animal Medicine, Volume 8*, edited by R. E. Miller and M. E. Fowler. W.B. Saunders, Philadelphia, Pennsylvania, United States.

Roll, U., T. Dayan, and N. Kronfeld-Schor, 2006 On the role of phylogeny in determining activity patterns of rodents. Evol. Ecol. 20: 479–490.

Sagar, P., P. Teotia, A. D. Sahlot, and H. C. Thakur, 2017 An analysis of silent flight of

owl. Mater. Today Proc. 4: 8571–8575.

Soberón, J., 2007 Grinnellian and Eltonian niches and geographic distributions of species. Ecol. Lett. 10: 1115–1123.

Vitti, J. J., S. R. Grossman, and P. C. Sabeti, 2013 Detecting Natural Selection in Genomic Data. Annu. Rev. Genet. 47: 97–120.

Walls, G. L., 1942 *The vertebrate eye and its adaptive radiation*. The Cranbrook Institute of Science. Hafner Publishing Company, Bloomfield Hills, Michigan.

Weinstein, K., 1990 The Owl in art, myth, and legend. 144 p.

Wink, M., A.-A. El-Sayed, H. Sauer-Gürth, and J. Gonzalez, 2009 Molecular Phylogeny of Owls (Strigiformes) Inferred from DNA Sequences of the Mitochondrial Cytochrome *b* and the Nuclear *RAG-1* gene. Ardea 97: 581–591.

Wu, Y., E. A. Hadly, W. Teng, Y. Hao, W. Liang *et al.*, 2016 Retinal transcriptome sequencing sheds light on the adaptation to nocturnal and diurnal lifestyles in raptors. Sci. Rep. 6: 1–11.

Yang, Z., 2007 PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24: 1586–1591.

Yang, Z., and R. Nielsen, 2000 Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol. Biol. Evol. 17: 32–43.

Chapter 1

Genomic Evidence for Sensorial Adaptations to a Nocturnal Predatory Lifestyle in Owls

Pamela Espíndola-Hernández[1], Jakob C. Mueller[1], Martina Carrete[2], Stefan Boerno[3], and Bart Kempenaers[1]

[1]Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany
[2]Department of Physical, Chemical and Natural Systems, Universidad Pablo de Olavide, Sevilla, Spain
[3]Sequencing Core Facility, Max Planck Institute for Molecular Genetics, Berlin, Germany

Abstract

Owls (Strigiformes) evolved specific adaptations to their nocturnal predatory lifestyle, such as asymmetrical ears, a facial disk, and a feather structure allowing silent flight. Owls also share some traits with diurnal raptors and other nocturnal birds, such as cryptic plumage patterns, reversed sexual size dimorphism, and acute vision and hearing. The genetic basis of some of these adaptations to a nocturnal predatory lifestyle has been studied by candidate-gene approaches but rarely with genome-wide scans. Here, we used a genome-wide comparative analysis to test for selection in the early history of the owls. We estimated the substitution rates in the coding regions of 20 bird genomes, including 11 owls of which five were newly sequenced. Then, we tested for functional over-representation across the genes that showed signals of selection. In the ancestral branch of the owls, we found traces of positive selection in the evolution of genes functionally related to visual perception, especially to phototransduction, and to chromosome packaging. Several genes that have been previously linked to acoustic perception, circadian rhythm, and feather structure also showed signals of an accelerated evolution in the origin of the owls. We discuss the functions of the genes under positive selection and their putative association with the adaptation to the nocturnal predatory lifestyle of the owls.

Published as:

Espíndola-Hernández, Pamela, Jakob C. Mueller, Martina Carrete, Stefan Boerno, and Bart Kempenaers. "Genomic evidence for sensorial adaptations to a nocturnal predatory lifestyle in owls." Genome biology and evolution 12, no. 10 (2020): 1895-1908.

Genomic Evidence for Sensorial Adaptations to a Nocturnal Predatory Lifestyle in Owls

Pamela Espíndola-Hernández[1,*], Jakob C. Mueller[1], Martina Carrete[2], Stefan Boerno[3], and Bart Kempenaers[1]

[1]Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany

[2]Department of Physical, Chemical and Natural Systems, Universidad Pablo de Olavide, Sevilla, Spain

[3]Sequencing Core Facility, Max Planck Institute for Molecular Genetics, Berlin, Germany

*Corresponding author: E-mail: pespindola@orn.mpg.de.

Accepted: 5 August 2020

Data deposition: Raw sequences have been deposited at the short-read archive (SRA) of the NCBI db under the BioProject PRJNA592858.

Abstract

Owls (Strigiformes) evolved specific adaptations to their nocturnal predatory lifestyle, such as asymmetrical ears, a facial disk, and a feather structure allowing silent flight. Owls also share some traits with diurnal raptors and other nocturnal birds, such as cryptic plumage patterns, reversed sexual size dimorphism, and acute vision and hearing. The genetic basis of some of these adaptations to a nocturnal predatory lifestyle has been studied by candidate gene approaches but rarely with genome-wide scans. Here, we used a genome-wide comparative analysis to test for selection in the early history of the owls. We estimated the substitution rates in the coding regions of 20 bird genomes, including 11 owls of which five were newly sequenced. Then, we tested for functional over-representation across the genes that showed signals of selection. In the ancestral branch of the owls, we found traces of positive selection in the evolution of genes functionally related to visual perception, especially to phototransduction, and to chromosome packaging. Several genes that have been previously linked to acoustic perception, circadian rhythm, and feather structure also showed signals of an accelerated evolution in the origin of the owls. We discuss the functions of the genes under positive selection and their putative association with the adaptation to the nocturnal predatory lifestyle of the owls.

Key words: night-active, raptor, genome-wide analysis, comparative genomics, positive selection, Strigiformes.

Significance

Beneficial mutations are fixed by positive selection, and the process can be analyzed by comparing genome sequences of different related species. Here, we aim to trace signals of positive selection in the early history of owls. The owls are the only nocturnal raptors among birds with specific adaptations such as acute vision and hearing and silent flight. The genetic basis of these adaptations has been studied in single candidate genes but rarely with genome-wide scans. We found traces of positive selection in the early evolution of owls mostly in genes that are functionally related to visual and acoustic perception.

Introduction

The owls (Strigiformes) are the only avian lineage of nocturnal raptors. They separated from their sister group, the diurnal Coraciimorph clade, in the Paleocene (Prum et al. 2015), and divided into two families, Strigidae and Tytonidae (Wink et al. 2009; Ponder and Willette 2015). Presumably, the past diversification of owls was associated with a concurrent radiation of small mammals, which led to an expansion of prey availability in the nocturnal niche (Feduccia 1999). Owls evolved an interesting set of raptorial adaptations to the nocturnal niche. Some of those adaptations are shared with other diurnal raptors, whereas others are shared with nocturnal bird species that are not raptors.

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

Like other raptors, owls have cryptic plumage coloration, reversed sexual size dimorphism as well as acute vision and hearing (del Hoyo et al. 1999; Duncan 2013). With other nonraptor nocturnal birds, such as kiwis and oilbirds, owls share an enhanced visual sensitivity but lost color discrimination to some extent (Martin et al. 2004; Corfield et al. 2015; Emerling 2018). Owls have binocular vision, large tubular eyes, and a duplex retina dominated by rods that characterize a typical nocturnal eye (Walls 1942; Fite 1973). Owls also have unique traits that are clearly adaptive for nocturnal raptors. For instance, many species have asymmetrical ears and a facial disk, which improves their ability to find prey in darkness by hearing (Payne 1971). Additionally, the feathers of owls have a serrated leading edge, a fringe trailing edge, and very fine barbules compared with other birds (Sagar et al. 2017). These features make the feathers softer and allow silent flight (Kopania 2016; Sagar et al. 2017), which presumably also improves hunting success.

Independent of timing of activity, a raptorial lifestyle may involve adaptations for hunting, including visual acuity and forward-looking eyes, claws, and curved beaks. It is likely that these adaptations have been retained among landbirds (Telluraves) from their common raptorial ancestor (Hackett et al. 2008; Jarvis 2014; Prum et al. 2015; McClure et al. 2019). Expected adaptations of diurnal raptors are likely related to the maintenance of the visual system and photoresponse recovery (Wu et al. 2016), blood circulation, nervous system development, olfaction, and beak development (Zhan et al. 2013; Wu et al. 2016; Zhou et al. 2019).

A nocturnal lifestyle generally involves adaptations related to the sensory system, circadian rhythms, and plumage color patterns. For example, previous studies reported associations between diel activity patterns and eye shape and size (Hall and Ross 2007; Lisney et al. 2012), size of olfactory bulbs (Healy and Guilford 1990), neural visual pathways (Gutiérrez-Ibáñez et al. 2013), and iris coloration (Passarotto et al. 2018). In birds, the genetic basis for nocturnal adaptations has mostly been studied in the visual system of two nocturnal species, the kiwi *Apteryx mantelli* and the barn owl *Tyto alba* (Borges et al. 2015, 2019; Le Duc et al. 2015; Emerling 2018). Le Duc et al. (2015) showed that adaptations to nocturnality in kiwis are associated with an increase in the olfactory receptor repertoire and an accumulation of evolutionary changes in genes related to color vision, mitochondrial function, and energy expenditure. The avian visual system is characterized by tetrachromatic vision and dense retinas (Yokoyama 2000; Bowmaker 2008; Davies et al. 2012) and relatively large eyes (Howland et al. 2004; Hall and Ross 2007). The avian retinas have six classes of photoreceptor cells: one rod, four single cones, and one double cone (Hart and Hunt 2007; Bowmaker 2008). The membranes of these photoreceptors contain specific photopigments, that is, light-sensitive molecules formed by an opsin and a chromophore. The opsins can be divided into five subfamilies: visual opsins, melanopsins, pineal opsins, vertebrate nonvisual opsins, and photoisomerases (Terakita 2005; Lamb et al. 2007). The visual opsins trigger the phototransduction cascade after light stimulation in the membrane of photoreceptor cells. Cones and rods use different sets of opsins and phototransduction molecules and are specialized in photopic (bright light conditions) and scotopic (dim light conditions) vision, respectively (Lamb et al. 2016). Thus, the cones provide acute and color vision, and the rods are highly sensitive to light.

Diel activity patterns are highly constrained by phylogenetic history (Anderson and Wiens 2017). The majority of the extant avian species are diurnal, but the diel activity pattern of the common ancestor of all birds is unknown. Two hypotheses have been proposed. The first hypothesis is that the avian common ancestor had a diurnal lifestyle, which is supported by a vast amount of morphological and genetic evidence (Schmitz and Motani 2011; Anderson and Wiens 2017). For instance, the ancestral bird probably had similar color discrimination as the diurnal modern birds, because there is no evidence for any global loss or gain of genes related to color vision among birds (Zhang et al. 2014; Borges et al. 2015). The second hypothesis proposes that the common ancestor was nocturnal, with a transition to cathemeral (active during day and night), similar to mammals (Wu and Wang 2019). Assuming a diurnal common ancestor, nocturnality evolved many times independently in parrots, kiwis, oilbirds, nightjars, and owls (Ericson et al. 2006; Hackett et al. 2008; Braun and Huddleston 2009; Le Duc et al. 2015). This should be paralleled by the accumulation of genetic changes related to nocturnality on each of the ancestral branches of the nocturnal clades.

The genetic basis of adaptations to a nocturnal and raptorial lifestyle has been studied by candidate- gene approaches but rarely with genome-wide scans, and the results have been mostly related to the visual system. Thus, how evolution shaped the specific combination of traits observed in the owls remains poorly understood. Here, we aim to answer the following questions. 1) What is the general role of positive selection in the early adaptive history of Strigiformes? 2) Which genes and associated functions evolved under positive selection in the owls? 3) Are the positively selected genes associated with adaptation to the night-active or the predatory lifestyle of the owls? We used substitution rates to test for selection in the early history of Strigiformes in a genome-wide comparative analysis, using 20 species of birds including 11 owls of which five were newly sequenced for this study. Complementing the search for single, genome-wide significant genes, we used overrepresentation analyses to functionally interpret groups of genes that showed any signal (including weak signals) of positive selection.

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

Materials and Methods

Study Species and Reference Genome

This study includes genomes from 20 bird species that were selected to produce a balanced tree around the ancestral branch of the owls: 11 Strigiformes, two Accipitriformes, four Coraciimorphae, one Falconiform, one Passeriform, and one Galliform. In contrast to the mostly nocturnal Strigiformes, all other species included in this study are diurnal. We included Coraciimorphae as the sister group to the owls, Accipitriformes and Falconiformes as diurnal raptors, and the Passeriformes and Galliformes because of their high genome sequence quality (Zhang et al. 2014).

All genomes included in this study were aligned using the assembly of *Athene cunicularia* (burrowing owl, assembly athCun1) as reference genome (Mueller et al. 2018). The burrowing owl is a peculiar species among the owls, being diurnal and gregarious, which implies that its genome may contain some unique features and may lack some of the genes that are present in the rest of the owls. However, drastic gene loss is unlikely considering the short evolutionary history of burrowing owls. Moreover, we used a selection test ("ω test") that is based on the codon sequences that are common among all the compared species, including non-owls. Therefore, the important criteria to avoid loss of information are the assembly completeness and the continuity of the annotated gene sequences to construct the multispecies codon alignment for the selection test (see below). We therefore used athCun1 as the reference, because it is the highest-quality owl genome assembly that was available in terms of completeness and N50 criterion: athCun1 has longer scaffolds, the assembly is more continuous and more complete. The assembly contained 94.8% of complete Benchmarking Universal Single-Copy Orthologs (BUSCO v. 4.0.6) based on the avian database of 8,338 genes (BUSCO summary in supplementary file 1, Supplementary Material online) (Simão et al. 2015). The reference genome was annotated by the NCBI Eukaryotic Genome Annotation Pipeline (NCBI *Athene cunicularia* Annotation Release 100; NCBI Assembly Accession GCA_003259725.1).

Protocol A: Sequencing and Read Mapping to Reference

The following eight owl genomes were sequenced and mapped to athCun1: *Bubo scandiacus* (snowy owl), *Strix uralensis* (ural owl), *Strix nebulosa* (great gray owl), *Athene noctua* (little owl), *Surnia ulula* (northern hawk-owl), *Bubo bubo* (Eurasian eagle-owl), *Asio otus* (long-eared owl), and *Asio flammeus* (short-eared owl). The DNA was obtained from blood samples stored in ethanol. For the majority of the samples, we extracted the DNA using the QuickPure kit (Macherey-Nagel) applying a predigestion with Proteinase K in Digsol buffer. After initial quality control, we used the Kapa HyperPrep DNA kit (Roche) to prepare 200 to 300 bp insert paired-end libraries. Then, the majority of the samples were sequenced with an Illumina HiSeq4000 in paired-end, 150 bp mode (Sequencing Core Facility of the Max Planck Institute for Molecular Genetics, Berlin, Germany), yielding between 74 and 141 million fragments (read pairs) mapped per individual sample (15.2–26.8× genome coverage). The samples of *Athene noctua*, *Asio otus*, and *Asio flammeus* were extracted using the phenol–chloroform method; the libraries were prepared using Illumina's TruSeq DNA protocol and sequenced on an Illumina HiSeq2500. After sequencing, we used the aligner software BWA-MEM v.0.7.17-r1188 (Li 2013) to map the reads of each species against the reference genome. Parameters are detailed in section 1.1 of the supplementary file 1, Supplementary Material online.

Protocol B: Genome-Scale Sequence Mapping to Reference

For the following species, we downloaded the genome assemblies from NCBI (accession numbers and details in supplementary table S1 in file 1, Supplementary Material online): *Strix occidentalis* (spotted owl), *Tyto alba* (barn owl), *Falco peregrinus* (peregrine falcon), *Taeniopygia guttata* (zebra finch), *Picoides pubescens* (downy woodpecker), *Apaloderma vittatum* (bar-tailed trogon), *Leptosomus discolor* (cuckoo roller), *Colius striatus* (speckled mousebird), *Cathartes aura* (Turkey vulture), *Haliaeetus leucocephalus* (bald eagle), and *Gallus gallus* (red junglefowl). We downloaded the assemblies as FASTA files and aligned them to the reference genome using LAST v. 921 (Kiełbasa et al. 2011). Parameters are detailed in section 1.1 of the supplementary file 1, Supplementary Material online. The overlapping regions were resolved with SingleCov2 (Multiz-tba.012109, Blanchette et al. 2004) with default parameters, and the final alignment was created using maf-convert (LAST v. 921) and samtools (Li et al. 2009).

The pairwise sequence alignments produced by both protocols, a and b, were similar among owls in terms of gaps and percentage of the reference genome covered (supplementary table S1 in file 1, Supplementary Material online).

Multispecies Codon Alignment

After the alignment of each species to the reference, the general workflow consisted of six steps to produce a multispecies codon alignment for each annotated gene in the reference genome (supplementary fig. S1 in file 1, Supplementary Material online). 1) Piling up the reads in the coding regions using samtools. 2) Variant calling with bcftools (Danecek and McCarthy 2017). 3) Producing the consensus sequence using default parameters with bcftools, choosing the allele with more reads or better mapping quality in case of heterozygous sites. 4) Masking all the sites with zero read coverage. Note that the species with lower read coverage or those more distantly related to the reference had a higher

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

percentage of masked sites (see supplementary table S1 in file 1, Supplementary Material online). 5) Extracting the sequence of each gene from the consensus sequence of each species and concatenate all in a single, multispecies FASTA file with bedtools (Quinlan and Hall 2010; Dale et al. 2011). 6) Running a multispecies codon alignment for each gene using MACSE (Ranwez et al. 2011). We used MACSE because it aligns protein-coding gene sequences correcting for potentially erroneous frameshifts (e.g., indels smaller than triplets) without disrupting the underlying codon structure.

Finally, we inferred the percentage of low-quality regions of each multispecies gene alignment using BMGE v. 1.12 (Criscuolo and Gribaldo 2010). After removing codon sites with missing data (gaps) in one of the species (similar to the procedure of the ω tests; see below) we identified sites with a smoothed entropy-score higher than 0.5. These highly variable regions were considered as low-quality regions, potentially caused by misalignments or sequencing errors (small indels). Genes (multispecies alignments [MSAs]) with any low-quality region were excluded for further analyses ($N = 10$ genes). We used AMAS (Borowiec 2016) to quantify the percentage of variable sites.

Phylogenetic Tree

The phylogenetic tree of the selected species was based on information from the phylogeny of all birds (Prum et al. 2015) and the phylogeny of owls (Wink et al. 2009). The subset of species was extracted, keeping the topology and ignoring the branch lengths, using the software Mesquite version 3.40 (Maddison W and Maddison D 2018). Figure 1 shows the unrooted tree used for the selection tests in CodeML.

Selection Tests

To assess past selection on each gene at the ancestral branch of the owls, we estimated the nonsynonymous to synonymous substitution rate ratio ($\omega = dN/dS$; for a review, see Nielsen [2005]). This ratio measures the direction and magnitude of selection on protein-coding genes. In the rest of the text, we simply refer to it as the "ω test" (for the test) and the "ω value" (for the estimated value).

The ω value of each gene can be calculated for specific branches of a phylogenetic tree and reflects the evolutionary history of that branch, with $\omega < 1$ indicating purifying selection, $\omega = 1$ neutral evolution, and $\omega > 1$ positive selection. We tested ω for the ancestral branch of Strigiformes using a maximum-likelihood method implemented in the CodeML program in PAML 4.9h (Yang 2007), using the branch model (Yang 1998) and the branch-site model (Yang and Nielsen 2002; Zhang et al. 2005; Yang and dos Reis 2011). For both models, we excluded all the sites with missing data in the MSA, and we defined the ancestral branch of Strigiformes as the foreground (label "#" in fig. 1).

When a gene had a ω value >1 in the ancestral branch of the owls, we considered this as evidence of strong positive selection, that is, the nucleotide changes in this gene were adaptive for the ancestral owl. When a gene has a ω value that lies between the background value and 1 ($\omega_{background} < \omega_{foreground} < 1$), two interpretations are possible. First, it can indicate relaxed purifying selection, suggesting a loss of function of that gene. Second, it can indicate weak positive selection acting only in a few sites or for a short period. We cannot distinguish between these two options. The majority of the protein-coding sequences are conserved during most of their evolutionary history ($\omega < 1$), but positive selection acting only in few sites in the foreground branch would increase the average ω value of the foreground above the background (Toll-Riera et al. 2011; Nery et al. 2013). We identified genes with ω values <1, but with a significantly higher value on the ancestral branch of the owls than in the background in a separate category of "weak positive selection or relaxed purifying selection." We also used the term "accelerated substitution rate" to concisely describe the ω values of genes in this category in combination with the category of "strong positive selection."

The branch model tests a null hypothesis (H_0), assuming all branches of the phylogenetic tree have the same ω ratio, against an alternative hypothesis (H_1), where the labeled branch of interest ("foreground") has a different ω ratio (ω_1) than all other branches of the phylogenetic tree (ω_0, "background").

The branch-site model tests for positive selection among codon sites in the ancestral branch of the owls. In this model, ω is allowed to vary between foreground and background branches as well as among sites on each gene, under both the null (H_0) and the alternative hypothesis (H_1). This represents a more realistic and complex scenario where different codon sites of the same gene can evolve under different selection, and selection can also differ among the branches in the phylogeny. The model estimates the proportion of sites that have certain combinations of ω values for the foreground and background. The estimated foreground and background ω values for each site are then divided into three categories: $\omega < 1$, $\omega = 1$, and $\omega > 1$ (referred to as ω_0, ω_1, and ω_2). Under the H_0, no ω is allowed to be larger than 1, both in the foreground and the background, meaning that positive selection is not allowed at any site. Under the (H_1) , some ω values (at some sites) can be larger than one in the foreground branch, representing the category of positively selected sites (for a summarized explanation of this model see https://selectome.unil.ch/cgi-bin/methods.cgi, last accessed August 19, 2020). Thus, this model tests a null hypothesis (H_0), where the foreground cannot have positive selection at any site, against an alternative hypothesis (H_1), where the foreground lineage is allowed to have a proportion of sites evolving under positive selection (Yang and Nielsen 2002; Yang and dos Reis 2011).

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

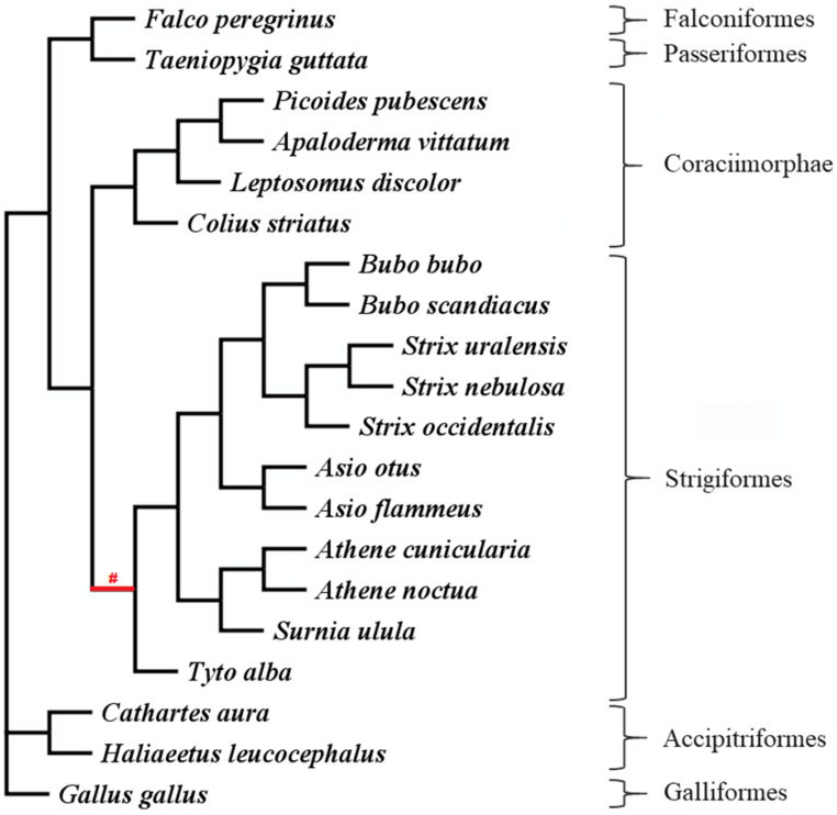


Fig. 1.—Unrooted species tree without branch lengths extracted from Prum et al. (2015) and Wink et al. (2009). The ω tests were based on this tree, whereby the red symbol "#" indicates the foreground branch in contrast to the rest of the branches (background).

For each model, we tested whether the alternative hypothesis is more likely than the null hypothesis using the likelihood ratio test (LRT) statistic, that is, twice the difference in log-likelihoods between the two hypotheses ($LRT = 2 \times [\ln L_{H1} - \ln L_{H0}]$), which is compared with a χ^2 distribution with one degree of freedom. Hence, the LRT was considered significant when >3.8415. We excluded genes with significant LRT values but estimated foreground ω values >500 (24% of genes for the branch model and 40% of genes for the branch-site model) from further analyses, because such high ω values in CodeML indicate a synonymous substitution rate estimate close to 0, which means that ω cannot be reliably calculated (Yang and dos Reis 2011). The results of all genes with nominal significant ω tests and of all a priori defined candidate genes (including nonsignificant results) are in supplementary file 2, Supplementary Material online.

We identified 27,746 annotated isoforms for the protein-coding genes in the athCun1 reference genome. We applied filters to these annotated isoforms before and after the tests, to select gene sequences that fulfilled the requirements for the ω tests and further functional analyses. First, we selected the longest isoform for each protein-coding gene (13,841 unique genes) with at least 20 codons in the MSAs and with at least one variant site, and without regions of potential misalignment problems as measured by high-entropy blocks. This yielded 12,160 genes for the ω tests. Second, we filtered out the genes with estimated ω values >500 on the ancestral branch of the owls ($N = 629$ genes for the branch model and 231 for the branch-site model). After the filters, we applied a false-discovery-rate (FDR) correction for multiple testing for each model. The raw and corrected P values are included in supplementary file 2, Supplementary Material online. For further analyses, we considered three categories of genes with a significant ω test: i) those showing strong positive selection signals according to the branch model, ii) those with weak positive or relaxed purifying selection according to the branch model, and iii) those with positive selection according to the branch-site model. We refer to these as list i, ii, and iii.

As an alternative to the branch-site model, we applied the aBSREL model (Smith, 2015) of the HyPhy package to all nominal significant genes from the branch-site model (list iii) to search for selection signals specific for owls. We ran this model with two settings: with and without the a priori specified foreground. The first setting is similar to the one used in CodeML and we used it to compare the CodeML results. The second setting explores all the branches of the tree and then selects the genes that have a significant signal in the ancestral branch of the owls, but not in any other branch. The significance threshold is corrected for the number of branches tested.

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

Multinucleotide mutations within codons are known to cause false inferences of the branch-site model (Venkat et al. 2018). Thus, we quantified the proportion of codons with multiple differences (CMDs) between owls and chicken and used this measure as a proxy for codons with multiple substitutions in the ancestral owl branch. First, we read each MSA as a matrix using R 3.5.0 (R Core Team 2018) and the R package "ape" (Paradis et al. 2004; Paradis and Schliep 2019). Using the MSA as a matrix allows us to compare the nucleotide sites between each owl species and the chicken sequence. Starting from the first nucleotide, every three consecutive nucleotides are considered a codon site. A codon site of the alignment that contained more than one difference for any of the owl species with the chicken sequence was counted as one CMD. We tested whether there is an effect of the proportion of CMDs on the outcome of the ω tests (*t*-test) and estimated ω values (correlation) for the branch-site model (supplementary fig. S2*a* and *b* in file 1, Supplementary Material online). We applied a two-sample permutation *t*-test using the R package "Deducer" (Fellows 2012) and a Kendall's rank correlation test using the R package "base." We found an effect on the significance level (Welsh *t*-statistic, $t = -9.084$, *P* value < 0.001), but no effect on the estimated ω values ($\tau_B = 0.0096$, *P* value $= 0.22$) in the branch-site model (supplementary fig. S2, Supplementary Material online).

Overrepresentation Analysis

We performed overrepresentation analyses to test whether the gene sets of the three ω test categories (list i–iii) and the genome-wide category were enriched for a particular biological function or a metabolic pathway. We used two software packages for this analysis: ClueGO v2.5.4 plug-in (Bindea et al. 2009) in Cytoscape (Shannon et al. 2003) and the R package GOfuncR (Grote 2019). Both packages use a standard candidate-versus-background hypergeometric enrichment test with a custom functional annotation database as the background gene list. We made this custom annotation database for all genes occurring in the reference athCun1, combining human (org.Hs.eg.db) and chicken (org.Gg.eg.db) annotations of gene ontologies (GOs) and KEGG pathways (Huber et al. 2015; Pagès et al. 2019).

ClueGO reduces the redundancy among the GO terms by grouping the significantly enriched GO terms based on the shared genes (Bindea et al. 2009). Each functional group in the graph has a leading GO term, which is the most significant term. ClueGO uses a Bonferroni correction for multiple testing by using the number of genes in the gene sets as a proxy for the number of tests. GofuncR uses a more conservative method for multiple testing correction, using family-wise error rates (FWERs) based on 10,000 random permutations of the gene-associated variables (candidate-versus-background genes). In general, the results were consistent between the two methods.

Candidate Genes Related to the Nocturnal Predatory Lifestyle of Owls

In addition to the data-driven, genome-wide approach, we specifically tested for positive selection on an a priori defined list of candidate genes that are likely related to the nocturnal predatory lifestyle of owls based on previous studies (information-driven or candidate-gene approach). We used 1) genes proposed as candidates by previous studies (Le Duc and Schöneberg 2016; Wu et al. 2016) and 2) genes found by key-word searching on GO terms in AmiGO2 (Carbon et al. 2009). The included keywords were: vision, eye, ear, hearing, vestibular (because the vestibular system is part of the inner ear and brings balance and spatial orientation), circadian, sleep (pooled together with the circadian genes), and keratin (as feathers are made of β-keratins). We identified 253 candidate genes in the reference assembly athCun1, listed in supplementary table S2 of the file 1, Supplementary Material online, in the following four categories: vision ($N = 104$), hearing ($N = 69$), circadian rhythm ($N = 67$), and feather keratin ($N = 13$).

The vision category includes the opsin genes. We searched in athCun1 for all genes in the opsin gene family that are annotated in *Gallus gallus* (galGal5), using BLAST reciprocal best hits with the web-based tool Galaxy (https://usegalaxy.eu/, last accessed July 17, 2020) (Afgan et al. 2018). Nine opsin genes were found in athCun1: *RHO*, *OPN1MSW*, *OPN3* (opsin-3, or encephalopsin), *OPN4* (melanopsin), *OPN4*-1 (melanopsin-like), *OPN5* (opsin-5, or neuropsin), *OPNVA* (vertebrate ancient opsin), *RGR* (retinal G-protein-coupled receptor), and *RRH* (retinal pigment epithelium-derived rhodopsin homolog).

We based our interpretation and discussion of the functions of the relevant candidate genes and the genes that are part of the networks from the overrepresentation analysis on information found in the following databases: NCBI (https://www.ncbi.nlm.nih.gov/search/, last accessed July 17, 2020), GeneCards (Safran et al. 2010) (https://www.genecards.org/, last accessed Juyl 17, 2020), AmiGO2 (Carbon et al. 2009) (http://amigo.geneontology.org/amigo, last accessed July 17, 2020), and reactome (Fabregat et al. 2018).

Results

Data-Driven Approach: Genes with Genome-Wide Significant Selection Signals in the Owl Ancestor

After correction for multiple testing across all tested genes, 21 genes of the branch model (list i and ii) and two genes of the branch-site model (list iii) were significant with a 5% FDR (supplementary table S3 in file 2, Supplementary Material online). Considering this set of genes (22 in total from lists i–iii),

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

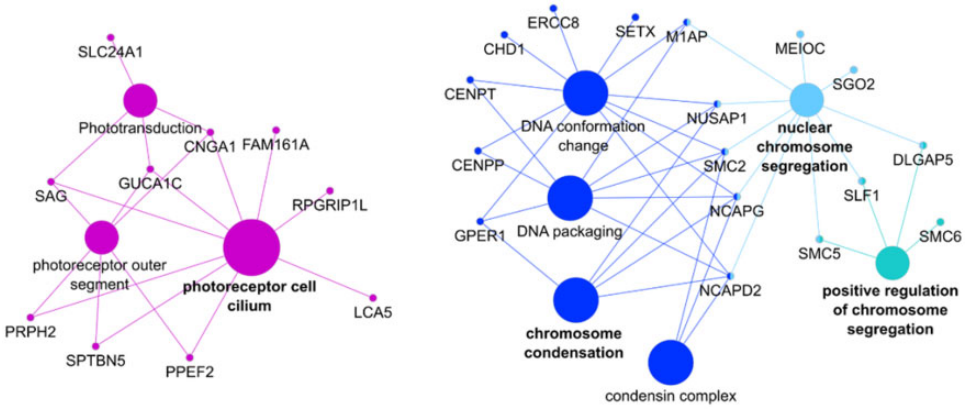


Fig. 2.—Functional overrepresentation of GO terms and KEGG pathways among the genes with signals of strong positive selection (list i). The GO terms were clustered in four groups by the ClueGO software (shown in different colors). Each group can contain several GO terms with shared genes. There are two major groups: ten genes are related to the visual system (purple group) with "photoreceptor cell cilium" as leading GO term, and 17 genes mostly related to "chromosome condensation" (blue, light blue, and turquoise).

one GO term, "detection of stimulus involved in sensory perception," was significantly enriched (FWER = 0.001).

Data-Driven Approach: Genome-Wide Functional Overrepresentation of Genes with Nominal Significant Selection Signals in the Owl Ancestor

The ω test based on the branch model was nominal significant for 486 out of 11,613 tested genes (supplementary table S4 in file 2, Supplementary Material online). We differentiated between genes with a signal of strong positive selection on the foreground (list i, with $\omega_0 \leq 1 < \omega_1$; $N = 199$ genes), and genes with a signal of weak positive or relaxed purifying selection on the foreground (list ii, with $\omega_0 < \omega_1 < 1$; $N = 287$ genes). The branch-site model identified 123 genes with a signal of positive selection on specific codon sites in the foreground (list iii, with $\omega_2 > 1$; supplementary table S5 in file 2, Supplementary Material online). Supplementary tables S4 and S5 in file 2, Supplementary Material online, show the raw and FDR corrected *P* values of all nominal significant results included in the overrepresentation analysis. We identified 42 genes that were in common between both models (28 shared genes in lists i and iii, and 14 in lists ii and iii). The tests based on the aBSREL model are significant for 59% of the significant branch-site tests (73 tests out of the 123 tested genes in list iii when the model was run with the a priori specified foreground; supplementary table S6 in file 2, Supplementary Material online). After running the model without the a priori specified foreground, we found nine genes with a positive selection signal specific for the ancestral branch of the owls (supplementary table S6 in file 2, Supplementary Material online). The genes with a signal of strong positive selection on the foreground (list i) showed enrichment in four functional groups (fig. 2). These groups form two major networks functionally associated with photoreceptor cells and chromosome condensation. The GO terms that were overrepresented among the genes that evolved under weak positive or relaxed purifying selection in the foreground (list ii) clustered into seven functionally enriched groups (fig. 3). Most of these groups formed a highly connected large network associated with functions of sensory perception (visual and auditory) and plasma membrane bounded cell projection. Another smaller isolated group related to the function of DNA conformation change. The GO terms that were overrepresented among the genes that evolved under positive selection in specific sites of the foreground (list iii) clustered in four functional groups (fig. 4). Most of the functions of these groups are associated with microtubules, including "mitotic nuclear division" and "sperm flagellum." The detailed results of all overrepresentation analyses with statistical support values are shown in supplementary file 1, Supplementary Material online. Functional groups and single GO/KEGG terms from the overrepresentation analysis by ClueGO are listed in supplementary tables S7 and S8, Supplementary Material online, respectively; supplementary table S9, Supplementary Material online, shows the results of the overrepresentation analysis by GOfuncR. Irrespective of the different multiple testing correction methods, ClueGO and GofuncR produced consistent results.

Information-Driven Approach: Selection Signals in A Priori Defined Candidate Genes Related to the Nocturnal Predatory Lifestyle of Owls

From the 253 identified candidate genes in the annotation of the reference athCun1, 40 genes had significant ω tests, of which 37 were based on the branch model and three on the branch-site model (significant results are in table 1 and results for all candidate genes are in supplementary table S10 of the file 2, Supplementary Material online). Only one candidate gene (*RP1*) showed evidence for selection in both models (lists ii and iii). The total number of significant results is more than

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

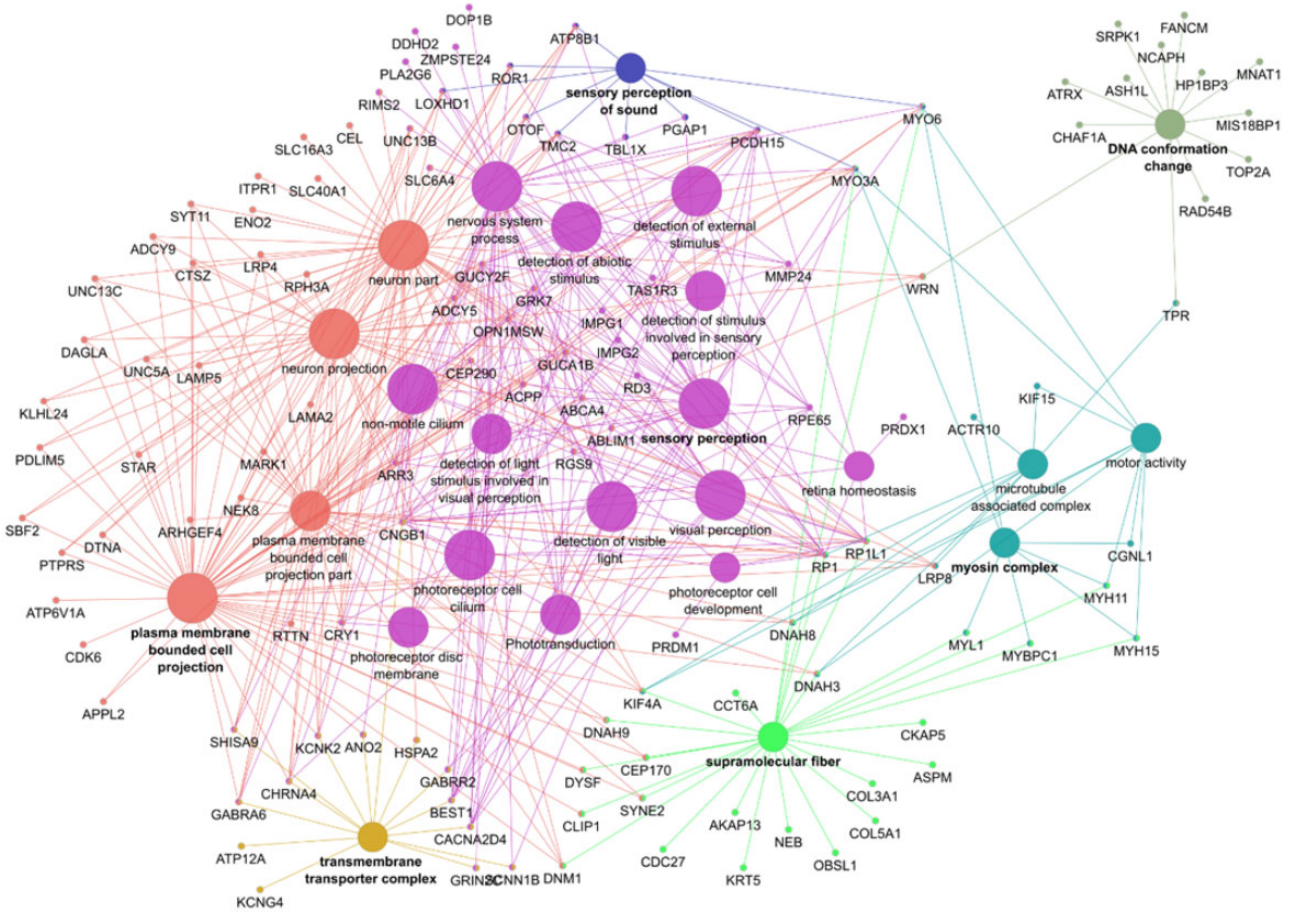


Fig. 3.—Functional overrepresentation of GO terms and KEGG pathways among the genes with a signal of weak positive or relaxed purifying selection (list ii). The GO terms were clustered in seven groups by the ClueGO software (shown in different colors). Each group can contain several GO terms with shared genes. The groups "plasma membrane bounded cell projection" (salmon) and "sensory perception" (purple) form the main part of the network with 87 genes. This main part also overlaps in several genes with the groups "sensory perception of sound" (blue), "transmembrane transporter complex" (gold), "myosin complex" (turquoise), and "supramolecular fiber" (lime). The functional group "DNA conformation change" with 13 genes forms another, more isolated cluster.

expected by chance ($253 \times 0.05 \sim 13$ expected significant tests).

Twenty-one candidate genes related to vision, ten to hearing, two to feather keratin, and five to circadian rhythm showed a higher ω value on the ancestral branch of the owls compared with the background (branch model; table 1, lists i and ii). Three candidate genes had a significant signal of positive selection at specific sites on the ancestral branch of the owls (branch-site model; table 1, list iii): one from the feather keratin category, and the other two from the visual system.

Discussion

Genome-Wide Significant Selection on Single Genes in the Owl Ancestor

Our study detected only 22 single genes with genome-wide significant signals of selection at the origin of the owls. The relatively low number of genes is expected because correction for multiple testing is strong in genome analyses and hence only the strongest single signals will pass as significant. These 22 genes encode mostly components of the membrane and are functionally associated with sensory perception (vision and sound), DNA condensation, and lipid metabolism. A formal functional enrichment analysis identified a single significant GO term "detection of stimulus involved in sensory perception" which contains the genes *TMC2*, *PCDH15*, *PPEF2*, and *CACNA2D4*. The first two genes play a role in auditory perception and the latter three in visual perception (NCBI gene db, GeneCards, and AmiGO2). *TMC2* is involved in mechanotransduction in cochlear hair cells of the inner ear, and *PCDH15* participates in the maintenance of normal retinal and cochlear function (GeneCards). *PPEF2* is expressed specifically in photoreceptors and the pineal gland and participates in phototransduction (GeneCards). The protein encoded by *CACNA2D4* plays an important role in the normal functioning of the retina and cardiac tissue because it is involved in

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

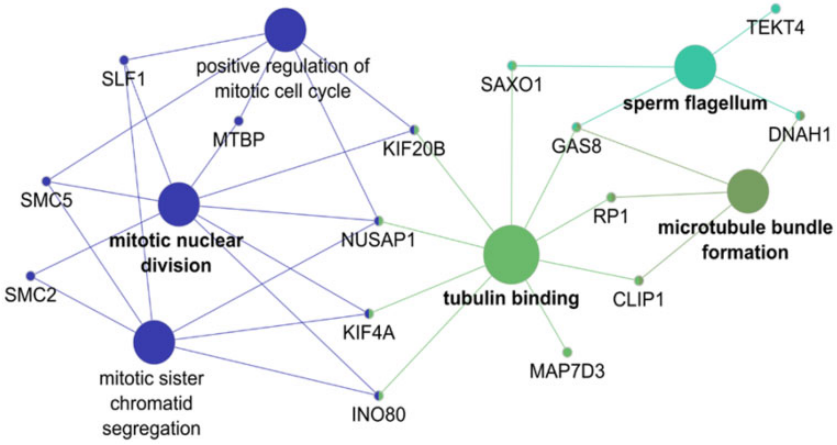


Fig. 4.—Functional overrepresentation of GO terms and KEGG pathways among the genes that show signals of positive selection on specific sites of the ancestral branch of the owls (list iii, branch-site model). The GO terms were clustered in four groups by the ClueGO analysis (shown as different colors). Each group can contain several GO terms with shared genes. The two major groups are related to "mitotic nuclear division" (blue) and to functions linked to microtubules and tubulin, including sperm flagellum (all other colors). Some genes, such as *RP1* (see also fig. 3), also participate in the development and maintenance of photoreceptors.

transmembrane transport of calcium (GeneCards). The description of the function and related diseases in humans for all 22 genome-wide significant genes is provided in supplementary table S3 in file 2, Supplementary Material online.

Genome-Wide Functional Overrepresentation of Genes with Nominal Selection Signals in the Owl Ancestor

Further analysis of all nominal significant signals (irrespective of their genome-wide significance and potentially comprising weaker selection signals) for enrichment of functionally related gene sets is recommended, because it can be informative if co-selection of functions or pathways is suspected (Mooney et al. 2014; Mueller et al. 2020). The major functional groups consistently found among the different tests were related to the processes of sensory perception (vision and hearing) and chromosome conformation.

Functional Overrepresentation Related to Vision

We found a strong and consistent enrichment of genes related to functions in photoreceptors among genes with an accelerated substitution rate in the origin of the owls (lists i and ii). Several of these genes are relevant for light perception, the first steps in phototransduction, dim-light vision, or the development and maintenance of the retina. Besides being part of the overrepresented visual-related functional groups, three of these genes are also genome-wide significant (*CACNA2D4*, *PCDH15*, and *PPEF2*), and one gene has an owl-specific signal of positive selection (*RP1*). The gene network related to functions in the plasma membrane (list ii, fig. 3) is highly connected to the network of photoreceptor functions (as shown by the many shared genes, fig. 3). This is probably because sensory perception depends on the transduction of the stimuli through reaction cascades on the plasma membrane of the photoreceptors.

The overrepresented functional group linked to photoreceptors comprises ten genes with evidence for strong positive selection ($\omega_{foreground} > 1$). Three of these genes have also been identified in previous studies on raptors (*CNGA1*, *SAG*, and *SLC24A1*), are expressed in rods, and play a role in phototransduction and recovery of the rod photoreceptors (Wu et al. 2016). The other seven genes with evidence for strong positive selection (*FAM161A*, *GUCA1C*, *LCA5*, *PPEF2*, *PRPH2*, *RPGRIP1L*, and *SPTBN5*) have not been described before as genes that may have played a role in the early diversification of the owls. Interestingly, Wu et al. (2016) did not find the gene *GUCA1C* in the transcriptome of owls. This gene encodes for a cone-specific protein that participates in photoresponse recovery and the authors suggested that this gene might have been lost or has become a nonfunctional pseudo-gene in the Strigidae. However, our results indicate that this gene is present in owls and has evolved under positive selection in the ancestral branch.

The overrepresented functional group linked to sensory perception includes 50 genes that evolved faster in the owl ancestor ($\omega_{background} < \omega_{foreground} < 1$). We found confirmatory evidence that four of these genes (*CNGB1*, *ABCA4*, *PCDH15*, and *BEST1*) have evolved faster in the owl ancestor, as reported in previous studies (Wu et al. 2016; Cho et al. 2019). The gene *RP1* is also present in the functional network of genes showing positive selection on specific sites (list iii) and links to a function for microtubules, which might be associated with the development and maintenance of photoreceptors.

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

Table 1
Candidate Genes that Evolved under Positive Selection or Relaxed Purifying Selection in the Ancestral Branch of the Owls

Gene Symbol	List	Candidate Gene Category	No. Codons Tested	% of Reference Gene Tested	Branch Model			Branch-Site Model			
					Alternative Hypothesis			Alternative Hypothesis			
					ω_0	ω_1	LRT Statistic	ω_0	ω_1	ω_2	LRT Statistic
ABCA4	ii	Vision	2,236	95.9	0.25	0.66	6.92	0.09	1	1.00	<0.01
ARR3	ii	Vision	305	77.6	0.06	0.85	4.19	0.03	1	1.00	<0.01
ATP8B1	ii	Vision	825	65.5	0.09	0.49	9.13	0.04	1	2.50	0.20
BEST1	ii	Vision	743	97.3	0.09	0.33	4.56	0.03	1	1.00	0.00
CACNA2D4	ii	Vision	963	87.3	0.06	0.71	29.83	0.03	1	3.70	0.96
CNGA1	i	Vision	605	93.7	0.15	1.63	8.73	0.03	1	5.03	0.45
CNGB1	ii	Vision	596	48.1	0.19	0.50	5.62	0.04	1	8.24	1.03
CNGB3	iii	Vision	738	94.6	0.31	0.28	0.02	0.06	1	245.42	9.44
GABRR2	ii	Vision	479	98	0.14	0.73	9.60	0.06	1	3.49	0.56
GRK7	ii	Vision	550	100	0.28	0.72	5.32	0.04	1	2.33	0.45
GUCA1B	ii	Vision	198	100	0.03	0.25	4.66	0.02	1	5.99	0.18
GUCA1C	i	Vision	190	100	0.16	2.87	9.19	0.04	1	3.99	0.34
GUCY2F	ii	Vision	1,115	97.9	0.25	0.51	4.00	0.06	1	1.99	0.12
OPN1MSW	ii	Vision	254	71.5	0.05	0.46	13.65	0.04	1	1.07	<0.01
PCDH15	ii	Vision	2,105	95.7	0.12	0.60	19.63	0.03	1	19.14	1.08
PRPH2	i	Vision	354	100	0.10	1.58	11.20	0.03	1	3.65	0.61
RGS9	ii	Vision	453	93.4	0.12	0.96	8.58	0.03	1	2.99	0.30
RP1	ii and iii	Vision	1,950	92.1	0.42	0.97	5.51	0.16	1	39.69	7.44
RPE65	ii	Vision	514	93.3	0.02	0.11	9.46	0.01	1	4.21	0.87
RRH	i	Vision	334	100	0.11	52.01	6.47	0.05	1	256.52	0.38
SAG	i	Vision	388	95.6	0.26	1.27	8.14	0.06	1	1.00	<0.01
SLC24A1	i	Vision	615	92.1	0.22	4.57	15.09	0.04	1	7.33	0.77
LOXHD1	ii	Hearing	2,236	96.6	0.11	0.24	5.77	0.03	1	2.20	0.09
MYO3A	ii	Hearing	1,697	96.4	0.18	0.88	10.08	0.02	1	5.88	0.66
MYO6	ii	Hearing	1,215	96	0.05	0.13	4.03	0.02	1	1.00	<0.01
OTOF	ii	Hearing	1,401	70.1	0.05	0.14	7.75	0.02	1	1.00	<0.01
PGAP1	ii	Hearing	750	98.6	0.30	0.90	5.30	0.12	1	3.31	<0.01
ROR1	ii	Hearing	815	91	0.01	0.21	7.12	0.01	1	1.00	<0.01
SCRIB	i	Hearing	656	94.3	0.06	1.31	9.68	0.02	1	4.73	0.39
TBL1X	ii	Hearing	514	98.3	0.03	0.42	5.59	0.02	1	1.00	<0.01
TMC2	ii	Hearing	903	97.2	0.14	0.65	16.89	0.04	1	1.00	<0.01
TMPRSS3	i	Hearing	472	99	0.12	1.30	8.67	0.05	1	4.69	0.41
GPER1	i	Feather kerat.	357	100	0.03	1.28	8.19	0.01	1	1.00	<0.01
KRT5	ii	Feather kerat.	768	60.5	0.05	0.29	6.15	0.02	1	9.26	0.07
TCHP	iii	Feather kerat.	230	78.8	0.35	0.12	0.99	0.10	1	87.68	3.94
CPT1A	i	Circadian rhythm	742	96.4	0.09	2.53	19.58	0.03	1	9.58	0.60
CRY1	ii	circadian rhythm	457	98.9	0.05	0.53	5.69	0.02	1	3.23	−0.78
OPN4-1	ii	Circadian rhythm	482	82.4	0.22	0.64	4.02	0.08	1	1.58	<0.01
SLC6A4	ii	Circadian rhythm	660	98.4	0.08	0.50	6.23	0.04	1	4.02	0.19
STAR	ii	Circadian rhythm	124	42.2	0.05	0.61	4.82	0.04	1	1.00	<0.01

NOTE.—Genes are classified by functional category (vision, hearing, feather keratin, and circadian rhythm) and sorted alphabetically. List refers to the significant ω test categories, whereby list i includes genes with a signal of strong positive selection ($\omega_1 > 1$, branch model), list ii includes genes with a signal of weak positive or relaxed purifying selection ($\omega_0 < \omega_1 < 1$, branch model), and list iii includes genes with a signal of site-specific positive selection in the foreground branch ($\omega_2 > 1$, branch-site model).

Functional Overrepresentation Related to Hearing

Several species that are adapted to darkness or dim-light conditions have enhanced hearing or olfaction capabilities, complementing the visual cues by auditory or olfactory information. In birds, the kiwi and the barn owl are well-studied cases. The kiwis are the only nocturnal ratite relying more on olfaction than vision for foraging and this group has evolved an extended repertoire of odorant receptors (Le Duc et al. 2015). Barn owls possess acute hearing and an ability to localize their prey in darkness (Payne 1971). They have several special traits that improve their hearing, such as a facial disk, asymmetrical position of the ears, and resistance to hearing loss by aging (Krumm et al. 2017).

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

The GO term "sensory perception of sound" is overrepresented among genes with signals of weak positive or relaxed purifying selection in the owl ancestor. Considering the well-developed auditory system of the owls, it seems likely that the elevated ω values reflect positive selection either for a short period or with low intensity in the owl ancestor. The ten genes associated with this GO term are *ATP8B1*, *LOXHD1*, *MYO3A*, *MYO6*, *OTOF*, *PCDH15*, *PGAP1*, *ROR1*, *TBL1X*, and *TMC2*. From these genes, *PCDH15* and *TMC2* were also genome-wide significant (see above) and are described in supplementary table S3 in file 2, Supplementary Material online. The other genes are involved in inner ear receptor cell development and nerve formation or related to the cytoskeleton and may thus function in mechanotransduction of sound stimuli (NCBI gene db, GeneCards, and AmiGO2). Mutations in *LOXHD1*, *OTOF*, *PCDH15*, *TBL1X*, and *TMC2* have been associated with hereditary disorders of balance, deafness or hearing loss in humans (NCBI gene db and GeneCards).

Overrepresentation in Other Functional Categories

We found consistent evidence that 32 genes (lists i–iii) related to DNA conformation change, chromosome condensation, and chromatid segregation have an accelerated substitution rate in the origin of the owls. From these genes, *ATRX*, *SMC2*, and *SMC5* had also genome-wide significant selection signals, and the latter two had nominal significant selection signals from both models (i.e., are in lists i and iii). This group of genes suggests that owls might have evolved a special type of DNA packaging in the retina, similar to what has been found in the rods of nocturnal mice and primates (Solovei et al. 2009). Nocturnal mammals show an unusual radially inverted pattern of hetero- and euchromatin in the nuclei of the rod photoreceptor cells, which acts as a collecting lens channeling the light efficiently toward the light-sensing outer segments, thereby increasing light availability in the deep layers of the retina (Solovei et al. 2009; Joffe et al. 2014; Tan et al. 2019).

The genes with a positive selection signal at specific sites (branch-site model) are enriched in functional categories related to microtubules, including "mitotic nuclear division" and "sperm flagellum" (fig. 4). Of note, microtubules also play an important role in the visual signal transduction cascade of the photoreceptor sensory cilium. The functional overrepresentation associated with the "sperm flagellum" is somewhat unexpected, because owls seem to be strictly genetically monogamous (Lawless et al. 1997; Müller et al. 2001; Rodriguez-Martínez et al. 2014). Genetic monogamy does not promote sperm competition and selection on sperm morphology (Lifjeld et al. 2010; Rowe et al. 2015; Carballo et al. 2019). However, the results from the branch-site model should be interpreted cautiously due to the potential influence of CMDs (Venkat et al. 2018).

Selection in A Priori Defined Candidate Genes Related to the Nocturnal Predatory Lifestyle of Owls

Candidate Genes Related to Vision

The gene *RP1* is the only candidate gene that was significant in both the branch and the branch-site model (lists ii and iii). Furthermore, *RP1* was also significant according to the model aBSREL, indicating a signal of positive selection that is specific for the ancestral branch of the owls. *RP1* encodes a retinal-specific protein related to photosensitivity and the outer segment morphogenesis of rod photoreceptors and is essential for nocturnal vision. *RP1* is also a microtubule-associated protein, required for correct stacking of the outer segment disks.

Our finding that the genes *RGS9*, *BEST1*, *RRH*, *RDH8*, *RPE65*, *PDE6B*, and *ALCAM* evolved faster in the ancestral branch of owls than in the background branches, partially confirm previous results for nocturnal birds and raptors (Wu et al. 2016; Cho et al. 2019; Zhou et al. 2019). These genes are functionally related to visual perception, photoreceptor activity, phototransduction cascades, regeneration of visual pigments, and retina development, and some of them have been linked to genetic diseases related to vision in humans. Our MSA also confirmed the two owl-specific missense mutations in *ALCAM* first reported by Zhou et al. (2019, fig. 3d), which presumably change the charge of a relevant region of the protein surface from neutral to negative.

We found evidence for relaxed purifying selection in the opsin gene *OPN1MSW* on the ancestral branch of the owls, which fits the described pseudogenization of this gene in tytonids (Borges et al. 2015; Wu et al. 2016; Hanna et al. 2017). The opsin genes *OPN1LW* and *SWS2* were not found in the burrowing owl assembly. This is likely an assembly or annotation error because previous studies showed that owls have retained these two cone opsin genes (Borges et al. 2015; Wu et al. 2016; Hanna et al. 2017). Moreover, Wu et al. 2016 found signals of positive selection for both genes at the ancestral branch of owls and suggested that this might be adaptive for crepuscularity. The opsin genes *RHO* and *RGR* were detected and tested, but the ω tests were not significant.

Candidate Genes Related to Hearing

We found evidence for an accelerated substitution rate at the ancestral branch of the owls in the hearing-related candidate genes *LOXHD1*, *MYO3A*, *MYO6*, *OTOF*, *PGAP1*, *ROR1*, *SCRIB*, *TBL1X*, *TMPRSS3*, and *TMC2*. *SCRIB*, *TMPRSS3*, and *TMC2* showed the strongest signal of positive selection, the first two in terms of the ω value (table 1) and the third in terms of *P* value. *SCRIB* is involved in different aspects of polarized cell differentiation, regulating epithelial and neuronal morphogenesis. *TMPRSS3* is expressed in the fetal cochlea, probably participating in the development and maintenance of the inner ear. Mutations in *TMPRSS3* are associated with

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

congenital deafness in humans (NCBI gene db, GeneCards, and AmiGO2).

Candidate Genes Related to Circadian Rhythm

The genes involved in the molecular mechanism behind the circadian rhythm, for example, those coding for nonvisual photopigments, are mostly conserved across mammals and birds (Yoshimura et al. 2000; Bhadra et al. 2017). Our results show an accelerated substitution rate at the ancestral branch of the owls in five candidate genes related to circadian rhythm and sleep: *OPN4-1*, *CRY1*, *CPT1A*, *STAR*, and *SLC6A4*. Our finding of *OPN4-1*, a nonvisual opsin, as a candidate gene is consistent with previous studies on nocturnal birds (Borges et al. 2015; Le Duc et al. 2015; Cho et al. 2019). *CRY1* is a central component of the circadian clock (Griffin et al. 1999). *CPT1A* encodes a key protein for the mitochondrial oxidation of long-chain fatty acids and is linked to the GO term "circadian rhythm" (GeneCards and AmiGO2) and the "circadian clock" pathway (GeneCards and reactome: reactome.org/PathwayBrowser/#/R-HSA-400253) in humans. The protein encoded by *STAR* plays a role in the regulation of steroid hormone synthesis by mediating the transport of cholesterol through the mitochondrial membrane and is linked to the GO terms "circadian rhythm" and "circadian sleep/wake cycle, REM sleep" in humans (GeneCards and AmiGO2). *SLC6A4* regulates synaptic concentrations of serotonin, indirectly influencing perception and anxiety-related behavior. *SLC6A4* and *CRY1* have been related to sleep disorders in humans (Carskadon et al. 2012; Patke et al. 2017).

Cho et al. (2019) found a burrowing owl-specific amino-acid variant in *SLC51A*. Cho et al. associate this amino-acid variant with the diurnality of this species, because the gene is associated with bile acid transmembrane transporter activity and has an indirect effect on the circadian rhythm. We did not find evidence for selection on *SLC51A*, but we confirmed this variant for the burrowing owl and its congeneric, the little owl, indicating that this pattern might be associated with the *Athene* taxon, but not necessarily with diurnality.

Candidate Genes Related to Feather Structure

The feathers of owls have a special noise absorption structure that allows them to fly silently while hunting, and this feature has been studied morphologically and acoustically (Kopania 2016; Sagar et al. 2017; Weger and Wagner 2017). However, the genetic correlates of this adaptation in owls remain unclear.

Here, we present evidence for positive selection in the ancestral branch of the owls for three candidate genes related to feather production: *GPER1*, *TCHP*, and *KRT5*. *GPER1* and *TCHP* are related to keratin filament development and production. The gene *KRT5* belongs to the keratin gene family; it is co-expressed during differentiation of simple and stratified epithelial tissues and is important for keratinization, cornification, and epidermis development (GeneCards and AmiGO2).

Conclusions

We conducted a genome-wide comparative analysis focusing on the early history of Strigiformes. Our study suggests novel candidate genes whose role in the evolution of owls can be further explored. Our study also contributes the raw genome sequencing data of eight owl species (NCBI BioProject PRJNA592858).

Our results support that owls—similar to other nocturnal birds—early on evolved sensory adaptations that allowed them to cope with dim light. In particular, phototransduction in the rods, enhanced motion detection and retina repair, but also acoustic perception seem to be important for the owls. We also found evidence for functional overrepresentation associated with chromosome packaging. This suggests a role of chromatin packaging for enhanced light channeling in photoreceptor cells as a target of adaptation in the owl ancestor. The information-driven approach also supports the idea that genes involved in feather development and circadian rhythm have evolved under positive selection in the ancestral branch of the owls.

In agreement with the diurnal ancestry of raptorial landbirds, our results show the accumulation of genetic changes in several genes functionally associated with nocturnal hunting, indicating the independent adaptive history of owls as nocturnal birds of prey.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank Francisco Salinas García, Iulia Darolti, Fidel Botero-Castro, Henryk Milewski, and Jan Drosd for their kind and helpful support at different stages of the bioinformatic work. We are grateful to the following persons for kindly providing the biological samples of the owls: Jean-Michel Hatt (ZooZurich, Switzerland), James Duncan and Charlene Berkvens (Discover Owls, Manitoba, Canada), Guillermo Blanco (MNCN-CSIC, Spain), José A. Sánchez Zapata and Juan M. Pérez (UMH, Spain), Philippe Helsen (Zoo Antwerp, Belgium), and Alexandre Roulin (University of Lausanne, Switzerland). This work was supported by the Max Planck Society (to B.K.), Acuerdo Bilateral DAAD/BECAS-Chile scholarship (to P.E.-H.), and MINECO, Spain (projects CGL2012-31888 and CGL2015-71378-P), Fundación Repsol, and a Severo Ochoa microproyecto award from Estación Biológica de Doñana (to M.C.). Finally, we thank two anonymous reviewers for their constructive comments.

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

Literature Cited

Afgan E, et al. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. 46(W1):W537–W544.

Anderson SR, Wiens JJ. 2017. Out of the dark: 350 million years of conservatism and evolution in diel activity patterns in vertebrates. Evolution 71(8):1944–1959.

Bhadra U, Thakkar N, Das P, Pal Bhadra M. 2017. Evolution of circadian rhythms: from bacteria to human. Sleep Med. 35:49–61.

Bindea G, et al. 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25(8):1091–1093.

Blanchette M, et al. 2004. Aligning multiple genomic sequences with the threaded blockset aligner. Genome Res. 14(4):708–715.

Borges R, et al. 2015. Gene loss, adaptive evolution and the co-evolution of plumage coloration genes with opsins in birds. BMC Genomics. 16(1):751.

Borges R, et al. 2019. Avian binocularity and adaptation to nocturnal environments: genomic insights from a highly derived visual phenotype. Genome Biol Evol. 11(8):2244–2255.

Borowiec ML. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4:e1660.

Bowmaker JK. 2008. Evolution of vertebrate visual pigments. Vision Res. 48(20):2022–2041.

Braun MJ, Huddleston CJ. 2009. A molecular phylogenetic survey of caprimulgiform nightbirds illustrates the utility of non-coding sequences. Mol Phylogenet Evol. 53(3):948–960.

Carballo L, et al. 2019. Sperm morphology and evidence for sperm competition among parrots. J Evol Biol. 32(8):856–867.

Carbon S, et al. 2009. AmiGO: online access to ontology and annotation data. Bioinformatics 25(2):288–289.

Carskadon MA, Sharkey KM, Knopik VS, McGeary JE. 2012. Short sleep as an environmental exposure: a preliminary study associating 5-HTTLPR genotype to self-reported sleep duration and depressed mood in first-year university students. Sleep 35(6):791–796.

Cho YS, et al. 2019. Raptor genomes reveal evolutionary signatures of predatory and nocturnal lifestyles. Genome Biol. 20(1):1–11.

Corfield JR, Parsons S, Harimoto Y, Acosta ML. 2015. Retinal anatomy of the New Zealand Kiwi: structural traits consistent with their nocturnal behavior. Anat Rec. 298(4):771–779.

Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 10(1):210.

Dale RK, Pedersen BS, Quinlan AR. 2011. Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. Bioinformatics 27(24):3423–3424.

Danecek P, McCarthy SA. 2017. BCFtools/csq: haplotype-aware variant consequences. Bioinformatics 33(13):2037–2039.

Davies WIL, Collin SP, Hunt DM. 2012. Molecular ecology and adaptation of visual photopigments in craniates. Mol Ecol. 21(13):3121–3158.

del Hoyo J, Elliott A, Sargatal J, Cabot J. 1999. Handbook of the birds of the world. Barcelona (Spain): Lynx Edicions.

Duncan JR. 2013. Owls of the world: their lives, behavior and survival. Baltimore: Johns Hopkins University Press.

Emerling CA. 2018. Independent pseudogenization of CYP2J19 in penguins, owls and kiwis implicates gene in red carotenoid synthesis. Mol Phylogenet Evol. 118:47–53.

Ericson PGP, et al. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett. 2(4):543–547.

Fabregat A, et al. 2018. Reactome diagram viewer: data structures and strategies to boost performance. Bioinformatics 34(7):1208–1214.

Feduccia A. 1999. The origin and evolution of birds. New Haven (CT): YaleUniversity Press.

Fellows I. 2012. Deducer: a data analysis GUI for R. J Stat Softw. 49:1–15.

Fite KV. 1973. Anatomical and behavioral correlates of visual acuity in the great horned owl. Vision Res. 13(2):219–230.

Griffin EA, Staknis D, Weitz CJ. 1999. Light-independent role of CRY1 and CRY2 in the mammalian circadian clock. Science 286(5440):768–771.

Grote S. 2019. GOfuncR: gene ontology enrichment using FUNC. https://git.bioconductor.org/packages/GOfuncR, last accessed Aug 19, 2020.

Gutiérrez-Ibáñez C, Iwaniuk AN, Lisney TJ, Wylie DR. 2013. Comparative study of visual pathways in Owls (Aves: Strigiformes). Brain Behav Evol. 81(1):27–39.

Hackett SJ, et al. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320(5884):1763–1768.

Hall MI, Ross CF. 2007. Eye shape and activity pattern in birds. J Zool. 271(4):437–444.

Hanna ZR, et al. 2017. Northern spotted owl (*Strix occidentalis caurina*) genome: divergence with the barred owl (*Strix varia*) and characterization of light-associated genes. Genome Biol Evol. 9(10):2522–2545.

Hart NS, Hunt DM. 2007. Avian visual pigments: characteristics, spectral tuning, and evolution. Am Nat. 169(S1):S7–S26.

Healy S, Guilford T. 1990. Olfactory-bulb size and nocturnality in birds. Evolution 44(2):339–346.

Howland HC, Merola S, Basarab JR. 2004. The allometry and scaling of the size of vertebrate eyes. Vision Res. 44(17):2043–2065.

Huber W, et al. 2015. Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods. 12(2):115–121.

Jarvis ED, et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science 346(6215):1320–1331.

Joffe B, Peichl L, Hendrickson A, Leonhardt H, Solovei I. 2014. Diurnality and nocturnality in primates: an analysis from the rod photoreceptor nuclei perspective. Evol Biol. 41(1):1–11.

Kiełbasa SM, Wan R, Sato K, Horton P, Frith MC. 2011. Adaptive seeds tame genomic sequence comparison. Genome Res. 21(3):487–493.

Kopania J. 2016. Acoustics parameters the wings of various species of owls. Inter-Noise 2016:2868–2876.

Krumm B, Klump G, Köppl C, Langemann U. 2017. Barn owls have ageless ears. Proc R Soc B Biol Sci. 284(1863):20171584.

Lamb TD, Collin SP, Pugh EN. 2007. Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. Nat Rev Neurosci. 8(12):960–976.

Lamb TD, et al. 2016. Evolution of vertebrate phototransduction: cascade activation. Mol Biol Evol. 33(8):2064–2087.

Lawless SG, Ritchison G, Klatt PH, Westneat DF. 1997. The mating strategies of eastern screech-owls: a genetic analysis. Condor 99(1):213–217.

Le Duc D, Schöneberg T. 2016. Adaptation to nocturnality—learning from avian genomes. BioEssays 38(7):694–703.

Le Duc D, et al. 2015. Kiwi genome provides insights into evolution of a nocturnal lifestyle. Genome Biol. 16(1):147.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN].

Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078–2079.

Lifjeld JT, Laskemoen T, Kleven O, Albrecht T, Robertson RJ. 2010. Sperm length variation as a predictor of extrapair paternity in passerine birds. PLoS One 5(10):e13456.

Lisney TJ, Iwaniuk AN, Bandet MV, Wylie DR. 2012. Eye shape and retinal topography in owls (Aves: Strigiformes). Brain Behav Evol. 79(4):218–236.

Maddison W, Maddison D. 2018. Mesquite: a modular system for evolutionary analysis. Version 3.51. Available from: http://www.mesquiteproject.org, last accessed August 19, 2020.

Martin G, Rojas LM, Ramírez Y, McNeil R. 2004. The eyes of oilbirds (*Steatornis caripensis*): pushing at the limits of sensitivity. Naturwissenschaften 91(1):26–29.

McClure CJW, et al. 2019. Commentary: defining raptors and birds of prey. J Raptor Res. 53(4):419.

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

Mooney MA, Nigg JT, McWeeney SK, Wilmot B. 2014. Functional and genomic context in pathway analysis of GWAS data. Trends Genet. 30(9):390–400.
Mueller JC, et al. 2018. Evolution of genomic variation in the burrowing owl in response to recent colonization of urban areas. Proc R Soc B Biol Sci. 285: 1–9.
Mueller JC, et al. 2020. Genes acting in synapses and neuron projections are early targets of selection during urban colonization. Mol Ecol. 15451: 1–10.
Müller W, Epplen JT, Lubjuhn T. 2001. Genetic paternity analyses in little owls (*Athene noctua*): does the high rate of paternal care select against extra-pair young? J Ornithol. 142(2):195–203.
Nery MF, Arroyo JI, Opazo JC. 2013. Accelerated evolutionary rate of the myoglobin gene in long-diving whales. J Mol Evol. 76(6):380–387.
Nielsen R. 2005. Molecular signatures of natural selection. Annu Rev Genet. 39(1):197–218.
Pagès H, Carlson M, Falcon S, Li N. 2019. AnnotationDbi: manipulation of SQLite-based annotations in Bioconductor. Version 1. Available from: https://bioconductor.org/packages/AnnotationDbi/, last accessed August 19, 2020.
Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20(2):289–290.
Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35(3):526–528.
Passarotto A, Parejo D, Penteriani V, Avilés JM. 2018. Colour polymorphism in owls is linked to light variability. Oecologia 187(1):61–73.
Patke A, et al. 2017. Mutation of the human circadian clock gene CRY1 in familial delayed sleep phase disorder. Cell 169(2):203–215.e13.
Payne RS. 1971. Acoustic location of prey by barn owls (*Tyto alba*). J Exp Biol. 54(3):535–573.
Ponder JB, Willette MM. 2015. Strigiformes. In: Miller RE, Fowler ME, editors. Fowler's zoo and wild animal medicine. Vol. 8. Philadelphia (PA): W.B. Saunders. p. 189–198.
Prum RO, et al. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526(7574):569–573.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841–842.
R Core Team. 2018. R: A Language and Environment forStatistical Computing. Vienna. Austria: R Foundation for Statistical Computing.
Ranwez V, Harispe S, Delsuc F, Douzery EJP. 2011. MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. PLoS One 6(9):e22594.
Rodriguez-Martínez S, Carrete M, Roques S, Rebolo-Ifrán N, Tella JL. 2014. High urban breeding densities do not disrupt genetic monogamy in a bird species. PLoS One 9(3):e91314.
Rowe M, et al. 2015. Postcopulatory sexual selection is associated with accelerated evolution of sperm morphology. Evolution 69(4):1044–1052.
Safran M, et al. 2010. GeneCards version 3: the human gene integrator. Database (Oxford) 2010:baq020.
Sagar P, Teotia P, Sahlot AD, Thakur HC. 2017. An analysis of silent flight of owl. Mater Today Proc. 4(8):8571–8575.
Schmitz L, Motani R. 2011. Nocturnality in dinosaurs inferred from scleral ring and orbit morphology. Science 332(6030):705–708.
Shannon P, et al. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13(11):2498–2504.
Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31(19):3210–3212.
Smith MD, et al. 2015. Less Is More: An Adaptive Branch-Site Random Effects Model for Efficient Detection of Episodic Diversifying Selection. Mol Biol Evol. 32(5):1342–1353.
Solovei I, et al. 2009. Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. Cell 137(2):356–368.
Tan L, Xing D, Daley N, Xie XS. 2019. Three-dimensional genome structures of single sensory neurons in mouse visual and olfactory systems. Nat Struct Mol Biol. 26(4):297–307.
Terakita A. 2005. The opsins. Genome Biol. 6(3):213.
Toll-Riera M, Laurie S, Albá MM. 2011. Lineage-specific variation in intensity of natural selection in mammals. Mol Biol Evol. 28(1):383–398.
Venkat A, Hahn MW, Thornton JW. 2018. Multinucleotide mutations cause false inferences of lineage-specific positive selection. Nat Ecol Evol. 2(8):1280–1288.
Walls GL. 1942. The vertebrate eye and its adaptive radiation. Bloomfield Hills (MI): The Cranbrook Institute of Science, Hafner Publishing Company.
Weger M, Wagner H. 2017. Distribution of the characteristics of barbs and barbules on barn owl wing feathers. J Anat. 230(5):734–742.
Wink M, El-Sayed A-A, Sauer-Gürth H, Gonzalez J. 2009. Molecular phylogeny of owls (Strigiformes) inferred from DNA sequences of the mitochondrial cytochrome *b* and the nuclear *RAG-1* gene. Ardea 97(4):581–591.
Wu Y, Wang H. 2019. Convergent evolution of bird-mammal shared characteristics for adapting to nocturnality. Proc R Soc B 286(1897):20182185.
Wu Y, et al. 2016. Retinal transcriptome sequencing sheds light on the adaptation to nocturnal and diurnal lifestyles in raptors. Sci Rep. 6: 1–11.
Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol. 15(5):568–573.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24(8):1586–1591.
Yang Z, dos Reis M. 2011. Statistical properties of the branch-site test of positive selection. Mol Biol Evol. 28(3):1217–1228.
Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol. 19(6):908–917.
Yokoyama S. 2000. Molecular evolution of vertebrate visual pigments. Prog Retin Eye Res. 19(4):385–419.
Yoshimura T, et al. 2000. Molecular analysis of avian circadian clock genes. Mol Brain Res. 78(1–2):207–215.
Zhan X, et al. 2013. Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. Nat Genet. 45(5):563–566.
Zhang G, et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. Science 346(6215):1311–1320.
Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol. 22(12):2472–2479.
Zhou C, et al. 2019. Comparative genomics sheds light on the predatory lifestyle of accipitrids and owls. Sci Rep. 9(1):2249.

Associate editor: Judith Mank

Downloaded from https://academic.oup.com/gbe/article/12/10/1895/5889951 by pespindola@orn.mpg.de on 18 October 2020

Supplementary File 1

Genomic evidence for sensorial adaptations to a nocturnal predatory lifestyle in owls

Pamela Espíndola-Hernández[1*], Jakob C. Mueller[1], Martina Carrete[2], Stefan Boerno[3], Bart Kempenaers[1]

[1] Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany
[2] Department of Physical, Chemical and Natural Systems, Universidad Pablo de Olavide, Sevilla, Spain
[3] Sequencing Core Facility, Max Planck Institute for Molecular Genetics, Berlin, Germany
[*] Corresponding author: pespindola@orn.mpg.de

1 Extended Materials and Methods

Table S1. Samples and pairwise sequence alignment information.
Table S2. Candidate gene list by functional categories.
Figure S1. Workflow to test selection on the ancestral branch of owls.

1.1 Commands
Protocol a: Read mapping to reference
Protocol b: Genome-scale sequence mapping to reference
Multi-species Codon Alignment
Quality assessment of multispecies alignments
Selection test
Newick format of the unrooted species tree
Overrepresentation analyses

2 Extended Results

BUSCO summary
Figure S2. **a, b.** Proportion of CMDs and the Branch-Site Model
Table S3. (in Supplementary File 2*) Description of genes with genome-wide significant ω tests.
Table S4. (in Supplementary File 2*) Significant results from the branch model.
Table S5. (in Supplementary File 2*) Significant results from the branch-site model.
Table S6. (in Supplementary File 2*) Significant results from the aBSREL model if two rate classes were estimated.
Table S7. Overrepresented functional GO-groups of gene list i, ii, iii by ClueGO.
Table S8. Overrepresented functional GO-terms of gene list i, ii, iii by ClueGO.
Table S9. Overrepresented functional GO-terms of gene list i, ii, iii with family-wise error rate (FWER) < 0.05 identified by GOfunR based on 10000 permutations.
Table S10. (in Supplementary File 2*) Results of the ω test for all candidate genes, including non-significant results.

1 Extended Materials and Methods

Table S1. Samples and pairwise sequence alignment information. The assembly of *Athene cunicularia* (Burrowing owl, assembly athCun1) was used as a reference genome with a total length of 1,157,069,517 bp, a codon sequence length of 26,451,662 bp, and an N50 value of 42,147,404 bp. The table lists basic information for each sample or genome data and pairwise alignment characteristics of each species (gaps and percentage of the reference genome covered, where "N" represents a site with zero coverage in relation to the reference. Separate information is given for the genome and for the coding region (CDS).

								All genome			CDS	
Scientific name	Common name	Data/sample origin	Publication	GenBank assembly accession	Sex	Average read Depth	"N" count (bp)	Gaps count	% covered	% of "N" in the alignment	% Covered	% of "N" in the alignment
Athene noctua	Little owl	Blood in ethanol, UMH, Spain	This study	PRJNA592858[1]	male	17.64	25,094,316	1,186,824	97.83	2.17	99.21	0.79
Surnia ulula	Northern hawk-owl	Blood in ethanol, Zoo Zurich, Switzerland	This study	PRJNA592858[1]	male	26.19	34,668,645	1,459,917	97.01	2.99	99.04	0.96
Bubo scandiacus	Snowy owl	Liver in buffer, Zoo Antwerp, Belgium	This study	PRJNA592858[1]	female	26.80	48,146,487	1,732,063	95.85	4.15	98.63	1.37
Bubo bubo	Eurasian eagle-owl	Blood in ethanol, UMH, Spain	This study	PRJNA592858[1]	male	24.95	48,541,302	1,844,811	95.81	4.19	98.58	1.42
Strix uralensis	Ural owl	Blood in ethanol, Zoo Zurich	This study	PRJNA592858[1]	female	21.84	53404,103	1,839,472	95.65	4.35	98.46	1.54
Strix nebulosa	Great grey owl	Blood in ethanol, Zoo Zurich	This study	PRJNA592858[1]	female	18.84	52,167,639	1,950,101	95.50	4.50	98.40	1.60
Asio otus	Long-eared owl	Blood in ethanol, UMH, Spain	This study	PRJNA592858[1]	male	15.21	50,328,877	2,457,902	95.65	4.35	98.27	1.73
Asio flammeus	Short-eared owl	Blood in ethanol, UMH, Spain	This study	PRJNA592858[1]	male	15.30	51,462,491	2,358,876	95.55	4.45	97.51	2.49
Strix occidentalis	Spotted owl	NCBI	Hanna et al. 2017	GCA_002372975.1	female		65,347,625	2,755,189	94.35	5.65	97.41	2.59
Tyto alba	Barn owl	NCBI	Zhang et al. 2014	GCA_000687205.1	female		113,335,374	3,771,159	90.20	9.80	96.85	3.15
Haliaeetus leucocephalus	Bald eagle	NCBI	Zhang et al. 2014	GCA_000737465.1	male		132,070,523	4,106,734	88.59	11.41	94.43	5.57
Falco peregrinus	Peregrine falcon	NCBI	Zhan et al. 2013	GCA_000337955.1	male		150,796,022	5,811,744	86.97	13.03	94.14	5.86
Leptosomus discolor	Cuckoo roller	NCBI	Zhang et al. 2014	GCA_000691785.1	male		156,539,425	5,197,363	86.47	13.53	92.06	7.94
Gallus gallus (Ggallus5)	Red junglefowl	NCBI	International Chicken Genome Consortium 2015	GCA_000002315.3	female		469,639,409	6,240,954	59.41	40.59	91.78	8.22
Cathartes aura	Turkey vulture	NCBI	Zhang et al. 2014	GCA_000699945.1	female		132,420,611	3,549,849	88.56	11.44	90.55	9.45
Picoides pubescens	Downy woodpecker	NCBI	Zhang et al. 2014	GCA_000699005.1	female		433,822,859	5,781,688	62.51	37.49	88.86	11.14
Colius striatus	Speckled mousebird	NCBI	Zhang et al. 2014	GCA_000690715.1	male		272,689,826	6,596,761	76.43	23.57	88.28	11.72
Taeniopygia guttata (taeGut2)	Zebra finch	NCBI	Warren et al. 2010	GCA_000151805.2	male		362,730,059	7,279,320	68.65	31.35	88.20	11.80
Apaloderma vittatum	Bar-tailed trogon	NCBI	Zhang et al. 2014	GCA_000703405.1	male		253,133,073	5,972,504	78.12	21.88	87.95	12.05

Raw sequences have been submitted to the short read archive (SRA) of the NCBI db under BioProject PRJNA592858.

Hanna ZR et al. 2017. Northern spotted owl (*Strix occidentalis caurina*) genome: Divergence with the barred owl (*Strix varia*) and characterization of light-associated genes. Genome Biol. Evol. 9:2522–2545. doi: 10.1093/gbe/evx158.

Warren WC et al. 2010. The genome of a songbird. Nature. 464:757–762. doi: 10.1038/nature08819.

Zhan X et al. 2013. Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. Nat Genet. 45(5):563-6

Zhang G et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 346:1311–1320. doi: 10.1126/science.1251385.

Table S2. Candidate gene list by functional categories**.**

	Gene	candidate category		Gene	candidate category
1	AANAT	circadian rhythm	128	SCRIB	hearing
2	ABCC9	circadian rhythm	129	SEC24B	hearing
3	ADA	circadian rhythm	130	SLC26A5	hearing
4	ADORA2A	circadian rhythm	131	SLC9A3R1	hearing
5	AHCY	circadian rhythm	132	SOBP	hearing
6	ARNTL	circadian rhythm	133	SOD1	hearing
7	ARNTL2	circadian rhythm	134	SPRY2	hearing
8	BHLHE40	circadian rhythm	135	STRC	hearing
9	BTBD9	circadian rhythm	136	TBL1X	hearing
10	CACNA1C	circadian rhythm	137	TECTA	hearing
11	CACNA1I	circadian rhythm	138	TJP1	hearing
12	CIPC	circadian rhythm	139	TMC1	hearing
13	CLOCK	circadian rhythm	140	TMC2	hearing
14	CPT1A	circadian rhythm	141	TMIE	hearing
15	CREB1	circadian rhythm	142	TMPRSS3	hearing
16	CRHR1	circadian rhythm	143	TRIOBP	hearing
17	CRY1	circadian rhythm	144	TUB	hearing
18	CSNK1D	circadian rhythm	145	USH1C	hearing
19	CSNK2A2	circadian rhythm	146	USH1G	hearing
20	DIO2	circadian rhythm	147	USH2A	hearing
21	DLAT	circadian rhythm	148	WDPCP	hearing
22	FBXL3	circadian rhythm	149	WHRN	hearing
23	FOS	circadian rhythm	150	ABCA4	vision
24	FSHB	circadian rhythm	151	ALCAM	vision
25	GHRL	circadian rhythm	152	ARR3	vision
26	GNB3	circadian rhythm	153	ATP8A2	vision
27	GRIA3	circadian rhythm	154	ATP8B1	vision
28	GRIN2A	circadian rhythm	155	BBS4	vision
29	HCRTR2	circadian rhythm	156	BEST1	vision
30	HDAC3	circadian rhythm	157	BHLHE23	vision
31	HTR2A	circadian rhythm	158	CACNA2D4	vision
32	HTR7	circadian rhythm	159	CACNB2	vision
33	ID2	circadian rhythm	160	CACNB4	vision
34	IL18	circadian rhythm	161	CCDC66	vision
35	MAOA	circadian rhythm	162	CLN5	vision
36	MTNR1A	circadian rhythm	163	CLN6	vision
37	NAMPT	circadian rhythm	164	CLN8	vision
38	NCOR1	circadian rhythm	165	CNGA1	vision
39	NFIL3	circadian rhythm	166	CNGA3	vision
40	NOCT	circadian rhythm	167	CNGB1	vision
41	NPAS2	circadian rhythm	168	CNGB3	vision
42	NPSR1	circadian rhythm	169	COL11A1	vision
43	NR0B2	circadian rhythm	170	CRABP1	vision
44	NR1D2	circadian rhythm	171	CRYBA1	vision
45	NRIP1	circadian rhythm	172	DMD	vision
46	NT5E	circadian rhythm	173	DNAJC19	vision
47	OPN4	circadian rhythm	174	EPAS1	vision
48	OPN4-1	circadian rhythm	175	EPHB2	vision
49	OXTR	circadian rhythm	176	EYS	vision
50	PCSK2	circadian rhythm	177	GABRR2	vision
51	PER2	circadian rhythm	178	GJD2	vision
52	PER3	circadian rhythm	179	GLRA1	vision
53	PPARGC1A	circadian rhythm	180	GLRB	vision
54	PPP1CB	circadian rhythm	181	GNAT1	vision
55	PRKG1	circadian rhythm	182	GNAT2	vision
56	PROK1	circadian rhythm	183	GNB1	vision
57	PROK2	circadian rhythm	184	GNB5	vision
58	RAI1	circadian rhythm	185	GNGT2	vision
59	SLC29A1	circadian rhythm	186	GRK7	vision
60	SLC29A3	circadian rhythm	187	GUCA1A	vision

61	SLC6A4	circadian rhythm	188	GUCA1B	vision
62	SRD5A1	circadian rhythm	189	GUCA1C	vision
63	SRRD	circadian rhythm	190	GUCY2F	vision
64	STAR	circadian rhythm	191	HCN1	vision
65	TEF	circadian rhythm	192	ISL1	vision
66	TH	circadian rhythm	193	KCNA2	vision
67	TRIB1	circadian rhythm	194	LAMC3	vision
68	CSNK1A1	feather keratin	195	LUM	vision
69	EPPK1	feather keratin	196	MYO5A	vision
70	FAM83H	feather keratin	197	MYO7A	vision
71	FBF1	feather keratin	198	NAV2	vision
72	GPER1	feather keratin	199	NOB1	vision
73	KRT14	feather keratin	200	NR2E1	vision
74	KRT5	feather keratin	201	NRP1	vision
75	KRT6A	feather keratin	202	NRP2	vision
76	KRT7	feather keratin	203	NTRK2	vision
77	KRT71	feather keratin	204	NYX	vision
78	KRT75	feather keratin	205	OPA1	vision
79	KRT8	feather keratin	206	OPN1MSW	vision
80	TCHP	feather keratin	207	OPN3	vision
81	ALG10	hearing	208	OPN5	vision
82	ATP6V0A4	hearing	209	OPTC	vision
83	CACNA1D	hearing	210	OPTN	vision
84	CDC14A	hearing	211	PCDH15	vision
85	CDH23	hearing	212	PDCL	vision
86	CDKN1B	hearing	213	PDE5A	vision
87	CEMIP	hearing	214	PDE6B	vision
88	CHD7	hearing	215	PDE6C	vision
89	CHRNA9	hearing	216	PDE6D	vision
90	CLIC5	hearing	217	PDE6G	vision
91	CLRN1	hearing	218	PDE6H	vision
92	COCH	hearing	219	PHOX2B	vision
93	CRYM	hearing	220	PLXNA4	vision
94	DCDC2	hearing	221	PPT1	vision
95	DNER	hearing	222	PRPH2	vision
96	EPYC	hearing	223	RARB	vision
97	EYA1	hearing	224	RBP4	vision
98	EYA4	hearing	225	RDH10	vision
99	FBXO11	hearing	226	RDH8	vision
100	FZD4	hearing	227	REEP6	vision
101	GABRA5	hearing	228	RGR	vision
102	GABRB2	hearing	229	RGS9	vision
103	GABRB3	hearing	230	RGS9BP	vision
104	GPX1	hearing	231	RHO	vision
105	HEXA	hearing	232	RORB	vision
106	HEXB	hearing	233	Rp1	vision
107	HOMER2	hearing	234	RPE65	vision
108	KCNQ4	hearing	235	RPGR	vision
109	KIT	hearing	236	RRH	vision
110	LHFPL5	hearing	237	RS1	vision
111	LOXHD1	hearing	238	SAG	vision
112	LRIG1	hearing	239	SALL1	vision
113	LRIG2	hearing	240	SEMA3A	vision
114	LRP2	hearing	241	SEMA3F	vision
115	MARVELD2	hearing	242	SIX4	vision
116	MKKS	hearing	243	SLC1A3	vision
117	MYO3A	hearing	244	SLC24A1	vision
118	MYO6	hearing	245	SLC24A2	vision
119	NR4A3	hearing	246	SLITRK6	vision
120	OTOF	hearing	247	TFAP2A	vision
121	OTOGL	hearing	248	THY1	vision
122	OTOS	hearing	249	TMEM126A	vision
123	PDZD7	hearing	250	TRPM1	vision

124	PGAP1	hearing	251	TULP1	vision
125	PTPRQ	hearing	252	VSX1	vision
126	RIPOR2	hearing	253	WFS1	vision
127	ROR1	hearing			

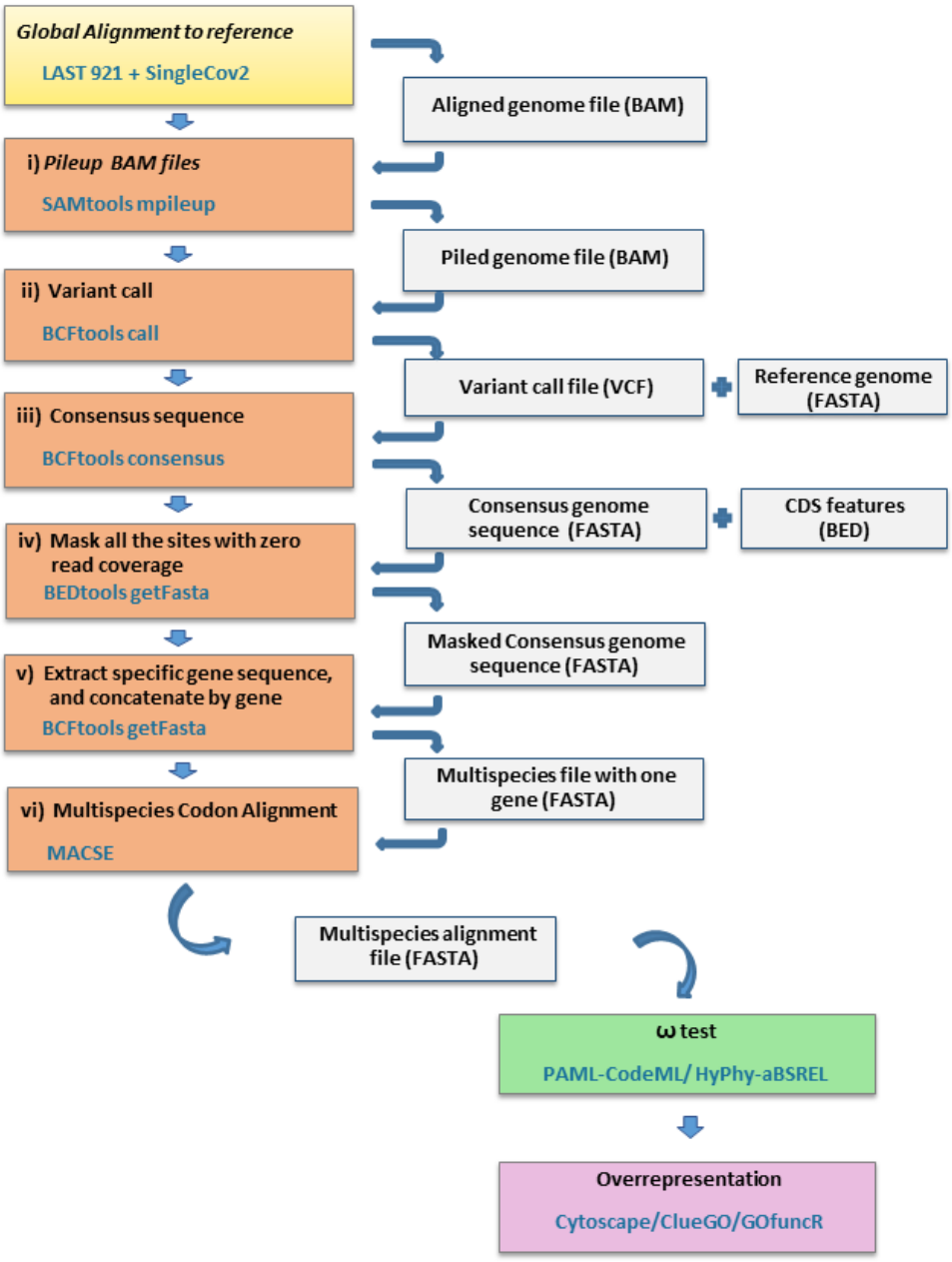


Fig. S1. Workflow to test selection on the ancestral branch of owls. After global genome alignment to the reference genome (yellow), six steps produce the multispecies codon alignment of each gene (orange). Finally, selection tests by ω=dN/dS estimations (green) and overrepresentation analyses (purple) are performed. The names of the software used on each step are given in blue within boxes.

1.1 Commands

The parameter used with each software are detailed below, excluding the infile/outfile.

Protocol a: Read mapping to reference

The reads were mapped against the reference genome using bwa (alignment via Burrows-Wheeler transformation), version: 0.7.17-r1188

bwa[1] mem -M -R

Protocol b: Genome-scale sequence mapping to reference

We aligned species genome assemblies to the reference using LAST v. 921:

lastdb[2] -uMAM8 -cR11
lastal[2] -E0.001 -i3G -m100
SingleCov2[3]
maf-convert[2] sam
samtools[4] view -bS

Multi-species Codon Alignment

i) Piling up the reads or genome sequences in the coding regions

Protocol a:

samtools[4, a] mpileup -u -f athCun1.fa -l CDS.bed -I --output-tags AD,INFO/AD,DP,SP

Protocol b:

samtools [b] mpileup -u -f athCun1.fa -l CDS.bed -I -A --output-tags AD,INFO/AD,DP,SP

ii) Variant calling,

Protocol a:

bcftools[5, a] call -m

protocol b:

bcftools [b] call -m –A

iii) Producing consensus sequences :

bcftools[5] consensus -f

iv) Masking all the sites with zero coverage:

bedtools[6] genomecov bga
bedtools[6] maskfasta

v) Producing one multispecies FASTA file for each gene using BEDtools.

bedtools[6] getfasta

vi) Multispecies codon alignment for each gene using MACSE[7]:

macse -prog alignSequences
fas2phy.R[8]

---

[1]bwa: http://bio-bwa.sourceforge.net/
[2]LAST: http://last.cbrc.jp/
[3]Multiz-tba.012109: https://www.bx.psu.edu/miller_lab/
[4]SAMtools: http://www.htslib.org/doc/samtools.html
[5]BCFtools: http://www.htslib.org/doc/bcftools.html
[6]BEDtools: https://bedtools.readthedocs.io/en/latest/
[7]MACSE: https://bioweb.supagro.inra.fr/macse/
[8]fas2phy, Converts FASTA files into PHYLIP format: https://github.com/fmichonneau/chopper/tree/master/R

Quality assessment of multispecies alignments

Two steps trimming with BMGE (version 1.12) [9]

Step 1: Remove the gaps ignoring entropy:

```
BMGE -t CODON -h 1 -w 1 -g 0
```

Step 2: Remove blocks of codons with high entropy:

```
BMGE -t CODON -m BLOSUM62
```

Summary statistics of multispecies alignments

```
AMAS.py summary -f fasta -d dna -c 40
```

[9]BMGE: https://bioweb.pasteur.fr/packages/pack@BMGE@1.12

Selection tests

We tested for accelerated ω on the ancestral branch of Strigiformes using a maximum-likelihood method implemented in the CodeML program in PAML 4.9h [10] using the following settings in the control files:

Branch model

Null hypothesis (H_0)

```
model = 0        * models for codons:  0: one ω ratio for all branches, 1: one ω ratio for
                   each branch, 2: 2 or more ω ratio for branches
NSsites = 0      * 0: one estimated ω; 1: Nearly neutral; 2: Positive selection
fix_kappa= 0     * 1: kappa fixed, 0: kappa to be estimated
kappa = 2        *  initial or fixed kappa value
fix_omega= 0     * 1: omega or omega_1 fixed, 0: estimate
omega = 1        *  initial or fixed omega value
cleandata = 1    *  remove sites with ambiguity data (1:yes, 0:no)
```

Alternative hypothesis (H_1)

```
model = 2
NSsites = 0
fix_kappa = 0
kappa = 2
fix_omega = 0
omega = 1
cleandata = 1    *  remove sites with ambiguity data (1:yes, 0:no)
```

Branch-Site

Null hypothesis (H_0)

```
model = 2        * models for codons:   0: one ω ratio for all branches, 1: one ω ratio for
                   each branch, 2: 2 or more ω ratio for branches
NSsites = 2      * 0: one estimated ω; 1: NearlyNeutral; 2: Positive selection
fix_kappa = 0
kappa = 2
fix_omega = 1
omega = 1
cleandata = 1    * remove sites with ambiguity data (1:yes, 0:no)
```

Alternative hypothesis (H_1)

```
model = 2
NSsites = 2
fix_kappa = 0
kappa = 2
fix_omega = 0
omega = 1.3
cleandata = 1    * remove sites with ambiguity data (1:yes, 0:no)
```

[10] PAML: http://abacus.gene.ucl.ac.uk/software/paml.html

Newick format of the unrooted species tree used for selection test in CodeML

(((Falco peregrinus,Taeniopygia guttata),((((Picoides pubescens Apaloderma vittatum), Leptosomus discolor), Colius striatus),(((((Bubo bubo,Bubo scandiacus),((Strix uralensis,Strix nebulosa),Strix occidentalis)),(Asio otus,Asio flammeus)),((Athene cunicularia,Athene noctua),Surnia ulula)),Tyto alba)#1)),(Cathartes aura,Haliaeetus leucocephalus),Gallus gallus);

HyPhy/HYPHYMPI[11]

- > with *a priori* specified foreground

HYPHYMPI absrel --branches Foreground

- > without *a priori* specified foreground

HYPHYMPI absrel

Overrepresentation analyses

ClueGO[12] **v2.5.4 plug-in for Cytoscape**[13]

Statistical Test = Enrichment (Right-sided hypergeometric test)
Correction Method for multiple testing = Bonferroni
GO Fusion = true
GO Group = true
Kappa Score Threshold = 0.4
Min GO Level = 3, Max GO Level = 8
Over View Term = SmallestPValue
Group By Kappa Statistics = true

GOfuncR[14]

results = go_enrich(infile, annotations=CustomAnnot_AthCun, n_randset=10000)

* infile is the list of genes with significant ω values, *i.e.*: list i, ii, iii or the list of genome-wide significant genes.

[11]HyPhy: https://stevenweaver.github.io/hyphy-site/methods/selection-methods/
[12]ClueGO: http://apps.cytoscape.org/apps/cluego
[13]Cytoscape: https://cytoscape.org/
[14]GOfuncR: https://bioconductor.org/packages/release/bioc/vignettes/GOfuncR/inst/doc/GOfuncR.html

2 Extended Results

BUSCO summary

BUSCO version is: 4.0.6
The lineage dataset is: aves_odb10 (Creation date: 2019-11-20)
Summarized benchmarking in BUSCO notation for file athCun1.fa
BUSCO was run in mode: genome

C:94.8%[S:94.6%,D:0.2%],F:1.6%,M:3.6%,n:8338

7905	Complete BUSCOs (C)
7887	Complete and single-copy BUSCOs (S)
18	Complete and duplicated BUSCOs (D)
131	Fragmented BUSCOs (F)
302	Missing BUSCOs (M)
8338	Total BUSCO groups searched

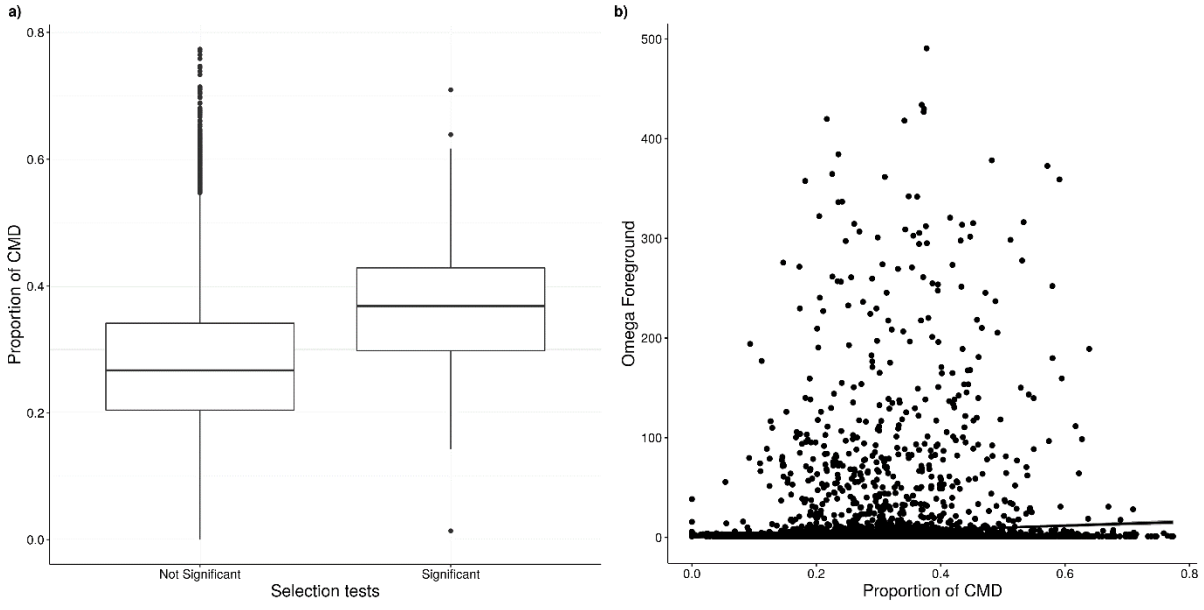


Fig. S2. Proportion of CMDs and the Branch-Site Model. a) Box-plots comparing the proportion of CMDs among the significant tests and the non-significant tests of the Branch-Site model of CodeML. b) Scatter plot of foreground ω values of the alternative hypothesis in relation to the proportion of CMDs for all genes.

Table S7. Overrepresented functional GO-groups of gene list i,ii,iii by ClueGO.

Group	Function by leading GO term	Genes in the functional group	Group p-value	Bonferroni corrected group p-Value
List i: Genes evolving under strong positive selection on the foreground (with ω0 ≤ 1 < ω1). N* = 199 genes				
3	photoreceptor cell cilium	CNGA1, FAM161A, GUCA1C, LCA5, PPEF2, PRPH2, RPGRIP1L, SAG, SLC24A1, SPTBN5	<0.001	<0.001
1	positive regulation of chromosome segregation	DLGAP5, SLF1, SMC5, SMC6	<0.001	<0.001
4	chromosome condensation	CENPP, CENPT, CHD1, ERCC8, GPER1, M1AP, NCAPD2, NCAPG, NUSAP1, SETX, SMC2	<0.001	<0.001
2	nuclear chromosome segregation	DLGAP5, M1AP, MEIOC, NCAPD2, NCAPG, NUSAP1, SGO2, SLF1, SMC2, SMC5	0.001	0.003
List ii: Genes evolving under relaxed purifying or weak positive selection on the foreground (with ω0 < ω1 < 1), N* = 287 genes				
7	sensory perception	ABCA4, ABLIM1, ACPP, ADCY5, ANO2, ARR3, ATP8B1, BEST1, CACNA2D4, CEP290, CHRNA4, CNGB1, CRY1, DDHD2, DOP1B, GABRA6, GABRR2, GRIN2C, GRK7, GUCA1B, GUCY2F, IMPG1, IMPG2, KCNK2, LOXHD1, MMP24, MYO3A, MYO6, OPN1MSW, OTOF, PCDH15, PGAP1, PLA2G6, PRDM1, PRDX1, RD3, RGS9, RIMS2, ROR1, RP1, RP1L1, RPE65, SCNN1B, SHISA9, SLC6A4, TAS1R3, TBL1X, TMC2, UNC13B, ZMPSTE24	<0.001	<0.001
6	plasma membrane bounded cell projection	ABCA4, ABLIM1, ACPP, ADCY5, ADCY9, ANO2, APPL2, ARHGEF4, ARR3, ATP6V1A, ATP8B1, CDK6, CEL, CEP170, CEP290, CHRNA4, CLIP1, CNGB1, CRY1, CTSZ, DAGLA, DNAH3, DNAH8, DNAH9, DNM1, DTNA, DYSF, ENO2, GABRA6, GRK7, GUCA1B, GUCY2F, HSPA2, ITPR1, KCNK2, KIF4A, KLHL24, LAMA2, LAMP5, LOXHD1, LRP4, LRP8, MARK1, MYO3A, MYO6, NEK8, OPN1MSW, OTOF, PCDH15, PDLIM5, PTPRS, RGS9, RIMS2, ROR1, RP1, RP1L1, RPH3A, RTTN, SBF2, SHISA9, SLC16A3, SLC40A1, SLC6A4, STAR, SYNE2, SYT11, TMC2, UNC13B, UNC13C, UNC5A, WRN	<0.001	<0.001
5	myosin complex	ACTR10, CGNL1, DNAH3, DNAH8, KIF15, KIF4A, LRP8, MYBPC1, MYH11, MYH15, MYL1, MYO3A, MYO6, RP1, TPR	<0.001	<0.001
4	sensory perception of sound	ATP8B1, LOXHD1, MYO3A, MYO6, OTOF, PCDH15, PGAP1, ROR1, TBL1X, TMC2	<0.001	0.001
1	DNA conformation change	ASH1L, ATRX, CHAF1A, FANCM, HP1BP3, MIS18BP1, MNAT1, NCAPH, RAD54B, SRPK1, TOP2A, TPR, WRN	<0.001	0.001
2	supramolecular fiber	AKAP13, ASPM, CCT6A, CDC27, CEP170, CKAP5, CLIP1, COL3A1, COL5A1, DNAH3, DNAH8, DNAH9, DNM1, DYSF, KIF4A, KRT5, MYBPC1, MYH11, MYH15, MYL1, MYO3A, MYO6, NEB, OBSL1, RP1, RP1L1, SYNE2	<0.001	0.001
3	transmembrane transporter complex	ANO2, ATP12A, BEST1, CACNA2D4, CHRNA4, CNGB1, GABRA6, GABRR2, GRIN2C, HSPA2, KCNG4, KCNK2, SCNN1B, SHISA9	<0.001	0.001
List iii: Genes evolving under positive selection on specific sites of the foreground branch, N* = 123 genes				
2	tubulin binding	CLIP1, GAS8, INO80, KIF20B, KIF4A, MAP7D3, NUSAP1, RP1, SAXO1	<0.001	0.001
3	sperm flagellum	DNAH1, GAS8, SAXO1, TEKT4	0.002	0.007
1	microtubule bundle formation	CLIP1, DNAH1, GAS8, RP1	0.003	0.011
4	mitotic nuclear division	INO80, KIF20B, KIF4A, MTBP, NUSAP1, SLF1, SMC2, SMC5	0.003	0.012

* Number of genes after all filtering steps.

A complete list of the GO-terms identified for each list of genes and their statistical support are in Table S8.

Table S8. Overrepresented functional GO-terms of gene list i,ii,iii by ClueGO.

GO ID	GO Term	Associated genes	Ontology category	Associated GO group	% associated genes	Nº associated genes	Raw p-Value	Bonferroni corrected p-Value
List i: Genes evolving under strong positive selection on the foreground (with ω0 ≤ 1 < ω1). N* = 199 genes								
GO:0097733	photoreceptor cell cilium	CNGA1, FAM161A, GUCA1C, LCA5, PPEF2, PRPH2, RPGRIP1L, SAG, SPTBN5	CC	3	10.59	9	<0.001	<0.001
GO:0030261	chromosome condensation	GPER1, NCAPD2, NCAPG, NUSAP1, SMC2	BP	4	22.73	5	<0.001	0.001
GO:0071103	DNA conformation change	CENPP, CENPT, CHD1, ERCC8, GPER1, M1AP, NCAPD2, NCAPG, NUSAP1, SETX, SMC2	BP	4	6.47	11	<0.001	0.002
GO:0006323	DNA packaging	CENPP, CENPT, GPER1, M1AP, NCAPD2, NCAPG, NUSAP1, SMC2	BP	4	8.70	8	<0.001	0.003
GO:0000796	condensin complex	NCAPD2, NCAPG, SMC2	CC	4	42.86	3	<0.001	0.005
GO:0051984	positive regulation of chromosome segregation	DLGAP5, SLF1, SMC5, SMC6	BP	1	17.39	4	<0.001	0.015
KEGG:04744	Phototransduction	CNGA1, GUCA1C, SAG, SLC24A1	KEGG	3	17.39	4	<0.001	0.015
GO:0001750	photoreceptor outer segment	CNGA1, GUCA1C, PPEF2, PRPH2, SAG, SPTBN5	CC	3	9.09	6	<0.001	0.020
GO:0098813	nuclear chromosome segregation	DLGAP5, M1AP, MEIOC, NCAPD2, NCAPG, NUSAP1, SGO2, SLF1, SMC2, SMC5	BP	2	5.21	10	0.001	0.029
List ii: Genes evolving under relaxed purifying or weak positive selection on the foreground (with ω0 < ω1 < 1), N* = 287 genes								
GO:0007600	sensory perception	ABCA4, ABLIM1, ACPP, ARR3, ATP8B1, BEST1, CACNA2D4, CHRNA4, CNGB1, GABRR2, GRK7, GUCA1B, GUCY2F, IMPG1, IMPG2, LOXHD1, MMP24, MYO3A, MYO6, OPN1MSW, OTOF, PCDH15, PGAP1, RD3, RGS9, ROR1, RP1, RP1L1, RPE65, SCNN1B, TAS1R3, TBL1X, TMC2	BP	7	9.32	33	<0.001	<0.001
GO:0050877	nervous system process	ABCA4, ABLIM1, ACPP, ADCY5, ARR3, ATP8B1, BEST1, CACNA2D4, CHRNA4, CNGB1, DDHD2, DOP1B, GABRA6, GABRR2, GRIN2C, GRK7, GUCA1B, GUCY2F, IMPG1, IMPG2, KCNK2, LOXHD1, MMP24, MYO3A, MYO6, OPN1MSW, OTOF, PCDH15, PGAP1, PLA2G6, RD3, RGS9, RIMS2, ROR1, RP1, RP1L1, RPE65, SCNN1B, SHISA9, SLC6A4, TAS1R3, TBL1X, TMC2, UNC13B, ZMPSTE24	BP	7	6.61	45	<0.001	<0.001
GO:0007601	visual perception	ABCA4, ABLIM1, ARR3, BEST1, CACNA2D4, CNGB1, GABRR2, GRK7, GUCA1B, GUCY2F, IMPG1, IMPG2, MYO3A, OPN1MSW, PCDH15, RD3, RGS9, RP1, RP1L1, RPE65	BP	7	12.99	20	<0.001	<0.001
GO:0009584	detection of visible light	ABCA4, BEST1, CACNA2D4, CNGB1, GRK7, GUCA1B, GUCY2F, OPN1MSW, RP1, RPE65	BP	7	19.61	10	<0.001	<0.001
GO:0009581	detection of external stimulus	ABCA4, BEST1, CACNA2D4, CNGB1, GRK7, GUCA1B, GUCY2F, MMP24, OPN1MSW, PCDH15, RP1, RPE65, TMC2	BP	7	13.00	13	<0.001	<0.001

GO:0097733	photoreceptor cell cilium	ABCA4, ARR3, CEP290, CNGB1, CRY1, GRK7, GUCA1B, GUCY2F, OPN1MSW, PCDH15, RP1, RP1L1	CC	7	14.12	12	<0.001	<0.001
GO:0009582	detection of abiotic stimulus	ABCA4, BEST1, CACNA2D4, CNGB1, GRK7, GUCA1B, GUCY2F, MMP24, OPN1MSW, PCDH15, RP1, RPE65, TMC2	BP	7	12.62	13	<0.001	<0.001
GO:0097730	non-motile cilium	ABCA4, ANO2, ARR3, CEP290, CNGB1, CRY1, GRK7, GUCA1B, GUCY2F, OPN1MSW, PCDH15, RP1, RP1L1	CC	7	12.38	13	<0.001	<0.001
GO:0120025	plasma membrane bounded cell projection	ABCA4, ABLIM1, ACPP, ADCY5, ADCY9, ANO2, APPL2, ARHGEF4, ARR3, ATP6V1A, ATP8B1, CDK6, CEP170, CEP290, CHRNA4, CLIP1, CNGB1, CRY1, CTSZ, DAGLA, DNAH3, DNAH8, DNAH9, DNM1, DTNA, DYSF, GABRA6, GRK7, GUCA1B, GUCY2F, KCNK2, KIF4A, KLHL24, LAMA2, LAMP5, LOXHD1, LRP4, LRP8, MARK1, MYO3A, MYO6, NEK8, OPN1MSW, PCDH15, PDLIM5, PTPRS, ROR1, RP1, RP1L1, RPH3A, RTTN, SBF2, SHISA9, STAR, SYNE2, SYT11, TMC2, UNC13B, UNC13C, UNC5A, WRN	CC	6	4.16	61	<0.001	<0.001
GO:0097458	neuron part	ABCA4, ADCY9, ARR3, ATP8B1, CEL, CEP290, CHRNA4, CNGB1, CRY1, CTSZ, DAGLA, DNM1, DTNA, ENO2, GABRA6, GRK7, GUCA1B, GUCY2F, HSPA2, ITPR1, KCNK2, KIF4A, KLHL24, LAMA2, LAMP5, LOXHD1, LRP4, LRP8, MARK1, MYO3A, OPN1MSW, OTOF, PCDH15, PDLIM5, PTPRS, RGS9, RIMS2, ROR1, RP1, RP1L1, RPH3A, SBF2, SHISA9, SLC16A3, SLC40A1, SLC6A4, STAR, SYT11, TMC2, UNC13B, UNC13C, UNC5A, WRN	CC	6	4.38	53	<0.001	<0.001
GO:0043005	neuron projection	ABCA4, ADCY9, ARR3, ATP8B1, CEP290, CHRNA4, CNGB1, CRY1, CTSZ, DAGLA, DNM1, DTNA, GABRA6, GRK7, GUCA1B, GUCY2F, KCNK2, KIF4A, KLHL24, LAMA2, LAMP5, LOXHD1, LRP4, LRP8, MARK1, MYO3A, OPN1MSW, PCDH15, PDLIM5, PTPRS, ROR1, RP1, RP1L1, RPH3A, SBF2, SHISA9, STAR, SYT11, TMC2, UNC13B, UNC13C, UNC5A, WRN	CC	6	4.83	43	<0.001	<0.001
GO:0050906	detection of stimulus involved in sensory perception	BEST1, CACNA2D4, CNGB1, GUCY2F, MMP24, PCDH15, RPE65, TAS1R3, TMC2	BP	7	16.67	9	<0.001	0.001
KEGG:04744	Phototransduction	CNGB1, GRK7, GUCA1B, GUCY2F, OPN1MSW, RGS9	KEGG	7	26.09	6	<0.001	0.002
GO:0050908	detection of light stimulus involved in visual perception	BEST1, CACNA2D4, CNGB1, GUCY2F, RPE65	BP	7	35.71	5	<0.001	0.002
GO:0120038	plasma membrane bounded cell projection part	ABCA4, ADCY9, APPL2, ARHGEF4, ARR3, ATP8B1, CEP170, CEP290, CHRNA4, CNGB1, CRY1, CTSZ, DAGLA, DNAH3, DNAH8, DNAH9, DNM1, GABRA6, GRK7, GUCA1B, GUCY2F, KCNK2, KIF4A, LAMA2, LAMP5, LRP4, LRP8, MARK1, MYO3A, MYO6, NEK8, OPN1MSW, PCDH15, ROR1, RP1, RP1L1, RTTN, SHISA9, SYNE2, SYT11, TMC2, UNC13B, UNC13C, UNC5A	CC	6	4.32	44	<0.001	0.004
GO:0097381	photoreceptor disc membrane	ABCA4, CRY1, GRK7, GUCA1B, GUCY2F	CC	7	31.25	5	<0.001	0.005
GO:0007602	phototransduction	ABCA4, CNGB1, GRK7, GUCA1B, GUCY2F, OPN1MSW, RP1	BP	7	16.28	7	<0.001	0.010
GO:0016459	myosin complex	CGNL1, MYBPC1, MYH11, MYH15, MYL1, MYO3A, MYO6	CC	5	14.89	7	<0.001	0.018

GO:0001895	retina homeostasis	ABCA4, CNGB1, PCDH15, PRDX1, RP1, RP1L1, RPE65	BP	7	14.58	7	<0.001	0.021
GO:0007605	sensory perception of sound	ATP8B1, LOXHD1, MYO3A, MYO6, OTOF, PCDH15, PGAP1, ROR1, TBL1X, TMC2	BP	4	9.80	10	<0.001	0.022
GO:0042461	photoreceptor cell development	CEP290, CNGB1, PRDM1, RP1, RP1L1, RPE65	BP	7	17.65	6	<0.001	0.023
GO:0003774	motor activity	CGNL1, DNAH3, DNAH8, KIF15, KIF4A, MYH11, MYH15, MYO3A, MYO6	MF	5	10.84	9	<0.001	0.023
GO:0071103	DNA conformation change	ASH1L, ATRX, CHAF1A, FANCM, HP1BP3, MIS18BP1, MNAT1, NCAPH, RAD54B, SRPK1, TOP2A, TPR, WRN	BP	1	7.65	13	<0.001	0.027
GO:0005875	microtubule associated complex	ACTR10, DNAH3, DNAH8, KIF15, KIF4A, LRP8, RP1, TPR	CC	5	11.76	8	<0.001	0.031
GO:0099512	supramolecular fiber	AKAP13, ASPM, CCT6A, CDC27, CEP170, CKAP5, CLIP1, COL3A1, COL5A1, DNAH3, DNAH8, DNAH9, DNM1, DYSF, KIF4A, KRT5, MYBPC1, MYH11, MYH15, MYL1, MYO3A, MYO6, NEB, OBSL1, RP1, RP1L1, SYNE2	CC	2	4.73	27	<0.001	0.045
GO:1902495	transmembrane transporter complex	ANO2, ATP12A, BEST1, CACNA2D4, CHRNA4, CNGB1, GABRA6, GABRR2, GRIN2C, HSPA2, KCNG4, KCNK2, SCNN1B, SHISA9	CC	3	6.86	14	<0.001	0.047
List iii: Genes evolving under positive selection on specific sites of the foreground branch, N* = 123 genes								
GO:0015631	tubulin binding	CLIP1, GAS8, INO80, KIF20B, KIF4A, MAP7D3, NUSAP1, RP1, SAXO1	MF	2	4.25	9	<0.001	0.002
GO:0140014	mitotic nuclear division	INO80, KIF20B, KIF4A, MTBP, NUSAP1, SLF1, SMC2, SMC5	BP	4	4.00	8	0.001	0.006
GO:0000070	mitotic sister chromatid segregation	INO80, KIF4A, NUSAP1, SLF1, SMC2, SMC5	BP	4	5.22	6	0.001	0.007
GO:0036126	sperm flagellum	DNAH1, GAS8, SAXO1, TEKT4	CC	3	7.27	4	0.002	0.017
GO:0001578	microtubule bundle formation	CLIP1, DNAH1, GAS8, RP1	BP	1	6.45	4	0.003	0.026
GO:0045931	positive regulation of mitotic cell cycle	KIF20B, MTBP, NUSAP1, SLF1, SMC5	BP	4	4.46	5	0.004	0.039

* Number of genes after all filtering steps

Table S9. Overrepresented functional GO-terms of gene list i, ii, iii with family-wise error rate (FWER) < 0.05 identified by GOfunR based on 10000 permutations.

GO ID	GO Term	Ontology category	Raw p-Value	FWER corrected p-Value
List i: Genes evolving under strong positive selection on the foreground (with ω0 ≤ 1 < ω1). N* = 199 genes				
GO:0097733	photoreceptor cell cilium	CC	<0.001	0.002
GO:0097731	9+0 non-motile cilium	CC	<0.001	0.002
GO:0097730	non-motile cilium	CC	<0.001	0.009
GO:0030261	chromosome condensation	BP	<0.001	0.029
GO:0000796	condensin complex	CC	<0.001	0.038
GO:0007076	mitotic chromosome condensation	BP	<0.001	0.039
List ii: Genes evolving under relaxed purifying or weak positive selection on the foreground (with ω0 < ω1 < 1), N* = 287 genes				
GO:0007600	sensory perception	BP	<0.001	<0.001
GO:0050877	nervous system process	BP	<0.001	<0.001
GO:0003008	system process	BP	<0.001	<0.001
GO:0007601	visual perception	BP	<0.001	<0.001
GO:0050953	sensory perception of light stimulus	BP	<0.001	<0.001
GO:0009584	detection of visible light	BP	<0.001	<0.001
GO:0001750	photoreceptor outer segment	CC	<0.001	<0.001
GO:0097733	photoreceptor cell cilium	CC	<0.001	<0.001
GO:0097731	9+0 non-motile cilium	CC	<0.001	<0.001
GO:0097730	non-motile cilium	CC	<0.001	<0.001
GO:0120025	plasma membrane bounded cell projection	CC	<0.001	0.001
GO:0009581	detection of external stimulus	BP	<0.001	0.001
GO:0043005	neuron projection	CC	<0.001	0.001
GO:0097458	neuron part	CC	<0.001	0.001
GO:0009582	detection of abiotic stimulus	BP	<0.001	0.001
GO:0042995	cell projection	CC	<0.001	0.002
GO:0009583	detection of light stimulus	BP	<0.001	0.002
GO:0097381	photoreceptor disc membrane	CC	<0.001	0.007
GO:0050906	detection of stimulus involved in sensory perception	BP	<0.001	0.009
GO:0044463	cell projection part	CC	<0.001	0.009
GO:0120038	plasma membrane bounded cell projection part	CC	<0.001	0.009
GO:0098590	plasma membrane region	CC	<0.001	0.011
GO:0042623	ATPase activity, coupled	MF	<0.001	0.021
GO:0016887	ATPase activity	MF	<0.001	0.022
GO:0016459	myosin complex	CC	<0.001	0.026
GO:0050908	detection of light stimulus involved in visual perception	BP	<0.001	0.028
GO:0050962	detection of light stimulus involved in sensory perception	BP	<0.001	0.028
GO:0003774	motor activity	MF	<0.001	0.037
GO:0005875	microtubule associated complex	CC	<0.001	0.049
List iii: Genes evolving under positive selection on specific sites of the foreground branch, N* = 123 genes				
GO:0008017	microtubule binding	MF	<0.001	0.016
GO:0070701	mucus layer	CC	<0.001	0.035

CC=cellular_component, BP=biological_process, MF=molecular_function

***** Number of genes after all filtering steps

Chapter 2

Genomic signatures of the evolution of a diurnal lifestyle in Strigiformes

Pamela Espíndola-Hernández[*], Jakob C. Mueller[*], Bart Kempenaers[*]

[*] Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany

Supplemental material:
SupplementalFile_1_SupplementalFigures_andExtendedMethods.pdf
SupplementalFile_2_Tables.xlsx (online)
SupplementalFile_3_Owl-specificCNEEs.bed (online)

Supplemental data:
Uploaded to https://figshare.com/s/1eff03514ab6d1570cf1

Abstract

Understanding the targets of selection associated with changes in behavioral traits represents an important challenge of current evolutionary research. Owls (Strigiformes) are a diverse group of birds, most of which are considered nocturnal raptors. However, a few owl species independently adopted a diurnal lifestyle in their recent evolutionary history. We searched for signals of accelerated rates of evolution associated with a diurnal lifestyle using a genome-wide comparative approach. We estimated substitution rates in coding and non-coding conserved regions of the genome of seven owl species, including three diurnal species. Substitution rates of the non-coding elements were more accelerated than those of protein-coding genes. We identified new, owl-specific conserved non-coding elements as candidates of parallel evolution during the emergence of diurnality in owls. Our results shed light on the molecular basis of adaptation to a new niche and highlight the importance of regulatory elements for evolutionary changes in behavior. These elements were often involved in the neuronal development of the brain.

Published as:

Espíndola-Hernández, Pamela, Jakob C. Mueller, and Bart Kempenaers. "Genomic signatures of the evolution of a diurnal lifestyle in Strigiformes." G3 12, no. 8 (2022): jkac135.

G3, 2022, 12(8), jkac135
https://doi.org/10.1093/g3journal/jkac135
Advance Access Publication Date: 30 May 2022
Investigation

Genomic signatures of the evolution of a diurnal lifestyle in Strigiformes

Pamela Espíndola-Hernández ,* Jakob C. Mueller , Bart Kempenaers

Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany

*Corresponding author: Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany. Email: pamela.dola@gmail.com

Abstract

Understanding the targets of selection associated with changes in behavioral traits represents an important challenge of current evolutionary research. Owls (Strigiformes) are a diverse group of birds, most of which are considered nocturnal raptors. However, a few owl species independently adopted a diurnal lifestyle in their recent evolutionary history. We searched for signals of accelerated rates of evolution associated with a diurnal lifestyle using a genome-wide comparative approach. We estimated substitution rates in coding and noncoding conserved regions of the genome of seven owl species, including three diurnal species. Substitution rates of the noncoding elements were more accelerated than those of protein-coding genes. We identified new, owl-specific conserved noncoding elements as candidates of parallel evolution during the emergence of diurnality in owls. Our results shed light on the molecular basis of adaptation to a new niche and highlight the importance of regulatory elements for evolutionary changes in behavior. These elements were often involved in the neuronal development of the brain.

Keywords: parallel evolution; comparative genomics; diel-activity pattern; diurnality; adaptation; CNEEs; protein-coding genes

Introduction

Even though owls are considered one of the most iconic nocturnal birds, species vary considerably in their diel activity patterns. The spectrum of phenotypes ranges from exclusively nocturnal owls (family Tytonidae) to diurnal ones (the snowy owl *Bubo scandiacus*, the northern hawk owl *Surnia ulula* and the burrowing owl *Athene cunicularia*), with many intermediate activity patterns (e.g. crepuscular or cathemeral) (del Hoyo *et al.* 1999; König and Weick 2008; Duncan 2018). Diurnality in owls is absent in the family Tytonidae, but has emerged independently at least twice among the family Strigidae (König and Weick 2008; Wink *et al.* 2009; Salter *et al.* 2020). This provides an opportunity to study genomic signatures of a recent case of parallel evolution in birds.

The owls belong to the clade of the Afroaves and presumably evolved from an ancestral diurnal landbird with raptorial features (Ericson *et al.* 2006; Hackett *et al.* 2008; Jarvis *et al.* 2014; Prum *et al.* 2015; McClure *et al.* 2019). Currently, 250 species of owls live in a variety of ecosystems around the world (del Hoyo *et al.* 1999; König and Weick 2008). Their diversification from the rest of the Afroaves was probably fostered by increasing opportunities to hunt small nocturnal mammals, which experienced a rapid radiation during the Eocene (56–33 Ma) (Feduccia 1995, 1999, 2003). Many of the owls' early adaptations to nocturnality have been shaped by positive selection on genes functionally associated with visual perception, including phototransduction and chromatin packaging (Espíndola-Hernández *et al.* 2020). However, little is known about the mechanisms and targets of selection that shaped the more recent shift into a diurnal activity pattern observed in some owls.

The diurnal owls, as well as their cathemeral relatives, have been described as "time-shifter" species. Despite phylogenetic constraints on the evolution of diel activity patterns (Roll *et al.* 2006; Anderson and Wiens 2017), the "time-shifter" species might have changed their activity pattern in response to competition for food (Schoener 1974; Jaksić 1982; Carothers and Jaksić 1984). Shorter nights during summer and interference competition might have been the main drivers of diurnality in the owls included in this study (Pei *et al.* 2018).

Modern evolutionary biology tries to understand whether genetic correlates of parallel evolution of novel phenotypic traits exist, and of which type these are. The emergence of diurnality in different owl clades is a case of parallel evolution (Gould 2002; Pearce 2012; Rosenblum *et al.* 2014), and has likely occurred from similar genomic elements of the common Strigidae ancestor. A general distinction is often made between regulatory and structural changes, and evidence for both exists. In birds, for instance, loss of flight evolved independently in different clades and has been linked to protein-coding genes (Burga *et al.* 2017; Pan *et al.* 2019), as well as to noncoding elements (Sackton *et al.* 2019).

Nonsynonymous changes in protein-coding regions affect the structure of the gene product and, therefore, the function of the protein itself. Because of the supposed strong phenotypic effect, these structural modifications have been considered as major evolutionary factors. However, nonsynonymous changes are relatively rare, and closely related species are often almost identical

Received: October 22, 2021. **Accepted:** May 17, 2022
© The Author(s) 2022. Published by Oxford University Press on behalf of Genetics Society of America.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

in protein-coding regions of the genome. Thus, King and Wilson (1975) suggested that the phenotypic differences observed between closely related species, such as human and chimpanzee, are likely due to mutations in regulatory regions of the genome. Many studies have now shown that changes in the regulation of gene expression contribute to differences in a multitude of phenotypic traits (Wray 2007; Rubinstein and de Souza 2013; Stern 2013; Hill *et al.* 2021). Conserved nonexonic elements (CNEEs), which have been used as markers for avian phylogenomic inferences (Edwards *et al.* 2017; Tiley *et al.* 2020), are usually located in the *cis*-regulatory domain of genes, and mutations in these regions have been linked to a wide variety of phenotypic changes that often constitute evolutionary innovations (Wray 2007; Rubinstein and de Souza 2013). In birds, CNEEs have been used to study the evolution of the development of avian limbs and flight feathers (Seki *et al.* 2017), convergent evolution associated with the loss of flight in ratites (Sackton *et al.* 2019), and the diversification of bill shape (Yusuf *et al.* 2020).

Here, we report on a search for signals of accelerated evolution linked to the emergence of diurnality. We compared the substitution rates in the genomes of seven owl species, of which four are strictly nocturnal and 3 are consistently diurnal. To obtain the maximum contrast in diel activity patterns, we did not include species with intermediate or cathemeral phenotypes. We used a genome-wide comparative approach to estimate substitution rates in conserved coding regions (CDS: coding sequences) and noncoding regions (CNEE) of the respective genomes. Our study aims to answer the following questions. (1) Are there CDS and CNEEs that evolved under accelerated substitution rates among diurnal owls? (2) Is there an enrichment of functions linked with these CDS and CNEEs, and therefore with a diurnal lifestyle? (3) Are these genomic signatures predominantly structural (CDS) or regulatory (CNEEs)?

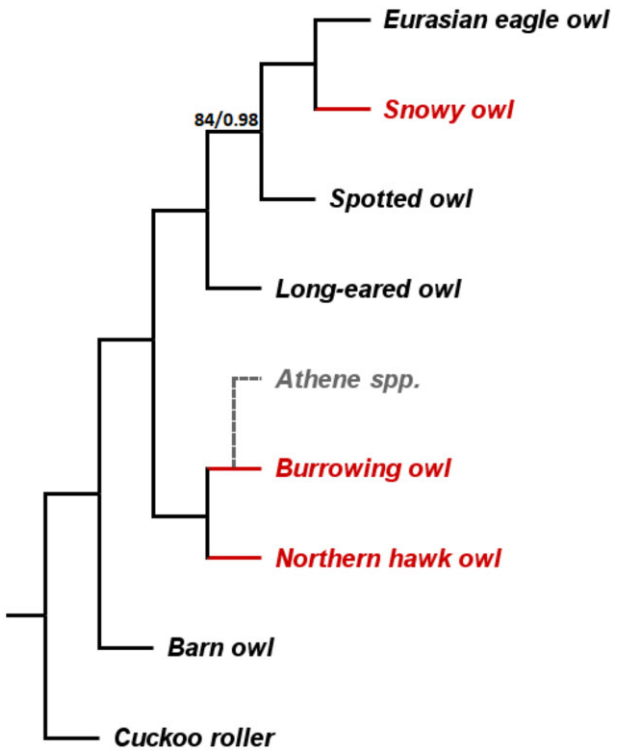


Fig. 1. Phylogenetic topology of the included owl species. A maximum likelihood and Bayesian inference analyses (Salter *et al.* 2020) showed that all nodes received 100% bootstrap support except for the ancestral node of *Strix* and *Bubo* spp., which is labeled with the exact value (ML bootstrap support/Bayesian posterior probability). The tip branches of the three diurnal owl species, where the transition to diurnality occurred, were tested for accelerated evolution (in red). The non-included sister species of the burrowing owl (the rest of the *Athene* spp.) are mostly nocturnal (in gray).

Materials and methods

Study species, reference genome, and multispecies alignment

We used the genome assembly and annotation of *A. cunicularia* (Burrowing owl) as reference for the studied species (Mueller *et al.* 2018). The reference genome was annotated by the NCBI Eukaryotic Genome Annotation Pipeline (NCBI *A. cunicularia* Annotation Release 100; NCBI Assembly Accession GCA_003259725.1 of athCun1).

The genomes of *Asio otus* (long-eared owl), *Bubo bubo* (Eurasian eagle owl), *B. scandiacus* (snowy owl), and *S. ulula* (Northern hawk owl) have been sequenced and mapped to the reference for a previous study (Espíndola-Hernández *et al.* 2020). The genome assemblies of *Strix occidentalis* (spotted owl, Hanna *et al.* 2017), *Tyto alba* (barn owl, Ducrest *et al.* 2020), and *Leptosomus discolor* (cuckoo roller, Zhang *et al.* 2014, used as outgroup) were downloaded from NCBI and mapped to the reference genome using LAST v. 921 (Kiełbasa *et al.* 2011). Despite some ambiguity in the phylogeny of owls, the topological relationships among the included owl species is well established and remained the same in studies using different markers (mitochondrial and ultra-conserved genome-wide markers) (Wink *et al.* 2009; Salter *et al.* 2020). We used the consensus topology of these phylogenetic trees with the Cuckoo roller as the outgroup for all analyses (Fig. 1). We used an unrooted tree that is a modified version of the same topology for the acceleration rate tests in coding genes (see Extended Methods section of the Supplementary File 1). The Supplementary Material provides the accession numbers of the downloaded genomes (Supplementary Table 1 in Supplementary File 2), a general workflow diagram of the analyses (Supplementary Fig. 1 in Supplementary File 1), and a more detailed description of the pipelines and parameters.

We produced a single genome-wide, reference-mapped sequence for each species in four steps. (1) Compilation ("piling up") of all the reads or sequences of the whole genome using samtools (Li *et al.* 2009). (2) Variant calling with bcftools (Danecek and McCarthy 2017). (3) Producing the reference-mapped, species-specific sequence with bcftools, choosing the allele with more reads or better mapping quality in case of heterozygous sites. (4) Soft-masking (change to lowercase) of the repetitive regions (based on the reference genome), and hard-masking (change to "N") of sites with zero-read coverage (per species).

To produce multispecies alignments, we first extracted the sequence of each element (gene or CNEE; see below) from the reference-mapped sequence of each species using bedtools (Quinlan and Hall 2010; Dale *et al.* 2011). We then concatenated the extracted sequences of all species in a single, multispecies FASTA file and ran a multispecies aligner for each element, either using MACSE (Ranwez *et al.* 2011) for the genes, or PRANK (Löytynoja 2014) for the CNEEs. We used MACSE for protein-coding gene sequences because it corrects for potentially erroneous frameshifts (e.g. indels smaller than triplets) without disrupting the underlying codon structure. Finally, we removed high-entropy regions and gaps with BMGE v. 1.12 (Criscuolo and Gribaldo 2010).

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

Avian-specific CNEEs and identification of owl-specific CNEEs

We used the 284,001 avian-specific CNEEs identified and described by Sackton et al. (2019), which are conserved among 35 species across the avian clade, are at least 50 bp long, and include a large fraction of known regulatory elements. The positions of these avian-specific CNEEs are publicly available in the coordinates of the Chicken 4.0 assembly (Sackton et al. 2019). We used MafFilter (Dutheil et al. 2014) to transfer ("liftover") the avian-specific CNEEs from the Chicken 4.0 coordinates to the Burrowing owl (athCun1) coordinates.

Additionally, we identified new owl-specific CNEEs that are shared among nocturnal owls. First, we used PhyloFit (from the software package PHAST: PHylogenetic Analysis with Space-Time models, Hubisz et al. 2011) to estimate a neutral model based on 4-fold degenerate sites (4d sites) of all coding regions of the four nocturnal-owl genomes (Eurasian eagle owl, long-eared owl, spotted owl, and barn owl). We used msa_view (from PHAST) to extract these 4d sites (Hubisz et al. 2011). Then, we used this neutral model (also referred to as the nonconserved model) as null model for the identification of the "most-conserved" regions in the noncoding regions of the four nocturnal-owl genomes with PhastCons (Siepel et al. 2005). We excluded all CNEE sequences <50 bp.

Test for accelerated substitution rates in coding sequences

We followed the method used in Espíndola-Hernández et al. (2020) to test for accelerated rates of evolution in CDS. In brief, we estimated the nonsynonymous to synonymous substitution rate ratio ($\omega = dN/dS$; for a review, see Nielsen 2005) to measure the direction and magnitude of selection on protein-coding genes. A value of $\omega < 1$ indicates purifying selection, $\omega = 1$ neutral evolution, and $\omega > 1$ positive selection. We used the maximum-likelihood method implemented in the CodeML program of PAML 4.9h (Yang 2007), based on the branch model (Yang 1998) and the branch-site model (Yang and Nielsen 2002; Zhang et al. 2005; Yang and dos Reis 2011). For both models, we used a preset unrooted tree topology with the branches of the diurnal owls labeled as the foreground (see Supplementary File 1), and the rest of the tree branches as background.

We consider as an "accelerated substitution rate" each case where the alternative model had a significant better fit to the data and had a $\omega_{foreground} > \omega_{background}$, which mostly indicates positive selection at specific sites, but might also include cases of relaxed purifying selection (Espíndola-Hernández et al. 2020). We applied the false discovery rate (FDR) correction to control for multiple testing for the CodeML models. To complement the selection test results based on CodeML, we used the aBSREL model (Smith et al. 2015), implemented in the HyPhy package, to test for selection signals that are specific for the diurnal owls. This test implements a modified version of the branch-site model to test for selection exclusively in the foreground branches. In this test, we included all the genome-wide significant protein-coding genes according to the CodeML tests. To control for multiple testing, we used the Holm–Bonferroni sequential rejection procedure from the HyPhy package (Smith et al. 2015).

Test for accelerated substitution rates in CNEEs

We tested different evolutionary models to identify an accelerated substitution rate in the foreground branches leading to the three diurnal owls using the Bayesian approach implemented in PhyloAcc (Hu et al. 2019). PhyloAcc uses a hierarchical Bayesian phylogenetic model to identify branches on a phylogeny on which particular genomic elements change their substitution rate, from a conserved or neutral to an accelerated substitution rate (Hu et al. 2019). The conservation or acceleration is estimated in relation to a neutral model. The neutral model was first built using PhyloFit (Hubisz et al. 2011) based on the 4d sites of all coding regions from the complete set of eight bird genomes used in this study (similar to the model used for the detection of owl-specific CNEEs except for the set of species). PhyloAcc considers that the elements have initially evolved at a neutral rate ($r_0 = 1$, having the same substitution rate as the initial neutral model), and then become conserved at the root or some other branch on the phylogeny ($r_1 < 1$, having a lower substitution rate than the neutral model). The elements might then evolve with an accelerated rate ($r_2 > r_1$, having a higher substitution rate than the conserved state) (Hu et al. 2019). The program PhyloAcc restricts the possible shift patterns in 3 nested models: the null model (M_0), where the substitution rate in any branch is not allowed to shift to an accelerated rate; the lineage-specific model (M_1), where the substitution rates of the diurnal owls are allowed to shift to an accelerated rate; and the full model (M_2), where the substitution rate of any branch is allowed to shift. The marginal likelihood of the data under each model is compared by two Bayes factors, BF1 and BF2 (Hu et al. 2019). Briefly, BF1 is the ratio of the marginal likelihoods of the data under M_1 and M_0, indicating how much the data support M_1 in relation to M_0. BF2 is the ratio of the marginal likelihoods of the data under M_1 and M_2, indicating how much the data support M_1 in relation to M_2. To identify DNA elements accelerated exclusively in target lineages, Hu et al. (2019) recommend considering only cases with high values in both Bayes factors. Thus, we considered a CNEE as a candidate for parallel accelerated evolution during the emergence of diurnality in owls when the following conditions were met: $\log_{BF1} \geq 10$, $\log_{BF2} \geq 1$, and the posterior probability for accelerated evolution under $M_2 > 0.8$ for at least two of the three diurnal owl species. The distribution of both Bayes factors across all tested CNEEs is shown in Supplementary Fig. 2 in Supplementary File 1.

Functional overrepresentation analysis

Within each group of elements (genes, avian-specific and owl-specific CNEEs), we ranked the elements according to the strength of evidence for accelerated evolution in the diurnal owls, and applied a Wilcoxon rank-sum test with the R package GOfuncR (Grote 2018). We used a custom-made gene ontology (GO) annotation database made for all annotated athCun1 genes, combining human (org.Hs. e.g. db) and chicken (org.Gg. e.g. db) annotations to GOs. We ranked the genes by the test statistic (log-likelihood ratio value) of the branch or branch-site test, and included only the genes with an accelerated substitution rate in the diurnal owls ($\omega_{background} < \omega_{foreground}$) in the case of the branch model. The CNEEs were ranked by a custom-made parameter based on the posterior probability of acceleration (pp) along the phylogenetic tree. For each branch, we estimated the probability of acceleration relative to that of the respective ancestral branch ($pp_{branch} - pp_{ancestral\ branch}$). Then, we summed these relative probabilities for all diurnal owls and for the other branches from the nocturnal species. The custom-made parameter for the CNEEs is then the difference between the sum from the diurnal species and the sum from the nocturnal species (see the formula in Supplementary File 1).

For the ranked genes, the gene-GO annotation file was used directly, while for the ranked CNEEs, we produced a CNEE-GO

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

annotation file using the GOs of nearby genes. We linked CNEEs to genes by intersecting the CNEEs with the putative "Gen Regulatory Domain Region" of "One Closest" genes established by GREAT (McLean *et al.* 2010). The "One Closest" option of GREAT determines for each gene a potential regulatory domain that extends maximally 1 Mb from the Transcription Start Site (TSS) in both directions until the mid-point between this TSS and the TSS of the adjacent gene (McLean *et al.* 2010). To account for multiple testing and for potential clustering of CNEEs around genes, we used the family-wise error rate (FWER) estimation procedure of GOfuncR, which permutes the ranking parameter while the annotations of CNEEs or genes to GO categories stay fixed and re-estimates the statistics for every GO term (Grote 2018).

Comparison of evolutionary rates between protein-coding genes and CNEEs

The comparison between rates of evolution in coding and non-coding regions of the genome is not straightforward. The codon structure of genes adds another level of complexity in evolutionary models, due to the different constraints of substitutions for each of the nucleotide positions in a codon. The sites in the non-coding CNEEs apparently do not show systematic patterns of evolutionary constraints. However, depending on the definition of CNEEs, they likely also include neutrally evolving and more or less conserved sites. Thus, we compared the acceleration rates of evolution of CDS and CNEEs by a simple substitution model without considering the codon structure, using PhyloP (Pollard *et al.* 2010). The scale estimates indicate the rate of evolution relative to the neutral model (same model as for the PhyloAcc analysis described above). To this end, we ran the likelihood ratio test of PhyloP with the lineage-specific option to compare a null model having one single scale parameter with an alternative model having two estimated scale parameters: one scale for the branches leading to the diurnal owls (foreground scale) and a second scale for all remaining branches (background scale). We compared the distributions of the estimated subscale (ratio between foreground and background scale in the alternative model) between coding genes and CNEEs. We used ggplot2 to visualize these distributions (Wickham 2016). To account for variation in sequence length of the tested elements (a potential confounder for scale estimates), we plotted subscale for different length intervals. This allows the comparison of subscale values between protein-coding genes and CNEEs of similar length.

Results

Accelerated evolutionary rates in coding genes and CNEEs and their functional enrichment

Of the 12,298 tested protein-coding genes, 69 showed a significantly higher ω-value in the diurnal owls compared to the background of nocturnal species (branch model: FDR $\leq$ 0.05, Supplementary Table 2 in Supplementary File 2), and 15 showed accelerated substitution rates on specific sites of the diurnal owl sequences (branch-site model: FDR $\leq$ 0.05, Supplementary Table 2 in Supplementary File 2). Seven of these genes showed evidence for positive selection at specific sites in at least two diurnal owls and not in any other species (*IKZF2*, *SOX18*, *JPH2*, *WNT4*, *CAMK1D*, *GIT2*, and *CASP8*), according to the aBSREL model (Supplementary Table 7 in Supplementary File 2).

Based on the branch-model tests, we found no evidence for functional enrichment among the high-ranked genes. For the branch-site model tests, high-ranked genes were significantly enriched for the GO term "HAUS complex" (Wilcoxon rank-sum test, FWER = 0.001, Supplementary Table 3 in Supplementary File 2). According to the complementary aBSREL model, the functions of the significantly accelerated 7 genes are predominantly related to regulatory functions, including transcription regulation (Supplementary Table 7 in Supplementary File 2).

Among the 265,599 tested avian-specific CNEEs (Sackton *et al.* 2019), 113 elements showed significantly accelerated rates of evolution in diurnal owls based on the Bayes factor thresholds, 13 of these were accelerated in 2 diurnal owl species, and only two showed evidence for accelerated evolution in all three diurnal owl species according to the threshold of the posterior probability of acceleration in the full model (Fig. 2 and Supplementary Table 4 in Supplementary File 2). The high-ranked avian-specific CNEEs were significantly enriched for elements linked to one GO term associated to the axolemma, the plasma membrane of the neurons' axon (Wilcoxon rank-sum test, FWER < 0.05, Supplementary Table 5 in Supplementary File 2). There are 629 avian-CNEEs in the putative regulatory domain of 12 genes (*ADORA1*, *ADORA2A*, *ANK1*, *CNTNAP2*, *EPB41L3*, *KCNC2*, *KCNJ11*, *MAPT*, *MYO1D*, *ROBO2*, *SPTBN1*, and *THY1*) related to this GO term (Supplementary Table 5 in Supplementary File 2). We manually annotated these genes using public gene databases and found that most of them are functionally linked to neuronal development and connectivity (Supplementary Table 6 in Supplementary File 2).

We identified 2,364 new owl-specific CNEEs present among all four nocturnal owl species. From these, 31 showed evidence for accelerated evolution in at least one of the diurnal owl species based on the Bayes factor thresholds. Only three of them had a posterior probability of accelerated evolution above the threshold in the full model in two diurnal owl species (Fig. 2), and none in all three diurnal species. Twenty-eight showed evidence for accelerated evolution in the snowy owl only. There was no genome-wide significant functional enrichment of GO terms among the ranked owl-specific CNEEs.

The genome-wide detected genes (CDS) and the genes associated with the CNEEs with evidence for accelerated evolution in at least two diurnal owl species do not have elements in common. This is true for both, owl-specific CNEEs and avian-specific CNEEs.

Comparison of acceleration rates between genes and CNEEs

According to the LRT from PhyloP and after correction for multiple testing, 2.3% of the genes (278 out of 12,298 genes tested), 2.8% of the owl-specific CNEEs (67 out of 2,364 CNEEs tested), and 0.1% of the avian-specific CNEEs (329 out of 265,599 CNEEs tested) showed evidence for accelerated evolution in the diurnal owls (PhyloP results with subscale > 1 and FDR corrected *P*-value < 0.05, i.e. genome-wide significance).

Among the genome-wide significantly accelerated elements, the sub-scale values of CNEEs were generally higher than those of the protein-coding genes (Figs. 3 and 4). The CNEEs also showed outlier groups of extreme values. To account for the fact that CNEEs are on average shorter than protein-coding genes, we compared subscale values within intervals of sequence lengths (Fig. 4). Most of the elements with extreme subscale values had also extreme sequence lengths. Those cases were excluded from Fig. 4, which only shows the length intervals for which data from all three categories of elements were available (see legend).

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

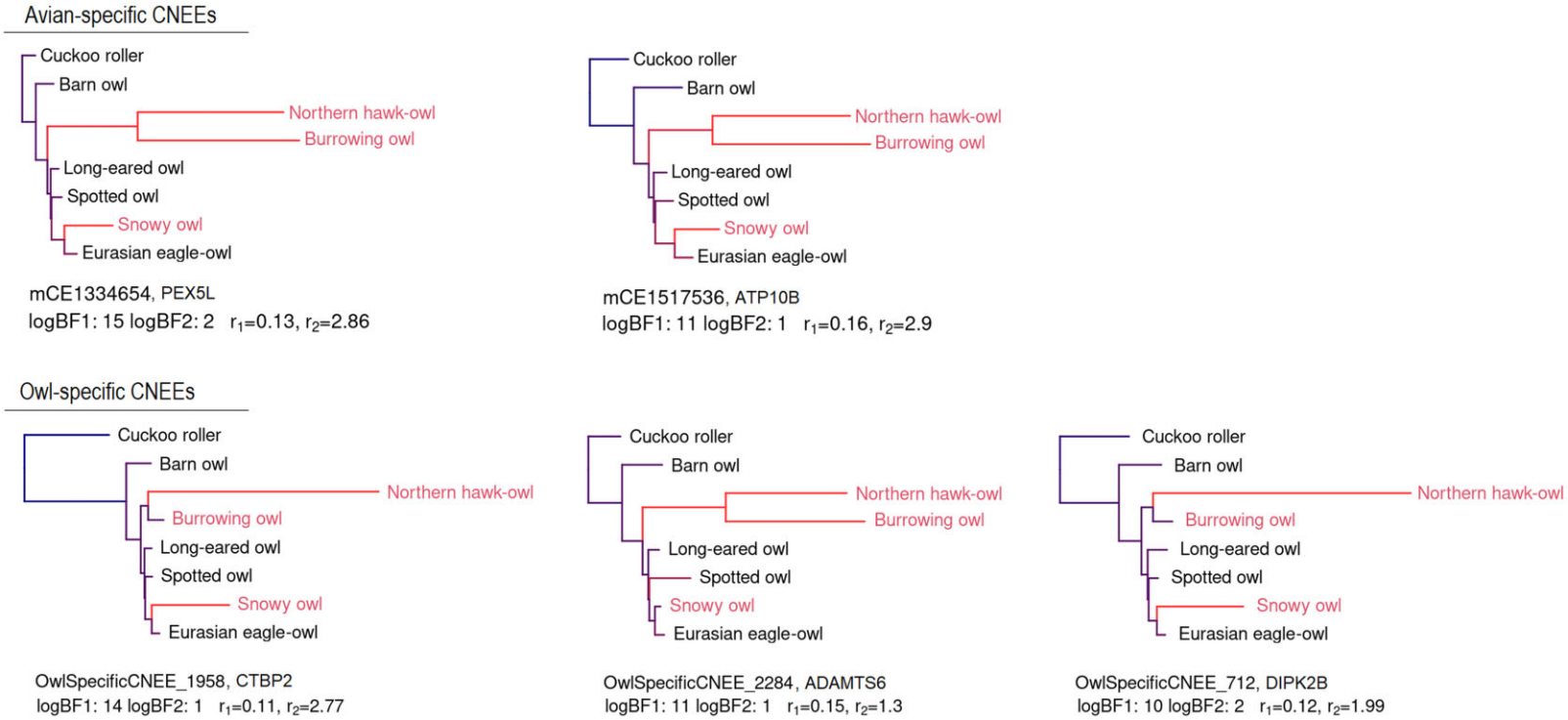


Fig. 2. Avian- and owl-specific CNEEs with evidence of accelerated rates of evolution in diurnal owls. The phylogenetic tree illustrates the shift in substitution rates under the full model [according to Hu *et al.* (2019) and Sackton *et al.* (2019)]. Diurnal species are indicated in red. The branch lengths are proportional to the posterior mean substitution rate. The line below each tree shows the name of the CNEE and its associated gene, the values of the two log-BFs, and the conserved (r_1) and accelerated substitution rate (r_2). The sequence data support a parallel shift from a conserved CNEE indicating purifying selection (blue) to an accelerated substitution rate (red) in two or three diurnal owl species.

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

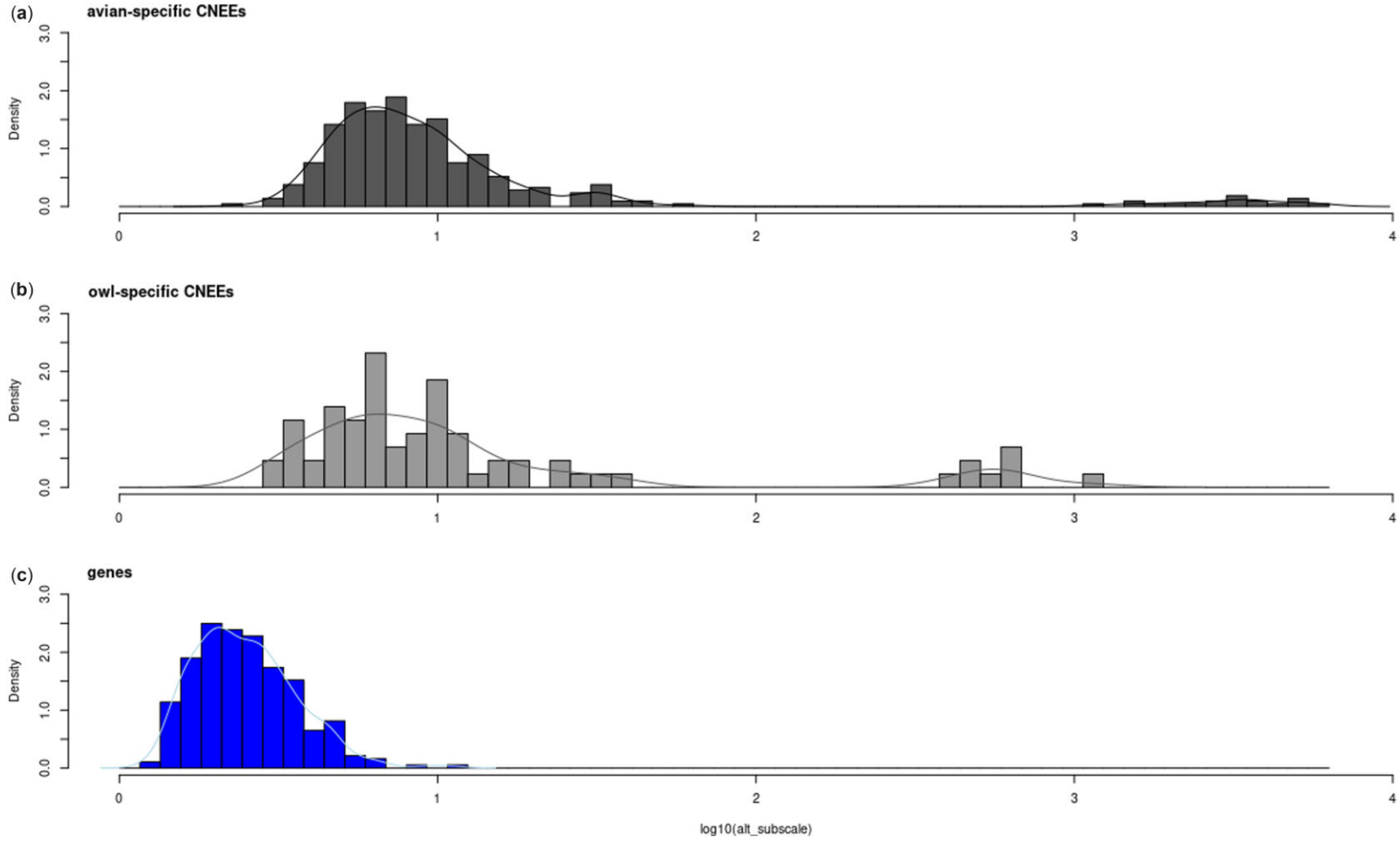


Fig. 3. Distributions of subscale values in the alternative model (alt_subscale) from significantly accelerated elements (PhyloP results with subscale > 1 and FDR-corrected *P*-value < 0.05). The plot compares the histograms and density curves of $\log_{10}$-transformed subscale values of avian-specific CNEEs (a, black, $N = 329$), owl-specific CNEEs (b, gray, $N = 67$), and genes (c, blue, $N = 286$). Most of the elements with values on the tail ends of these distributions have also extreme sequence lengths (genes longer than 1,000 bp, or CNEEs shorter than 200 bp; see Fig. 4).

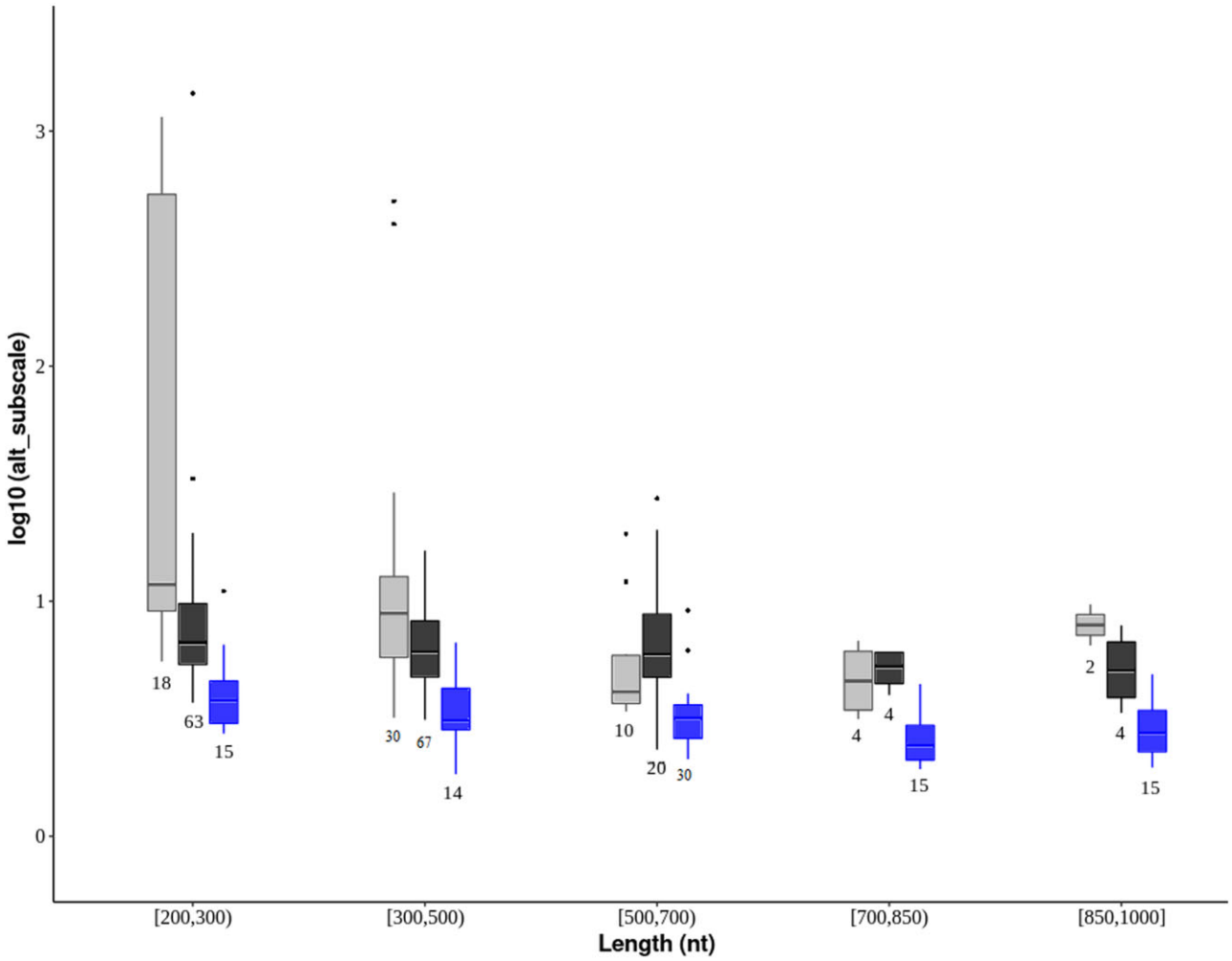


Fig. 4. Comparison of subscale values between owl-specific CNEEs (gray), avian-specific CNEEs (black), and genes (blue) in relation to sequence length (intervals). Only elements with genome-wide significance are included (PhyloP results with subscale >1 and FDR-corrected *P*-value < 0.05). Shown are box plots with sample sizes (number of elements). We excluded genes longer than 1,000 bp, and avian-specific CNEEs shorter than 200 bp, because there were no owl-specific CNEEs outside of the range from 200 to 1,000 bp.

Discussion

Our study aimed at detecting genomic signals of selection linked to the evolution of a diurnal lifestyle in owls, whereby we searched for accelerated substitution rates in protein-coding and noncoding elements of the genome. Our results showed that accelerated substitution rates during the evolution of diurnality in owls occurred in both coding and noncoding regions of the genome. The absolute number of significantly accelerated elements was comparable between protein-coding genes and CNEEs. However, among those elements with evidence for accelerated evolution, the magnitude of acceleration (subscale value) was larger in the CNEEs than in the protein-coding genes. Our comparison between these genomic regions is based on a general substitution model without considering the codon structure of the protein-coding genes or the expected variable evolutionary constraints among the sites in CNEEs. Hence, our approach can only serve as a rough average comparison across all elements. Further, most of the protein-coding genes with signals of positive selection at specific sites exclusively in the diurnal owls are associated with regulatory processes of gene expression. This functional association with regulatory processes and the higher magnitude of acceleration in potentially *cis*-regulatory elements (CNEEs) suggest that regulatory evolution might have been more relevant than structural evolution during the shift to a diurnal lifestyle in owls.

Structural and regulatory changes as mechanisms for adaptation have long been discussed. Several papers have reviewed evidence about which part of the genome plays a more relevant role in adaptative evolution (Macintyre 1982; Carroll 2005; Wray 2007; Hoekstra and Coyne 2007; Romero *et al.* 2012; Hill *et al.* 2021). Different levels of pleiotropy are important for evolutionary hypotheses about why genetic substitutions might occur more frequently in regulatory noncoding regions than in structural protein-coding genes. Many of the CNEEs are in *cis*-regulatory genomic regions with modular organization, such that they regulate the expression of only one nearby gene and are affected by only a single transcription factor. This implies that a mutation in one of the many regulatory modules might selectively affect only one aspect of the gene-expression network, e.g. only in a specific tissue (Wray 2007; Molodtsova *et al.* 2014). Another important aspect regarding the evolution of noncoding regions is their functional redundancy. In some cases, regulatory elements share functions and this redundancy acts as a buffer against genetic

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

disturbances, allowing genetic changes without compromising essential biological functions. The buffering conferred by the *cis*-regulatory redundancy might mediate the recruitment of novel regulatory binding sites from existing ones and eventually the achievement of novel gene regulation pathways (Hong *et al.* 2008; Frankel *et al.* 2010; Perry *et al.* 2010; Wittkopp and Kalay 2012; Rubinstein and de Souza 2013). Many behavioral traits are inherently dynamic and this might need "fine tuning" by regulatory responses to a dynamic environment (Macintyre 1982; Wray 2007). Diel activity is such a behavioral trait that might require a dynamic and flexible control, and therefore is expected to predominantly evolve through regulatory mutations affecting specific gene regulatory network interactions.

The significance of regulatory evolution has been highlighted in other recent studies in birds. For instance, Seki *et al.* (2017) found that birds have a higher proportion of conserved elements in the non-coding part of the genome in comparison to mammals. The authors showed that these avian-specific, highly conserved elements in the noncoding region are associated with genes that participate in the development of avian limbs and flight feathers. Their results support the hypothesis that changes in noncoding regulatory sequences might have played an important role in the emergence of avian evolutionary innovations (Seki *et al.* 2017). Additional support for this hypothesis came from a comparative study among palaeognathous species (Sackton *et al.* 2019). This study showed that noncoding elements with accelerated rates of evolution were overrepresented near key limb developmental genes. They further proved the *cis*-regulatory activity of the CNEEs through their effect of open chromatin states during embryonic development. Thus, the study suggested that convergent morphological evolution and loss of flight in ratites were more strongly associated with changes in the regulatory noncoding part of the genome than in protein-coding genes (Sackton *et al.* 2019). In another study, Yusuf *et al.* (2020) identified candidate loci related to macro-evolutionary shifts in bird beak shape evolution across distantly related avian taxa, and studied whether those morphological shifts were explained by shifts in molecular rates of coding and noncoding genomic regions. The study found that signals in the noncoding regions were more often associated with avian bill shape diversification.

Each identified signal of selection or accelerated evolution provides a candidate element, either coding or noncoding, for further study of parallel evolution of diurnality in owls. We attempted to interpret and summarize these signals using functional enrichment analyses of GO terms. Among the protein-coding elements, the high-ranked genes showed a significant association with the GO term "HAUS complex" (*HAUS1*, *HAUS2*, *HAUS3*, *HAUS6*, and *HAUS8*; Supplementary Table 3 in Supplementary File 2). This GO term refers to a microtubule-binding complex involved in the generation of the mitotic spindle (Goshima *et al.* 2008), that also plays a key role in neuronal migration, polarization, and development through local regulation of the cytoskeleton in axons and dendrites (Cunha-Ferreira *et al.* 2018). Due to its effects on the development of neuronal connectivity in the brain, it might play a role in the evolution of behavior (Mueller *et al.* 2020), and consequently in adaptations to a diurnal lifestyle in owls.

Among the avian-CNEEs, only *PEX5L* and *ATP10B* showed accelerated substitution rates in all three diurnal species. These two genes are involved in the organization and maintenance of organelles in the cytoplasm, especially in the brain cells. In particular, they function in the cortical neurons (*PEX5L* in the peroxisomes and *ATP10B* in the maintenance of lysosome membrane integrity). Considering the difference between the diurnal and nocturnal species in terms of their probability of shift to acceleration (ranking parameter), there was an overrepresentation of high-ranked avian-CNEEs placed around genes functionally linked to the axolemma GO term. We inspected the functions of the genes in this GO term, using information from GeneCards (www.genecards.org, last accessed: 11.03.2022), NCBI gene (www.ncbi.nlm.nih.gov/search/, last accessed: 11.03.2022), and amiGO2 (http://amigo.geneontology.org/amigo, last accessed: 11.03.2022). Most of these genes are involved in interactions between the intra- and extra-cellular environment through the plasma membrane, especially in the brain cells, and several of these genes were related to the development of neurons and the regulation of membrane potentials in the neurons (*ADORA1*, *ADORA2A*, *CNTNAP2*, *EPB41L3*, *KCNC2*, *KCNJ11*, *MAPT*, *MYO1D*, *ROBO2*, and *THY1*; Supplementary Table 6 in Supplementary File 2). Four of these genes are related to human phenotypes that involve a variety of abnormalities in the development of eyes and ears (*ANK1*, *MAPT*, *MYO1D*, and *SPTBN1*; Supplementary Table 6 in Supplementary File 2), and two of these genes are related to regulation of the circadian rhythm and sleep (*ADORA1*, and *ADORA2A*; Supplementary Table 6 in Supplementary File 2). These genes therefore seem to be good candidates in the context of adaptation to a diurnal lifestyle in the owls.

In addition to the avian-specific CNEEs (Sackton *et al.* 2019), we identified 2,364 new owl-specific CNEEs among the nocturnal owls. These owl-specific CNEEs are candidates for regulatory elements during the evolution of owls. Only three of these elements are strong candidates for regulatory changes during the evolution of diurnality in owls, showing accelerated substitution rates in at least two of the three diurnal species. These three owl-specific CNEEs are linked to the genes *ADAMTS6*, *CTBP2*, and *DIPK2B*. *ADAMTS6* is generally involved in proteolysis, and kidney and heart development, but also encodes two isoforms that are upregulated by tumor necrosis factor alpha (TNFα) in retinal pigment epithelial cells (Bevitt *et al.* 2003; Lu *et al.* 2013). One of the isoforms encoded by *CTBP2* (ribeye) is a major component of specialized synapses known as synaptic ribbons. These specialized synapses are involved in visual (Schmitz *et al.* 2000) and auditory perception (West and McDermott 2011), as well as circadian timing and the pupillary light reflex (Hannibal and Fahrenkrug 2006; Østergaard *et al.* 2007). Mutations in the human *CTBP2* have been linked to retinitis pigmentosa, night blindness, and deafness (GeneCards, www.genecards.org, last accessed: 11.03.2022). *DIPK2B* (*DIA1R*) encodes signal peptides for protein targeting in the secretory pathway, and is expressed in embryonic and adult brain tissues.

We found no genes in common between those identified as showing evidence for genome-wide accelerated substitution rates and those associated with the CNEEs that showed such evidence in at least two diurnal owls. This result is in line with another comparative study (Yusuf *et al.* 2020), which showed 2 different sets of genes associated with signals of accelerated evolution in coding and noncoding regions, even though both were implicated in beak development.

In summary, our results showed that accelerated evolution occurs in coding and noncoding conserved genomic regions during the emergence of diurnality in owls. Acceleration rates were higher in the noncoding elements than in the protein-coding genes, and accelerated protein-coding genes in diurnal owls are functionally associated with regulation of gene expression. Our results suggest that regulatory evolution might have played a

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

predominant role in the shift to a diurnal lifestyle in owls. In addition, as expected for a shift to a diurnal lifestyle with sensory and behavioral adaptations, several accelerated noncoding and coding elements are functionally linked to nervous system development and brain connectivity.

Data availability

Sequence data are publicly available, and their references and accession numbers are listed in Supplementary Table 1 in Supplementary File 2. The multispecies alignments of all elements are available in repository https://doi.org/10.25387/g3.19369118.

Supplemental material is available at G3 online.

Acknowledgments

We thank Meng-Ching Ko and Yifan Pei for constructive feedback during the early stages of this study.

Funding

This work was funded by the Max Planck Society (to BK), and by a scholarship for doctoral studies in Germany cofunded by the DAAD and BecasChile (to PE-H).

Conflicts of interest

The authors declare to have no competing interests.

Literature cited

Anderson SR, Wiens JJ. Out of the dark: 350 million years of conservatism and evolution in diel activity patterns in vertebrates. Evolution. 2017;71(8):1944–1959.

Bevitt DJ, Mohamed J, Catterall JB, Li Z, Arris CE, Hiscott P, Sheridan C, Langton KP, Barker MD, Clarke MP, *et al.* Expression of ADAMTS metalloproteinases in the retinal pigment epithelium derived cell line ARPE-19: transcriptional regulation by TNFα. Biochim Biophys Acta. 2003;1626(1–3):83–91.

Burga A, Wang W, Ben-David E, Wolf PC, Ramey AM, Verdugo C, Lyons K, Parker PG, Kruglyak L. A genetic signature of the evolution of loss of flight in the Galapagos cormorant. Science. 2017;356(6341):eaal3345.

Carothers JH, Jaksic FM. Time as a niche difference: the role of interference competition. Oikos. 1984;42(3):403–406.

Carroll SB. Evolution at two levels: on genes and form. PLoS Biol. 2005;3(7):e245.

Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010;10:210.

Cunha-Ferreira I, Chazeau A, Buijs RR, Stucchi R, Will L, Pan X, Adolfs Y, van der Meer C, Wolthuis JC, Kahn OI, *et al.* The HAUS complex is a key regulator of non-centrosomal microtubule organization during neuronal development. Cell Rep. 2018;24(4):791–800.

Dale RK, Pedersen BS, Quinlan AR. Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. Bioinformatics. 2011;27(24):3423–3424.

Danecek P, McCarthy SA. BCFtools/csq: haplotype-aware variant consequences. Bioinformatics. 2017;33(13):2037–2039.

del Hoyo J, Elliott A, Sargatal J, Cabot J. Handbook of the Birds of the World. Barcelona: Lynx Edicions; 1999.

Ducrest A-L, Neuenschwander S, Schmid-Siegert E, Pagni M, Train C, Dylus D, Nevers Y, Warwick Vesztrocy A, San-Jose LM, Dupasquier M, *et al.* New genome assembly of the barn owl (*Tyto alba alba*). Ecol Evol. 2020;10(5):2284–2298.

Duncan JR. Owls of the World: Their Lives, Behavior and Survival. Baltimore (MD): Johns Hopkins University Press; 2018.

Dutheil JY, Gaillard S, Stukenbrock EH. MafFilter: a highly flexible and extensible multiple genome alignment files processor. BMC Genomics. 2014;15:53.

Edwards SV, Cloutier A, Baker AJ. Conserved nonexonic elements: a novel class of marker for phylogenomics. Syst Biol. 2017;66(6):1028–1044.

Ericson PGP, Anderson CL, Britton T, Elzanowski A, Johansson US, Källersjö M, Ohlson JI, Parsons TJ, Zuccon D, Mayr G, *et al.* Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett. 2006;2(4):543–547.

Espíndola-Hernández P, Mueller JC, Carrete M, Boerno S, Kempenaers B. Genomic evidence for sensorial adaptations to a nocturnal predatory lifestyle in owls. Genome Biol Evol. 2020;12(10):1895–1908.

Feduccia A. Explosive evolution in tertiary birds and mammals. Science. 1995;267(5198):637–638.

Feduccia A. The Origin and Evolution of Birds. New Haven (CT): Yale University Press; 1999.

Feduccia A. 'Big bang' for tertiary birds? Trends Ecol. Evol. 2003;18(4):172–176.

Frankel N, Davis GK, Vargas D, Wang S, Payre F, Stern DL. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature. 2010;466(7305):490–493.

Goshima G, Mayer M, Zhang N, Stuurman N, Vale RD. Augmin: a protein complex required for centrosome-independent microtubule generation within the spindle. J Cell Biol. 2008;181(3):421–429.

Gould SJ. The Structure of Evolutionary Theory. Cambridge (MA): Harvard University Press; 2002.

Grote S. 2018. GOfuncR: gene ontology enrichment using FUNC. R package version 1.2.0.

Hackett SJ, Kimball RT, Reddy S, Bowie RCK, Braun EL, Braun MJ, Chojnowski JL, Cox WA, Han K-L, Harshman J, *et al.* A phylogenomic study of birds reveals their evolutionary history. Science. 2008;320(5884):1763–1768.

Hanna ZR, Henderson JB, Wall JD, Emerling CA, Fuchs J, Runckel C, Mindell DP, Bowie RCK, DeRisi JL, Dumbacher JP, *et al.* Northern spotted owl (*Strix occidentalis caurina*) genome: divergence with the barred owl (*Strix varia*) and characterization of light-associated genes. Genome Biol Evol. 2017;9(10):2522–2545.

Hannibal J, Fahrenkrug J. Neuronal input pathways to the brain's biological clock and their functional significance. Adv Anat Embryol Cell Biol. 2006;182:1–71.

Hill MS, Vande Zande P, Wittkopp PJ. Molecular and evolutionary processes generating variation in gene expression. Nat Rev Genet. 2021;22(4):203–215.

Hoekstra HE, Coyne JA. The locus of evolution: evo devo and the genetics of adaptation. Evolution. 2007;61(5):995–1016.

Hong JW, Hendrix DA, Levine MS. Shadow enhancers as a source of evolutionary novelty. Science. 2008;321(5894):1314.

Hu Z, Sackton TB, Edwards SV, Liu JS. Bayesian detection of convergent rate changes of conserved noncoding elements on phylogenetic trees. Mol Biol Evol. 2019;36(5):1086–1100.

Hubisz MJ, Pollard KS, Siepel A. Phast and Rphast: phylogenetic analysis with space/time models. Brief Bioinform. 2011;12(1):41–51.

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

Jaksić FM. Inadequacy of activity time as a niche difference: the case of diurnal and nocturnal raptors. Oecologia. 1982;52(2):171–175.
Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT, *et al.* Whole-genome analyses resolve early branches in the tree of life of modern birds. Science. 2014; 346(6215):1320–1331.
Kiełbasa SM, Wan R, Sato K, Horton P, Frith MC. Adaptive seeds tame genomic sequence comparison. Genome Res. 2011;21(3): 487–493.
King M, Wilson a C. Evolution at two levels in humans and chimpanzees. Science. 1975;188(4184):107–116.
König C, Weick F. Owls of the World. London: Christopher Helm. A & C Black; 2008.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–2079.
Löytynoja A. Phylogeny-aware alignment with PRANK. In: D Russell, editor. Multiple Sequence Alignment Methods. Totowa (NJ): Humana Press; 2014. p. 155–170.
Lu Y, Vitart V, Burdon KP, Khor CC, Bykhovskaya Y, Mirshahi A, Hewitt AW, Koehn D, Hysi PG, Ramdas WD, *et al.*; NEIGHBOR Consortium. Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus. Nat Genet. 2013;45(2):155–163.
Macintyre RJ. Regulatory genes and adaptation. In: MK Hecht, B Wallace, and GT Prance, editors. Evolutionary Biology. Boston (MA): Springer; 1982. p. 247–285.
McClure CJW, Schulwitz SE, Anderson DL, Robinson BW, Mojica EK, Therrien J-F, Oleyar MD, Johnson J. Commentary: defining raptors and birds of prey. J. Raptor Res 2019;53(4):419.
McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010;28(5):495–501.
Molodtsova D, Harpur BA, Kent CF, Seevananthan K, Zayed A. Pleiotropy constrains the evolution of protein but not regulatory sequences in a transcription regulatory network influencing complex social behaviours. Front Genet 2014;5:431–437.
Mueller JC, Carrete M, Boerno S, Kuhl H, Tella JL, Kempenaers B. Genes acting in synapses and neuron projections are early targets of selection during urban colonization. Mol Ecol. 2020;29(18): 3403–3412.
Mueller JC, Kempenaers B, Kuhl H, Boerno S, Tella JL, Carrete M. Evolution of genomic variation in the burrowing owl in response to recent colonization of urban areas. Proc R Soc B Biol Sci. 2018; 285:1–9.
Nielsen R. Molecular signatures of natural selection. Annu. Rev. Genet. 2005;39:197–218.
Østergaard J, Hannibal J, Fahrenkrug J. Synaptic contact between melanopsin-containing retinal ganglion cells and rod bipolar cells. Invest Ophthalmol Vis Sci. 2007;48(8):3812–3820.
Pan S, Lin Y, Liu Q, Duan J, Lin Z, Wang Y, Wang X, Lam SM, Zou Z, Shui G, *et al.* Convergent genomic signatures of flight loss in birds suggest a switch of main fuel. Nat Commun. 2019;10:1–11.
Pearce T. Convergence and parallelism in evolution: a neo-Gouldian account. Br J Philos Sci. 2012;63(2):429–448.
Pei Y, Valcu M, Kempenaers B. Interference competition pressure predicts the number of avian predators that shifted their timing of activity. Proc R Soc B. 2018;285(1880):20180744.
Perry MW, Boettiger AN, Bothma JP, Levine M. Shadow enhancers foster robustness of Drosophila gastrulation. Curr Biol. 2010; 20(17):1562–1567.
Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 2010;20(1):110–121.
Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature. 2015; 526(7574):569–573.
Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841–842.
Ranwez V, Harispe S, Delsuc F, Douzery EJP. MACSE: multiple alignment of coding SEquences accounting for frameshifts and stop codons. PLoS One. 2011;6(9):e22594.
Roll U, Dayan T, Kronfeld-Schor N. On the role of phylogeny in determining activity patterns of rodents. Evol Ecol. 2006;20(5): 479–490.
Romero IG, Ruvinsky I, Gilad Y. Comparative studies of gene expression and the evolution of gene regulation. Nat Rev Genet. 2012; 13(7):505–516.
Rosenblum EB, Parent CE, Brandt EE. The molecular basis of phenotypic convergence. Annu Rev Ecol Evol Syst. 2014;45(1): 203–226.
Rubinstein M, de Souza FSJ. Evolution of transcriptional enhancers and animal diversity. Philos Trans R Soc B Biol Sci. 2013;368: 20130017.
Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, Wheeler NE, Gardner PP, Clarke JA, Baker AJ, Clamp M, *et al.* Convergent regulatory evolution and loss of flight in paleognathous birds. Science. 2019; 364(6435):74–78.
Salter JF, Oliveros CH, Hosner PA, Manthey JD, Robbins MB, Moyle RG, Brumfield RT, Faircloth BC. Extensive paraphyly in the typical owl family (Strigidae). Auk. 2020;137:1–15.
Schmitz F, Königstorfer A, Südhof TC. RIBEYE, a component of synaptic ribbons: a protein's journey through evolution provides insight into synaptic ribbon function. Neuron. 2000; 28(3):857–872.
Schoener TW. Resource partitioning in ecological communities. Science. 1974;185(4145):27–39.
Seki R, Li C, Fang Q, Hayashi S, Egawa S, Hu J, Xu L, Pan H, Kondo M, Sato T, *et al.* Functional roles of Aves class-specific cis-regulatory elements on macroevolution of bird-specific features. Nat Commun. 2017;8:1–14.
Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005;15(8):1034–1050.
Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. Less is more: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. Mol Biol Evol. 2015;32(5):1342–1353.
Stern DL. The genetic causes of convergent evolution. Nat Rev Genet. 2013;14(11):751–764.
Tiley GP, Pandey A, Kimball RT, Braun EL, Burleigh JG. Whole genome phylogeny of Gallus: introgression and data-type effects. Avian Res. 2020;11(1):15.
West MC, McDermott BM. Ribeye a-mCherry fusion protein: A novel tool for labeling synaptic ribbons of the hair cell. J Neurosci Methods. 2011;197(2):274–278.
Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag. ISBN 978-3-319-24277-4.
Wink M, El-Sayed A-A, Sauer-Gürth H, Gonzalez J. Molecular phylogeny of owls (Strigiformes) inferred from DNA sequences of the

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

mitochondrial cytochrome *b* and the nuclear *RAG-1* gene. Ardea. 2009;97(4):581–591.

Wittkopp PJ, Kalay G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat Rev Genet. 2012;13(1):59–69.

Wray GA. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet. 2007;8(3):206–216.

Yang Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol. 1998; 15(5):568–573.

Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24(8):1586–1591.

Yang Z, dos Reis M. Statistical properties of the branch-site test of positive selection. Mol Biol Evol. 2011;28(3):1217–1228.

Yang Z, Nielsen R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol. 2002;19(6):908–917.

Yusuf L, Heatley MC, Palmer JPG, Barton HJ, Cooney CR, Gossmann TI. Noncoding regions underpin avian bill shape diversification at macroevolutionary scales. Genome Res. 2020;30(4):553–565.

Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, *et al.*; Avian Genome Consortium. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 2014;346(6215):1311–1320.

Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol. 2005;22(12):2472–2479.

Communicating editor: A. Rokas

Downloaded from https://academic.oup.com/g3journal/article/12/8/jkac135/6595023 by guest on 03 August 2022

Supplemental File 1

Genomic signatures of the evolution of a diurnal lifestyle in Strigiformes

Pamela Espíndola-Hernández[1*], Jakob C. Mueller[1], Bart Kempenaers[1]

[1] Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany
[*] Corresponding author: pamela.dola@gmail.com

Figures

Figure S1. General workflow

Figure S2. Distributions of Bayes factors

Extended Methods

I. Reference-mapping of sequences of the whole genome for each species
II. Multi-species aligning and trimming of conserved coding and non-coding sequences (CDS and CNEEs)
III. Producing the non-conserved (neutral) model
    1) Extracting the 4d-sites from the MSA of genes
    2) Producing the nonconserved-4d model with PhyloFit
IV. Identifying owl-specific CNEEs with PhastCons (4-Nocturnal Owls)
V. Test on "CNEEs" with phyloAAC
VI. Test on "CDSs" with CodeML
VII. Test on "CDSs" with aBSREL
VIII. GO annotation by Regulatory Domains from GREAT and GO overrepresentation analysis by GOfuncR
    1) GREAT getting the gen regulatory domain region
    2) Intersect the CNEEs with the gen regulatory domains from GREAT
    3) GOfuncR Wilcoxon rank-sum test
IX. Comparison CDS v/s CNEE by PhyloP

Figures

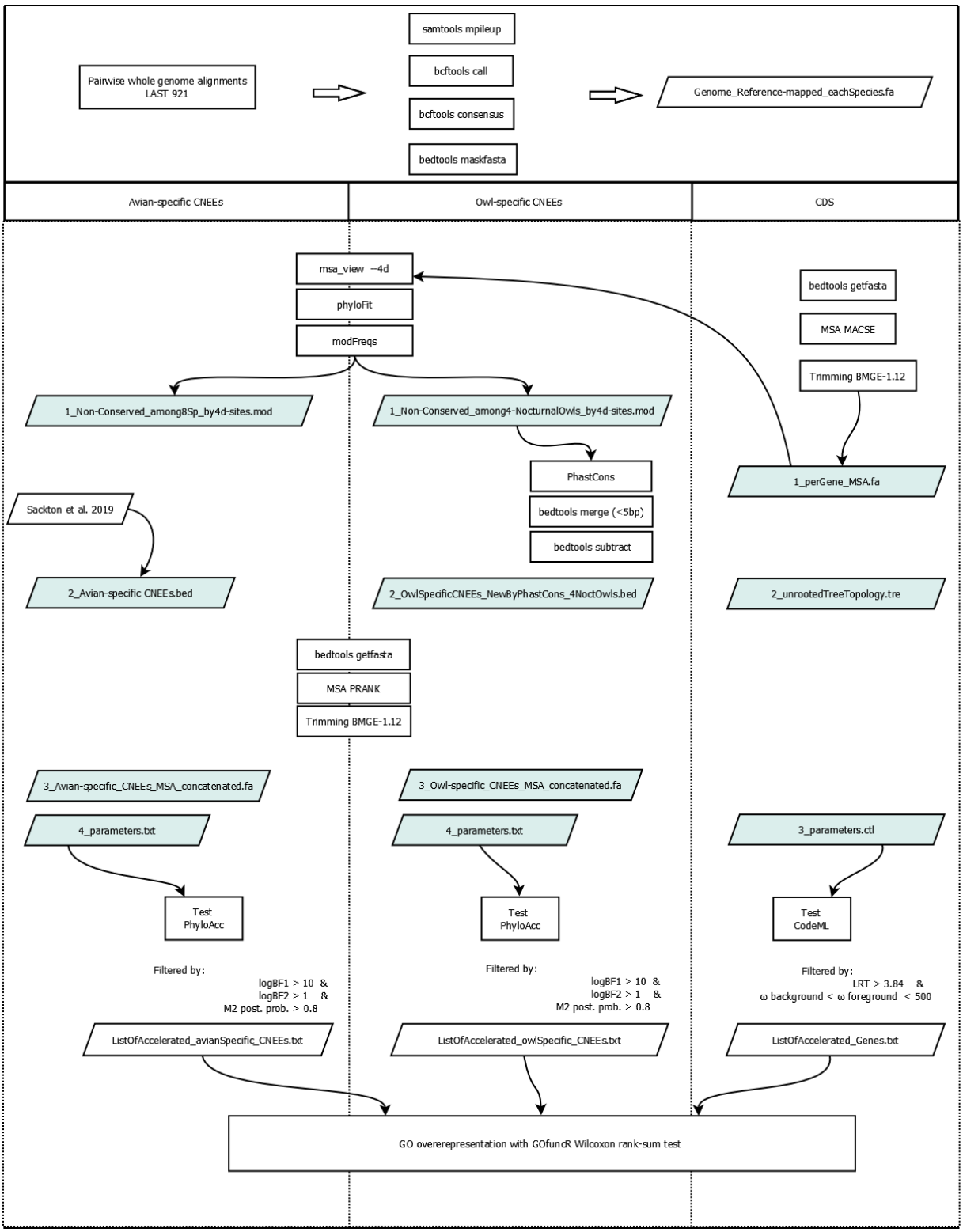


Figure S1. General workflow. The boxes in light-blue are inputs for the selection tests.

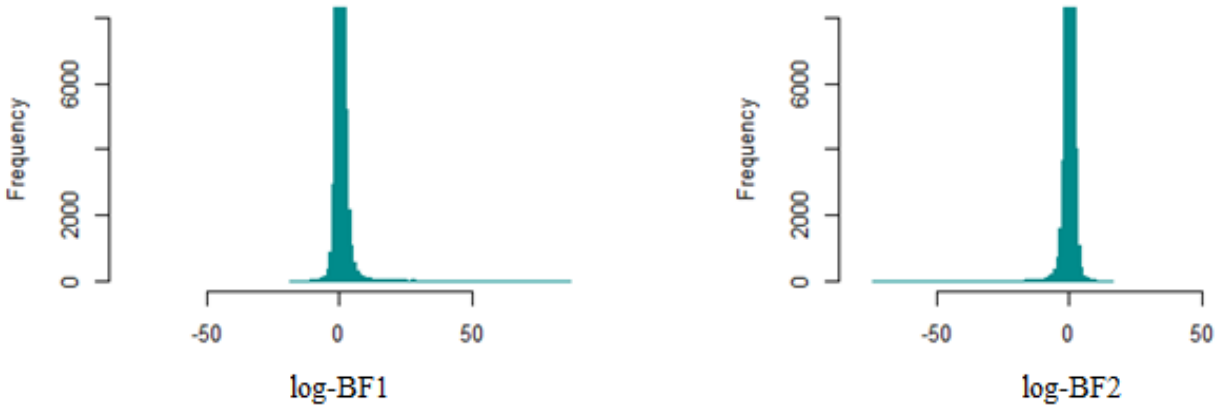


Figure S2. Distributions of Bayes factors

Extended Methods

All the steps, commands, and parameters used for the analyses, excluding the infile/outfile name are listed.

I. Reference-mapping of sequences of the whole genome for each species

This pipeline is a modification of the previous work detailed in Espíndola-Hernández *et al*. 2020.

i) Mapping to reference

Protocol a: Read mapping to reference
The reads were mapped against the reference genome using bwa (alignment via Burrows-Wheeler transformation), version: 0.7.17-r1188

bwa[1] mem -M -R

Protocol b: Genome-scale sequence mapping to reference
We aligned species genome assemblies to the reference using LAST v. 921:

lastdb[2] -uMAM8 -cR11
lastal[2] -E0.001 -i3G -m100
SingleCov2[3]
maf-convert[2] sam
samtools[4] view -bS

ii) Piling up the reads or genome sequences of the whole genomes.
Protocol a:
samtools[4, a] mpileup -u -I --output-tags AD,INFO/AD,DP,SP
Protocol b:
samtools [b] mpileup -u -I -A --output-tags AD,INFO/AD,DP,SP

iii) Variant calling,
Protocol a:
bcftools[5, a] call -m
Protocol b:
bcftools [b] call -m –A

iv) Producing reference-mapped sequences:
bcftools[5] consensus -f

v) Masking all the sites with zero coverage:
bedtools[6] genomecov bga
bedtools[6] maskfasta

vi) Extracting conserved coding and non-coding sequences, and concatenate them (CDS and CNEEs)
bedtools[6] getfasta

---

[1] bwa: http://bio-bwa.sourceforge.net/
[2] LAST: http://last.cbrc.jp/
[3] Multiz-tba.012109: https://www.bx.psu.edu/miller_lab/
[4] SAMtools: http://www.htslib.org/doc/samtools.html
[5] BCFtools: http://www.htslib.org/doc/bcftools.html
[6] BEDtools: https://bedtools.readthedocs.io/en/latest/

II. Multi-species aligning and trimming of conserved coding and non-coding sequences (CDS and CNEEs)

i) Multi-species aligning of CDS with MACSE[7]

macse -prog alignSequences
macse -prog refineAlignment

Trimming blocks of codons with high entropy using BMGE:
BMGE.jar -t CODON -m BLOSUM65 -g 1
BMGE.jar -t CODON -h 1 -w 1 -g 0.01
msa_view --in-format PHYLIP --out-format FASTA
fas2phy.R[8]

ii) Multi-species aligning of CNEEs with PRANK

prank -F -DNA -once -t

Rooted species tree topology (Newick format), used for PRANK:
((((((*Bubo bubo*, *Bubo scandiacus*), *Strix occidentalis*), *Asio otus*),(*Athene cunicularia*, *Surnia ulula*)), *Tyto alba*), *Leptosomus discolor*);

Trimming blocks of the alignment that contains gaps, ignoring entropy

BMGE.jar -t DNA -h 1 -w 1 -g 0.01
msa_view --in-format PHYLIP --out-format FASTA

seqkit concat -w 70 -j 10 --quiet
seqkit fx2tab -n -l -g

III. Producing the non-conserved (neutral) model

1) Extracting the 4d-sites from the MSA of genes

msa_view --4d --features
msa_view --in-format SS --out-format SS --tuple-size 1

2) Producing the non-conserved-4d model with PhyloFit

i) Determining GC content with AMAS to control for GC frequencies in order to maintain reversibility of the neutral model

AMAS.py summary -f fasta -d dna --cores 20

ii) phyloFit

-for PhastCons & PhyloP
4 Nocturnal owls
phyloFit --tree "(((*Bubo bubo*, *Strix occidentalis*), *Asio otus*),TytoA)" --msa-format SS

8 Species
phyloFit --tree "((((((*Bubo bubo*, *Bubo scandiacus*), *Strix occidentalis*), *Asio otus*),(*Athene cunicularia*, *Surnia ulula*)), *Tyto alba*), *Leptosomus discolor*)" --msa-format SS

---

[7]MACSE: https://bioweb.supagro.inra.fr/macse/
[8]fas2phy, Converts FASTA files into PHYLIP format: https://github.com/fmichonneau/chopper/tree/master/R

-for PhyloAcc use:
phyloFit --subst-mod SSREV --precision HIGH --init-random --sym-freqs --log

iii) modFreqs ---> Change background frequencies of reversible tree model in such a way that reversibility is maintained.

modFreqs <G+Cfreq>

iv) Naming ancestors with tree_doctor in phast (http://compgen.cshl.edu/phast/help-pages/tree_doctor.txt):

tree_doctor --name-ancestors

IV. Identifying owl-specific CNEEs with PhastCons (4-Nocturnal Owls)

i) Extracting sufficient statistics (SS) from a FASTA file for a complete chromosome.

msa_view --soft-masked --collapse-missing --in-format FASTA --out-format SS

ii) PhastCons

phastCons --target-coverage 0.4 --expected-length 45 --score --rho 0.2

iii) Post-processing of phastCons outputs for comparing

parallel -j20 'wig2bed`
parallel -j2 'gff2bed`

iv) sort
sort -k1,1 -k2,2n

v) bedtools merge (when closer than 5bp, and excluding overlap with exons and avian-specific CNEEs)

bedtools merge -i -d 5
bedtools subtract -nonamecheck -a -b

V. Test of CNEEs with PhyloAcc

PhyloAcc parameters:

BURNIN 400
MCMC 1000
CHAIN 1
TARGETSPECIES Snowy;AthCun;Surnia
OUTGROUP CuckooR
CONSERVE Bbubo;StrixOcc;Aotus;TytoA
NUM_THREAD 80

VI. Test on "CDSs" with CodeML

To prepare the sequence alignment inputs see sections I, II, and III. The unrooted tree (Newick format), used for the selection tests in CodeML was:

(((((*Bubo bubo*, *Bubo scandiacus* #1), *Strix occidentalis*), *Asio otus*),(*Athene cunicularia* #1, *Surnia ulula* #1)), *Tyto alba*, *Leptosomus discolor*);

We tested for accelerated ω on the diurnal owls (labeled as the foreground species in the tree above by "#1") using a maximum-likelihood method implemented in the CodeML program in PAML 4.9h [9] using the following settings in the control files:

Branch model

Null hypothesis (H_0)

```
model = 0        * models for codons:  0: one ω ratio for all branches, 1: one ω ratio for   each branch, 2: 2 or
                       more ω ratio for branches
NSsites = 0      * 0: one estimated ω; 1: Nearly neutral; 2: Positive selection
fix_kappa= 0     * 1: kappa fixed, 0: kappa to be estimated
kappa = 2        *  initial or fixed kappa value
fix_omega= 0     * 1: omega or omega_1 fixed, 0: estimate
omega = 1        *  initial or fixed omega value
cleandata = 1    *  remove sites with ambiguity data (1:yes, 0:no)
```

Alternative hypothesis (H_1)

```
model = 2
NSsites = 0
fix_kappa = 0
kappa = 2
fix_omega = 0
omega = 1
cleandata = 1    *  remove sites with ambiguity data (1:yes, 0:no)
```

Branch-Site

Null hypothesis (H_0)

```
model = 2        * models for codons:   0: one ω ratio for all branches, 1: one ω ratio for   each branch, 2: 2 or
                       more ω ratio for branches
NSsites = 2      * 0: one estimated ω; 1: NearlyNeutral; 2: Positive selection
fix_kappa = 0
kappa = 2
fix_omega = 1
omega = 1
cleandata = 1    * remove sites with ambiguity data (1:yes, 0:no)
```

Alternative hypothesis (H_1)

```
model = 2
NSsites = 2
fix_kappa = 0
kappa = 2
fix_omega = 0
omega = 1.3
cleandata = 1    * remove sites with ambiguity data (1:yes, 0:no)
```

[9] PAML: http://abacus.gene.ucl.ac.uk/software/paml.html

VII. Test on "CDSs" with aBSREL

All the genes with genome-wide significant results by CodeML were additionally tested with aBSREL:

parallel -j15 'hyphy absrel --alignment {} --tree ../Tree.tre ' ::: ./*.fa

Where the Tree.tre file contain the rooted topology of the included species:
((((((*Bubo bubo*, *Bubo scandiacus*), *Strix occidentalis*), *Asio otus*),(*Athene cunicularia*, *Surnia ulula*)), *Tyto alba*), *Leptosomus discolor*);

VIII. GO annotation of regulatory domains of genes using GREAT and GO overrepresentation analysis by GOfuncR

1) Getting the gene regulatory domain region with GREAT[10], using default parameters and providing two input files: one with the sizes of the chromosomes of the reference and other with the TSS of each gene.

   createRegulatoryDomains oneClosest

2) Intersect the CNEEs with the gen regulatory domains from GREAT

   sort -k1,1 -k2,2n

   bedtools intersect -a InFile_A.bed -b InFile_B.bed -wa -wb

3) GOfuncR[11] Wilcoxon rank-sum test

We used a custom-made gene ontology (GO) annotation database made for the reference (CustomAnnotation). The input of this test is a text file (Wilcoxon_input) with a list of elements ranked according to the "Ranking parameter". The genes were ranked by the LRT value. The CNEEs were ranked by a custom-made parameter calculated as:

pp: posterior probability
Anc: ancestral branch of respective species

```
P_Nocturnal     <-    (pp_Bbubo- pp_Anc_BuboSnowy)  +
                       (pp_StrixOcc- pp_Anc_BuboStrix)  +
                       (pp_Aotus- pp_Anc_BuboAsio)      +
                       (pp_Anc_BuboSnowy  - pp_Anc_BuboStrix )   +
                       (pp_Anc_BuboStrix  - pp_Anc_BuboAsio)      +
                       (pp_Anc_BuboAsio   - pp_Anc_BuboAthene)  +
                      (pp_Anc_AtheSurnia - pp_Anc_BuboAthene)  +
                      (pp_Anc_BuboAthene - pp_Anc_allOwls  )    +
                       (pp_Tyto - pp_Anc_allOwls)
```

[10] McLean, C. Y., D. Bristor, M. Hiller, S. L. Clarke, B. T. Schaar et al., 2010 GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. 28: 495–501.

[11] GOfuncR: https://bioconductor.org/packages/release/bioc/vignettes/GOfuncR/inst/doc/GOfuncR.html

Ranking parameter <- ((pp_Surnia- pp_Anc_AtheSurnia) +
(pp_Snowy- pp_Anc_AtheSurnia) +
(pp_AthCun -pp_Anc_AtheSurnia)) - P_Nocturnal

go_enrich(Wilcoxon_input, test = 'wilcoxon', annotations=(CustomAnnotation), n_randset=1000)

IX. Comparison CDS v/s CNEE by PhyloP

Estimate the acceleration/conservation/neutrality on each element (--features)

phyloP -m LRT -i FASTA --mode CONACC --branch Snowy,Surnia,AthCun --features

Chapter 3:

Comparing two methods for the identification of the W chromosomes of owls

Pamela Espíndola-Hernández[*], Jakob C. Mueller[*], Judith Mank[**], Bart Kempenaers[*]

[*] Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany
[**]Department of Zoology, University of British Columbia, V6T 1Z4 Vancouver, Canada

[*] Corresponding author: pamela.dola@gmail.com

Abstract

The understanding of the underlying causes and mechanisms of the vast majority of animals having separate sexes is a central aim of evolutionary biology. Having separate sexes is the starting point for the consequent divergent fitness interests, and then the *sexual conflict*. At the genetic level, genes might have different effects on each one of the sexes. These effects sometimes can be completely contrary, as in the s*ex antagonistic genes* that have alleles with advantageous effects in one sex but disadvantageous effects in the other. This *sexual conflict* at the genetic level can be solved by stopping recombination in those genetic regions. This solution causes the differentiation and degradation of the so-called *sex chromosomes*. The sex chromosomes were homologous autosomes, which evolved structurally and functionally after stopping the recombination between them. They have evolved independently many times, generating an opportunity to study convergent evolution. However, several challenges to their assembly have produced the exclusion of sex-restricted chromosomes from most genetic studies. Here, we have compared two methods for the identification of the W-chromosome in a non-model organism. Our results highlight the importance of a *de-novo* assembly for the sex chromosomes of each studied species.

Unpublished manuscript

Introduction

Sex chromosomes have evolved many times independently (Bachtrog *et al.* 2014), and in addition to carrying the sex-determining gene, they often also contain genes related to additional sex-specific characteristics (Mank 2009). In many cases, sex chromosomes originated from homologous autosomes, and have diverged from each other structurally and functionally once recombination between them was halted. The lack of recombination that led the evolution of sex chromosomes, has occurred in different taxa independently and could be an effect of selective pressure on keeping *sex antagonistic genes* separated from each other (Rice 1987; Mank and Ellegren 2009; Wright *et al.* 2016).

New bioinformatic approaches have made it possible to study many aspects of sex chromosomes that were not previously possible. In particular, the sex-limited Y and W chromosomes have been difficult to sequence, due to their highly repetitive and heterochromatic nature (Kapusta and Suh 2017). Although some methods of Y and W chromosome assembly exist (Tomaszkiewicz *et al.* 2016; Bellott *et al.* 2017), they all require some form of long-read sequence data to span the repetitive elements. Moreover, assemblies of the sex-limited chromosome always require heavy manual curation to improve the contiguity of the assemblies.

Birds share a female heterogametic sex chromosome system that arose from a pair of homologous autosomes in the recent common ancestor of all extant avian species (Zhou *et al.* 2014), where male birds have two Z chromosomes (homogametic sex), and female birds have one Z and one W chromosome (heterogametic sex). The avian W chromosome is characterized by a small effective population size (Ne) and the sensitivity to selection that follows from absence of recombination (outside the PAR region) and exposure of

recessive mutations as well as the accumulation of transposable elements (TEs) and tandem repeats (Zhou *et al.* 2014), as well as co-inheritance with the mitochondrial genome DNA (Berlin *et al.* 2007). In the avian linage, recombination suppression has spread independently, with new strata forming independently in different clades (Smeds *et al.* 2015). However, W-chromosomes are frequently excluded from genomic studies on birds, mainly because of the methodological difficulty of assembling DNA with a high content of repetitive sequences. In owls, there has been some attempts to assemble their sex chromosomes but mostly with a focus on the identification of the pseudo-autosomal region (PAR). These studies have reported a relative small PAR in the Barn owl (Zhou *et al.* 2014), the Burrowing owl (Mueller *et al.* 2018) and the Spotted owl (Fujito *et al.* 2021).

Here we identify the W-chromosome of two species of owls with two different methods. One method relies on mapping reads to a reference and the other method requires a genome-wide *de-novo* assembly. As the latter method is more complicated and time-consuming, our aim was to evaluate whether using a genome-wide *de-novo* assembly is worthy for future comparative studies of W-chromosomes among owl species.

Materials and Methods

Study species and reference genome

The genome assemblies of *Strix occidentalis* (Spotted owl) (Hanna *et al.* 2017; Fujito *et al.* 2021), *Tyto alba* (Barn owl) (Jarvis *et al.* 2014), and *Athene cunicularia* (Burrowing owl) (Mueller *et al.* 2018) were downloaded from NCBI. We extracted and sequenced the DNA from four blood samples from two male and two female Barn owls. The paired-end DNA sequencing was performed on an Illumina HiSeq4000 by the Sequencing Core Facility of the Max Planck Institute for Molecular Genetics (Berlin, Germany). Supplementary Tables S1 and S2 (Supplementary File 1) provide information of the sequenced samples and the accession numbers of the downloaded genome samples. A more detailed description of the pipeline, commands, and parameters can also be found in the Supplementary File 1. After sequencing, we followed four methodological steps described below.

i) De-novo assembly

We started by performing a *de-novo* assembly of one female (ZW) per species. First, we assessed the read quality of the sequencing using FastQC v0.11.9 (Andrews *et al.* 2010). Second, we trimmed adaptors with Trimmomatic v0.36 (Bolger *et al.* 2014). Third, we performed a *de-novo* assembly with ABySS (Simpson *et al.* 2009).

ii) Identification method based on de-novo assembly and fold differences of mapped reads

We used Bowtie2 (Langmead and Salzberg 2012) with default parameters, to map reads from one male and one female for each owl species, to the *de-novo* genome reference of their respective species. Finally, we estimated the depth in every contig or scaffold with

samtools (Li *et al.* 2009), considering only their concordant and properly paired reads. Finally, we corrected by length of the contig and by the total of mapped reads on each sample.

As the female owls have one W chromosome and the male owls have none, it is expected that the samples from males should have no reads in W-specific regions of the reference. We estimated the fold difference between sexes by the log2 of the ratio between female depth to male depth at each contig or scaffold. We summed 0.0001 to all depth values to shift all the values from zero to avoid the issues about dividing by zero. Then, we used ggplot2 R-package to visualize it (Wickham 2016). We consider a contig or scaffold as a "W-candidate" when the female-biased fold difference in depth is higher than twice the standard deviation from the mean of all contigs, while excluding contigs or scaffolds shorter than 500 nucleotides length.

iii) Identification method based on mapping reads to a closely related reference

We mapped reads from the whole genome of the two female owls that we *de-novo* assembled, starting from the same raw files (*fastq*) that we used to construct the genome references described in the previous step (ii). In order to construct a reference-mapped sequence, we used Bowtie2 (Langmead and Salzberg 2012) with default parameters to map the reads against the Spotted owl genome as a reference (Fujito *et al.* 2021). Then we compiled ("piling up") reads of the whole genome using samtools (Li et al. 2009), performed variant calling and produced a reference-mapped sequence for both species with bcftools (Danecek *et al.* 2016; Danecek and McCarthy 2017).

iv) Estimating the overlap between both methods

We estimated the overlap between both methods by mapping the regions (W candidate contigs) identified by the first method to the one identified by the second method using dc-megablast (discontiguous megablast used to find more distant (e.g., interspecies) sequences) in galaxy.eu platform with default parameters (Camacho *et al.* 2009; Cock *et al.* 2015). We consider BLAST matches with a minimum percentage identity of 70% and a minimum percentage query coverage of 50%.

Results

We produced a whole-genome *de-novo* assembly of two female owls and then used these to identify putative W-linked scaffolds by comparing read depth between females and males. The *de-novo* draft assembly of the female Burrowing owl contained 3,460,473 scaffolds, with a total length of 1.2 Gb. The scaffold N50 was 586 bp and the average scaffold length was 357.75 bp, where the largest scaffold reached 31.3 Kb (Table S3, Supplementary file 1). The *de-novo* draft assembly of the female Barn owl contained 2,311,401 scaffolds and a total length of 1.2 Gb. The scaffold N50 was 1.1 Kb and the average scaffold length was of 536.83 bp, where the largest scaffold reached 39.3 Kb (Table S3, Supplementary file 1).

Identification method based on de-novo assembly and fold differences of mapped reads

From the Burrowing owl genome, we identified 6,417 W-candidate contigs or scaffolds with a threshold of 0.78 fold difference (twice the standard deviation from the mean of log2 f:m read depth across all contigs), which summed up to a total length of 4.9Mb. From the Barn owl genome, we identified 23,517 W-candidate contigs or scaffolds with

a threshold of 1.62 fold difference (twice the standard deviation from the mean of log2 f:m read depth across all contigs), summing up to a total length of 19.1Mb.

Identification method based on mapping reads to a closely related reference

The publicly available Spotted owl W chromosome that we used as reference has 44 contigs (some are also small scaffolds) and a total length of 8.6Mb (Fujito *et al.* 2021). By mapping the Burrowing owl reads to the Spotted owl genome, we could identify and reconstruct fragments from 42 of the 44 contigs in the reference. In the Barn owl case, we could identify and reconstruct fragments from 32 of the 44 contigs in the reference.

The overlap between both methods

From the *de-novo* assembled Burrowing owl genome, 680 contigs can be found in the W-Chromosome of the Spotted owl (Fujito *et al.* 2021), producing a 10.60% overlap of contigs above the 0.78 threshold (Figure 1, Figure S1.a in Supplementary file 1). From the *de-novo* assembled Barn owl genome, 125 W-candidate contigs or scaffolds are also identifiable by mapping to the reference W-chromosome of the Spotted owl (Fujito *et al.* 2021). Producing a 0.53% overlap of contigs above the 1.62 threshold (Figure 2, Figure S1.b in Supplementary file 1).

The inset plots of Figures 1 and 2 show all contigs of the owls' genomes sorted by the log2 f:m read depth. The region highlighted in the inset plot correspond to the W-candidate contigs and is shown in more detail in the larger plot (Figures 1 and 2). In Figure 1 and 2, we highlighted (red dots) the W-candidate contigs that can be found by both identification methods, the one based on the *de-novo* assembly of an individual from the same species and the one based on mapping to the reference W-chromosome of the Spotted owl (Fujito *et al.* 2021).

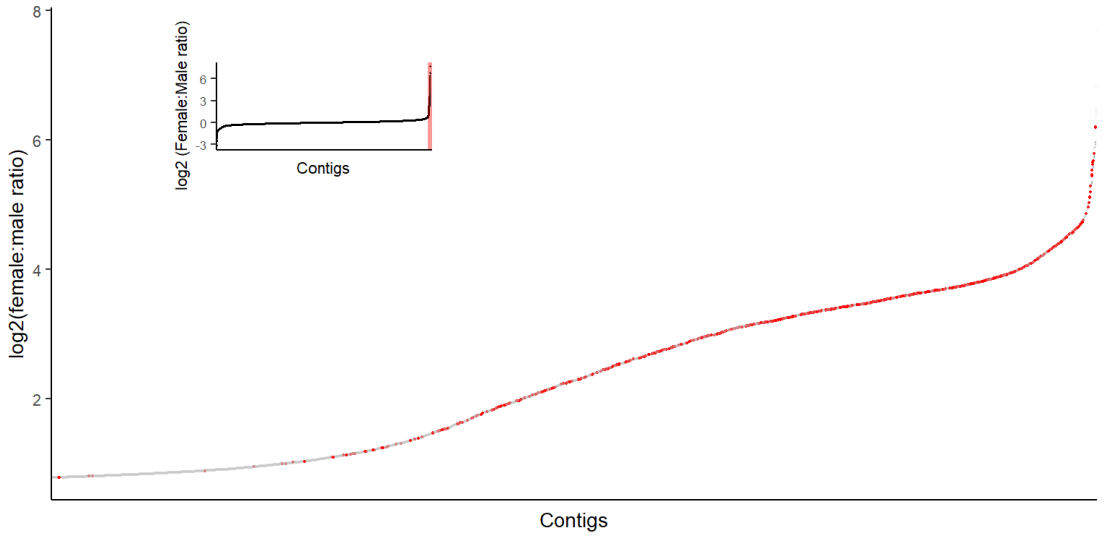


Figure 1: Burrowing owl (*Athene Cunicularia*) W-candidates. Contigs or scaffolds with more than 500 nucleotides length and log2 of the ratio between female depth to male depth >0.78 (twice the standard deviation from the mean of all contigs). The red dots represent the overlap between both identification methods, that is contigs or scaffolds that can be considered W-candidates based on the de-novo assembly and on the mapping to the (Spotted owl) reference method.

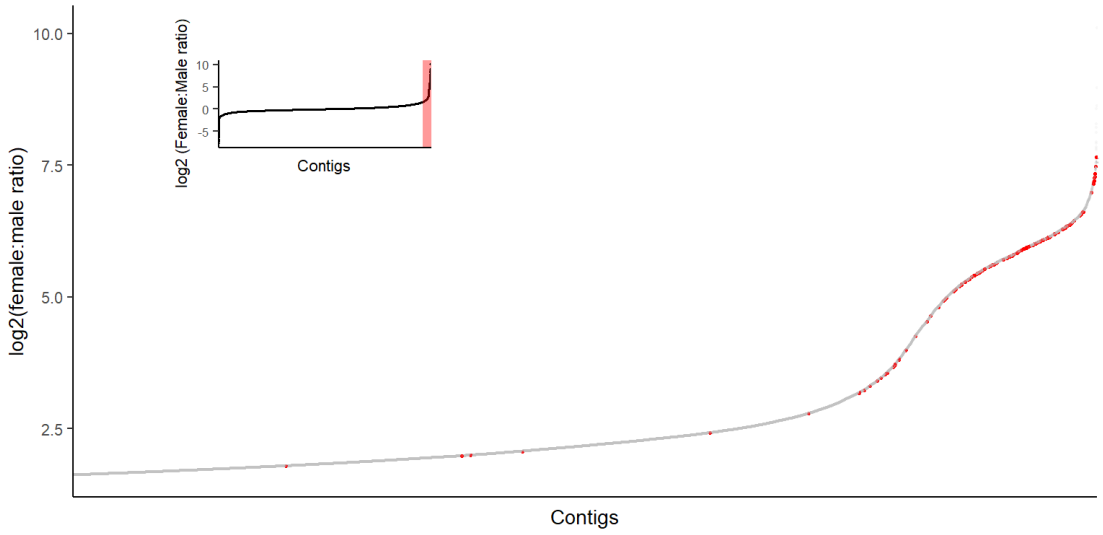


Figure 2: Barn owl (*Tyto alba*) W-candidates. Contigs or scaffolds with more than 500 nucleotides length and log2 of the ratio between female depth to male depth >1.62 (twice the standard deviation from the mean of all contigs). The red dots represent the overlap between both identification methods, that is contigs or scaffolds that can be considered W-candidates based on the de-novo assembly and on the mapping to the (Spotted owl) reference method.

Discussion

Here we compared the power of two methods for the identification of W chromosomes in two owl species. In both species, the number of W-candidates were higher when we used the identification method based on the *de-novo* assembly of short-reads of a female sample with remapping from same species male and female reads, than when we used the one based on mapping reads to a closely related reference. However, we observed large differences in the overlap, with 10.60% in Burrowing owls and only 0.53% in Barn owls. The Spotted owl is phylogenetically closer to the Burrowing owl than to the Barn owl (Wink *et al.* 2009; Salter *et al.* 2020), and this relatedness difference might explain the difference we observe. However, it is also possible that the major differences in N50 scores in our assemblies also influence our ability to detect putative W contigs and scaffolds in the Spotted owl reference W chromosome.

In general, methods based on *de-novo* genome assembly require more time and computational resources than methods relying on mapping reads to a genome reference, in this case a pre-existing W chromosome reference. The downside of the latter is having less power in the identification of unknown parts of the genome, because these methods require sequence homology in the genome reference (Narzisi and Mishra 2011; Henson *et al.* 2012; Khan *et al.* 2018). Then, the *de-novo* assembly is the best approach for characterizing structural variation among genomes and for the study of non-model organisms which might not have a genome reference, among other applications. The assembly of the sex-restricted chromosomes present special challenges, mainly due to their characteristic highly repetitive nature (Kapusta and Suh 2017), and for that reason, they are frequently excluded from genomic studies.

The degree of sex chromosomes differentiation is very diverse across the tree of life. In birds, it is very diverse as well due to the lineage-specific recombination suppressions and different tempo of W chromosome degeneration (Zhou *et al.* 2014), and in owls this is totally unknown.

Previous studies have reported a small pseudo-autosomal region (PAR) in three owl species, *i.e.*, the Barn owl (Zhou *et al.* 2014), the Burrowing owl (Mueller *et al.* 2018) and the Spotted owl (Fujito *et al.* 2021), while the non-recombining region in owls' genomes still being a mystery. Here, we have compared two methods for the identification of the W-chromosome and our results suggest the need of a *de-novo* assembly for the sex chromosomes of each owl species. The best characterized owl W-chromosome is from the Spotted owl (Fujito *et al.* 2021), which has been an important contribution for population studies of the same species. However, here we have shown that future comparative studies cannot rely on mapping reads from owls across different clades. Considering also the most recently published genome assemblies that includes the sex chromosomes, a combination of *de-novo* assembly based on short and long reads, plus Hi-C sequencing and manual curation is recommended (Darolti *et al.* 2022; Kirkpatrick *et al.* 2022; Jasonowicz *et al.* 2022).

Data availability

Sequence data are publicly available, and their references and accession numbers are listed in Supplementary Table 1 and 2 in Supplementary File 1. The assemblies and annotation files are available in the repository <available by publication date>.

Acknowledgments

We would like to thank Iulia Darolti and Francisco Salinas for their helpful suggestions at different stages of the bioinformatic work. The biological samples of the owls were provided by Martina Carrete (Estación Biológica de Doñana, Spain), and Alexandre Roulin and Celine Simon (University of Lausanne, Switzerland). The sequencing process was carried out by Dr. Stefan Börno at the Sequencing Core Facility of the Max Planck Institute for Molecular Genetics (Berlin, Germany).

Funding

This work was funded by the Max Planck Society (to B.K.), and by a scholarship for doctoral studies in Germany co-funded by the DAAD and BecasChile (to P.E.-H.).

ORCID IDs:

Pamela Espíndola-Hernández https://orcid.org/0000-0002-9771-6907
Jakob C. Mueller https://orcid.org/0000-0001-6676-7595
Judith E. Mank https://orcid.org/0000-0002-2450-513X
Bart Kempenaers https://orcid.org/0000-0002-7505-5458

Conflicts of Interest

The authors declare to have no competing interests.

References

Andrews, S., F. Krueger, A. Segonds-Pichon, L. Biggins, C. Krueger *et al.*, 2010 FASTQC. A quality control tool for high throughput sequence data.

Bachtrog, D., J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto *et al.*, 2014 Sex Determination: Why So Many Ways of Doing It? PLoS Biol. 12: 1–13.

Bellott, D. W., H. Skaletsky, T. J. Cho, L. Brown, D. Locke *et al.*, 2017 Avian W and mammalian y chromosomes convergently retained dosage-sensitive regulators. Nat. Genet. 49: 387–394.

Berlin, S., D. Tomaras, and B. Charlesworth, 2007 Low mitochondrial variability in birds may indicate Hill-Robertson effects on the W chromosome. Heredity (Edinb). 99: 389–396.

Bolger, A. M., M. Lohse, and B. Usadel, 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114–2120.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos *et al.*, 2009 BLAST+: Architecture and applications. BMC Bioinformatics 10:.

Cock, P. J. A., J. M. Chilton, B. Grüning, J. E. Johnson, and N. Soranzo, 2015 NCBI BLAST+ integrated into Galaxy. Gigascience 4:.

Danecek, P., and S. A. McCarthy, 2017 BCFtools/csq: haplotype-aware variant consequences (I. Birol, Ed.). Bioinformatics 33: 2037–2039.

Danecek, P., S. Schiffels, and R. Durbin, 2016 Multiallelic calling model in bcftools (-m). 1–2.

Darolti, I., P. Almeida, A. E. Wright, and J. E. Mank, 2022 Comparison of methodological approaches to the study of young sex chromosomes: A case study in Poecilia. J. Evol. Biol. 00: 1–13.

Fujito, N. T., Z. R. Hanna, M. Levy-Sakin, R. C. K. Bowie, P. Y. Kwok *et al.*, 2021

Genomic Variation and Recent Population Histories of Spotted (Strix occidentalis) and Barred (Strix varia) Owls. Genome Biol. Evol. 13: 1–12.

Hanna, Z. R., J. B. Henderson, J. D. Wall, C. A. Emerling, J. Fuchs *et al.*, 2017 Northern spotted owl (strix occidentalis caurina) genome: Divergence with the barred owl (strix varia) and characterization of light-associated genes. Genome Biol. Evol. 9: 2522–2545.

Henson, J., G. Tischler, and Z. Ning, 2012 Next-generation sequencing and large genome assemblies. Pharmacogenomics 13: 901–915.

Jarvis, E. D., S. Mirarab, A. J. Aberer, B. Li, P. Houde *et al.*, 2014 Whole-genome analyses resolve early branches in the tree of life ofmodern birds. Science (80-.). 346: 1320–1331.

Jasonowicz, A. J., A. Simeon, M. Zahm, C. Cabau, C. Klopp *et al.*, 2022 Generation of a chromosome-level genome assembly for Pacific halibut (Hippoglossus stenolepis) and characterization of its sex-determining genomic region. Mol. Ecol. Resour. 00: 1–16.

Kapusta, A., and A. Suh, 2017 Evolution of bird genomes—a transposon's-eye view. Ann. N. Y. Acad. Sci. 1389: 164–185.

Khan, A. R., M. T. Pervez, M. E. Babar, N. Naveed, and M. Shoaib, 2018 A Comprehensive Study of De Novo Genome Assemblers: Current Challenges and Future Prospective. Evol. Bioinform. Online 14: 1176934318758650.

Kirkpatrick, M., J. M. Sardell, B. J. Pinto, G. Dixon, C. L. Peichel *et al.*, 2022 Evolution of the canonical sex chromosomes of the guppy and its relatives. G3 Genes|Genomes|Genetics 12:.

Langmead, B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. Nat. Methods 9: 357–359.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence

Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Mank, J. E., 2009 Sex chromosomes and the evolution of sexual dimorphism: Lessons from the genome. Am. Nat. 173: 141–150.

Mank, J. E., and H. Ellegren, 2009 Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. Evolution (N. Y). 63: 1464–1472.

Mueller, J. C., B. Kempenaers, H. Kuhl, S. Boerno, J. L. Tella *et al.*, 2018 Evolution of genomic variation in the burrowing owl in response to recent colonization of urban areas. Proc. R. Soc. B Biol. Sci. 285: 1–9.

Narzisi, G., and B. Mishra, 2011 Comparing De Novo genome assembly: The long and short of it. PLoS One 6:.

Rice, W. R., 1987 The Accumulation of Sexually Antagonistic Genes as a Selective Agent Promoting the Evolution of Reduced Recombination between Primitive Sex Chromosomes. Evolution (N. Y). 41: 911.

Salter, J. F., C. H. Oliveros, P. A. Hosner, J. D. Manthey, M. B. Robbins *et al.*, 2020 Extensive paraphyly in the typical owl family (Strigidae). Auk 137: 1–15.

Simpson, J. T., K. Wong, S. D. Jackman, J. E. Schein, S. J. M. Jones *et al.*, 2009 ABySS: A parallel assembler for short read sequence data. Genome Res. 19: 1117–1123.

Smeds, L., V. Warmuth, P. Bolivar, S. Uebbing, R. Burri *et al.*, 2015 Evolutionary analysis of the female-specific avian W chromosome. Nat. Commun. 6: 1–10.

Tomaszkiewicz, M., S. Rangavittal, M. Cechova, R. C. Sanchez, H. W. Fescemyer *et al.*, 2016 A time- and cost-effective strategy to sequence mammalian Y chromosomes: An application to the de novo assembly of gorilla Y. Genome Res. 26: 530–540.

Wickham, H., 2016 ggplot2: Elegant Graphics for Data Analysis.

Wink, M., A.-A. El-Sayed, H. Sauer-Gürth, and J. Gonzalez, 2009 Molecular Phylogeny of Owls (Strigiformes) Inferred from DNA Sequences of the Mitochondrial Cytochrome *b* and the Nuclear *RAG-1* gene. Ardea 97: 581–591.

Wright, A. E., R. Dean, F. Zimmer, and J. E. Mank, 2016 How to make a sex chromosome. Nat. Commun. 7: 1–8.

Zhou, Q., J. Zhang, D. Bachtrog, N. An, Q. Huang *et al.*, 2014 Complex evolutionary trajectories of sex chromosomes across bird taxa. Science (80-.). 346: 1246338.

Supplemental File 1

Comparing two methods for the identification of the W-chromosomes of owls

Pamela Espíndola-Hernández[*], Jakob C. Mueller[*], Judith Mank[**], Bart Kempenaers[*]

[*] Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany
[**]Department of Zoology, University of British Columbia, V6T 1Z4 Vancouver, Canada

[*] Corresponding author: pamela.dola@gmail.com

Tables

Information about the included samples and mapping statistics.

Table S1. Samples that were used as a reference.

Scientific name	Common name	Sex	GenBank assembly /accession	Associated publication
Athene cunicularia	Burrowing owl	female	SRR6670247	Mueller *et al.* 2018*
Gallus gallus	Red junglefowl	female	GCA_000002315.3	International Chicken Genome Consortium
Strix occidentalis	Spotted owl	female	GCA_002372975.1	Fujito *et al.* 2021
Tyto alba	Barn owl	female	-	This study**

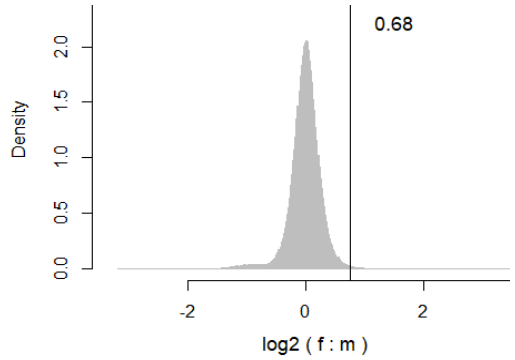
Table S2. Samples that were mapped to the reference.

Scientific name	Common name	Sex	GenBank assembly /accession	Associated publication
Athene Cunicularia	Burrowing owl	female	SRR6670165	Mueller *et al.* 2018*
Athene Cunicularia	Burrowing owl	male	SRR6670159	Mueller *et al.* 2018*
Tyto alba	Barn owl	female	GCA_000687205.1	Jarvis *et al.* 2014
Tyto alba	Barn owl	male	-	This study**

Table S3. Assembly stats

Information	Burrowing owl	Burrowing owl
scaffolds	3460473	2311401
Total scaffold length	1238000441	1240839481
Average scaffold length	357.75	536.83
Scaffold N50	586	1161
Scaffold L50	454631	281353
Largest scaffold	31314	39321
Total gap length in scaffolds	70892	7022
Average gap length in scaffolds	2.21	2.18
Gap N50 in scaffolds	12	12
Gap L50 in scaffolds	1112	111
Largest gap in scaffolds	104	89
GC content %	41.95	41.94

a)



b)

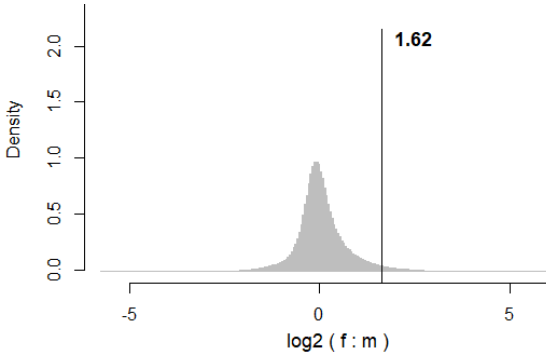


Figure S1. Density plots of the log2 f:m distribution. a. Density plots of Burrowing owl genome, where the black line show the threshold in 0.78 log2 f:m value, calculated as the mean plus two times the standard deviation. b. Density plots of Barn owl genome, where the black line show the threshold in 1.62 log2 f:m value, calculated as the mean plus two times the standard deviation.

Extended Methods

All the commands and parameters used in the analyses, excluding the infile/outfile name are listed.

I. Assembly de novo of "the best female of each species"

i. Abyss assembly (Simpson *et al.* 2009)

```
export TMPDIR=/var/tmp
abyss-pe name= <species name>  k=96 B=25G in='reads1.fa reads2.fa'
```

ii. Calculate assembly contiguity statistics:

```
abyss-fac input.fa
```

II. Genome-scale sequence mapping to reference of one male and one female to each de-novo Assembly to identify w-chromosome regions.

Using: bowtie2-2.4.5-linux-x86_64 (Langmead and Salzberg 2012)

i. bowtie2 index

```
bowtie2-build -f <reference_Input.fasta> <Output_Index>
```

ii. Check the Index

```
bowtie2-inspect [options] <bt2_base> -o <report_Output.txt>

bowtie2-inspect <INDEX_Bowtie2> -o Report.txt
```

iii. Mapping reads to assembly using bowtie2 and process with samtools (Li *et al.* 2009) version 1.14: http://www.htslib.org/doc/samtools.html

```
bowtie2 -p 10  --align-paired-reads  -x <INDEX_Bowtie2> -b
reads_for_mapping.bam | samtools view -@ 10 -bS -o output_mapped.bam


nohup parallel -j10 'samtools sort -@ 10 {} -o {.}_sorted.bam'  :::
*_mapped.bam &

nohup parallel -j2 'samtools index -@ 2 {} '  ::: *_mapped_sorted.bam
```

iii) Check the mapping quality

```
parallel -j10 'samtools stats -@ 5 {} -o {.}_stats.txt'  ::: *_mapped_sorted.bam
```

III. Coverage

```
samtools coverage input.bam  >  output_coverage.txt
```

REFERENCES

* Samples provided by Martina Carrete (Estación Biológica de Doñana, Spain).
** Samples provided by Alexandre Roulin and Celine Simon (University of Lausanne, Switzerland).

Fujito, N. T., Z. R. Hanna, M. Levy-Sakin, R. C. K. Bowie, P. Y. Kwok *et al.*, 2021 Genomic Variation and Recent Population Histories of Spotted (Strix occidentalis) and Barred (Strix varia) Owls. Genome Biol. Evol. 13: 1–12.

Jarvis, E. D., S. Mirarab, A. J. Aberer, B. Li, P. Houde *et al.*, 2014 Whole-genome analyses resolve early branches in the tree of life ofmodern birds. Science (80-.). 346: 1320–1331.

Langmead, B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. Nat. Methods 9: 357–359.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Mueller, J. C., B. Kempenaers, H. Kuhl, S. Boerno, J. L. Tella *et al.*, 2018 Evolution of genomic variation in the burrowing owl in response to recent colonization of urban areas. Proc. R. Soc. B Biol. Sci. 285: 1–9.

Simpson, J. T., K. Wong, S. D. Jackman, J. E. Schein, S. J. M. Jones *et al.*, 2009 ABySS: A parallel assembler for short read sequence data. Genome Res. 19: 1117–1123.

General Discussion

In the three chapters of this dissertation, I used a comparative genomics approach to investigate the genomic fingerprints of the evolution of owls (Strigiformes). First, I assessed the accelerated substitution rates that can be inferred for the ancestral branch of the owls to identify the selection signals of their adaptations as nocturnal raptors (Chapter 1). Second, I searched for signals of accelerated substitution rates in the genome of three diurnal owl species, with the aim of understanding the respective role of coding and non-coding genomic elements in the emergence of a diurnal lifestyle in the owls' clade (Chapter 2). Third, I contrasted the contribution of *de-novo* and reference-guided assembly strategies, in particular to the study of highly repetitive sex-restricted genomic regions, such as the W chromosomes of owls (Chapter 3).

The owls' tale

The history of owls dates back to the Paleocene (60 Ma) when they diversified from the rest of the Afroaves (Feduccia 1995, 1999, 2003; del Hoyo *et al.* 1999; König and Weick 2008), and separated from their sister group, the diurnal Coraciimorph clade (Prum *et al.* 2015). At that time, small mammals experienced rapid radiation which was probably a key hunting opportunity for owls. By colonizing the nocturnal niche, those small mammals were successful in avoiding predation from diurnal hunters and increasing new opportunities for the nocturnal ones. The ancestral of all owls was probably a diurnal landbird with raptorial features (Ericson *et al.* 2006; Hackett *et al.* 2008; Jarvis *et al.* 2014; Prum *et al.* 2015; McClure *et al.* 2019), and the results that I presented in Chapter 1 show that their early adaptations to a nocturnal lifestyle have

been shaped by positive selection on genes functionally associated with visual perception, including phototransduction and chromatin packaging (Espíndola-Hernández *et al.* 2020). The results from Chapter 1 show an accumulation of genetic changes at the ancestral branch of the owls in genes that are functionally associated with nocturnal hunting. Altogether, these results support the hypothesis of the diurnal ancestor of all living birds and an independent adaptive history of owls as nocturnal birds of prey (Espíndola-Hernández *et al.* 2020).

Currently, owls have occupied almost all continents (except Antarctica) with around 250 owl species currently alive (del Hoyo *et al.* 1999; König and Weick 2008; Duncan 2018). They have also reached urban regions. For example, in Bahia Blanca, Argentina, they are widely established (Rebolo-Ifrán *et al.* 2017; Mueller *et al.* 2018; Luna *et al.* 2019). The Burrowing owl (*Athene cunicularia*) is gregarious, active during the day, and has colonized urban areas, making them a very particular species among the rest of the owls. The Burrowing owl is one of the three diurnal owls that I studied in Chapter 2, where I consider the phenotypic variation among owl species in terms of their diel activity patterns, contrasting the three diurnal owls against a background of nocturnal owl genomes. The results of Chapter 2 suggest that regulatory evolution might have played a predominant role in the shift to a diurnal lifestyle in owls. This conclusion is based on the higher acceleration of substitution rates at noncoding elements, and on the functional association of the accelerated protein-coding genes in diurnal owls with the regulation of gene expression (Espindola-Hernandez *et al.* 2022).

The sexual issues among owls

When organisms of the same species have different sexes (separate sexes), the fitness optimization strategy of each sex might diverge. If the fitness interests diverge the so-called *sexual conflict* emerges. The definition of *sexual conflict* is not straightforward. Parker defined sexual conflict as "a conflict between the evolutionary interests of individuals of the two sexes" (Parker 1979). The *sexual conflict* at the genetic level can be solved by stopping recombination in those genetic regions. This solution might cause the differentiation and degradation of the so-called sex chromosomes. The sex chromosomes were homologous autosomes, which evolved structurally and functionally after stopping the recombination between them (Bachtrog *et al.* 2011; Abbott *et al.* 2017). However, several challenges to their assembly have produced the exclusion of sex-restricted chromosomes from most genetic studies, mainly due to their characteristic highly repetitive nature (Kapusta and Suh 2017). The recently published analyses by Darolti et al. (2022), shown that in systems where the recombination between the sex chromosomes has stopped recently, the sex chromosomes share sequences with high similarity and, as consequence, many reads can be mapped to the wrong sex chromosome, which obscure the delimitation of the sex-restricted elements, especially in young sex chromosomes (Darolti *et al.* 2022).

In birds, males have two Z chromosomes (homogametic sex), and females have one Z and one W chromosome (heterogametic sex). Their sex chromosomes evolved from a pair of homologous autosomes in a recent common ancestor of all living avian species (Zhou *et al.* 2014). The recombination suppression has spread independently across the avian phylogeny, forming new strata independently (Smeds *et al.* 2015), representing an opportunity to study convergent evolution.

The W chromosome of owls have been studied mostly for population studies, as a marker of the matrilineal genetic inheritance (Nelson *et al.* 2000; Keis *et al.* 2013; Smeds *et al.* 2015). The avian Z chromosome has shown "Fast-Z effect" (Dean *et al.* 2015). In other words, there is an increased rate of functional change in the avian Z chromosome compared to autosomes (Charlesworth *et al.* 1987). The part of the sex chromosomes that recombine during meiosis is the Pseudo-Autosomal Region (PAR), which represents a relatively small portion of the sex chromosomes of the Barn owl (Zhou *et al.* 2014), the Burrowing owl (Mueller *et al.* 2018) and the Spotted owl (Fujito *et al.* 2021). In chapter 3, I compared two methods for identifying W chromosome in owls. The results highlight the higher power of a *de-novo* assembly approach, compared to the strategies based on reference, for the identification of unknown parts of the genome, such the W chromosomes of the owls.

In birds, social monogamy is very common (Birkhead and Møller 1995; Neodorf 2004). Social monogamy implies keeping a social unique partner, at least, during the breeding season but engage in extra-pair copulation. This results in extra-pair paternity. Just as other birds of prey, owls are solitary during most of the year and form monogamous pairs to breed. Owls are monogamous, present very low sexual dimorphism in terms of plumage coloration, but they show reversed sexual size dimorphism (del Hoyo *et al.* 1999; König and Weick 2008; Duncan 2018). Furthermore, the owls' monogamy is very strict. It has been seen that the owls are genetically monogamous even in new environment with high conspecific density, as it has happened in some cases of urban colonization (Rodriguez-Martínez *et al.* 2014)

The substitution rates as genomic signals of selection

In general, changes in a genetic sequence can be neutral (no functional effect on the organism), deleterious, or advantageous. All these mutations can accumulate in populations, but the accumulation of deleterious mutations is counteracted by negative selection, whereas the accumulation of advantageous mutations is promoted by positive selection. Understanding the targets of selection associated with changes in behavioral traits represents an important challenge of current evolutionary research. I used substitution rates (i) to test for selection in the early history of Strigiformes in a genome-wide comparative analysis, (ii) to compare rates of evolution between coding and non-coding regions, and (iii) to test for selection in the recent shift to a diurnal lifestyle in three owl species.

The estimation of substitutions rate from genome assemblies is a powerful and, nowadays, accessible approach for studying genomic signatures of selection. However, it is not exempt of limitations. There are some advisable considerations for applying these methods based on the estimation of substitution rates. First, it is crucial to invest especial effort in avoiding systematic errors in the sequencing and post-processing of the genome sequencing data because the main point is the identification of variable sites across genomes, which must be error-free to be meaningful. Second, the selection of the genome aligner software is very relevant, being phylogeny-aware aligner the best option. Third, the estimated substitution rates are specific for the pool of species included in the respective phylogeny, and methods that average the evolutionary rate of the target species branches might overlook when a single species is responsible for the majority of the variation.

In summary

In accordance with the review by Le Duc and Schöneberg (2016), the diel activity patterns in birds are a suitable study model that can bring insight about the underlying genetic mechanisms of adaptation. Comparative approaches and the use of *omics* sciences contribute to the understanding of classic questions in evolutionary biology. Nowadays, evolutionary biologist can enjoy the possibility of testing evolutionary models, that emerged decades ago from classic population genetics and statistics, with real data provided by the sequencing technologies. There was a rapid and irrevocable change in the scientific approach to answer evolutionary biology questions (revised by Schuster in 2007) with the technological transition from Sanger sequencing (Sanger et al. 1977) to "next-generation sequencing (NGS)", and the current “high-throughput sequencing”. The advancement in the accessibility of bioinformatic tools is opening opportunities for the study of the evolution in non-model organisms.

The acceleration in the substitution rates at the ancestral branch of the owls is a fingerprint of the shift to a nocturnal lifestyle in the ancestral owls that lived more than 50 million years ago. At the branches leading to the three diurnal owl species currently alive, the acceleration found in the noncoding elements and the functional association of the accelerated protein-coding genes with the regulation of gene expression suggest that regulatory evolution might have played a predominant role in the emergence of a diurnal lifestyle in the owls’ clade.

References

Abbott, J. K., A. K. Nordén, and B. Hansson, 2017 Sex chromosome evolution: Historical insights and future perspectives. Proc. R. Soc. B Biol. Sci. 284.

Bachtrog, D., M. Kirkpatrick, J. E. Mank, S. F. McDaniel, J. C. Pires *et al.*, 2011 Are all sex chromosomes created equal? Trends Genet. 27: 350–357.

Birkhead, T. R., and A. P. Møller, 1995 Extra-pair copulation and extra-pair paternity in birds. Anim. Behav. 49: 843–848.

Charlesworth, B., J. A. Coyne, and N. H. Barton, 1987 The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. 130: 113–146.

Darolti, I., P. Almeida, A. E. Wright, and J. E. Mank, 2022 Comparison of methodological approaches to the study of young sex chromosomes: A case study in Poecilia. J. Evol. Biol. 00: 1–13.

Dean, R., P. W. Harrison, A. E. Wright, F. Zimmer, and J. E. Mank, 2015 Positive selection underlies Faster-Z evolution of gene expression in birds. Mol. Biol. Evol. 32: 2646–2656.

Le Duc, D., and T. Schöneberg, 2016 Adaptation to nocturnality – learning from avian genomes. BioEssays 38: 694–703.

Duncan, J. R., 2018 *Owls of the world: their lives, behavior and survival*. Johns Hopkins University Press, Baltimore, United States.

Ericson, P. G. P., C. L. Anderson, T. Britton, A. Elzanowski, U. S. Johansson *et al.*, 2006 Diversification of Neoaves: Integration of molecular sequence data and fossils. Biol. Lett. 2: 543–547.

Espíndola-Hernández, P., J. C. Mueller, M. Carrete, S. Boerno, and B. Kempenaers, 2020 Genomic evidence for sensorial adaptations to a nocturnal predatory lifestyle in owls. Genome Biol. Evol. 12: 1895–1908.

Espindola-Hernandez, P., J. C. Mueller, and B. Kempenaers, 2022 Genomic signatures of the evolution of a diurnal lifestyle in Strigiformes. G3 Genes, Genomes, Genet. 12.

Feduccia, A., 2003 ‘Big bang’ for tertiary birds? Trends Ecol. Evol. 18: 172–176.

Feduccia, A., 1995 Explosive Evolution in Tertiary Birds and Mammals. Science 267: 637–638.

Feduccia, A., 1999 *The origin and evolution of birds*. Yale University Press, United States.

Fujito, N. T., Z. R. Hanna, M. Levy-Sakin, R. C. K. Bowie, P. Y. Kwok *et al.*, 2021 Genomic Variation and Recent Population Histories of Spotted (Strix occidentalis) and Barred (Strix varia) Owls. Genome Biol. Evol. 13: 1–12.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun *et al.*, 2008 A phylogenomic study of birds reveals their evolutionary history. Science. 320: 1763–1768.

del Hoyo, J., A. Elliott, J. Sargatal, and J. Cabot, 1999 *Handbook of the birds of the world*. Lynx Edicions, Barcelona.

Jarvis, E. D., S. Mirarab, A. J. Aberer, B. Li, P. Houde *et al.*, 2014 Whole-genome analyses resolve early branches in the tree of life ofmodern birds. Science 346: 1320–1331.

Kapusta, A., and A. Suh, 2017 Evolution of bird genomes—a transposon's-eye view. Ann. N. Y. Acad. Sci. 1389: 164–185.

Keis, M., J. Remm, S. Y. W. Ho, J. Davison, E. Tammeleht *et al.*, 2013 Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia (A. Phillimore, Ed.). J. Biogeogr. 40: 915–927.

König, C., and F. Weick, 2008 *Owls of the World*. Christopher Helm. A & C Black, London.

Luna, Á., A. Palma, A. Sanz-Aguilar, J. L. Tella, and M. Carrete, 2019 Personality-dependent breeding dispersal in rural but not urban burrowing owls. Sci. Rep. 9: 2886.

McClure, C. J. W., S. E. Schulwitz, D. L. Anderson, B. W. Robinson, E. K. Mojica *et al.*, 2019 Commentary: Defining Raptors and Birds of Prey. J. Raptor Res. 53: 419.

Mueller, J. C., B. Kempenaers, H. Kuhl, S. Boerno, J. L. Tella *et al.*, 2018 Evolution of genomic variation in the burrowing owl in response to recent colonization of urban areas. Proc. R. Soc. B Biol. Sci. 285: 1–9.

Nelson, W. S., T. Dean, and J. C. Avise, 2000 Matrilineal history of the endangered Cape Sable seaside sparrow inferred from mitochondrial DNA polymorphism. Mol. Ecol. 9: 809–813.

Neodorf, D. L. H., 2004 Extrapair Paternity in Birds: Understanding Variation Among Species. Auk 121: 302–307.

Parker, G. A., 1979 SEXUAL SELECTION AND SEXUAL CONFLICT. Sex. Sel. Reprod. Compet. Insects 123–166.

Prum, R. O., J. S. Berv, A. Dornburg, D. J. Field, J. P. Townsend *et al.*, 2015 A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526: 569–573.

Rebolo-Ifrán, N., J. L. Tella, and M. Carrete, 2017 Urban conservation hotspots: Predation release allows the grassland-specialist burrowing owl to perform better in the city. Sci. Rep. 7: 1–9.

Rodriguez-Martínez, S., M. Carrete, S. Roques, N. Rebolo-Ifrán, and J. L. Tella, 2014 High urban breeding densities do not disrupt genetic monogamy in a bird species (N. Johnson, Ed.). PLoS One 9: e91314.

Smeds, L., V. Warmuth, P. Bolivar, S. Uebbing, R. Burri *et al.*, 2015 Evolutionary analysis of the female-specific avian W chromosome. Nat. Commun. 6: 1–10.

Zhou, Q., J. Zhang, D. Bachtrog, N. An, Q. Huang *et al.*, 2014 Complex evolutionary trajectories of sex chromosomes across bird taxa. Science. 346: 1246338.

Acknowledgments

... and so he said, "if you would like to be like her[*]*, you should start by getting a PhD in Germany".*

I am very grateful to my advisors Bart Kempenaers and Jakob Mueller for their support and the amazing opportunity of working in Seewiesen these years. I highlight the incredible patience and courage of Jakob, who accepted the challenge of introducing me to the field of bioinformatics. I am also very grateful to all members of my Thesis Advisory Committee (TAC/PAC) for their guidance, constructive feedback, and all the conversations that undoubtedly contributed to the development of my thoughts and knowledge, and to the building of the framework of this dissertation, *i.e.*, Bart Kempenaers, Jakob Mueller, John Parch, Judith Mank, and Martina Carrete.

I feel very grateful for all the people that lend a hand to this project in so many ways. Thanks to all the ones that provided the biological samples from the owls: James Duncan and Charlene Berkvens (Discover Owls, Manitoba, Canada), Jean-Michel Hatt (ZooZurich, Switzerland), Guillermo Blanco (MNCN-CSIC, Spain), Jose A. Sanchez Zapata and Juan M. Perez (UMH, Spain), Philippe Helsen (Zoo Antwerp, Belgium), and Alexandre Roulin (University of Lausanne, Switzerland). I am really grateful that most of them used non-invasive sampling techniques, ensuring low stress and survival for the owls. Thanks to everyone who has shared their knowledge with me, taught me, and saved

[*] Professor Dr. Carolina Villagran Moraga. Departamento de Biología, Facultad de Ciencias, Universidad de Chile.

me (and my laptop… and probably the server as well) from all kinds of informatic disasters: Francisco Salinas García♥, Meng-Ching Ko, Iulia Darolti, Fidel Botero-Castro, Henryk Milewski, Jan Drosd, and the users-support (Super-) team from gwdg in Göttingen.

I would like to thank all Seewiseners, my colleagues from the Kempenaers department, and my neighbors in *The Ghetto* for their friendly company during these years. Especially to Kim, Yifan, and Stefan for catsitting Gioco and Mika (I know it is hard work), and all passengers of the 951 bus who made those hours nicer and transformed them into a chance for culinary, linguistic, and philosophical colloquia.

Even when I know that the list is endless and that I am profoundly aware and grateful for the "little" details that have kept me going, I will use a few lines to thank my colleagues, friends, and family. Thank every person who has been with me during all these years, especially those who have been in online meetings at inconvenient hours (for them) to talk about science, cats, birds, owls, sweet recipes, cats, home-made bread, semantical issues, cats, movies, and other essentials. I thank my parents, my younger "little" brother, and all my family for also sharing those essential conversations with me, and I acknowledge their unstoppable intelligence and kindness, which taught me that these virtues are totally inexpensive and, therefore, can be enjoyed at any table. I thank Pancho♥ for his holistic company and support, for being my partner, my friend, and such a terrific peer reviewer.

♥ same person

Funding and facilities

My doctoral work was financed by a co-funded scholarship from the DAAD and BecasChile. Additional funding support came from the IMPRS, and the Max Planck Society through Kempenaers Department budget.

The Max Plank Institute for Ornithology provides the facilities and IT support.

This work used the Scientific Compute Cluster at GWDG, the joint data center of Max Planck Society for the Advancement of Science (MPG) and University of Göttingen.

Pamela A. Espíndola Hernández

pamela.dola@gmail.com

RESEARCH EXPERIENCE

09.2022 – present **Visiting Researcher.**
Department of Environmental Microbiology, Helmholtz Centre for Environmental Research UFZ. Permoserstr. 15, 04318 Leipzig, Germany
Supervisor: Dr. Ulisses Nunes da Rocha
Description: Functional annotation of birds' microbiomes using Machine Learning.

04.2017 – present **PhD Student**
Max Planck Institute for Ornithology, Seewiesen, Germany
Supervisors: Dr. Bart Kempenaers, and Dr. Jakob Mueller
Description: Genome-wide comparative study of selection signatures associated with adaptation to the nocturnal lifestyle of owls.

03.2015 - 08.2016 **Technical Assistant**
Departamento de Ecología y Biodiversidad, Univ. Andrés Bello, Chile.
Supervisor: Dr. Rodrigo A. Vásquez, and Dr. Verónica Quirici
Description: Fieldwork (experiments and sampling for behavioral studies of birds), telomere length quantification by qPCR, and statistical analyses.

12.2012 - 01.2015 **Technical Assistant**
Terrestrial Ecology Laboratory, Faculty of Science, Univ. de Chile, Chile
Supervisor: Dr. Rodrigo A. Vásquez
Description: Fieldwork (experiments and sampling for behavioural studies of birds and mammals), DNA extraction and PCR, and logistical support.

EDUCATION

04.2017 – present PhD Student.
LMU and Max Planck Institute for Ornithology, Germany.

2014 -2016 Master in Biological Sciences.
Universidad de Chile, Santiago, Chile.

2009 -2014 Bachelor's degree, with professional title: Biologist with a major in Environment.
Universidad de Chile. Santiago, Chile

2006 -2008 Bachelor's degree in natural sciences.
Universidad de Chile. Santiago, Chile

LANGUAGES

- Spanish (native)
- English (advance)
- German (intermediate)

AWARDS, GRANTS, AND SCHOLARSHIPS

10.2022	IMPRS Annual Best Paper Award, for Espíndola-Hernández *et. al.* 2022
09.2022	HIDA Visiting Reasercher Grant. Three-month project funded by the HIDA - Helmholtz Information & Data Science Academy.
03.2017	DAAD-BecasChile Scholarship for doctoral studies in Germany.
03.2006	Juan Gómez Millas Scholarship (Beca Juan Gómez Millas). Partially found bachelor's degrees of lower-income students with academic excellence.

PUBLICATIONS

* Genomics related:

Espíndola-Hernández, P., Mueller, J. C., & B. Kempenaers **(2022)** Genomic signatures of the evolution of a diurnal lifestyle in Strigiformes. *G3: Genes, Genomes, Genetics*. doi:10.1093/g3journal/jkac135

Mueller, J. C., Botero-Delgadillo, E., **Espíndola-Hernández, P.**, Gilsenan, C., Ewels, P., Gruselius J., & B. Kempenaers. (2022) Local selection signals in the genome of Blue tits emphasize regulatory and neuronal evolution. *Molecular Ecology*. doi:10.1111/mec.16345

Espíndola-Hernández, P., Mueller, J. C., Carrete, M., Boerno, S., & B. Kempenaers (2020) Genomic evidence for sensorial adaptations to a nocturnal predatory lifestyle in owls. *Genome Biology and Evolution*. doi:10.1093/gbe/evaa166

* Behavioural ecology related:

Poblete, Y., Botero-Delgadillo, E., **Espíndola-Hernández, P.**, & R. A. Vásquez. (2021) Risk-taking behaviour relates to timing of breeding in a sub-Antarctic rainforest bird. *Ibis*. doi:10.1111/ibi.12941

Poblete, Y., Botero-Delgadillo, E., **Espíndola-Hernández, P.**, Südel, G., & R. A. Vásquez. (2021) Female extra-pair behavior is not associated with reduced paternal care in Thorn-tailed Rayadito. *Ecology and Evolution*. doi: 10.1002/ece3.7232

Cuevas, E., Orellana-Peñailillo, C., Botero-Delgadillo, E., **Espíndola-Hernández, P**., Vásquez, R. A., & V. Quirici. (2021) Influence of the haemosporidian Leucocytozoon spp. over reproductive output in a wild Neotropical passerine, the Thorn-tailed Rayadito (*Aphrastura spinicauda*). *Ibis*. doi:10.1111/ibi.12934

Quirici, V., Botero-Delgadillo, E., González-Gómez, P. L., **Espíndola-Hernández, P.**, Zambrano, B., Cuevas, E., Wingfield, J. C. & R. A. Vásquez. (2021) On the relationship between baseline corticosterone levels and annual survival of the thorn-tailed rayadito. *General and Comparative Endocrinology*, 300. doi:10.1016/j.ygcen.2020.113635

Quirici, V., Hammers, M., Botero-Delgadillo, E., Cuevas, E., **Espíndola-Hernández, P.**, & R. A. Vásquez (2019). Age and terminal reproductive attempt influence laying date in the thorn-tailed rayadito. *Journal of Avian Biology*, 50(10). doi:10.1111/jav.02059

Espíndola-Hernández, P., Castaño-Villa, G. J., Vásquez, R. A., & V. Quirici (2017). Sex-specific provisioning of nutritious food items in relation to brood sex ratios in a non-dimorphic bird. *Behavioral Ecology and Sociobiology,* 71(4), 65. doi:10.1007/s00265-017-2294-4

Statutory declaration and statement

Eidesstattliche Erklärung
Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt worden ist.

München, den 02.11.2022 Pamela Espíndola Hernández

Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist. Des Weiteren erkläre ich, dass ich mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen habe oder ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich einer Doktorprüfung zu unterziehen.

München, den 02.11.2022 Pamela Espíndola Hernández