

Avian vocalizations and phylogenetic signal

(aves/phylogeny/vocalization/behavior)

KEVIN G. MCCracken*[†] AND FREDERICK H. SHELDON[‡]

*School of Forestry, Wildlife, and Fisheries, and [‡]Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803

Communicated by Charles G. Sibley, Santa Rosa, CA, February 6, 1997 (received for review September 27, 1996)

ABSTRACT The difficulty of separating genetic and ecological components of vocalizations has discouraged biologists from using vocal characters to reconstruct phylogenetic and ecological history. By considering the physics of vocalizations in terms of habitat structure, we predict which of five vocal characters of herons are most likely to be influenced by ecology and which by phylogeny, and test this prediction against a molecular-based phylogeny. The characters most subject to ecological convergence, and thus of least phylogenetic value, are first peak-energy frequency and frequency range, because sound penetration through vegetation depends largely on frequency. The most phylogenetically informative characters are number of syllables, syllable structure, and fundamental frequency, because these are more reflective of behavior and syringeal structure. Continued study of the physical principles that distinguish between potentially informative and convergent vocal characters and general patterns of homology in such characters should lead to wider use of vocalizations in the study of evolutionary history.

Anecdotal and scientific evidence suggest that avian vocalizations contain historical information. Field ornithologists are often able to predict taxonomic relationships on the basis of voice alone, and population biologists have used vocalizations to study the evolution of populations and species groups (1–6). Although avian vocalizations may contain information useful for constructing higher-level phylogeny (2, 3), this has not been seriously attempted because systematists studying vocalizations are confronted with several problems. Physical environment and other ecological factors play important roles in shaping vocalizations in most species, so that distantly related populations occupying similar habitats may possess vocalizations more similar than those of closely related populations in different habitats (7, 8). For example, vocalizations of species that live in dense vegetation tend to have lower frequencies and narrower frequency ranges than those of species that inhabit open areas. This is because longer wavelengths propagate energy more efficiently through vegetation than shorter wavelengths, which attenuate due to the scattering effects of leaves and branches (9–11). In addition, because vocalizations are signals, the frequency and energy of vocal components may vary according to their purpose as well habitat. As a result of Doppler-related effects, vocalizations meant to convey information about direction and distance have low frequencies and are usually of short duration, whereas alarm calls, which are more ventriloquial, have high frequencies and are generally of long duration (12). Use of avian vocalizations in phylogenetics may be confounded further by the problem of cultural evolution (13). In species that learn their songs or calls, acquired components may obscure genetic components. Finally, vocal-

izations are also constrained by syringeal morphology, which is the product of genetic and developmental influences.

These physical, ecological, behavioral, and morphological forces can cause vocal characters to be similar by convergent evolution or chance, thus limiting their usefulness for inferring phylogeny. Although these problems make systematic studies of avian vocalizations particularly difficult, they are simply homoplasies, which potentially affects all types of phylogenetic characters. Thus, recovering phylogenetic signal should be possible by careful cladistic analysis of vocal characters in taxa that have simple songs or calls that are not learned and whose habitat distributions are well understood.

With these issues and criteria in mind, we analyzed the phylogenetic information content in heron vocalizations. Herons (Ciconiiformes: Ardeidae) do not learn their vocalizations (2, 3), seem to have relatively conserved vocal repertoires (14), and inhabit a variety of open marshland and closed forest habitats (14). The phylogeny of the group has been estimated using DNA–DNA hybridization (15–17) and is reasonably well understood (18, 19). Thus, the elements for the first rigorous study of its kind are in place. As anticipated, we have found that some heron vocalization characters contain remarkably reliable phylogenetic information, even among distantly related taxa, whereas others are strongly influenced by ecological factors.

We analyzed 192 recordings of squawks, alarm calls, flight calls, and (in the case of bitterns) whistled songs from 14 heron species and an outgroup, glossy ibis (*Plegadis falcinellus*) (20, 21), and used the program CANARY version 1.1 (22) to create acoustic spectrograms. These recordings and program were provided by the Library of Natural Sounds, Cornell Laboratory of Ornithology. Spectrograms depict the frequency and energy of sound in time. Although herons are usually silent, their vocal repertoires nonetheless vary within and among species. Some species squawk, whereas others, such as bitterns, sing; some species deliver their calls in flight and others call from perches. We analyzed flight calls for species that call from the air, squawks for those that do not, and whistled songs for bitterns.

The question arises whether it is appropriate to compare a squawk in one species to a song or flight call in another? Certainly a phylogeneticist would not compare characters of the head in one species to those of forelimbs in another. Morphologists need to compare heads with heads and forelimbs with forelimbs to provide spatial reference for the identification of potentially homologous characters in different species. Homology is then tested by using the characters in a phylogenetic analysis (23). The identification of homology in vocal characters, however, proceeds by a different initial step. Squawks, songs, alarm calls, etc., are combinations of fundamental sounds, or phonics, just as words in language may be produced by combining syllables. These fundamental sounds are potential vocal homologies. To identify them, one might employ the strategy of morphology and compare similar types

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Copyright © 1997 by THE NATIONAL ACADEMY OF SCIENCES OF THE USA
0027-8424/97/943833-4\$2.00/0
PNAS is available online at <http://www.pnas.org>.

[†]To whom reprint requests should be addressed. e-mail: kmccrac@unix1.sncc.lsu.edu.

of vocalizations in different species (e.g. squawks) for similar fundamental sounds. But what defines a squawk (or song or call); these categories of vocalization often grade into one another. Also, what if a species does not have a squawk in its repertoire, but has a squawk-like sound in its song or flight call? Might not that sound be homologous to sound in a true squawk? An advantage of comparing vocalizations is that there is another method besides spatial reference to identify potentially homologous characters. Vocal characters are composed of quantitative features (e.g., wavelength and energy), which make it possible to relate sounds directly in different species. These characters may be simple noises that are largely a function of syringeal morphology or more complex sounds that feature a greater behavioral component. By observing these quantitative factors in spectrograms, it is possible to postulate homology of vocal characters among species, and then test hypotheses of homology by phylogenetic analysis.

Using this logic, we coded five characters: (i) mean number of syllables per vocalization, (ii) syllabic structure, (iii) fundamental frequency (kHz), (iv) first peak-energy (J) frequency (kHz), and (v) frequency range (kHz). These five characters are functionally independent and describe the tonal quality and structure of each vocalization in time (see Fig. 1 legend). Using the physics of sound energy propagation as criteria, we predicted that characters *iv* and *v* would be correlated with habitat parameters because species that live in densely vegetated habitats generally have lower peak frequencies and more narrow frequency ranges than species inhabiting more open areas (9–11). In contrast, characters *i–iii* should not be as readily influenced by ecological forces. Number of syllables (character *i*) should reflect genetic components of vocal behavior. Syllabic structure (character *ii*) and fundamental frequency (character *iii*) should likewise reflect vocal behavior, but are ultimately constrained by genetic components of syringeal morphology. Thus, *i–iii* should be more phylogenetically informative than *iv–v*. To test this prediction, we mapped these characters onto the DNA-hybridization estimate of heron phylogeny using the program MACCLADE (24) and performed a randomization test for phylogenetic conservatism (25).

When arranged by parsimonious optimization on the DNA-hybridization tree, the number of syllables, syllabic structure, and fundamental frequency are congruent with three lineages of herons: (i) the rufescent tiger-heron (*Tigrisoma lineatum*) and the boat-billed heron (*Cochlearius cochlearius*), (ii) bitterns, and (iii) day-herons and night-herons (Table 1). Moreover, the arrangement of characters into shared ancestral (symplesiomorphic) and shared derived (synapomorphic) states also corresponds strongly with the hierarchical arrangement of branches on the DNA-hybridization tree (Fig. 1A; $P < 0.0001$). The rufescent tiger-heron and the boat-billed heron form a basally branching clade that shares ancestral vocalization traits with the outgroup. Their syllables are produced in a rapid series, have relatively low fundamental frequencies (≤ 0.30 kHz), and consist of multiple harmonic overtones of the fundamental frequency (Fig. 1A). Bittern songs share two derived characters with their sister group the day-herons and night-herons; bittern syllables are not emitted in a rapid series

and have fundamental frequencies ($0.35 < f < 0.5$ kHz) intermediate between those of tiger-herons and typical herons (assuming ordered states for fundamental frequency). Unlike most herons, however, the bitterns sing, and thus their vocalizations are characterized by several autapomorphies. Their calls contain tonal, whistled notes, lack harmonics in all or some syllable parts, and possess syntax (26). Day-herons and night-herons are united by vocalizations that consist of one to three syllables (Fig. 1A). Day-herons and night-herons also lack pure whistled notes and have relatively high fundamental frequency vocalizations (≥ 0.5 kHz).

Although these three characters largely reflect phylogeny, they are not entirely free of homoplasy and must be interpreted with caution. For example, the call of the least bittern (*Ixobrychus exilis*), like that of the American bittern (*Botaurus lentiginosus*), contains harmonic notes. But it lacks the whistled notes characteristic of other members of the bittern clade, suggesting that the least bittern has either lost whistling or retained the ancestral condition. Nonetheless, least bittern vocalizations possess fewer and weaker harmonic overtones than do those of most other ardeids.

In contrast to vocal characters *i–iii*, peak-energy frequency and frequency range do not appear to contain substantial phylogenetic information. Their distribution among taxa is more consistent with physical environmental and ecological predictions (9–11) (Fig. 1B and C). For example, the rufescent tiger-heron and the zigzag heron (*Zebriulus undulatus*) inhabit forests and emit their vocalizations from within dense vegetation (14). They have unusually low peak-energy frequency vocalizations (≤ 0.6 kHz) and narrow frequency ranges (≤ 1.0 kHz). Conversely, species that inhabit savannas and grasslands, such as the whistling heron (*Syrigma sibilatrix*) and the cattle egret (*Bubulcus ibis*) (14), generally have higher peak-energy frequencies (≥ 2.5 kHz) and broader frequency ranges (≥ 2.5 kHz) than those of closely related species that live in habitats with intermediate vegetation densities, such as marshes.

There are, of course, exceptions to the expected habitat effect on peak-energy frequency and frequency range, just as there are for characters expected to be phylogenetically informative. Where a bird sings or calls may be more important than general habitat characteristics. For instance, unlike most other marsh herons, which generally deliver their calls in flight, the American bittern has a relatively low peak-energy frequency and narrow frequency range in the whistled portions of its call, probably because it vocalizes on the ground, deep within the marsh. Other bitterns and tiger-herons vocalize from perches in dense vegetation as opposed to edges of bushes or tops of trees. Such factors may vary on a fine ecological scale and potentially affect song characteristics.

In summary, the phylogenetic information content of the five vocal characters is highly predictable. The manner in which birds compile syllables, the structure of those syllables, and their fundamental frequencies are expected to be influenced mainly by cumulative forces of genetic history that have shaped syringeal morphology and singing/calling behavior. In contrast, harmonic modifications of the fundamental frequency are expected to be more plastic and to respond more readily to environmental structure and other ecological variables. The distinction, in this case, between potentially informative and uninformative vocal characters based on the simple physics of sound suggests that it should be relatively straightforward for systematists to identify and discard vocal characters most likely to be influenced by habitat. The ability to distinguish between potentially useful and unuseful characters in morphology is already well advanced (27), and there is no reason why this should not also be true of vocal characters and the behavioral components of vocalizations. To the extent that it is practicable, physical criteria should be used to assess the comparability, context, and meaning of different vocalizations both within and among clades. When applied consistently, such an ap-

Table 1. Phylogenetic signal contained in three vocal characters for three clades of herons and the glossy ibis

Species	No. syllables	Syllabic structure	Fundamental frequency
Outgroup (glossy ibis)	Series	Harmonic	Low
Tiger-heron, boat-billed heron	Series	Harmonic	Low
Bitterns	1, 2, or 5	Whistle	Intermediate
Day-herons, night-herons	1, 2, or 3	Harmonic	High

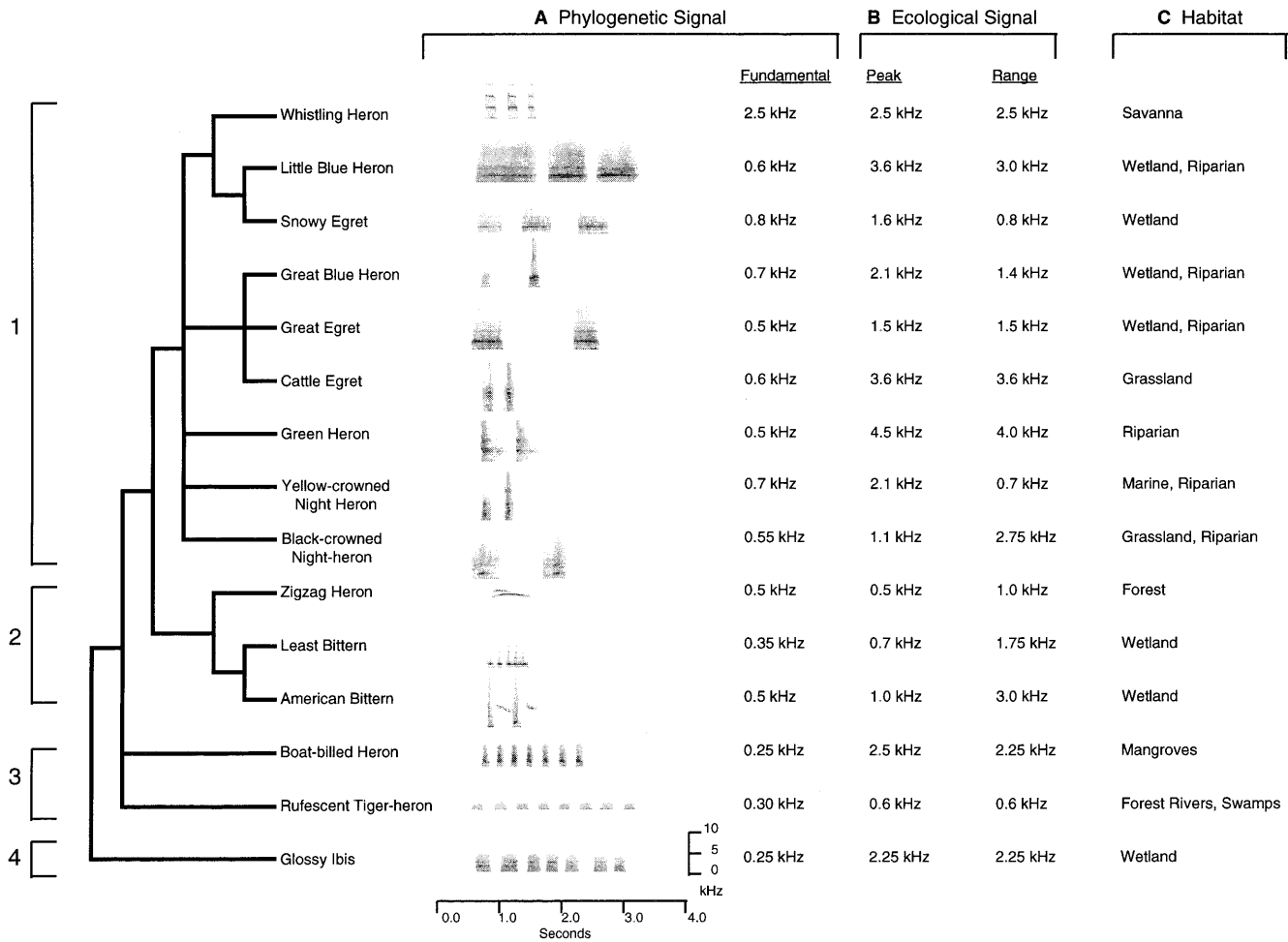


FIG. 1. Heron phylogeny, corresponding spectrograms, vocal characters, and habitat distributions. Branch topology represents the best estimate of heron phylogeny based on DNA-DNA hybridization (15-17) including night-herons and day-herons (1), bitterns (2), rufescent tiger-heron and boat-billed heron (3), and outgroup (4). (A) Phylogenetically informative vocal characters including: number of syllables, syllabic structure, and fundamental frequency were mapped onto the phylogeny using MACCLADE (24) yielding a tree of 10 steps with an ensemble consistency index (CI) = 0.8 (number of syllables, 6 steps, CI = 0.833; syllabic structure, 2 steps, CI = 0.5; fundamental frequency, 2 steps, CI = 1.0). (B) Ecologically informative vocal characters including peak-energy (J) frequency (kHz) and frequency range (kHz). (C) Habitat distributions suggest that species that inhabit open areas such as savannas, grasslands, and open marshes have higher peak-energy (J) frequencies (kHz) and broader frequency ranges (kHz) than do taxa inhabiting closed habitats such as forests. Number of syllables is the number most frequently produced. Ibises, tiger-herons, and boat-billed herons emit a rapid series of similar syllables; other heron vocalizations generally consist of singlets, doublets, or triplets. Syllabic structure may be tonal (i.e., pure whistled notes) or harmonic (i.e., possessing overtones; integral multiples of the base frequency). Fundamental frequency (kHz) is the base frequency of a syllable and is a function of syrinx morphology. All other notes are overtones. First peak-energy (J) frequency (kHz) is the frequency with the greatest energy amplitude. Frequency range (kHz) was calculated as the difference between the greater of the first or second peak energy frequency and the base frequency and describes the timbre or tonal quality of a syllable. Characters *i-ii* are discrete and were coded as unordered, whereas characters *iii-v* were ordered because they are continuous. Although they are not mathematically independent, fundamental frequency, peak-energy frequency, and frequency range (characters *iii-v*) can be considered functionally independent because the notes and range of notes comprising a song are ultimately determined by behavior and morphology. The randomization test consisted of randomly reshuffling number of syllables, syllabic structure, and fundamental frequency 10,000 times over the DNA-hybridization branch topology and comparing the distribution of tree lengths with the length of a single tree obtained by parsimoniously mapping the same three characters. The same test applied to peak-energy frequency and frequency range does not indicate phylogenetic conservativeness ($P < 0.7$).

proach has the potential to revitalize the study of vocal phylogenetics.

J. V. Remsen suggested that we compare heron vocalizations for phylogenetic inference. We also thank A. Afton, M. Cohn-Haft, P. Dunn, M. Hafner, D. Lane, E. Mayr, R. Payne, J. V. Remsen, C. Sibley, M. Stine, L. Whittingham, D. Winkler, and three anonymous reviewers for comments on the manuscript; the Library of Natural Sounds, Cornell Laboratory of Ornithology, Ithaca, NY; and the following recordists: P. P. Kellog, A. A. Allen, R. S. Little, G. B. Reynard, O. H. Hewitt, D. Minis, W. Belton, J. D. MacDonald, C. A. Sutherland, J. Priori, A. Priori, W. V. Ward, R. D. Bayer, D. S. McChesney, W. Y. Brockelman, J. W. Kimball, R. C. Stein, W. W. H. Gunn, L. Lipps, L. Payne, E. Peck, M. E. W. North, T. A. Parker, III, A. van den Berg, M. R. Plymire, D. S. Herr, T. H. Davis, G. F. Budney, S. R. Pantle,

L. I. Davis, F. Peck, L. F. Kibler, D. L. Ross, Jr., and W. R. Evans. Support for this project was provided by the Louisiana Cooperative Fish and Wildlife Research Unit, Louisiana State Board of Regents, National Science Foundation Grant BSR-9207991, and National Science Foundation/Louisiana Education Quality Support Fund Grant 1992-96-ADP-02.

- Gill, F. B. (1990) *Ornithology* (Freeman, New York).
- Payne, R. B. (1986) *Curr. Ornithol.* **3**, 87-126.
- Payne, R. B. (1981) in *Natural Selection and Social Behavior*, eds. Alexander, R. D. & Tinkle, D. W. (Chiron, New York), pp. 108-120.
- Kroodsma, D. E. (1977) *Am. Nat.* **111**, 995-1008.
- Catchpole, C. K. (1980) *Behaviour* **74**, 149-166.
- Payne, R. B. (1983) in *Social Behavior of Female Vertebrates*, ed. Wasser, S. K. (Academic, New York), pp. 55-91.

7. Nottebohm, F. (1975) *Am. Nat.* **109**, 605–624.
8. Hunter, M. L. & Krebs, J. R. (1979) *J. Anim. Ecol.* **48**, 759–785.
9. Chappius, C. (1971) *Terre Vie* **25**, 183–202.
10. Morton, E. S. (1975) *Am. Nat.* **109**, 17–34.
11. Wiley, R. H. & Richards, D. G. (1982) in *Acoustic Communication in Birds*, eds. Kroodsma, D. E. & Miller, E. H. (Academic, New York), Vol. 1D, pp. 132–181.
12. Marler, P. (1955) *Nature (London)* **176**, 6–8.
13. Payne, R. B. (1981) *Anim. Behav.* **29**, 688–697.
14. Hancock, J. & Elliot, H. (1978) *The Herons of the World* (Harper & Row, New York).
15. Sheldon, F. H. (1987) *Mol. Biol. Evol.* **4**, 56–69.
16. Sheldon, F. H. (1987) *Auk* **104**, 97–108.
17. Sheldon, F. H., McCracken, K. G. & Stuebing, K. G. (1995) *Auk* **112**, 672–679.
18. Bock, W. J. (1956) *Am. Mus. Nov.* **1779**, 1–49.
19. Payne, R. B. & Risley, C. J. (1976) *Misc. Publ. Univ. Mich. Mus. Zool.* **150**, 1–115.
20. Sibley, C. G. & Ahlquist, J. E. (1990) *Phylogeny and Classification of Birds* (Yale Univ. Press, New Haven, CT).
21. Sheldon, F. H. & Kinnarney, M. (1993) *Syst. Biol.* **42**, 32–48.
22. Cornell Laboratory of Ornithology (1993) CANARY: The Cornell Bioacoustics Workstation (Cornell Laboratory of Ornithology, Ithaca, NY), Version 1.1.
23. Patterson, C. (1982) in *Problems of Phylogenetic Reconstruction*, eds. Joysey, K. A. & Friday, A. E. (Academic, London), pp. 22–74.
24. Maddison, W. P. & Maddison, D. R. (1992) MACCLADE: Analysis of Phylogeny and Character Evolution (Sinauer, Sunderland, MA).
25. Maddison, W. P. & Slatkin, M. (1991) *Evolution (Lawrence, Kans.)* **45**, 1184–1197.
26. Spector, D. A. (1994) *J. Theor. Biol.* **168**, 373–381.
27. Marshall, C. R. (1992) *Mol. Biol. Evol.* **9**, 309–322.