Anim. Behav., 1996, **51,** 273–291

Hop, step and gape: do the social displays of the Pelecaniformes reflect phylogeny?

MARTYN KENNEDY*, HAMISH G. SPENCER* & RUSSELL D. GRAY†
**Department of Zoology, University of Otago*
†Department of Psychology, University of Auckland

(Received 6 December 1994; initial acceptance 24 January 1995; final acceptance 31 May 1995; MS. number: 4799)

Abstract. Several authors have recently revived the old ethological claim that behavioural characters could be used to provide accurate estimates of phylogeny. To test this contention we constructed a behavioural data set for the pelecaniforms based upon van Tets' (1965, *Ornithol. Monogr.*, **2,** 1–88) classic comparative study of their social behaviour. Parsimony analysis of the resulting 20 taxa, 37-character data set produced 12 shortest trees. These trees fitted the behavioural data well. A permutation tail probability test found that the behavioural data contained more structure than would be expected by chance alone. The behavioural trees were compared with previously published morphological and genetic estimates of pelecaniform phylogeny to test if they were more congruent than would be expected by chance. In all cases the behavioural trees were far more similar to the morphological and genetic trees than chance alone would predict. By mapping the distribution of the behavioural characters onto a best-estimate evolutionary tree constructed from the independent trees, it was possible to investigate the homology of the behavioural characters. All but three of the characters appeared to be homologous. The behavioural characters were, therefore, no more homoplasious than other types of characters. The best-estimate tree was used to predict the social displays of the pied cormorant, *Phalacrocorax varius*, and to evaluate hypotheses about the derivation of male advertising displays.

In his classic comparative study, van Tets (1965) investigated the distribution of social displays within the Pelecaniformes. One example of the displays he studied is the gape display of the cormorants and shags. The gape display is a bisexual recognition display of the 'in' (or sitting) bird which involves movement of the head in the median plane with the bill opened at the display's climax (van Tets 1965). Gaping, like some other recognition displays, may be derived from threat displays (van Tets 1965). van Tets also noted the specific position of the head during gaping, the number of times the bill was opened per display, and the vocalization during the display. He used displays such as gaping to construct what amount to informal evolutionary trees for the pelecaniforms (gaping is present in seven of the eight cormorant and shag species and unknown in the other; see van Tets 1965, Figures 47–49). Following the method of Lorenz (1941), van Tets constructed these diagrams by using common social attributes to group the taxa. He used, for example, the presence of the gaping display to group the cormorants and shags, repetitive gaping to group the pelagic shag, *Phalacrocorax pelagicus*, and red-faced shag, *P. urile*, and neck stretched during gaping to group the double-crested cormorant, *P. auritus*, and Neotropic cormorant, *P. olivaceus* (see van Tets 1965, Figure 49). The inference that can be taken from these diagrams is that taxa that share several social displays are closely related. van Tets compared the relationships implied by his diagrams with the proposed phylogenetic trees of Lanham (1947) and Sibley (1960) (based on anatomy and egg-white proteins, respectively). He argued that his analysis tended to support the tree of Lanham rather than that of Sibley (van Tets 1965). From his analysis, van Tets concluded that the distribution of displays such as gaping does

Correspondence: M. Kennedy, Department of Zoology, University of Otago, P.O. Box 56, Dunedin, New Zealand (email: martyn.kennedy@stonebow.otago.ac.nz). R. D. Gray is at the Department of Psychology, University of Auckland, Private Bag 92019, Auckland, New Zealand.

0003–3472/96/020273+19 $12.00/0

reflect the phylogenetic history of the pelecaniforms.

Despite the work of Lorenz (1941), van Tets (1965) and other early ethologists, behavioural traits are infrequently used to infer evolutionary relationships. Behavioural characters are commonly viewed as inferior to other types of characters in systematics (de Queiroz & Wimberger 1993). Typically, there have been two arguments used to support this perspective. The first is that the initial (non-phylogenetic) criteria for homologizing characters may be impossible to apply to behaviour (Atz 1970; Aronson 1981). Behavioural characters may lack the properties required to make good initial assessments of homology (de Queiroz & Weimberger 1993). The second argument is that behaviour would be a questionable indicator of evolutionary relationships because it is so liable to evolutionary change (Atz 1970). Both of these arguments imply that behavioural characters are likely to have a higher level of homoplasy (similar form not due to common descent which is manifested in a phylogenetic tree as character convergence or reversal) than other types of characters (de Queiroz & Wimberger 1993). Following the first argument, the homoplasy would be caused by a relatively large number of mistakes in assessing which behavioural characters are homologous. The second argument implies that even if behavioural characters can be assigned a priori as putative homologues, behaviour would still ultimately be more homoplasious than other types of characters (de Queiroz & Wimberger 1993).

Comparative biology has recently been revolutionized by advances in phylogenetic methods. Studies investigating adaptation (e.g. Ridley 1983; Coddington 1988), coevolution (e.g. Hafner & Nadler 1988; Brooks & McLennan 1991; Page 1993; Paterson et al. 1993; Hafner et al. 1994; Paterson & Gray, in press) and biogeography (e.g. Nelson & Platnick 1981; Page 1988, 1990; Cracraft 1994) have all benefited by using an explicit phylogenetic approach. There has also recently been renewed interest in the use of behaviour in phylogenetic studies. A growing number of studies that utilize these methods have shown that behavioural characters are no more convergent than other character types. Several studies have investigated behavioural homology by mapping the behavioural characters onto trees derived from either genetic or morphological characters (e.g. Beehler & Swaby 1991; Langtimm & Dewsbury 1991; Packer 1991; McKitrick 1992; Clayton & Harvey 1993; Sillen-Tullberg & Møller 1993; Winkler & Sheldon 1993). Other phylogenetic studies have used either behavioural characters or a mixture of behaviour and other types of characters to generate evolutionary trees (Cracraft 1985; McLennan et al. 1988; Arntzen & Sparreboom 1989; Prum 1990; Crowe et al. 1992; Proctor 1992; Paterson et al., in press). These studies have generally produced good estimates of phylogeny as assessed by low levels of homoplasy, or by congruence between behavioural and independently derived trees. de Queiroz & Wimberger (1993), for example, reviewed 22 data sets that contained both behavioural and morphological characters and showed that behavioural traits were no more homoplasious than morphological characters. Similarly, Paterson et al. (in press) found that molecular and behavioural characters gave equally good estimates of phylogeny in the penguins and petrels.

The classification of the pelecaniforms has been the subject of some debate. Traditionally, the pelecaniforms have been viewed as a monophyletic group (e.g. see Cracraft 1985). Sibley & Ahlquist (1990, page 502), however, concluded that the relationships suggested by their data differed so markedly from the traditional view of the Pelecaniformes that they 'may present the most complex and controversial questions in the avian phylogeny'. If behaviour reflects phylogeny in this group, behavioural displays may offer an alternative data source to address this type of problem. While van Tets (1965) did make inferences about the group's phylogenetic relationships based upon his diagrams, they cannot be considered a formal estimate of phylogeny. The inferences van Tets made are not based on an explicit analysis to find the best tree.

In our paper we address the question of whether pelecaniform social displays reflect phylogeny, using quantitative phylogenetic methods. We use a parsimony program to find the best tree or trees based on the behavioural information in van Tets (1965). If behaviour does reflect phylogeny, we should expect the data to provide a good fit to the tree(s) (high consistency and retention indices) and for the behavioural data to be no more homoplasious than other types of data. If the structure in the data is due to common evolutionary history, the behavioural tree(s) should be

congruent with other, independently derived trees (thus showing common phylogenetic signal).

ANALYSIS

Behavioural Data

The 20 pelecaniform taxa used in this study are those reported in van Tets (1965) Figures 47–49 [*Phaethon*, tropicbirds; *Fregata*, frigatebirds; *Pelecanus*, pelicans; *Morus capensis*, Cape gannet; *M. serrator*, Australasian gannet; *M. bassanus*, northern gannet; *Sula leucogaster*, brown booby; *S. dactylatra*, masked booby; *S. sula*, red-footed booby; *S. nebouxii*, blue-footed booby; *S. variegata*, Peruvian booby; *Anhinga*, anhingas or darters; *Phalacrocorax varius*, pied cormorant; *P. auritus*, double-crested cormorant; *P. olivaceus*, Neotropic (olivaceous) cormorant, subsequently synonymized with *P. brasilianus* by Browning (1989); *P. carbo*, great (black) cormorant; *P. penicillatus*, Brandt's cormorant; *P. pelagicus*, pelagic shag; *P. urile*, red-faced shag; *P. aristotelis*, European shag; the generic and specific names we used follow van Tets]. We collected the behavioural information for those taxa almost exclusively from van Tets' (1965) comprehensive monograph. In a few cases where van Tets (1965) provided no information about the presence or absence of a behavioural display in a species, additional data were taken from van Tets (1976). The exact reference for each character is given after the description of that character in Appendix 1. We generated 37 behavioural characters from van Tets (1965) (see Appendix 2). Virtually all of these characters were social displays (with the exception of locomotion on land, character 34). The characters ranged in complexity from, for example, vocalization prior to take off to the gaping display.

A potential problem with the behavioural data set is that a few of the characters are not independent. Characters 12–14 (head position during gape, repetitive gaping, vocalize during gape) are not independent of character 11 (gaping): for those taxa in which gaping is absent, it is obviously not possible to ask how the head is positioned during gaping, whether the gape is repeated or if vocalization also occurs. Similarly characters 29 and 30 are not independent of character 2, characters 6–8 are not independent of character 5 and character 4 is not independent of character 3. We included these characters, however, because if they were omitted substantial amounts of information used by van Tets to infer the relationships between the taxa would have been lost. Additionally the characters were included to allow the investigation of their evolution. A data set excluding the non-independent characters was also used to investigate whether they affected the resulting phylogeny. The non-independence of some of the characters does violate the assumptions of some of the statistical analyses we performed (e.g. the permutation tail probability test). The degree of support we appeared to find for the behaviour tree(s) may also be slightly artificially increased by the non-independence of some characters, and so our values of the consistency and retention indices should be interpreted with caution. The non-independence does not, however, affect the crucial congruence test between the behavioural and independently derived trees because these are comparisons of the trees, not the data.

Phylogenetic Analysis

Parsimony analysis of the behavioural data

Of the 37 characters generated, 28 were coded as either present or absent (i.e. binary characters: 1–3, 6, 7, 9–11, 13, 15–18, 20–26, 28–33, 36, 37). The remaining nine characters had more than two states (i.e. they were multi-state characters: 4, 5, 8, 12, 14, 19, 27, 34, 35). The multi-state characters were analysed as unordered characters (i.e. a change from any state to any other state within a character counts as one step, thus a change from state 0 to 1 or 0 to 3 both count as one step: see Maddison & Maddison 1992). Although it is likely that some behavioural characters will be more phylogenetically informative than others, intuitions about which characters are most informative may be misleading. Thus, in our analysis of the behavioural data, all characters were equally weighted.

The tropicbirds were specified as the outgroup because, within the pelecaniforms, they have been found to be the sister group of the remaining members of the group (e.g. Cracraft 1985; Sibley & Ahlquist 1990). The data were analysed using the 'branch and bound' option of PAUP 3.1.1 (Swofford 1993), which guarantees to find all the most parsimonious trees (Swofford 1993).

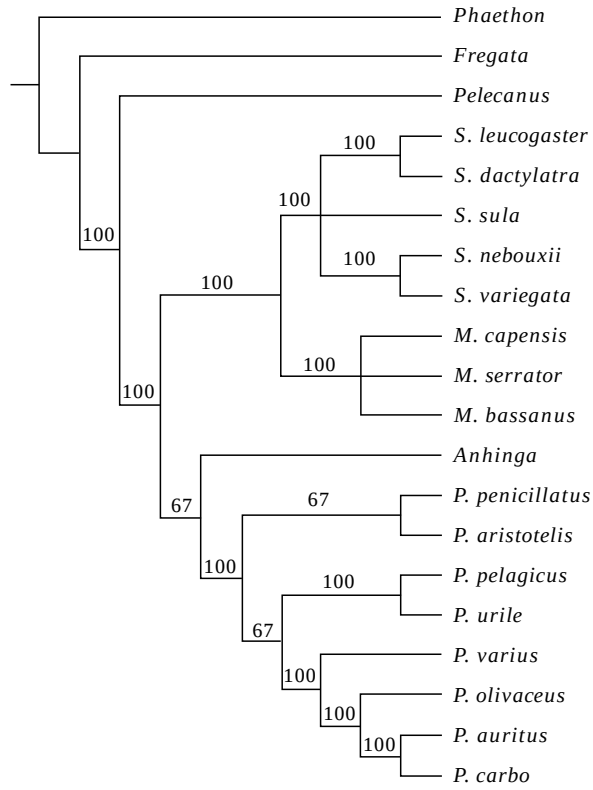

Figure 1. The majority rule consensus tree for van Tets' behaviour data. The percentages of the 12 shortest trees that include each branch are shown on the upperside of that branch.

Parsimony analysis of the behavioural data set produced 12 shortest (most parsimonious) trees of length 78. (The data set produced the same 12 shortest trees when the non-independent characters were excluded.) The treelength (TL) is the total number of evolutionary steps (i.e. character state changes from, for example, the absence to the presence of a behaviour) in the tree over all of the characters (Maddison & Maddison 1992). If the data fitted perfectly onto the trees, then the minimum possible number of steps for this data set would have been 58. The consistency index (CI) is a measure of how well the data do in fact fit the tree. The CI is calculated as the minimum possible treelength divided by the observed treelength (Kluge & Farris 1969). For the 12 shortest behavioural trees the CI was 0.744 [58/78; 0.740 = 57/77 when adjusted to remove the effect of the single uninformative character (1)]. The retention index (RI) for a tree is calculated as (maximum possible TL − actual TL)/(maximum possible TL − minimum possible TL) (Farris 1989; Maddison & Maddison 1992, page 271). The RI for the 12 most parsimonious trees was 0.836. For both the CI and RI a score of 1 represents a perfect fit, whereas a score approaching 0 signifies a complete lack of fit between the data and tree. A summary of the 12 trees is given as a majority rule consensus tree (Fig. 1). A majority rule consensus tree retains only the branches that occur in more than half of the shortest trees (Margush & McMorris 1981). Our behavioural consensus tree generally appears to be consistent with van Tets' figures. It groups the pelecaniform genera in the same way that van Tets' (1965) Figure 47 does. Within the cormorant and shag species the pelagic and red-faced shags are grouped as van Tets' Figure 49 suggests, as are the great, Neotropic and double-crested cormorants. The rest of the relationships within van Tets' Figure 49 are not well resolved and are thus difficult to interpret.

Although the CI and RI show a good fit between the data and the shortest behavioural trees, it is possible that these could be due to chance rather than any underlying signal in the data set. To test whether the behaviour data contain any more information than a random data set we used a permutation tail probability (PTP) test (Faith 1991; Faith & Cranston 1991). The PTP test compares the treelength of the 12 most parsimonious trees with a profile of shortest treelengths from randomized data sets. If the treelength of the shortest trees is significantly less than that of the randomized data sets, the data set contains significant structure. The data set was randomized by a shuffling process in which the states within each of the characters are randomly assigned to the different taxa. The results of the PTP test show that the behavioural data set contains more information than would be found by chance alone. The behavioural trees with a TL of 78 are much shorter than any of the 1000 random trees (mean TL = 135, shortest TL = 126). The difference between the lengths of the random trees and the shortest trees indicates that the behavioural trees contain significant information.

Construction of the independent trees

Four independent estimates of the pelecaniform phylogeny were extracted from the literature to compare with the behavioural trees we generated. The independent trees are based on the morphological and genetic data reported by Cracraft (1985), Sibley & Ahlquist (1990) and Siegel-Causey (1988).

To generate a tree from Cracraft's (1985) (predominantly) skeletal, morphological data set we reanalysed his published data matrix. (To differentiate between Cracraft's and Siegel-Causey's data we use the terms skeletal and osteological, respectively, as these were the terms they used.) Five of Cracraft's characters (8, 39, 42, 43 and 44) overlapped with the behavioural characters in our analysis. To avoid the circularity of comparing two estimates of phylogeny generated using some of the same characters, we removed those characters from the data set. We used all the taxa in our reanalysis of Cracraft's data, even though several of them would not be used in our comparisons. Swofford & Olsen (1990) have shown that better estimates of phylogeny are obtained by analysing a complete data set and then pruning the unnecessary taxa than by pruning the taxa before the analysis. Following Cracraft (1985) all the characters were ordered and the tree rooted via a hypothetical ancestor. Parsimony analysis of the reduced data set using the branch and bound option of PAUP 3.1.1 produced the same two shortest trees as Cracraft had found with his full data set. Both of these trees had the same subtree for the subset of taxa they have in common with the behavioural trees. Thus, for the comparison with the behavioural trees only one skeletal topology was used (see Fig. 2a).

To generate a tree from Siegel-Causey's osteological data we again reanalysed the published data matrix. Siegel-Causey's data were analysed in the same way as he had stated. All but four characters were ordered (8, 60, 62 and 97) and the taxa with multiple states within a character were coded as polymorphic (rather than uncertain). The number of taxa in the data set made it impractical to use the branch and bound option of PAUP 3.1.1 to find the most parsimonious tree(s). Instead we performed multiple heuristic searches using random addition sequences to seed 100 different searches. Multiple heuristic searches increase the chances of finding all the most parsimonious trees (see Maddison et al. 1992; Swofford 1993). This search produced the same subtree for the nine common taxa used for the comparison with the behavioural trees (see Fig. 2c).

We extracted two trees from Sibley & Ahlquist's (1990) analysis of DNA hybridization distances. The first is based on their Fitch analysis (Fitch & Margoliash 1967) and the tree is taken directly from Figure 339 on page 824 (see Fig. 2b). The second tree comes from Figures 356 (page 841) and 366 (page 851) and is based on Sibley & Ahlquist's UPGMA analysis of the DNA hybridization distances. By combining the two UPGMA figures we were able to obtain the maximum number of taxa for comparison (see Fig. 2d). While the DNA/DNA Fitch and DNA/DNA UPGMA trees are based on the same data set, the two different tree building algorithms did not generate identical trees. The DNA/DNA Fitch tree shares the same taxa as the Skeletal tree, but differs from it in the positioning of the pelicans and the frigatebirds (see Fig. 2a, b). Our composite of the DNA/DNA UPGMA trees differs from both the Skeletal and the DNA/DNA Fitch tree by grouping the pelicans and frigatebirds as sister taxa, and by grouping the gannets

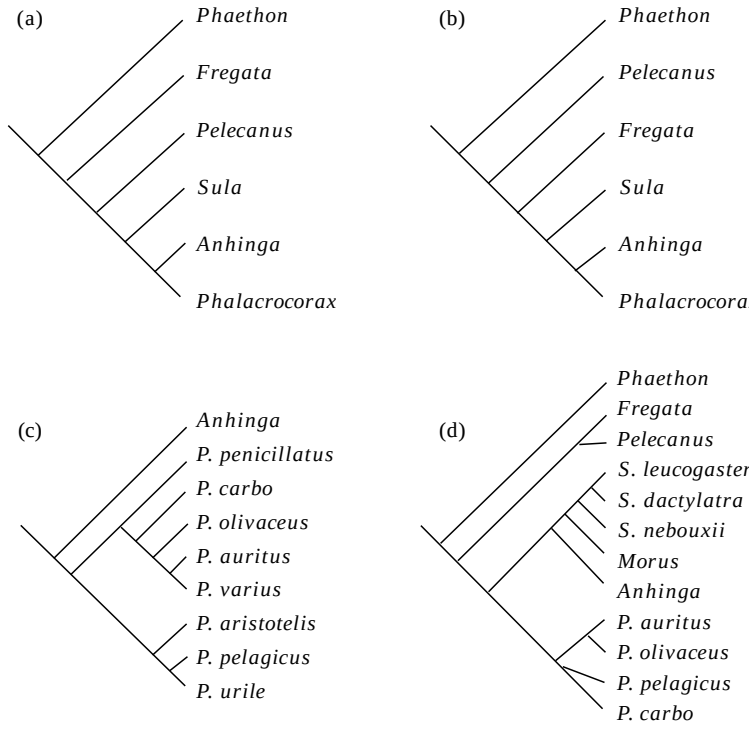

Figure 2. The pruned trees used for the comparisons. Each of the independently derived phylogenies has only one topology for the taxa it shares in common with the behaviour trees. For the 12 most parsimonious behavioural trees, there are only two pruned trees for the taxa they share with each of the independent phylogenies (although these differ depending on which taxa are being compared). Tree (a) is the Skeletal tree, tree (b) the DNA/DNA Fitch tree, tree (c) the Osteological tree, tree (d) the DNA/DNA UPGMA tree.

and boobies and the darters as sister taxa (see Fig. 2a, b, d).

Congruence between the independent and behavioural trees

The crucial test of whether behavioural characters contain phylogenetic information is the level of congruence between the behavioural trees and the independent trees. While the behavioural trees have significant structure, it is possible that this structure is a reflection of common ecological pressures rather than of evolutionary history. If the behavioural trees are congruent with the morphological and genetic trees, it is most likely to be due to common phylogenetic signal in the different data sets (Penny et al. 1982; Zink & Avise 1990).

The various trees differed in the number of taxa they shared. The Skeletal and DNA/DNA Fitch trees both shared six taxa with the behavioural data set, the Osteological tree nine taxa and the UPGMA tree 12 taxa. The behavioural trees were pruned to the appropriate taxa for the comparisons with the independently derived trees. It was necessary to collapse the behavioural trees to the generic level for the Skeletal and DNA/DNA Fitch comparisons. For the DNA/DNA UPGMA comparison the gannets were collapsed to the generic level. Collapsing the 12 most parsimonious behavioural trees for the six-taxa comparison with the Skeletal and DNA/DNA Fitch trees left just two topologies. For the nine-taxa comparison with the Osteological tree, again only two of the 12 behavioural trees' topologies were different. The difference between the remaining trees is the positioning of *P. aristotelis* with either the *P. pelagicus*, *P. urile* group, or with *P. penicillatus*. For the comparison with the DNA/DNA UPGMA tree the three *Morus* species

Table I. The mean measures of dissimilarity (symmetric difference of triplets, SDt), the range of the SDt distribution between the behavioural trees and trees derived from Cracraft (1985), Siegel-Causey (1988), Sibley & Ahlquist (1990) and the behavioural trees and 1000 random trees, and the overlap in the ranges of these comparisons

Behaviour versus	No. of taxa in common	Mean	Range	Overlap
DNA/DNA UPGMA	12	26.0	26–26	
Random		146.5	60–195	0 (0.00%)
DNA/DNA Fitch	6	3.5	3–4	
Random		13.4	3–20	15 (0.75%)
Skeletal	6	0.5	0–1	
Random		13.2	2–20	0 (0.00%)
Osteological	9	21.0	16–26	
Random		56.0	21–74	7 (0.35%)

Each comparison with an independent data set required a different comparison with the random trees due to the different number of taxa being compared. The overlap is the number (and percentage) of trees from the comparison between the behavioural trees and the random trees that had a SDt score within the range of the comparison between the behavioural and the independently derived trees (out of a possible 2000 for each comparison).

on the behavioural trees were replaced by a branch that had only the genus itself. This change did not affect the topologies of the behavioural trees in any other way as the *Morus* species were always monophyletic, but it did reduce the number of trees from 12 to six as the relative positions of the three species did differ. For the 12 taxa in common, the six remaining behavioural trees again collapsed to two topologies. These topologies differed only in whether the cormorants and shags are sister taxa of the darters, or the gannets and boobies (as the behavioural trees had for the six-taxa comparison).

To test for congruence between the behavioural trees and the four independent estimates of pelecaniform phylogeny we used the triplets tree comparison metric in the program COMPONENT 2.0 (Page 1992). This command measures the dissimilarity between two trees by comparing all the possible triplets between the trees (a triplet is the smallest informative subtree of a rooted tree and contains three terminal taxa; see Page 1992). COMPONENT calculates the symmetric difference of triplets (SDt) between the trees; this value can then be compared with a null distribution of SDt values. The triplets metric can be best explained by an example: we compare the Skeletal tree with the appropriate behavioural trees. The Skeletal tree (Fig. 2a) is identical to one of the pruned behavioural trees, thus containing all of the same triplets and so the SDt score is 0. The other pruned behavioural tree, however, differs from the Skeletal tree for the triplet containing *Anhinga*, *Sula* and *Phalacrocorax*. In the Skeletal tree *Anhinga* and *Phalacrocorax* are sister taxa within that triplet, while for the behavioural tree *Sula* and *Phalacrocorax* are sister taxa. For all of the other possible triplets the Skeletal and behavioural trees are identical. As there is only one triplet where the trees differ the SDt is 1; if the trees had differed in two triplets the SDt would have been 2 and so on. Thus, the comparison between the Skeletal tree and the behavioural trees gives SDts of 0 and 1 with a mean of 0.5 (see Table I).

To test if the observed degree of congruence between the trees could be due to chance, the value obtained (e.g. 0.5) can then be compared with the measure of dissimilarity between the behavioural trees and a set of randomly generated trees. We used the 'random trees' command and the 'labelled dendrogram' option on COMPONENT 2.0 (Page 1992) to produce 1000 random trees for each of the comparisons. This option generates dichotomous Markovian trees that assume a constant rate of speciation and no

extinction (Savage 1983; Page 1991), a commonly used null phylogenetic model (Losos & Adler 1995). It was necessary to generate different sets of 1000 random trees for the different comparisons as they had different numbers of taxa in common.

Our analysis revealed that the behavioural trees are more congruent with the other trees than would be expected by chance alone. The range reported (in Table I) for the comparisons between the behavioural trees and the independent trees represents both the values obtained (as only two comparisons were made: one independent tree versus two behavioural trees in each case). For the behavioural trees versus the random trees, however, the range is for the 2000 comparisons made (two behavioural trees versus 1000 random trees in each case). The level of overlap between the ranges of SDt values for the behavioural trees versus the independent trees and the behavioural trees versus the random trees varied from 0.00 to 0.75%.

Homology Testing

To investigate the homology of specific behavioural characters they should be mapped onto the tree that provides the best independent estimate of the pelecaniforms' phylogeny. To construct this best-estimate tree we combined the information from the available independently derived trees. Owing to differences in the types of data and levels of taxa we were unable to combine the data and generate a best estimate tree. At the generic level there is some disagreement between the Skeletal, DNA/DNA Fitch and the DNA/DNA UPGMA trees. Where there were points of disagreement we followed the structure of the Skeletal tree because it was constructed phylogenetically. The Skeletal tree is likely to be more reliable than the UPGMA derived tree as UPGMA generally fares worst when different methods of phylogenetic reconstruction are compared (see Charleston et al. 1994; Hillis et al. 1994). The Skeletal and DNA/DNA Fitch trees differed only in the relative positions of the frigatebirds and pelicans which would not affect the analyses greatly. We used information from the independent trees where the relationships were resolved in one of the trees but not the others. For example, for the cormorant and shag species we used the Osteological tree while for the species of gannets and boobies we used the DNA/DNA UPGMA tree as they were the only trees to provide information about these taxa. None of the independent trees provides information about the relative positions of the gannets or two of the booby species (*Morus*, *S. sula* and *S. variegata*). The positions of these taxa remained as they had been in the consensus behavioural tree. Figure 3 shows the best-estimate tree.

The best-estimate tree differs from the majority rule consensus behavioural tree in the positioning of the cormorant and shag species. When the behavioural data are mapped onto the best-estimate tree the resulting treelength is five steps longer than the majority rule consensus behaviour tree. Of the 37 behavioural characters, 21 (characters 3–5, 9–11, 13, 16–19, 24–26, 28–33, 35) fitted the best-estimate tree perfectly (RI = 1.0), while just five (1, 20, 22, 36, 37) did not fit at all (RI = 0.0).

While it is possible to use the RI of a character as an indication of how well it fits the trees, there are some instances where it may provide misleading information about the homology of a character. Bowing (character 2), for example, has an RI of 0.25 (see Table II) which would appear to indicate that it may not be an homologous character. If bowing is mapped onto the best-estimate tree, however, it can be seen that the low RI is due to three losses rather than multiple gains (Fig. 4a). The evolution of bowing, therefore, does appear to be an homologous character even though it has been lost three times. It is the absence of bowing that is not homologous; the RI shows that the character state is not homologous, not the character. Mapping the characters onto the best-estimate tree also allows the relative numbers of gains and losses required for each character to be calculated. Table II shows both the combination of gains and losses that require the fewest number of evolutionary events (when gains and losses are equally likely), and the number of losses required if the character had evolved only once. By mapping character 6 (vocalize during male advertising) onto the best-estimate tree it is possible to identify alternative explanations of the evolution of this character. In this example, the presence of the character may have evolved either twice, or once with five subsequent losses (Fig. 4b). When mapped onto the best-estimate tree it becomes apparent that 15 of the characters may have evolved more than once (5–8, 12, 14, 15, 19, 20, 22, 23, 27, 34, 36, 37). Some of these characters are discussed in more detail below. For the purposes of calculating the numbers of gains and

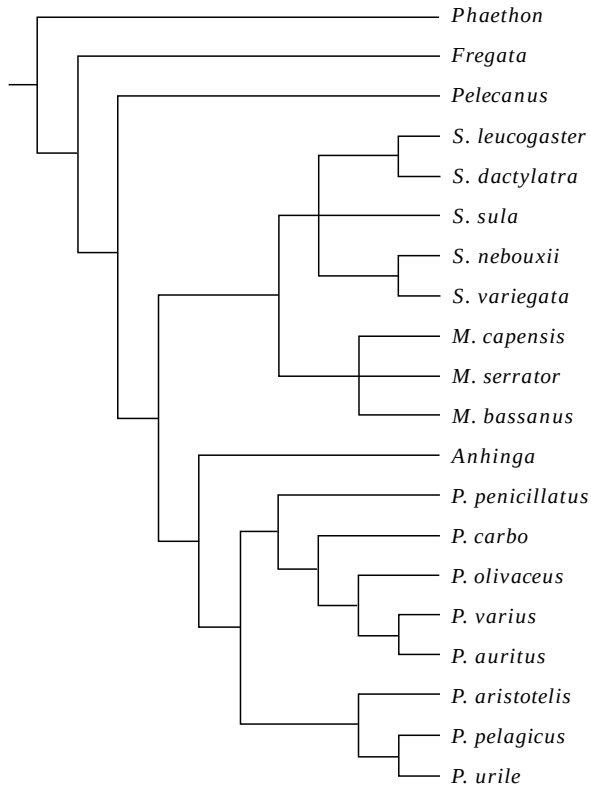

Figure 3. The best-estimate behavioural tree. The topology of this tree was constrained to reflect the topology of the independently derived trees.

losses required to explain the distribution of each character we recoded the multi-state characters as present/absent. This presence/absence coding provides information about the gains and losses of the character itself rather than of states within that character.

Evaluating the Origin of Displays

In his paper, van Tets (1965) suggested how some of the different displays may have evolved. He presented hypotheses for the derivatives of the pre-take-off, threat and nest-building displays (van Tets 1965, Figures 41–43). It is possible to evaluate such assumptions about character evolution by constructing a character state tree. A character state tree depicts graphically how transformations between the states are constrained to a tree-like graph (Maddison & Maddison 1992). For example, van Tets' (1965) phylogenetic tree for the pre-take-off derivatives follows a progression from a general intentional movement towards flying away, to the pre-take-off display of the gannets, to the sky-pointing display of the boobies which then evolves into both the wing-waving display of the darter (which waves alternate wings, see character 5) and the wing-waving display of the great, double-crested, Neotropic and pied cormorants (slow rate wing-waving of character 5). Slow rate wing-waving then, hypothetically, evolved into both rapid flutter wing-waving in Brandt's cormorant, and the pelagic and red-faced shags and into throwback in the European shag (van Tets 1965, Figure 41; Fig. 5). van Tets' hypothesis also suggests the pre-take-off displays of the boobies, darters, cormorants and shags are derived from intention movements towards flying away. This hypothesis can be coded as a character state tree using the type editor of MacClade 3.0 (Maddison & Maddison 1992) as shown in Fig. 6a.

Table II. Individual CIs and RIs for the behavioural characters mapped onto the best-estimated tree

Character	CI	RI	ARI	(a) minimum		(b) Minimum losses
				Gains	Losses	
1	1.00	0.00		1	0	0
				0	1	
2	0.25	0.25		1	3	3
5	1.00	1.00	0.80	1	1	1
				2	0	
6	0.50	0.86		2	0	5
7	0.33	0.33		1	2	2
				2	1	
				3	0	
8*	0.60	0.71	0.80	1	1	1
				2	0	
12*	0.75	0.75	1.00	1	0	0
14*	0.75	0.67	1.00	1	0	0
15	0.50	0.80		1	1	1
				2	0	
20	0.50	0.00		1	1	1
				2	0	
21	0.50	0.50		1	1	1
22	0.50	0.00		2	0	2
23	0.20	0.20		5	0	7
27*	0.75	0.50	0.75	1	1	1
				2	0	
34*†	0.67	0.50	—			
35†	1.00	1.00	—			
36	0.25	0.00		4	0	4
				3	1	
37	0.50	0.00		2	0	2

The adjusted RIs (ARI) are for those multi-state characters that differed from their initial RI when they were coded as binary characters (i.e. presence/absence). The number of times each behavioural character was gained and lost is shown for the assumption that (a) gains are as likely as losses, and the minimum number of times a behavioural character was lost is shown for the assumption that (b) each character arose only once. Those characters not present in the table had CIs and RIs of 1.00, one gain for (a) and no losses for either (a) or (b). For some of the characters alternative values of gains and losses (with an equal total of gains and losses) are given for (a).
*These multi-state characters have convergent character states.
†These multi-state characters could not be coded as present/absent.

A character state tree can be used to constrain the order of evolution of character states as van Tets (1965) hypothesized. For example, van Tets suggested that if the display sky-pointing is to evolve into rapid flutter wing-waving it must first evolve into slow rate wing-waving, and hence it takes two steps to evolve from sky-pointing to rapid flutter wing-waving. The character states can be mapped onto the best-estimate tree to see how well the hypothesis about their evolution fits. (To do so polytomies in the best-estimate tree had to be arbitrarily resolved because user-defined character states cannot be implemented by MacClade when they are present. This procedure made no difference to the results as the polytomies are within the gannets and the boobies, each of which have their own, separate states for the character.) Using the character-state tree as hypothesized by van Tets (1965) gave a length of eight steps (i.e. eight evolutionary changes). The number of steps can be compared with a null distribution for the numbers of steps generated

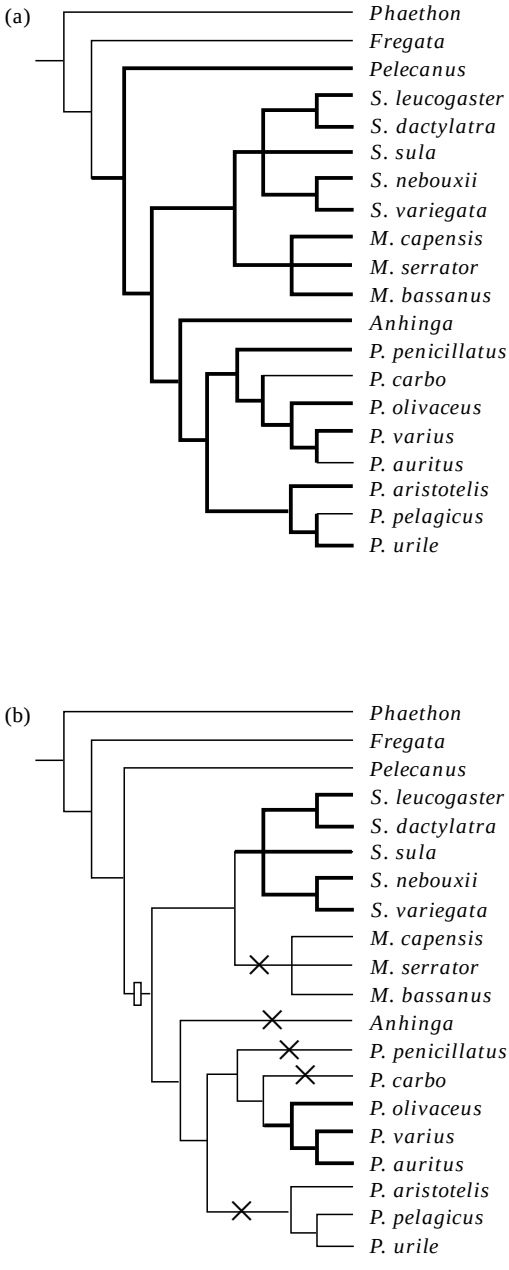

Figure 4. Two of the behavioural characters mapped onto the best-estimate tree. Character 2, bowing (a) arises once and is lost three times within the cormorants and shags (*P. carbo, P. auritus* and *P. pelagicus*). Character 6, vocalize during male advertising (b) either arises twice, or evolves once (open rectangle) and is lost a minimum of five times (crosses). The thick lines represent presence of the character while the thin lines represent absence.

from 1000 random character sets using MacClade 3.0 (Maddison & Maddison 1992; e.g. see Winkler & Sheldon 1993). MacClade effectively shuffles the states within 1000 character sets which have the same frequency of states as van Tets' hypothesis (i.e. three taxa have pre-take-off, five sky-pointing, one alternate wing-waving, four slow rate wing-waving, three rapid flutter wing-waving, and one throwback). Once the states have been randomized for the 1000 character sets, the distribution of the number of steps for the randomized character sets can be compared with the number of steps for van Tets' hypothesis. There was no overlap between the eight steps required to account for van Tets' hypothesis and the distribution of the 1000 randomized character sets (mean number of steps = 19.643, range 13–22). This comparison showed that the number of steps required for van Tets' hypothesis about the evolution of derivatives of the gannets' pre-take-off display is significantly less than would be expected by chance.

As there are seven states or displays in the character set for the pre-take-off derivatives, six is the minimum number of steps possible. Alternative character-state trees can be constructed that require only the minimum number of steps on the best estimate tree. For example, Fig. 6b shows a character state tree that fits the best-estimate tree with six steps. This example lacks the linear structure found in van Tets' hypothesis, with all but two of the displays evolving from intentional movements towards flying away (0). The two displays that are not hypothesized to have evolved from the 0 state both evolved from rapid flutter wing-waving (4). A second alternative (shown in Fig. 6c) also requires only six steps and has state 0 evolve into 1 and then 1 into 2 as suggested by van Tets, but differs from van Tets because state 0 also evolves into 4 which then evolves into both 3 and 5. In this example, state 6 could equally parsimoniously evolve from state 0 or 4. The implications of these alternatives for the origin and transformation of the pre-take-off derivatives are discussed below.

DISCUSSION

Our results strongly support both van Tets' conclusion that the social displays of the Pelecaniformes do reflect phylogeny, and his claims about pelecaniform relationships. The 12 most parsimonious trees generated by our analysis of the

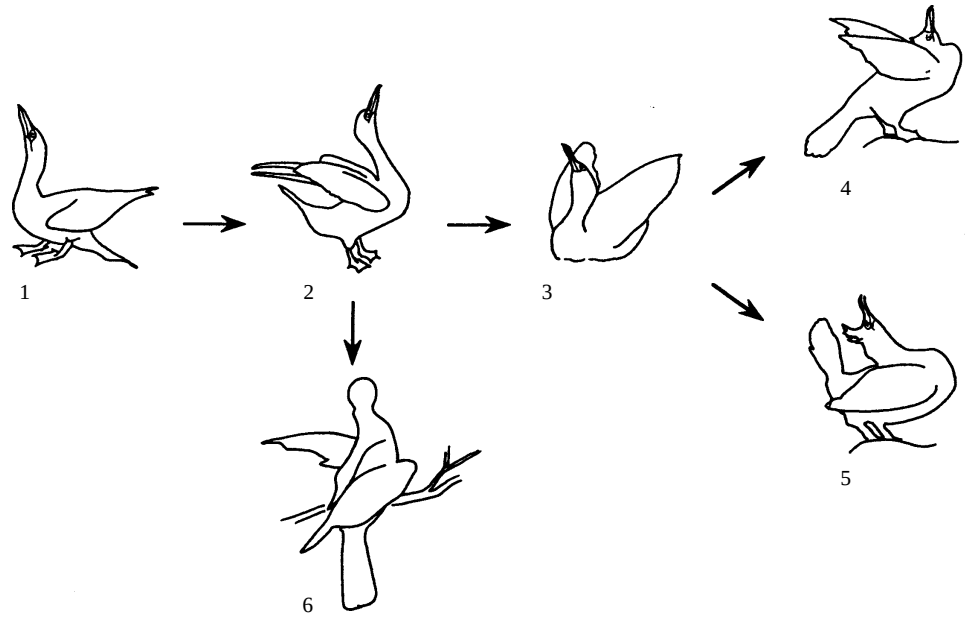

Figure 5. van Tets' hypothesis of the transformation of behavioural characters from the initial state: intention movements towards flying away; to (1) pre-take-off (Australasian gannet); to (2) sky-pointing (masked booby); to both (6) alternate wings waved (darter); and (3) slow rate wing-waving [great (black) cormorant]. Slow rate wing-waving then evolves into both (4) rapid flutter wing-waving (pelagic shag); and (5) throwback (European shag). (Redrawn from van Tets 1965.)

behavioural data had high CIs and RIs indicating that the behavioural characters fit well on these trees. Other studies (de Queiroz & Wimberger 1993; Paterson et al., in press) have found that behavioural characters are no more homoplasious than other character types. Hence, behavioural characters can be just as useful as morphological or molecular characters for the estimation of phylogeny.

For our 12 shortest trees the CI and RI are both quite high given the number of taxa involved. In a review of morphological and molecular phylogenetic studies Sanderson & Donoghue (1989) found that the CI is correlated with the number of taxa, with an increase in the number of taxa leading to a decrease in the CI. The CI we found (0.740, excluding uninformative characters) is higher than that found in any of the six studies with 20 taxa that Sanderson & Donoghue (1989) reviewed (mean = 0.54, range 0.35–0.68). Sanderson & Donoghue (1989) derived a formula to estimate the expected CI for any study from the number of taxa. For an analysis of 20 taxa this formula predicts a CI of 0.55, which suggests that our data provide a better than average fit. The CI we found can also be compared with that found in other phylogenetic studies of the pelecaniforms. Cracraft's (1985) cladistic analysis of the pelecaniforms and related taxa found a CI of 0.608 for the 13 taxa (Sanderson & Donoghue's 1989 formula predicts a CI of 0.65). Siegel-Causey's (1988) cladistic analysis of 36 species of cormorants and shags had a CI of 0.678 (Sanderson & Donoghue's 1989 formula predicts a CI of 0.384). These results suggest that while Cracraft's (1985) data fit about as well as would be expected to the trees he generated, both Siegel-Causey's (1988) and our data fit our respective trees substantially better than would be expected.

Our other results also suggest that there is phylogenetic information in the behavioural characters. The result of the PTP test revealed that the behavioural characters contained more structure than would be expected by chance. While demonstrating that there is some signal in the behavioural data, this test does not necessarily imply that signal is due to phylogeny. Instead of reflecting common evolutionary history, the signal in the

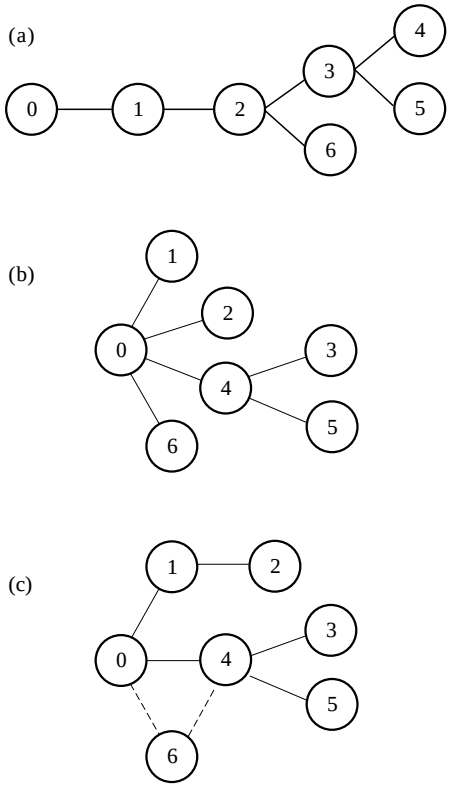

Figure 6. The transformation of behavioural characters for (a) van Tets' hypothesis, (b) and (c) two alternative hypotheses. The initial state is coded as 0, while the other behavioural characters are coded as in the legend for Fig. 5. The dashed line in (c) shows that state 6 could have evolved from either state 0 or 4 with no effect on the number of steps required.

behavioural data could be a manifestation of common ecological pressures that do not covary with phylogeny. But, the behavioural trees were more congruent with all of the independently derived trees than would be expected by chance. Thus, our results indicate that the behavioural characters do contain phylogenetic information about the pelecaniforms.

Given that we are confident that behaviour in the Pelecaniformes reflects their common evolutionary history, it is possible to investigate the evolution of their behaviour. Mapping the characters onto our best-estimate tree showed that up to 15 of the 37 behavioural characters may be convergent, assuming that evolutionary gains and losses are equally likely. If, however, we accept that for complex behavioural displays losses are more likely to occur over evolutionary time than gains, it is possible to explain seven of the 15 characters by either one convergence event (i.e. two gains) or by one gain and one or two losses (5, 7, 15, 20, 22, 27, 37). If losses are at least twice as likely as gains are to occur, then these behavioural characters would also appear to be homologous. The presence of four of the remaining eight characters can be explained by only one gain, but these multi-state characters have convergence events in their states (8, 12, 14, 34). While it is likely that complex social displays will not evolve independently many times, once a display has evolved there are constraints on the number of states that such a display may have. It is, therefore, plausible that displays with multiple forms would evolve once and then, within the different forms they may take, that there could be character-state convergence. Thus the presence of the display is homologous, but the displays have convergent states. For example, character 8 (male advertising, wing position) has the states, absent, wings folded, wings spread and wings not raised. The display can be explained by a single gain, but within its states the brown booby, *S. leucogaster*, and the European shag, *P. aristotelis*, are the only taxa to share 'wings not raised', which indicates a character-state convergence.

Three characters have either arisen at least twice or, if they had evolved only once, require four (36), five (6) and seven (23) losses to explain their distribution on the tree. From these findings it is possible to infer that these three characters are non-homologous as they have arisen on more than one occasion. According to our current data set: 'nest-worrying' (36) may have been gained independently by the great and double-crested cormorants, and the pelagic and European shags. Similarly, 'vocalization during male advertising' (6) may have arisen independently in the boobies, and the Neotropic, pied and double-crested cormorants; and 'pre-take-off vocalization' (23) may have evolved independently in the Cape and Australasian gannets and the great and double-crested cormorants and the pelagic and European shags. Nevertheless, it is remarkable that by far the majority of the 37 displays are homologous.

In the case of 'nest-worrying' (36) the information came from detail in van Tets (1965) about the presence of the trait in the four cormorants and shags in which it was known to occur. It is possible that, owing to lack of information, the absence of

this trait in the other species should not have been assumed and nest-worrying may actually be more widespread than our analysis of van Tets' data have shown. An investigation by Johnsgard (1993) suggests that nest-worrying is present in most or all species of cormorants and shags and possibly in both species of darter. With this additional information, nest-worrying can be considered homologous. Similarly with the 'pre-take-off vocalization character' (23), information about the presence of the character was taken from the text but it is possible that it is more widespread than van Tets had known at the time. The trait 'vocalization during male advertising' is present in the boobies and some of the cormorants while it was known to be absent in all of the other cormorants and shags. Relatively simple displays, such as vocalization displays (6 and 23), may convergently evolve more easily than the more complex displays.

Given that the majority of the displays appear to be homologous, it is possible to address the issue of missing data within the matrix. That is, it is possible to make predictions about the behaviour of a taxon depending on its position on the tree and the distribution of the behavioural characters on the tree. For example, the presence or absence of gaping (character 11) in the pied cormorant was unknown to van Tets (1965). If gaping is mapped onto the best-estimate tree the most parsimonious inference is that it evolved in the common ancestor of the cormorants and shags (since all the other cormorants and shags exhibit gaping). The position of the pied cormorant on the best-estimate tree thus suggests that this species probably does have this display. A survey of more recent literature confirms that the pied cormorant does gape (Marchant & Higgins 1990; Johnsgard 1993). In this case, the presence of gaping in the pied cormorant then leads to questions relating to the head position during gape, whether gaping is repetitive and the form of vocalization during gaping. By mapping these characters onto the best-estimate tree it is possible to predict that a pied cormorant should have its neck stretched upwards during the gape and it should not gape repetitively. Applying the principle of parsimony does not, however, allow a conclusive prediction about vocalization during gape. It is possible (equally parsimonious) that the male and female are similar or different in their vocalizations during gape, although the tree does suggest that it will not be only the male that vocalizes during the gape. This approach offers the opportunity to make predictions about the range of behavioural characters that may be found in a particular species. If, for example, the behaviour of the pied cormorant was poorly known, predictions such as the one about gaping could help direct studies of the species' behaviour by alerting researchers to which displays are likely and unlikely to be present.

It is also possible to evaluate hypotheses about the origin of the displays by mapping the distribution of the behavioural characters onto the best-estimate tree. van Tets' (1965) hypothesis about the transformation of the pre-take-off derivatives is more parsimonious than would be expected by chance. There are, however, alternative character state trees that account for the distribution of the displays in six rather than the eight evolutionary steps required for van Tets' hypothesis (Fig. 6b, c). These hypotheses, however, may not make biological sense. By evaluating the likely changes required for each of the character state trees it is possible to make judgments about the likelihood of the alternative hypotheses. For van Tets' hypothesis, possible positions for the steps are shown on Fig. 7(A–F). For this hypothesis, state 0 needs to evolve into state 1 and 1 to 2 (at A), state 2 needs to be lost (at B) and to evolve into states 6 (at C) and 3 (at D), state 3 then needs to evolve to state 4 (at both Es) and to state 5 (at F). Thus eight steps are taken with a loss (at B) and a convergence event (the two Es) causing the additional two steps required above the minimum six. State 1, the pre-take-off display, is maintained in some form by all of the species, while for the other states (which are male advertising displays) the change from one state to another is a replacement of the previous state.

For the first alternative hypothesis (Fig. 6b), state 0 evolves into states 2 (at I), 1 (at II), 6 (at III), and 4 (at IV), while state 4 evolves into both states 3 (at V) and 5 (at VI), thus taking six steps. The second alternative hypothesis (Fig. 6c), requires state 0 to evolve into state 1 (at i), state 1 to evolve into state 2 (at ii), state 0 must also evolve into state 4 (at either of the ivs, it is not possible to determine which one), while state 4 evolves into both states 3 (at iii) and 5 (at v). In this case, state 6 can evolve (at vi) from either state 0 or 4 depending on where state 4 evolved.

While both alternatives appear more parsimonious than van Tets' hypothesis they are not

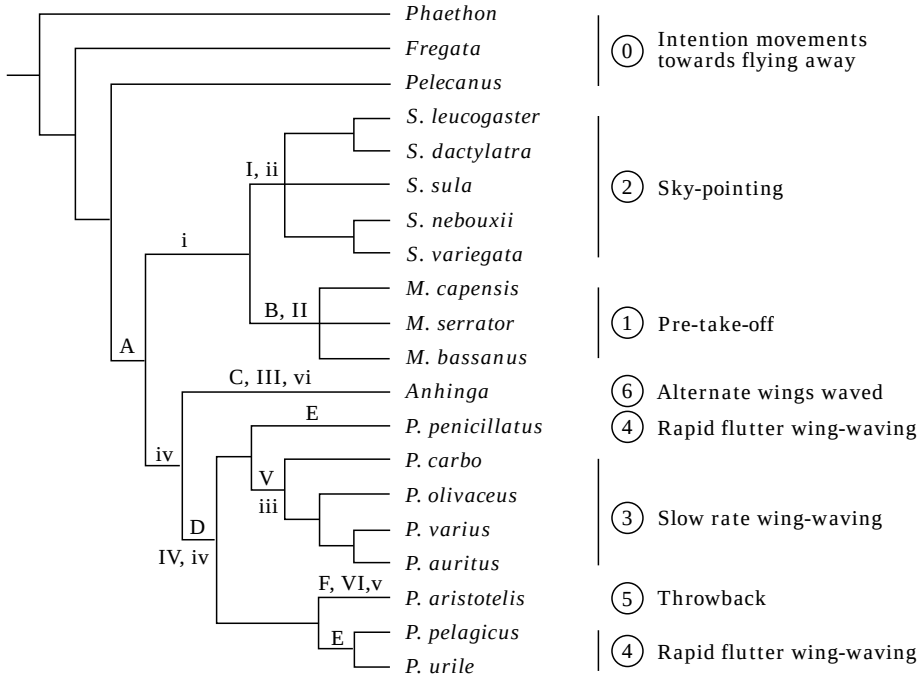

Figure 7. The distribution of the behavioural characters on the best-estimate tree and the points at which changes can be postulated to explain the distribution of the behavioural characters assuming the transformations of van Tets (Figs 5 and 6a) shown as letters A–F; the first alternative (Fig. 6b) shown as I–VI; and the second alternative (Fig. 6c) shown as numerals i–vi.

necessarily more plausible explanations for the pre-take-off derivatives. For the first alternative it seems highly unlikely that the intention movements towards flying away would evolve into all of: the pre-take-off display in the gannets, the sky-pointing display in the boobies, the wing-waving display in the darters and the different wing-waving displays of the cormorants and shags. Several major changes in the form and function of the displays appear far less likely to occur than the relatively small changes required for van Tets' transformations. For this hypothesis to be more plausible it would require the addition of the different pre-take-off displays as states to precede the changes to states 2, 4 and 6 from state 0 (i.e. 0 goes to the booby pre-take-off display before going to 2, 0 goes to the cormorant and shag pre-take-off display before going to 4, and 0 goes to the darter pre-take-off display before going to 6). Thus, this alternative is in fact a less parsimonious explanation than van Tets' hypothesis as it would require nine steps with the three additional character states. The second alternative allows that the pre-take-off display of the gannets evolved into the sky-pointing display of the boobies, but not that sky-pointing then evolved into the different wing-waving displays. Thus, the wing-waving displays must have also evolved (either once or twice) from the initial state of intention movements towards flying away. Again, it is more plausible for the sky-pointing and wing-waving displays to evolve separately from the initial state if wing-waving is preceded by the evolution of the cormorants and shags' pre-take-off display (and possibly if the darters' pre-take-off display evolves between state 0 and 6 if state 6 did indeed evolve directly from the initial state rather than from state 4). Both of the alternatives require large evolutionary changes and the intention to fly becoming different male advertising displays with very similar form independently in both the boobies and the darters, cormorants and shags. If the large steps are not considered plausible then additional intermediary steps are required which increase the number of steps required to explain the distribution of the displays. Thus, neither of the alternative hypotheses

appears to account for the transformation of the pre-take-off derivatives as well as van Tets'.

As van Tets' hypothesis is better than chance would predict and provides a more plausible explanation for the transformation of the pre-take-off derivatives, it is useful to investigate why it requires additional steps. The loss of sky-pointing as a male advertising display in the gannets is easy to accept, but the fact that it is more parsimonious to believe that sky-pointing evolved into rapid flutter wing-waving rather than slow-rate wing-waving is difficult to explain. The boobies make single wing movements when sky-pointing, and the darters, cormorants and shags repetitively wing-wave. Given the change from single wing movements to repetitive wing-waving, it seems more plausible that slow rate wing-waving would evolve before rapid flutter wing-waving. If the best-estimate tree had the Brandt's cormorant as a sister taxa to the pelagic and red-faced shags (i.e. if the taxa that share rapid flutter wing-waving were grouped together) the most parsimonious arrangement would be that slow-rate wing-waving preceded rapid flutter wing-waving. There is, in fact, some reason to question the position of Brandt's cormorant on the best-estimate tree. This species is the one major point of disagreement on the best-estimate tree between the behaviourally based taxonomy of van Tets (1976) and the morphologically based tree of Siegel-Causey (1988). It is possible that the structure imposed by using Siegel-Causey's (1988) tree to constrain the behavioural tree is not a true reflection of the evolutionary history of the group. The phylogenetic position of Brandt's cormorant thus needs corroboration from other sources.

Our study shows that behaviour does reflect phylogeny in the Pelecaniformes. The displays used provide significant signal and that signal is congruent with independent evolutionary trees. This congruence suggests that the signal in the behavioural data set is due to common evolutionary history. Why do the behavioural characters appear to be phylogenetically informative? To be informative they need to persist over millions of years. For example, fossil evidence suggests that the darters and cormorants and shags diverged at least 30 million years ago (Becker 1986). This date suggests that displays such as kink-throating and hopping have existed for over 30 million years. Lehrman (1953) pointed out that the phylogenetic persistence of a behavioural character does not indicate anything about its ontogenetic origins. The fact that the kink-throating and hop displays have persisted for over 30 million years does not mean that they develop without experiential input. All that a character requires to be phylogenetically informative is that the conditions for its development reliably reoccur down the lineage (Gray 1992; Griffiths & Gray 1994). As well as genetic inheritance, many ecological and social factors reliably reoccur down lineages (e.g. colonial nesting, foraging strategies, the form of chick feeding). Just as species-typical behaviour is the joint interactive product of species-typical experience and species-typical genetic inheritance (Gottlieb 1976; Johnston & Gottlieb 1990), so are the behavioural traces of deeper phylogenetic history.

APPENDIX 1

Behavioural characters used in the analysis. Unless otherwise stated the information comes from van Tets (1965). Some characters had additional information added (from van Tets 1976) to resolve the character state for one or more of the taxa. The page and or figure numbers referring to where the information originated from in van Tets (1965, except where stated) are shown in parentheses.

1. Presentation of nest-material: (0) absent; (1) present. (Page 42 and Figure 47.)
2. Bowing: (0) absent; (1) present. (Page 50 and Figure 47. Information on page 50 leads to the absence coding of *P. carbo*, *P. auritus* and *P. pelagicus* counter to Figure 47.)
3. Hop: (0) absent; (1) present. (Page 25 and Figure 47.)
4. Pre-hop, neck arched or not: (0) absent; (1) arched; (2) not arched. (Page 25 and Figure 49.)
5. Male advertising: (0) absent; (1) sky-pointing; (2) alternate wings waved; (3) slow rate wing-waving; (4) rapid flutter wing-waving; (5) darting and throwback. (Pp. 28–30, Figure 47 and pp. 122–123 of van Tets 1976.)
6. Vocalize during male advertising: (0) absent; (1) present. (Page 28 and Figure 49.)
7. Rump patch flash during male advertising: (0) absent; (1) present. (Page 30 and Figure 49.)
8. Male advertising, wing position (i.e. the position of primary with respect to the secondary feathers): (0) absent; (1) wings folded; (2) wings spread; (3) wings not raised. (Pp. 29–30 and Figure 48.)

9. Kink-throating: (0) absent; (1) present. (Pp. 58–59 and Figure 47.)
10. Head-wagging: (0) absent; (1) present. (Pp. 54–58 and Figure 48. *S. nebouxii* scored as present from the figure but not mentioned in the text.)
11. Gaping: (0) absent; (1) present. (Page 39 and in Figure 49.)
12. Head position during gape: (0) absent; (1) neck stretched upwards; (2) head swung back; (3) head swung forward. (Page 39 and Figure 49.)
13. Repetitive gaping: (0) absent; (1) present. (Page 39 and Figure 49.)
14. Vocalize during gape: (0) absent; (1) male and female similar; (2) male and female differ; (3) male only. (Page 39.)
15. Pointing: (0) absent; (1) present. (Pp. 40–41 and Figures 47 and 49.)
16. Pre-take-off body twisted: (0) absent; (1) present. (Page 22 and Figure 49.)
17. Pre-take-off sky-pointing: (0) absent; (1) present. (Figure 48.)
18. Pre-take-off wings folded (i.e. primary feathers folded behind the secondary feathers): (0) absent; (1) present. (Page 19 and Figure 48.)
19. Pre-take-off bill upward or forward: (0) absent; (1) upward; (2) forward. (Pp. 19–20 and Figures 48.)
20. Pre-take-off median crest-raising: (0) absent; (1) present. (Page 20 and Figure 49.)
21. Pre-take-off head expand and raise bill: (0) absent; (1) present. (Page 20.)
22. Pre-take-off head raise: (0) absent; (1) present. (Page 21.)
23. Vocalize prior to take off: (0) absent; (1) present. (Page 22.)
24. Post landing bill downwards: (0) absent; (1) present. (Page 22 and Figure 48.)
25. Post landing head-biting: (0) absent; (1) present. (Page 22.)
26. Post landing back-pawing: (0) absent; (1) present. (Page 22.)
27. Post landing call during display: (0) absent; (1) males only; (2) males and females identical; (3) males and females differ. (Page 25.)
28. Gargle-threat: (0) absent; (1) present. (Page 34 and Figure 49.)
29. Mutual wing-bowing: (0) absent; (1) present. (Page 45 and Figure 48.)
30. Quiver-bowing: (0) absent; (1) present. (Page 46.)
31. Mutual head wagging: (0) absent; (1) present. (Page 54 and Figure 48.)
32. Bill-up-face-away: (0) absent; (1) present. (Page 26 and Figure 48.)
33. Stepping: (0) absent; (1) present. (Page 26 and Figure 48.)
34. Locomotion on land: (0) irregularly walk with wings assisting; (1) high stepping gait or waddle with wings often spread for balance; (2) high stepping gait or waddle. (Pp. 16–17.)
35. Threat: (0) pecking and screaming: (1) pointing or wagging bills; (2) move towards intruder; (3) snapping and waving open bills; (4) raise, open and wave bills; (5) S-shaped expanded neck, open bill; (6) head repeatedly thrown forward; (7) repeated head darting followed by head shake and gargling sound. (Pp. 33–34.)
36. Nest-worrying: (0) absent; (1) present. (Page 50.)
37. Water begging: (0) absent; (1) present. (Page 52.)

APPENDIX 2

Data matrix of the behavioural characters for all of the taxa. Missing data are indicated by a '?'

Character	1 2 3 12345678901234567890123456789012345 67
Phaethon	00000000000000000000000000000000000000
Fregata	10000000000000000000000000000000000300
Pelecanus	11000000000000000000000000000000001400
M. capensis	11100000010000001010001?1000101002100
M. serrator	11100000010000001110001110001010021 00
M. bassanus	11100000010000001110000110001010021 00
S. leucogaster	11101103010000000020000100000101021 00
S. dactylatra	11101101010000000020000100000101021 00
S. sula	11101101010000000020000100000000021 00
S. nebouxii	11101102010000000?0000?0000000012100
S. variegata	11101102010000000?0000?0000000012100
Anhinga	11102001100000100000000000000000001200
P. varius	111?310110????100000000001?0000002500
P. auritus	10123101101301100000101001200000025 11
P. olivaceus	11123101101 30?100001100001?0000002500
P. carbo	10123011101102100001101001300000025 11
P. penicillatus	11114011101201100000010001000000026 00
P. pelagicus	10114011101212010000001001210000027 10
P. urile	11114011101212010000000001210000027 00
P. aristotelis	11115003101103000000011001100000026 10

ACKNOWLEDGMENTS

We thank Allen Rodrigo for kindly supplying a copy of the PTP test program, Adrian Paterson for discussion and comments, Tim Guilford and two anonymous referees for comments, and Ken Miller for drawing Fig. 5. This work was supported by grants from the Division of Sciences, University of Otago and the N.Z. Lotteries Board, Science Research Committee (to R.D.G. and H.G.S.). M.K. is supported by a University of Otago Postgraduate Scholarship. This paper is dedicated to the late G. F. van Tets.

REFERENCES

Arntzen, J. W. & Sparreboom, M. 1989. A phylogeny for the old world newts, genus *Triturus*: biochemical and behavioural data. *J. Zool., Lond.*, **219,** 645–664.

Aronson, L. R. 1981. Evolution of telencephalic function in lower vertebrates. In: *Brain Mechanisms of Behaviour in Lower Vertebrates* (Ed. by L. R. Aronson, E. Tabach, D. S. Lehrman & J. S. Rosenblatt), pp. 33–58. Cambridge: Cambridge University Press.

Atz, J. W. 1970. The application of the idea of homology to behaviour. In: *Development and Evolution of Behaviour: Essays in Memory of T. C. Schneirla* (Ed. by L. R. Aronson, E. Tobach, D. S. Lehrman & J. S. Rosenblatt), pp. 53–74. San Francisco: W. H. Freeman.

Becker, J. J. 1986. Reidentification of *Phalacrocorax subvolans* Brodkorb as the earliest record of Anhingidae. *Auk*, **103,** 804–808.

Beehler, B. M. & Swaby, R. J. 1991. Phylogeny and biogeography of the *Ptiloris* riflebirds (Aves: Paradisaeidae). *Condor*, **43,** 738–745.

Brooks, D. R. & McLennan, D. A. 1991. *Phylogeny, Ecology and Behavior: a Research Program in Comparative Biology*. Chicago: University of Chicago Press.

Browning, M. R. 1989. The correct name of the olivaceous cormorant, 'maigue' of Piso (1658). *Wilson Bull.*, **101,** 101–106.

Charleston, M. A., Hendy, M. D. & Penny, D. 1994. The effects of sequence length, tree topology, and number of taxa on the performance of phylogenetic methods. *J. computl Biol.*, **1,** 133–151.

Clayton, D. H. & Harvey, P. H. 1993. Hanging nests on a phylogenetic tree. *Curr. Biol.*, **3,** 882–883.

Coddington, J. A. 1988. Cladistic tests of adaptational hypotheses. *Cladistics*, **4,** 3–22.

Cracraft, J. 1985. Monophyly and the phylogenetic relationships of the Pelecaniformes: a numerical cladistic analysis. *Auk*, **102,** 834–853.

Cracraft, J. 1994. Species diversity, biogeography, and the evolution of biotas. *Am. Zool.*, **34,** 33–47.

Crowe, T. M., Harley, E. H., Jakutowicz, M. B., Komen, J. & Crowe, A. A. 1992. Phylogenetic, taxonomic and biogeographical implications of genetic, morphological, and behavioural variation in francolins (Phasianidae: *Francolinus*). *Auk*, **109,** 24–42.

Faith, D. P. 1991. Cladistic tests for monophyly and nonmonophyly. *Syst. Zool.*, **40,** 366–375.

Faith, D. P. & Cranston, P. S. 1991. Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. *Cladistics*, **7,** 1–28.

Farris, J. S. 1989. The retention index and the rescaled consistency index. *Cladistics*, **5,** 417–419.

Fitch, W. M. & Margoliash, E. 1967. Construction of phylogenetic trees. *Science*, **155,** 279–284.

Gottlieb, G. 1976. Conceptions of prenatal development: behavioural embryology. *Psychol. Rev.*, **83,** 215–234.

Gray, R. D. 1992. Death of the gene: developmental systems strike back. In: *Trees of Life* (Ed. by P. E. Griffiths), pp. 165–209. Dordrecht: Kluwer.

Griffiths, P. E. & Gray, R. D. 1994. Developmental systems and evolutionary explanation. *J. Phil.*, **91,** 277–304.

Hafner, M. S. & Nadler, S. A. 1988. Phylogenetic trees support the coevolution of parasites and their hosts. *Nature, Lond.*, **332,** 258–259.

Hafner, M. S., Sudman, P. D., Villablanca, F. X., Spradling, T. A., Demastes, J. W. & Nadler, S. A. 1994. Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science*, **265,** 1087–1090.

Hillis, D. M., Huelsenbeck, J. P. & Cunningham, C. W. 1994. Application and accuracy of molecular phylogenies. *Science*, **264,** 671–677.

Johnsgard, P. A. 1993. *Cormorants, Darters and Pelicans of the World*. Washington: Smithsonian Institution Press.

Johnston, T. D. & Gottlieb, G. 1990. Neophenogenesis: a developmental theory of phenotypic evolution. *J. theor. Biol.*, **147,** 471–495.

Kluge, A. G. & Farris, J. S. 1969. Quantitative phyletics and the evolution of the anurans. *Syst. Zool.*, **18,** 1–32.

Langtimm, C. A. & Dewsbury, D. A. 1991. Phylogeny and evolution of rodent copulatory behaviour. *Anim. Behav.*, **41,** 217–225.

Lanham, U. N. 1947. Notes on the phylogeny of the Pelecaniformes. *Auk*, **64,** 65–70.

Lehrman, D. S. 1953. A critique of Konrad Lorenz's theory of instinctive behaviour. *Q. Rev. Biol.*, **28,** 337–363.

Lorenz, K. Z. 1941. Vergleichende Bewegungsstudien an Anatinen. *J. Ornithol.*, **89,** 194–294.

Losos, J. B. & Adler, F. R. 1995. Stumped by trees? A generalized null model for patterns of organismal diversity. *Am. Nat.*, **145,** 329–342.

McKitrick, M. C. 1992. Phylogenetic analysis of avian parental care. *Auk*, **109,** 828–846.

McLennan, D. A., Brooks, D. R. & McPhail, J. D. 1988. The benefits of communication between comparative

ethology and phylogenetic systematics: a case study using gasterosteid fishes. *Can. J. Zool.*, **66,** 2177–2190.

Maddison, D. R., Ruvolo, M. & Swofford, D. L. 1992. Geographical origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. *Syst. Biol.*, **41,** 111–124.

Maddison, W. P. & Maddison, D. R. 1992. *MacClade: Analysis of Phylogeny and Character Evolution. Version 3.0*. Sunderland, Masssachusetts: Sinauer Associates.

Marchant, S. & Higgins, P. J. 1990. *Handbook of Australian, New Zealand and Antarctic Birds. Vol. 1. Ratites to Ducks*. Oxford: Oxford University Press.

Margush, T. & McMorris, F. 1981. Consensus *n*-trees. *Bull. math. Biol.*, **43,** 239–244.

Nelson, G. & Platnick, N. I. 1981. *Systematics and Biogeography: Cladistics and Vicariance*. New York: Columbia University Press.

Packer, L. 1991. The evolution of social behaviour and nest architecture in sweat bees of the subgenus *Evylaeus* (Hymenoptera: Halictidae): a phylogenetic approach. *Behav. Ecol. Sociobiol.*, **29,** 153–160.

Page, R. D. M. 1988. Quantitative cladistic biogeography: constructing and comparing area cladograms. *Syst. Zool.*, **37,** 254–270.

Page, R. D. M. 1990. Component analysis: a valiant failure? *Cladistics*, **6,** 119–136.

Page, R. D. M. 1991. Random dendrograms and null hypotheses in cladistic biogeography. *Syst. Zool.*, **40,** 54–63.

Page, R. D. M. 1992. *COMPONENT* User Manual (Release 2.0). London: Natural History Museum.

Page, R. D. M. 1993. Parasites, phylogeny and cospeciation. *Int. J. Parasitol.*, **23,** 499–506.

Paterson, A. M. & Gray, R. D. In press. Analytical tests of host/parasite cospeciation. In: *Coevolutionary Biology of Birds and Parasites* (Ed. by D. H. Clayton & J. Moore). Oxford: Oxford University Press.

Paterson, A. M., Gray, R. D. & Wallis, G. P. 1993. Parasites, petrels and penguins: does louse phylogeny reflect seabird phylogeny? *Int. J. Parasitol.*, **23,** 515–526.

Paterson, A. M., Gray, R. D. & Wallis, G. P. In press. Penguins, petrels and parsimony: does cladistic analysis of behaviour reflect seabird phylogeny? *Evolution*.

Penny, D., Foulds, L. R. & Hendy, M. D. 1982. Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. *Nature, Lond.*, **297,** 197–200.

Proctor, H. C. 1992. Sensory exploitation and the evolution of male mating behaviour: a cladistic test using water mites. *Anim. Behav.*, **44,** 745–752.

Prum, R. O. 1990. Phylogenetic analysis of the evolution of display behaviour in the neotropical mannakins (Aves: Pipridae). *Ethology*, **84,** 202–231.

de Queiroz, A. & Wimberger, P. H. 1993. The usefulness of behaviour for phylogeny estimation: levels of homoplasy in behavioural and morphological characters. *Evolution*, **47,** 46–60.

Ridley, M. 1983. *The Explanation of Organic Diversity*. Oxford: Oxford University Press.

Sanderson, M. J. & Donoghue, M. J. 1989. Patterns of variation in levels of homoplasy. *Evolution*, **43,** 1781–1795.

Savage, H. M. 1983. The shape of evolution: systematic tree topology. *Biol. J. Linn. Soc.*, **20,** 225–244.

Sibley, C. G. 1960. The electrophoretic patterns of avian egg-white proteins as taxonomic characters. *Ibis*, **102,** 215–284.

Sibley, C. G. & Ahlquist, J. E. 1990. *Phylogeny and Classification of Birds: a Study in Molecular Evolution*. New Haven: Yale University Press.

Siegel-Causey, D. 1988. Phylogeny of the Phalacrocoracidae. *Condor*, **90,** 885–905.

Sillen-Tullberg, B. & Møller, A. P. 1993. The relationship between concealed ovulation and mating systems in anthropoid primates: a phylogenetic analysis. *Am. Nat.*, **141,** 1–25.

Swofford, D. L. 1993. *PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1*. Champaign: Illinois Natural History Survey.

Swofford, D. L. & Olsen, G. J. 1990. Phylogeny reconstruction. In: *Molecular Systematics* (Ed. by D. M. Hillis & C. Moritz), pp. 411–501. Sunderland, Massachusetts: Sinauer Associates.

van Tets, G. F. 1965. A comparative study of some social communication patterns in the Pelecaniformes. *Ornithol. Monogr.*, **2,** 1–88.

van Tets, G. F. 1976. Australasia and the origin of shags and cormorants, Phalacrocoracidae. *Proc. int. ornithol. Congr.*, **XVI,** 121–124.

Winkler, D. W. & Sheldon, F. W. 1993. Evolution of nest construction in swallows (Hirundinidae): a molecular phylogenetic perspective. *Proc. natn. Acad. Sci. U.S.A.*, **90,** 5705–5707.

Zink, R. M. & Avise, J. C. 1990. Patterns of mitochondrial DNA and allozyme evolution in the avian genus *Ammodramus*. *Syst. Zool.*, **39,** 148–161.