

This is the peer reviewed version of the following article Paterson, A. M., Wallis, G. P., Kennedy, M. and Gray, R. D. (2014), Behavioural evolution in penguins does not reflect phylogeny. *Cladistics*, 30: 243–259. doi: 10.1111/cla.12040, which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1111/cla.12040/full> This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for self-archiving.

1 Behavioral evolution in penguins does not reflect phylogeny

2

3

4 Adrian M. Paterson<sup>a,\*</sup>, Graham P. Wallis<sup>b</sup>, Martyn Kennedy<sup>c</sup> and Russell D. Gray<sup>d</sup>

5

6

7 <sup>a</sup> Department of Ecology, Lincoln University

8 <sup>b</sup> Department of Zoology, University of Otago

9 <sup>c</sup> Allan Wilson Centre for Molecular Ecology and Evolution, Department of Zoology,

10 University of Otago

11 <sup>d</sup> Department of Psychology, University of Auckland

12

13

14 \* Correspondence:

15 Dr Adrian Paterson

16 Department of Ecology

17 Faculty of Agriculture and Life Sciences

18 Lincoln University

19 PO Box 84, Lincoln 7647

20 New Zealand

21

22 *E-mail address*        Adrian.Paterson@lincoln.ac.nz

23 *Phone*                    +64 3 325 2811

24 *Fax*                        +64 3 325 3844

25

26 Word count: ~5900 (excluding the references onwards)

27

28 **Abstract**

29           Over the last two decades, behavioral biologists and ecologists have made effective  
30 use of the comparative method but have often stopped short of adopting an explicitly  
31 phylogenetic approach. We examined 68 behavior and life history (BLH) traits of 15 penguin  
32 species to: 1) infer penguin phylogeny, 2) assess homology of behavioral characters, and 3)  
33 evaluate hypotheses about character evolution and ancestral states. Parsimony analysis of the  
34 BLH dataset found either two shortest trees (characters coded as unordered) or a single  
35 shortest tree (characters coded as a combination of unordered and dollo). The BLH data had  
36 significant structure. Kishino-Hasegawa tests indicated that BLH trees were significantly  
37 different from most previous estimates of penguin phylogeny. The BLH phylogeny generated  
38 from dollo characters appeared to be less accurate than the tree derived from the completely  
39 unordered dataset. Dividing BLH data into display and non-display traits resulted in no  
40 significant differences in level of homoplasy and no difference in the accuracy of phylogeny.  
41 Tests for homology of BLH traits were performed by mapping the characters onto a molecular  
42 tree. Assuming that independent gains are less likely than losses of character states, 65 of the  
43 68 characters were likely to be homologous across taxa and at least several characters  
44 appeared to have been stable since the origin of modern penguins around 30 million years.  
45 Finally, the likely BLH traits of the most recent common ancestor of extant penguins were  
46 reconstructed from character states along the internal branch leading to the penguins. This  
47 analysis suggested that the ‘proto-penguin’ probably had a similar life history to current  
48 temperate penguins but few ritualised behaviors. A southern, cool- temperate origin of  
49 penguins is suggested.

50

51

52

53           Despite a "flowering in phylogenetics" (Doyle, 1993), most comparative behavioral  
54 studies have taken the 'convergence approach' (Winkler, 2000) where correlations among  
55 similar events across lineages are explored using one of several methods (e.g., Gittleman and  
56 Kot, 1990; Harvey and Pagel, 1991). Fewer behavioral studies have taken the 'homology  
57 approach' (Winkler, 2000) where unique events within lineages are explored and similar traits  
58 in different species are assumed to be homologous. On this basis Lorenz (1958) and  
59 Tinbergen (1959) claimed that behavioral characters could be used to infer evolutionary  
60 relationships (see also Hinde and Tinbergen, 1965; Brooks and McLennan, 1991). Critics of  
61 the use of behavioral characters, however, pointed out that similarities among behavioral traits  
62 were not always indicative of homology. They claimed that behavior may be more prone to  
63 convergence than other types of traits (Atz, 1970; Aronson, 1981). As Tinbergen (1959)  
64 himself stated,

65           "I suggest that it is *because of the limited number of possibilities* of doing 'the  
66           opposite' to showing preparedness to attack, that we find some curious similarities  
67           in the defensive threat and appeasement postures of widely separated species."  
68           [author's italics].

69           Other often-cited problems with investigations of behavioral evolution are the  
70 supposed instability, evanescence, lack of character independence, and small likelihood of  
71 fossilisation of behavioral characters (Atz, 1970). Thus, a major challenge for those who  
72 championed use of behavioral characters was to demonstrate the homology of behavioral  
73 characters. Only characters that have arisen once can be considered to be homologous  
74 (Wagner, 1989), that is, traits are shared due to common ancestry (see Nixon and Carpenter,  
75 2012). As suggested by Japyassu and Machado (2010), and following Brower and  
76 Schawaroch (1996), assigning behavioural homology is a three-stage process where  
77 comparable features are discovered among taxa, then character variability is partitioned into  
78 subunits. These two steps identify phylogenetic homology (Nixon and Carpenter, 2012).

79 Finally, congruence between character state distribution and a phylogenetic hypothesis, such  
80 as a cladogram, is obtained either confirming the homology or identifying homoplasy (either  
81 through convergence or observational error (Nixon and Carpenter, 2012)). Using parsimony-  
82 based phylogenetic methods, character homology can be assessed by mapping a character onto  
83 an independent tree to assess whether it has arisen more than once (Brooks and McLennan,  
84 1991). The phylogenetic content of a behavioral dataset can also be evaluated by comparing  
85 trees based on behavior with independently derived trees.

86 In a renaissance of phylogenetic behavioral studies lead by Brooks and McLennan  
87 (1991), many studies concentrated on testing hypotheses of behavioral homology by mapping  
88 behaviors onto a genetic or morphological tree (e.g., Beehler and Swaby, 1991; Langtimm and  
89 Dewsbury, 1991; Packer, 1991; McKittrick, 1992; Sillen-Tullberg and Møller, 1993; Winkler  
90 and Sheldon, 1993; Price and Lanyon, 2002). Other studies used behavioral characters, or a  
91 mixture of behavioral and other characters, to reconstruct the phylogenetic relationships of  
92 organisms (e.g., Cracraft, 1985; McLennan et al., 1988; Arntzen and Sparreboom, 1989;  
93 Prum, 1990; Crowe et al., 1992; Proctor, 1992; Paterson et al., 1995; Kennedy et al., 1996;  
94 Slikas, 1998; Stuart and Hunter, 1998; McLennan and Mattern, 2001; Stuart and Currie, 2001;  
95 Noll, 2002; Mattern and McLennan, 2004; Robillard et al., 2006). These studies have  
96 generally shown relatively low levels of convergence (e.g., Prum, 1990; de Queiroz and  
97 Wimberger, 1993; Paterson et al., 1995; Kennedy et al., 1996; Slikas, 1998; McLennan and  
98 Mattern, 2001; Stuart and Currie, 2001), and high levels of congruence between trees derived  
99 from behavior and different character types (e.g., McLennan et al., 1988; Arntzen and  
100 Sparreboom, 1989; Prum, 1990; Paterson et al., 1995; Kennedy et al., 1996; Slikas, 1998;  
101 Noll, 2002; Stuart et al., 2002). Highly predictable, and often stereotypical, behaviors and life  
102 history traits present in all appropriate individuals are assumed to be inherited and, if found in  
103 different species, are potential homologues (Hapyassu and Machado, 2010). De Queiroz and  
104 Wimberger (1993) examined multiple datasets and demonstrated that there was no more

105 homoplasy in behavioral traits than morphological traits. Similarly, in a study examining  
106 multiple datasets for seabird species, Paterson et al. (1995) showed that retention indices of  
107 the molecular datasets were not significantly higher than that of behavioral data. Several  
108 techniques may assess whether congruence between behavioral and other trees is significant,  
109 from the strictly topological based (e.g., Page, 1990; Page, 1992) to those that examine the  
110 underlying uncertainty of the data generating the tree (e.g., Kishino and Hasegawa, 1989).

111 Here, we use behavior and life history (BLH) characters to derive a phylogeny for  
112 penguins (Order Sphenisciformes) and to assess the accuracy of this phylogeny by comparing  
113 with those constructed from other data types. Penguins are a homogenous group of seabirds  
114 that became extremely modified for swimming with the complete loss of aerial flight (Bertelli  
115 and Giannini, 2005). There has never been any doubt over monophyly of sphenisciforms,  
116 which are very distinct from all possible sister groups with an origin in excess of 60 million  
117 years ago (Clarke et al, 2007).

118 Congruence among trees generated from independent datasets will generally indicate  
119 common phylogenetic history unless there has been lineage sorting, hybridisation of taxa  
120 and/or non-independence of characters (Penny et al., 1982; Bledsoe and Raikow, 1990; Zink  
121 and Avise, 1990; Paterson et al., 1993; Miyamoto and Fitch, 1995). The degree of congruence  
122 among phylogenies generated from penguin BLH data and other character types will indicate  
123 how useful BLH characters are in reconstructing phylogeny. Each BLH character is examined  
124 for homology by calculating the number of times it is gained over a tree based on molecular  
125 data from several gene regions. Finally, evolution of BLH traits is examined by the  
126 reconstruction of ancestral BLH character states.

127

### 128 *Penguin phylogeny*

129 To assess the accuracy of the phylogeny derived from BLH characters, phylogenies  
130 derived from other character types are required. The relationships of penguins have long been

131 contentious. While it is generally accepted that the order is monophyletic and contains six  
132 extant genera, only a few studies have examined the relationships among these genera. The  
133 studies that have investigated the relationships of most extant penguin species include: a  
134 phenetic study of myology (Schreiweis, 1972), a study of behavior with no formal character  
135 analysis (Jouventin, 1982), a study of integumentary and breeding characters (Giannini and  
136 Bertelli, 2004), phenetic studies of both external morphology and skeletal measurements  
137 (Livezey, 1989), a molecular phylogenetic study of several gene regions (Baker et al., 2006),  
138 cladistic studies of general morphology and two gene regions (Bertelli and Giannini, 2005)  
139 and a parsimony analysis of skeletal traits (Ksepka and Clarke, 2010). Generic level  
140 relationships have been investigated using a cladistic study of myology (McKittrick, 1991), a  
141 cladistic study of osteology (O'Hara, 1989), a phenetic analysis of DNA-DNA hybridization  
142 (Sibley and Ahlquist, 1990) and a likelihood analysis of gene regions (Baker et al., (2006).  
143 Each of these analyses produces a different phylogeny for the penguin genera and species  
144 (Fig. 1). This lack of congruence is perhaps to be expected, as penguins are highly adapted to  
145 the marine environment and under intense convergent selection pressure in their morphology.  
146 In studies of morphology and behaviour, phylogenetic signal may thus be obscured by  
147 adaptation. Molecular traits, such as the gene regions examined by Baker et al. (2006) and  
148 BLH (particularly terrestrial) traits, may be more appropriate for reconstructing penguin  
149 phylogeny.

150

### 151 *Homology testing*

152 In addition to assessing homology of individual BLH characters it is useful to know  
153 whether certain types of BLH characters are, on average, more likely to show homoplasy than  
154 others. End-products of behavior may show more homoplasy (e.g. caddisfly cases; (Stuart and  
155 Currie, 2002), or be as reliable as other character types (e.g. ovenbird nests; (Zyskowski and  
156 Prum, 1999). Lorenz (1935; 1941) suggested that behavioral displays are often more useful as



157 phylogenetic characters than other BLH traits, although Tinbergen (1959) argued that, on  
158 average, displays would contain no more phylogenetic information than other BLH traits.  
159 Paterson et al. (1995) found that seabird agonistic and reproductive displays contained no  
160 more phylogenetic information than foraging behavior and life history characters. This lack of  
161 difference between display and non-display BLH traits in seabirds may also be the same for  
162 penguins. There have, however, been more detailed studies of penguin displays than of  
163 seabird displays (e.g., Jouventin, 1982), and this may influence the phylogenetic content of  
164 such information. Ksepka and Clarke (2010) also used five behaviour and life history traits in  
165 their large morphological analysis of penguins.

166

### 167 *Character evolution*

168 Characters states do not arise *de novo* but are assembled over evolutionary time in a  
169 particular sequence (McLennan, 1991). This is especially apparent in the evolution of  
170 behavior (Brooks and McLennan, 1991; McLennan, 1991; McKittrick, 1992). Most  
171 phylogenetic work on behavior has involved reconstructing evolutionary sequences and co-  
172 occurring displays. These studies include: courtship sequence of newts (Arntzen and  
173 Sparreboom, 1989), rodent copulation (Langtimm and Dewsbury, 1991), breeding behavior of  
174 three-spined sticklebacks (McLennan, 1991), nest architecture of sweat bees (Packer, 1991),  
175 avian parental care (McKittrick, 1992), cooperative breeding of perching birds (Edwards and  
176 Naeem, 1993), nest architecture of swallows (Winkler and Sheldon, 1993) and courtship  
177 displays in birds of paradise (Scholes, 2008).

178 Phylogenetic analysis can be used to infer likely BLH states found in ancestors of  
179 extant organisms by optimising or mapping characters onto an independent phylogeny  
180 (McLennan, 1991). Such analyses have not progressed significantly over the last decade since  
181 work by Omland (1999). There are several different approaches as characterised by Martins  
182 (1999): linear parsimony (Swofford and Maddison, 1987), sum of squared changes parsimony

183 (McArdle and Rodrigo, 1994), maximum likelihood (Schluter et al., 1997), and generalised  
184 least squares (Martins and Hansen, 1997). We attempt to reconstruct the mode of life and  
185 appearance of the most recent common ancestor (MRCA). Current views about the MRCA  
186 claim that it was a small *Eudyptula*-like diver that arose in the New Zealand region  
187 (Jouventin, 1982). Stonehouse (1975) suggested that the flightless condition arose in the  
188 MRCA which was no larger than the smallest living species (*Eudyptula*). In terms of its BLH  
189 traits, the most recent common ancestor for the extant clade probably walked and swam in a  
190 manner similar to extant penguins, was a general forager in productive shallow water habitats,  
191 had few predators, and no migration (Simpson, 1976). Early penguins are thought to have  
192 evolved in a broad southern zone where water temperatures (12–18°C) were warmer than  
193 today (Simpson, 1975). Fordyce and Jones (1990) have reported a penguin fossil (24 million  
194 years before present - mybp) morphologically similar to *Eudyptula minor*, although it is likely  
195 that this is due to convergence and does not represent relatedness (Clarke et al., 2007). Baker  
196 et al. (2006) suggested that the ancestor of most modern penguin groups moved out of  
197 Antarctica and successfully colonised lower latitudes. In as much as behavior and life history  
198 are linked to morphology we might expect some BLH traits to be as conservative over  
199 penguin evolution as their morphological traits appear to be.

200         In this study we obtain behavioral and life history characters for the penguins. We test  
201 the following predictions: that there is significant structure in these data and that it is largely  
202 phylogenetic signal, that individual characters are good hypotheses of homology and that  
203 different data types share similar levels of homology, and that ancestral states can be  
204 reconstructed. An estimate for the BLH character states for the most recent common ancestor  
205 will be made.

206

## 207 **Methods**

208 *Behavior and life history data*

209           Fifteen penguin species were analysed in this study (see Table 1). The unrepresented  
210 species were *Spheniscus humboldti* (Humboldt's Penguin) and *S. mendiculus* (Galapagos  
211 Penguin), which were excluded because of insufficient information (i.e. states for more than  
212 half of the characters were unknown). We used the gull *Larus dominicanus* (Order  
213 Charadriiformes) as the outgroup throughout. This order is thought to be relatively closely  
214 related to the Order Sphenisciformes (see Sibley and Ahlquist, 1990; McKittrick, 1991;  
215 Paterson et al., 1995). Behavioral information for the species was largely extracted from the  
216 Handbook of Australian, New Zealand and Antarctic Birds (Marchant and Higgins, 1990).  
217 This handbook is in a standardized format, with each species described by a specialist  
218 researcher, and summarizes virtually all of the literature to that date. Information for all  
219 species was also collected from other sources (Eggleton and Siegfried, 1979; del Hoyo et al.,  
220 1992). Additional information was collected for specific species: *S. magellanicus* from  
221 Stonehouse (1975), *S. demersus* from Eggleton and Siegfried (1979), *Eudyptula minor* from  
222 Waas (1991), *Eudyptes pachyrhynchus* from Warham (1974) and *L. dominicanus* from  
223 numerous sources (Tinbergen, 1959; Saunders, 1971; Watson, 1975; Soper, 1976; Nugent,  
224 1982; Cramp, 1983; Robertson, 1985; Ehrlich et al., 1988). While datasets compiled from the  
225 literature may be weaker than those derived from direct observation, we have ensured that our  
226 BLH information is of high quality by going back to primary sources where possible.

227           Our choice of characters came down to two basic kinds. Traits that were explicitly  
228 behavioral were those identified as stereotypical (often as fixed action patterns) displays and  
229 included characters 14–25, 30–35 and 27–40. Traits that were life history related were those  
230 that had measurable consequences as a result of behavioral decisions made reliably by all  
231 appropriate individuals of a species. Life history traits, such as which vegetation type a nest is  
232 located in, are the outcome of the interaction of several behavioral systems. They are no  
233 different to measuring other behavioral outcomes such as caddis fly cases (Stuart and Currie  
234 2001) or ovenbird nest architecture (Zykowski and Prum 1999). As the aim of this study was

235 to assess the overall information content of BLH characters, efforts were made to utilise all  
236 available penguin BLH information. We omitted only two types of characters prior to  
237 finalizing the dataset: characters that were invariant (across these taxa and a range of other  
238 outgroups) and characters with unknown states in more than half of the taxa. We used the  
239 principle of primary homology (Ignarsson and Coddington 2007), where homology is assumed  
240 based on similarity, in order to assign our characters. We used the principle of secondary  
241 homology, congruence with other phylogenetic hypotheses generated with other data, to test  
242 these character states (Ignarsson and Coddington 2007).

243

#### 244 *Phylogenetic analysis*

245       Of the 68 BLH characters (Appendix 1), 27 were multi-state and analysed as  
246 unordered. Numerical character states were partitioned where there were natural breaks in  
247 quantitative data. An important issue in the use of BLH characters to reconstruct phylogeny  
248 concerns coding and weighting of these characters. Some authors (Felsenstein, 1983;  
249 Templeton, 1983) have argued that strict parsimony, where a loss is as likely as a gain, may  
250 not be the optimal coding for characters, particularly where losses are more likely than  
251 independent gains. One such situation may be complex behavioral displays and actions such  
252 as head circling, found in *A. forsteri* and *A. patagonicus*. This behavior features birds  
253 throwing their heads back, then moving them forward in lateral, half circling motions while  
254 groaning. It would seem less likely that a complex behavior of this sort would arise  
255 independently than it would be lost, and should be coded as such. One way of factoring such  
256 directionality is dollo parsimony (Maddison and Maddison, 2005; Cruickshank and Paterson,  
257 2006), which allows convergence for character loss but assumes that a state gain can only  
258 occur once. Note that dollo parsimony makes no assumption about the length of time required  
259 for complex characters to evolve or be lost, only that losses are much more likely than gains.  
260 A less extreme form of dealing with the problem of complex behaviors may be to weight

261 gains heavier than losses (Paterson et al., 1995). Our first analysis of the BLH data coded all  
262 characters equally as unordered. A second analysis used the same characters, but designated  
263 15 of them (those that we hypothesised as less likely to have arisen more than once) as dollo  
264 characters (10, 15, 22, 32, 33, 37, 38, 40, 52, 53, 54, 56, 65, 66, 67). All characters were  
265 assumed to be independent and were weighted equally. It is probable that some characters  
266 have a more rapid rate of evolutionary change than others but there were no reasons *a priori*  
267 for selecting differential weighting (see Eernisse et al., 1992). The independence of characters  
268 was assessed by ensuring that all character states were mutually exclusive within and between  
269 characters. Just as ‘organisms are hierarchies of parts’ (McKittrick, 1994), we view complex  
270 behaviors as hierarchies of simple behaviors and, moreover, as diagnostic features for  
271 underlying neural structures and organization (see Japyassu and Machado, 2010). For  
272 example, a complex behaviour, such as nest building, was considered to be a suite of  
273 phylogenetically independent characters, i.e., made up of several less complex behaviors with  
274 their own individual, but congruent, evolutionary history.

275         The data were analysed using the branch and bound option of PAUP\* 4.0b10  
276 (Swofford, 2002) to generate BLH cladograms. The most fundamental question that can be  
277 asked about the data derived from the penguin BLH dataset is whether it contains any more  
278 cladistic information than a random dataset. To answer this question we employed a  
279 permutation tail probability (PTP) test (Faith, 1991; Faith and Cranston, 1991) in PAUP\* to  
280 test for significant cladistic structure and used the  $g_1$  statistic to evaluate the skew in the tree  
281 length distributions. The PTP test shows whether the structure in the most parsimonious tree  
282 could have arisen by chance alone. This test randomises character states without replacement  
283 across taxa for each character and derives the shortest tree-length for the randomised dataset.  
284 Tree-length reflects the degree to which shared characters co-vary upon a tree for a given  
285 character set. A value for the cladistic covariation was obtained by comparing the length of  
286 the most parsimonious tree with shortest tree-lengths derived from randomised datasets. If the

287 tree-length of the BLH shortest tree is not significantly smaller than shortest trees from the  
288 same dataset after randomisation then BLH data contains no significant phylogenetic  
289 structure. This study used 10000 randomised datasets to test for phylogenetic structure.

290 We assessed whether the incongruence between the BLH trees and the other estimates  
291 of penguin phylogeny (Fig. 1) could be due to sampling error or noise in the data by using the  
292 KH test (Kishino and Hasegawa, 1989) in PAUP\*. We used the SH test (Shimodaira and  
293 Hasegawa, 1999) implemented in PAUP\* (with a resample estimated log-likelihood [RELL]  
294 test distribution using 1000 bootstrap replicates) to compare the molecular tree of Baker et al.  
295 (2006) with the trees generated from other datasets.

296

#### 297 *Homology testing*

298 The homology of the BLH characters was assessed by mapping them onto the best  
299 estimate of penguin phylogeny - that derived from the concatenated sequences. Individual and  
300 overall measures of homoplasy, and instances of convergence, gains and losses, were  
301 calculated for BLH characters by mapping them onto the concatenated sequence tree. We took  
302 as our best measure of penguin phylogeny a tree derived from >5kb of mitochondrial and  
303 nuclear DNA (Baker et al., 2006). Other molecular studies, e.g., Bertelli and Giannini (2005)  
304 and Ksepka and Clarke (2010) provide largely similar placements of penguin taxa. This tree  
305 will hereafter be described as the molecular tree.

306

#### 307 *Display characters*

308 In order to assess relative phylogenetic value of display or behavioral (14–25, 30–35,  
309 37–40) compared to non-display or life history (1–13, 26–29, 36, 41–68) characters, these  
310 characters were mapped onto the molecular tree and measures of homoplasy (consistency  
311 index [CI] and retention index [RI]) and numbers of gains and losses were calculated for each  
312 character. Trees were constructed from display only and non-display only character datasets

313 using the branch and bound option of PAUP\*. The trees were then compared to the molecular  
314 tree.

315

### 316 *Character evolution*

317 The non-continuous BLH characters were optimized onto the molecular tree using a  
318 parsimony approach and character states were reconstructed to determine the likely states  
319 present in the most recent common ancestor of extant penguins (MRCA). There are several  
320 methods by which characters are optimized onto trees and ancestors reconstructed. The most  
321 appropriate method of optimising BLH characters is the Acctran option of PAUP\*. The  
322 Acctran option prefers reversals over parallelisms which approximates our assumption that  
323 complex behaviors are unlikely to continue to evolve in concert in different species. Retention  
324 indices were calculated for each character (when mapped onto the molecular tree), and these  
325 were compared for display versus non-display characters.

326

## 327 **Results**

### 328 *Phylogenetic analysis*

329 Parsimony analysis of the 68 unordered BLH characters (Table 1, Appendix 1)  
330 produced two shortest trees (Fig. 2; TL = 190, CI = 0.56 and RI = 0.59). Parsimony analysis of  
331 the same BLH dataset using dollo parsimony for 15 of the characters produced a single most-  
332 parsimonious tree (Fig. 3, TL = 196, CI = 0.54, RI = 0.67). The dollo tree differed from both  
333 unordered trees (although not significantly). For example, both unordered trees had  
334 *Megadyptes antipodes* as sister to a clade comprising the black-footed penguin together with  
335 all *Pygoscelis* and *Eudyptes*, whereas *M. antipodes* was placed within *Pygoscelis* in the dollo  
336 analysis.

337 The PTP tests (for both unordered and dollo parsimony) revealed that penguin BLH  
338 data contained significant cladistic structure ( $P = 0.0001$ ), i.e., the shortest trees were

339 substantially shorter than all of the trees generated by randomising the data. Similarly, the  
340 significantly skewed tree length distributions ( $g_1 = -0.911$  for unordered parsimony and  $g_1 = -$   
341  $0.629$  for dollo parsimony from 10000 random trees,  $P < 0.01$ ; Hillis and Huelsenbeck, 1992)  
342 indicates that the data contain significant signal. The BLH trees were generally incongruent  
343 with previous estimates of penguin phylogeny (Fig. 1), with some exceptions. Estimates were  
344 not significantly different for Jouventin, O'Hara, or Sibley and Ahlquist for the unordered  
345 data (Table 2); or for Livezey (external morphology), O'Hara, or Sibley and Ahlquist for the  
346 dollo data (Table 3). The SH tests showed that the molecular data could reject the BLH tree  
347 topologies and the other alternatives to the molecular tree, except that of Sibley and Ahlquist  
348 (Table 4).

349

#### 350 *Homology testing*

351 The BLH characters were mapped onto the molecular tree. Assessing the covariance of  
352 characters onto the trees found seventeen characters (2, 9, 11, 13, 17, 19, 21, 25, 29, 34, 37,  
353 38, 43, 47, 50, 65, 67) that had RIs of  $\geq 0.80$ , i.e. fitted well. Ten characters (4, 10, 14, 20, 23,  
354 26, 32, 33, 61, 64) showed a total lack of fit (RI = 0.00), one (27) with an RI of 0.17, and one  
355 (44) with an RI of 0.20. Nine characters (3, 5, 6, 28, 53, 55, 56, 57, 62) were autapomorphic,  
356 making their RIs irrelevant.

357

#### 358 *Display characters*

359 Mapping of display and non-display characters onto the molecular tree indicated that  
360 they had similar levels of homoplasy (see Fig. 4), i.e. the difference between the RIs was not  
361 significant (Mann-Whitney,  $U = 373.5$ ,  $df = 1$ ,  $P = 0.599$ ). This finding suggests that penguin  
362 display characters are no less homoplasious than the other BLH characters.

363 Parsimony analysis of the display characters alone produced 268 shortest trees (Fig. 5;  
364 TL = 56, CI = 0.66, RI = 0.70). Parsimony analysis of the non-display characters found 297



365 shortest trees (Fig. 6; TL = 130, CI = 0.53, RI = 0.569). The Kishino-Hasegawa tests indicated  
366 that the display dataset could reject, whereas the non-display dataset could not reject (just), the  
367 molecular tree (display: Ldiff = 6,  $t = 2.3238$ ,  $P$  (one tailed) = 0.0152; non-display: Ldiff = 8,  $t$   
368 = 1.6641,  $P$  (one-tailed) = 0.0515).

369

### 370 *Character evolution*

371 The distribution of BLH traits estimated for the MRCA are hypothesized as a pursuit  
372 diver that used crustaceans and/or fish as food, capturing food in short local trips in groups.  
373 The monogamous MRCA was a hillside nester, probably in arid conditions, had only a local  
374 migration phase, and bred in summer with the male contributing to nest and chick duties.  
375 Chicks were semi-altricial, nidicolous and creched soon after birth. Moults were complete and  
376 occurred in late summer. It performed few current reproductive displays, other than ecstatic  
377 and mutual bowing, and few current agonistic displays, except for primarily aggressive  
378 behaviors such as bite-nape fighting. Similar results were obtained if alternative outgroups of  
379 southern seabirds, Procellariiformes (*Pterodroma inexpectata*) and Pelecaniformes  
380 (*Stictocarbo punctatus*), were used.

381

### 382 **Discussion**

383 Our phylogenetic analyses indicate that penguin behavior and life history traits contain  
384 only limited phylogenetic information, unlike the results of several other similar studies of  
385 different taxa (e.g., McLennan et al., 1988; Arntzen and Sparreboom, 1989; Prum, 1990;  
386 Paterson et al., 1995; Kennedy et al., 1996; Slikas, 1998; Noll, 2002; Stuart et al., 2002). The  
387 PTP test and  $g_1$  statistics revealed that BLH characters contained more structure than expected  
388 due to chance, but the KH and SH tests showed that this signal was significantly different  
389 from the presumed correct molecular phylogeny derived from several gene regions. When  
390 trees from independent datasets are congruent it is assumed that the congruence is caused by a

391 common phylogenetic signal contained within each (Penny et al., 1982; Zink and Avise,  
392 1990). However, the KH tests were significant ( $P = 0.0432$  and  $0.0291$ , thus the BLH data can  
393 reject the molecular tree topology) as are the SH tests (all  $<0.001$ , thus the molecular data can  
394 rejects the BLH tree topologies), all of which implies that there is not a great deal in common  
395 between the BLH and molecular trees. Therefore, the structure found in the BLH data is not  
396 necessarily generated by phylogeny. This is readily apparent from even a cursory inspection of  
397 the BLH trees. In the unordered, dollo, display and non-display trees (Figs. 2–3, 5–6) only the  
398 position of *Aptenodytes* as sister to the rest of the penguins and the monophyly of *Eudyptes* is  
399 the same as the molecular tree. The sister relationship of *Aptenodytes* agreed with Schreiweis,  
400 Jouventin, Livezey, and Sibley and Ahlquist. Both unordered and dollo trees found that  
401 *Pygoscelis* and *Spheniscus* were polyphyletic, with Black-footed (*S. demersus*) as sister taxon  
402 to Yellow-eyed (*M. antipodes*) in the dollo tree. The unordered trees agreed with the  
403 molecular tree in the placement of most of the *Eudyptes* taxa. Only the placement of the Erect-  
404 crested (*E. sclateri*) within the genus differed from the molecular tree for one of the BLH  
405 trees. The dollo tree only found the Macaroni (*E. chrysolophus*) and Royal (*E. schlegeli*) pair  
406 within *Eudyptes*. There is no evidence from the other datasets that the pygoscelid penguins are  
407 a paraphyletic group.

408         Why is there a minimal amount of phylogenetic information in the BLH dataset? The  
409 display characters showed similar levels of homoplasy to the non-display characters, and, if  
410 anything, the resulting non-display based phylogeny was marginally less inaccurate (given that  
411 it was only just significantly different from the molecular phylogeny) than the display based  
412 phylogeny. Adaptation is particularly effective at erasing traces of phylogeny. It may be that  
413 penguin behavior and life history is itself severely constrained by the constraints of the marine  
414 environment (and changes in morphology to survive there) and that character states often  
415 represent convergence and parallelisms rather than homology. Such convergence has been  
416 observed in behavioral calls of lacewings probably because of the constraints on methods to

417 actually produce sounds (Henry et al., 1999). Another issue is that of our homology  
418 hypotheses. Although all of our characters satisfied primary homology requirements  
419 (characters were similar) and most satisfied secondary homology requirements (character  
420 states were congruent with an accepted phylogeny), there are difficulties in coding  
421 behavioural character states because there are multiple hypotheses that can be encoded. For  
422 example, Ksepka and Clarke (2010) have five reproductive characters in their data set that are  
423 similar to ours. However, where Ksepka and Clarke have a character for clutch size and one  
424 for size of first egg to second, we have one character where all of this information is encoded  
425 into multiple states. Ksepka and Clarke (2010) have one character on nest information  
426 whereas we have three. Ksepka and Clark (2010) use the ecstatic display as either present or  
427 absent whereas we have six different states. None of these differences are contradictory but  
428 they do represent different hierarchical views on the primary homology of penguin behavioral  
429 traits.

430

### 431 *Homology testing*

432         Studies of behavioral characters have shown that these characters may be no more  
433 homoplasious than other types of characters (de Queiroz and Wimberger, 1993; Paterson et  
434 al., 1995). We assumed this to be true of the BLH characters used in this study. In general this  
435 assumption was supported by the overall distribution of individual RIs. For instance, there  
436 were more characters that fitted the molecular tree well (17 with  $RI \geq 0.80$ ) than those that  
437 fitted poorly (12 with  $RI \leq 0.20$ , see Fig. 4), although these 12 introduced considerable  
438 homoplasy to estimates of phylogeny. Levels of homoplasy in non-display relative to display  
439 characters were similar. A more direct test for the homology of individual characters is to  
440 assess whether the behavior has arisen only once. Each display character was mapped on the  
441 molecular tree and numbers of gains and losses of each character were derived under two  
442 scenarios (Table 5). First, gains and losses were assumed to be equivalent in cost and, second,

443 independent gains were thought to be much less likely than independent losses (dollo).  
444 Thirteen characters appear to have clearly arisen once. For example, treading (34) arose deep  
445 in the lineage and has been retained by all descendent taxa (Fig. 7a). Seven display characters  
446 (14, 16, 18, 20, 25, 31, 40) may have arisen twice if gains and losses are considered equally  
447 likely and two (24, 35) may have arisen three times. Each of these nine characters was  
448 reassessed to calculate the likely number of losses implied if the display had arisen only once  
449 (Table 5). For example, the stare behavior (character 18) appears to have arisen twice: in *A.*  
450 *patagonicus* and in the ancestor of the other penguin genera. If this behavior is assumed to  
451 have arisen only once then a loss in the *A. forsteri* branch is required to explain the observed  
452 pattern (Fig. 7b). Only three characters (25: bill hiding, 31: allopreening, and 35: head shake  
453 and whine) posit fewer evolutionary events by assuming that the characters are not  
454 homologous. For example, bill-hiding (25) may have either evolved twice, once in the  
455 *Eudyptes* and once in the *Spheniscus* (two events, Fig. 7c), or once deeper in the tree but then  
456 lost once in *Eudyptula* and once in *Megadyptes* (three events). These BLH characters may  
457 represent the type of traits that Tinbergen (1959) hypothesized to be similar "because of the  
458 limited number of possibilities of doing the opposite".

459         Stability of behavior over evolutionary time is vital for behavioral characters to be  
460 useful in the construction of phylogeny. Penguins have a fossil history of more than 60 mybp  
461 (Slack et al., 2006). BLH characters appear to have persisted in penguin lineages over  
462 considerable evolutionary periods, e.g., pursuit diving (character 3: state 1), stare (18: 1), bite  
463 nape fight (20: 1), ecstatic (30: 1-5), mutual bowing (32: 1), treading (34: 1), and nesting  
464 behavior (54: 1). The most parsimonious inference is that many behaviors have persisted since  
465 the origin of the extant penguin species, i.e. the behavioral characters have persisted for at  
466 least 13 million years (Ksepka and Clarke, 2010) or possibly as much as 30 million years  
467 (Baker et al., 2006). Ethologists have often assumed that behaviors can be simply divided into  
468 those that are "innate" and those that are "learnt". Developmental researchers have argued that

469 this division is simplistic and misleading (Lehrman, 1953; Gray, 1992). One assumption that  
470 is typically thought to follow is that only behaviors considered to be "innate" contain  
471 phylogenetic information. This would exclude many behaviors that may have phylogenetic  
472 value. What is important for the persistence of behavioral characters down lineages is not how  
473 the behavior develops but that the conditions necessary for the development of the behavior  
474 reliably reoccur through the phylogenetic history of organisms (Gray, 1989; Gray, 1992).  
475 Many features of the penguin's physical (e.g. marine habitat, constant temperature, salinity)  
476 and social (e.g. predators, competitors, colonial nesting) environment, and morphology (e.g.  
477 flippers, monotypic feathers) have remained unaltered over millions of years. The persistence  
478 of these features may lead to repeated development of phylogenetically informative behavioral  
479 characters.

480

#### 481 *Character evolution*

482 By examining the distribution of BLH character states at the internal branch between  
483 the outgroup and penguins, we reconstructed the likely BLH traits of the [MRCA](#). The penguin  
484 ancestor has been claimed to be either a member of the genus *Eudyptula* or at least  
485 eudyptulid-like and inhabited the New Zealand region (Jouventin, 1982). Phylogenetic  
486 evidence from the molecular tree indicated that *Eudyptula* is not the basal genus of the  
487 penguins (or in any of the hypotheses of Fig.1). The Magellanic (*S. magellanicus*) and Gentoo  
488 (*P. papua*) were found to have most sequence composition in common with the hypothesised  
489 [MRCA](#). Most of the shared BLH character states between the [MRCA](#) and these species are  
490 those in common with the majority of the other penguin taxa. It appears unlikely, therefore,  
491 that the [most recent common ancestor of extant penguins](#) was more similar, in terms of  
492 behavior and life history, to the Little blue (*E. minor*) than to the other penguin species.

493 The [MRCA](#) had a similar life history to extant penguins but had few of the ritualised  
494 reproductive and agonistic behaviors. This agrees with the predictions made by Simpson

495 (1975; 1976), Jouventin (1982) and Fordyce and Jones (1990). One problem with assessing  
496 BLH character states of ancestors is that we obviously cannot reconstruct extinct behaviors.  
497 The reconstructed [MRCA](#) appears to have had few complex displays, or if it did, they were  
498 subsequently lost in lineages leading to extant species. The latter seems unlikely, as the  
499 likelihood of most of the [MRCA](#) complex behaviors completely disappearing to be replaced  
500 by new complex behaviors would require many evolutionary events. The idea of the  
501 [MRCA](#) with fairly simple BLH traits is also consistent with several ethological views on the  
502 evolution of behavior. For example, the ritualisation of behavior is the evolutionary change of  
503 a behavior in a direction that exaggerates the signal delivered to the receiver (Tinbergen,  
504 1959; Harper, 1991). A simple behavior may change over evolutionary time to become more  
505 complex (ritualised) to reduce signal ambiguity and manipulate and deceive the receiver  
506 (Harper, 1991). Alternatively, a simple behavior may become more complex over time by run-  
507 away sexual selection due to a Fisherian process or because the behavior is an advertisement  
508 for male quality (Kodric-Brown and Brown, 1984). All of these alternatives are consistent  
509 with an ancestor possessing less complex behaviors.

510 Did the [MRCA](#) inhabit Zealandia (Trewick et al., 2007)? New Zealand, and its  
511 surrounding islands, shows the greatest diversity of penguin species with five of the six  
512 penguin genera breeding or migrating throughout the region. New Zealand has, thus, been  
513 viewed as the likely area for origin of penguins (Jouventin, 1982; Fordyce and Jones, 1990).  
514 During the key phase of extant penguin diversification, the Zealandia region was largely  
515 inundated with only small island groups available for habitation (Landis et al., 2008). Baker et  
516 al. (2006) have suggested that extant penguins moved out of Antarctica during this period and  
517 colonised the more temperate regions. Both of the basal penguin genera, *Aptenodytes* and  
518 *Pygoscelis*, currently inhabit high southern latitudes. BLH traits of the [MRCA](#) are generally  
519 consistent with inhabiting this region with short summers in which to breed, (e.g. breeding in  
520 summer, and nesting in environments with little or no vegetation, short nest preparation

521 period, creching of chicks, moulting in late summer, when there is plentiful food). In the mid  
522 to late Cretaceous New Zealand occupied a more cold temperate and southerly position, 85–  
523 55°S (Fleming, 1979; Stevens, 1989; Cooper and Millener, 1993). Given that the [MRCA](#) is  
524 most similar in behavior and life history to the extant Magellanic (*S. magellanicus*) and  
525 Gentoo (*P. papua*) and that both of these species inhabit high latitude temperate regions the  
526 evidence is consistent with a hypothesis that penguins originated in the Zealandia region.

527         One particular confounding factor in this study may be the use of a single outgroup.  
528 The lack of several outgroups was largely the result of a difficulty in homologising behaviors  
529 in other possible bird groups. Penguins are a very distinct taxonomical group with a  
530 behavioral environment that differs from other birds. Close relatives, like Procellariiformes,  
531 share a common ancestor with penguins in excess of 60 million years ago (Ksepka and  
532 Clarke, 2010). We attempted to obtain traits from Procellariiformes and Pelecaniformes but  
533 found it difficult to make sensible homologies. Nixon and Carpenter (2012) suggest that it is  
534 more sensible not to hypothesize homology when analogy is more likely. We could sensibly  
535 find homologies between Charadriiformes species and the penguins but given that the origin  
536 of the extant penguins may be as recent as 12 million years ago (Ksepka and Clarke, 2010)  
537 there is no simple way to break the long branch to the outgroup. This is a weakness of using  
538 behaviour traits and may affect the rooting of the phylogeny and other aspects of topology.

539         A parsimony approach to reconstructing ancestral states is not without its flaws  
540 (Cunningham, 1999). Several other methods using maximum likelihood (Schluter et al., 1997;  
541 Maddison and Maddison, 2011) and generalised least squares (Martins, 1999; 2004) are more  
542 sophisticated approaches. For the most part, however, these work with continuous characters  
543 which make up only a few of those presented in this study or require models of evolution for  
544 traits that are currently unknown for behavior.

545         In summary, Brooks and McLennan (1991) have emphasized the productive links that  
546 are possible between behavioral, ecological and phylogenetic research. Our study indicates

547 that phylogenetic methods can 1) partially reconstruct penguin phylogeny from behavioral and  
548 life history data, 2) assess the homology of behavioral characters, and 3) make inferences  
549 about behavioral and ecological evolution. Why are BLH characters in penguins less  
550 phylogenetically informative than BLH traits in other bird groups, such as procellariiforms  
551 (Paterson et al., 1995), manakins (Prum, 1990), pelecaniforms (Kennedy et al., 1996) and  
552 storks (Slikas, 1998)? Homoplasy can be an indication that you have made an error in  
553 homologising traits (Nixon and Carpenter, 2012). Penguin BLH traits are either more difficult  
554 to operationally group into characters and states, or are more prone to convergence. Most of  
555 the behavioral traits appear to have persisted for a great length of time. It may be that, like  
556 morphology, much of penguin behavior and life history is subject to intense constraints from  
557 the marine environment and that this adaptive force tends to obscure phylogenetic signal.  
558 Identifying penguin BLH traits that are less prone to adaptive pressure from the marine  
559 environment, like reproductive displays, clutch size and nest-site fidelity, may provide a  
560 source of characters that contain primarily phylogenetic information.

561

## 562 **Acknowledgements**

563 Thanks to John Croxall, Steve Broni and Lloyd Davis for comments on an earlier  
564 version of our behavioral and life history characters. This work was funded by: a University of  
565 Otago Research Grant and Research Assistantship (GPW and RDG), Lincoln University  
566 Research Fund (AMP), and a University of Auckland Research Grant (RDG).

567

## 568 **References**

- 569 Agnarsson, I., Coddington, J. A. 2007. Quantitative test of primary homology. *Cladistics* 23,  
570 1–11.
- 571 Arntzen, J. W., Sparreboom, M. 1989. A phylogeny for the old world newts, genus *Triturus*:  
572 biochemical and behavioral data. *J. Zool* 219, 645–664.



- 573 Aronson, L. R. 1981. Evolution of telencephalic function in lower vertebrates. In: Laming, P.  
574 R. (Ed.) Brain mechanisms of behavior in lower vertebrates. Cambridge University  
575 Press, Cambridge, pp. 33–58.
- 576 Atz, J. W. 1970. The application of the idea of homology to behavior. In: Aronson, L. R.,  
577 Tobach, E., Lehrman, D. S., Rosenblatt, J. S. (Eds.), Development and evolution of  
578 behavior: essays in memory of T.C. Schneirla. W.H. Freeman, San Francisco, pp. 53–  
579 74.
- 580 Baker, A. J., Pereira, S. L., Haddrath, O. P., Edge, K.-A. 2006. Multiple gene evidence for  
581 expansion of extant penguins out of Antarctica due to global cooling. Proc. R. Soc.  
582 Lond. B. 273, 11–17.
- 583 Beehler, B. M., Swaby, R. J. 1991. Phylogeny and biogeography of the Ptiloris riflebirds  
584 (Aves: Paradisaeidae). Condor 43, 738–745.
- 585 Bertelli, S., Giannini, N. P. 2005. A phylogeny of extant penguins (Aves: Sphenisciformes)  
586 combining morphology and mitochondrial sequences. Cladistics 21, 209–239.
- 587 Bledsoe, A. H., Raikow, R. J. 1990. A quantitative assessment of congruence between  
588 molecular and nonmolecular estimates of phylogeny. J. Mol. Evol. 30, 247–259.
- 589 Brooks, D. R., McLennan, D. A. 1991. Phylogeny, ecology and behavior: a research program  
590 in comparative biology. University of Chicago Press, Chicago.
- 591 Brower, A. V. Z., Schawaroch, V. 1996. Three steps of homology assessment. Cladistics 12,  
592 265–272.
- 593 Clarke, J. A., Ksepka, D. T., Stucchi, M., Rubina, M., Giannini, N., Bertelli, S., Narvaez, Y.,  
594 Boyd, C. A. 2007. Paleogene equatorial penguins challenge the proposed relationship  
595 between biogeography, diversity, and Cenozoic climate change. Proceedings of the  
596 National Academy of Sciences 104, 11545–11550.
- 597 Cooper, R. A., Millener, P. R. 1993. The New Zealand biota: Historical background and new  
598 research. Trends Ecol. Evol. 8, 429–433.

599 Costa, D. P. 1991. Reproductive and foraging energetics of high latitude penguins, albatrosses  
600 and pinnipeds: Implications for life history patterns. *Am. Zool.* 31, 111–131.

601 Cracraft, J. 1985. Monophyly and the phylogenetic relationships of the Pelecaniformes: A  
602 numerical cladistic analysis. *Auk* 102, 834-853.

603 Cramp, S. 1983. *The birds of the Western Palearctic*. Oxford University Press, Oxford.

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

607 Cruickshank, R. C., Paterson, A. M. 2006. The Great Escape: Do Parasites Break Dollo's  
608 Law? *Trends in Parasitology* 22, 509–515.

609 Cunningham, C. W. 1999. Some limitations of ancestral character-state reconstruction when  
610 testing evolutionary hypotheses. *Systematic Biology* 48, 665–674.

611 de Queiroz, A., Wimberger, P. H. 1993. The usefulness of behaviour for phylogeny  
612 estimation: Levels of homoplasy in behavioural and morphological characters.  
613 *Evolution* 47, 46-60.

614 del Hoyo, J., Elliot, A., Sargatal, J. (Eds.) 1992. *Handbook of the birds of the world*. Lynx  
615 Edicions, Barcelona.

616 Doyle, J. J. 1993. DNA, phylogeny, and the flowering of plant systematics. *Bioscience* 43,  
617 380–389.

618 Edwards, S. V., Naeem, S. 1993. The phylogenetic component of cooperative breeding in  
619 perching birds. *American Naturalist* 141, 754-789.

620 Eernisse, D. J., Albert, J. S., Anderson, F. E. 1992. Annelida and Arthropoda are not sister  
621 taxa: A phylogenetic analysis of spiralian metazoan morphology. *Systematic Biology*  
622 41, 305-330.

623 Eggleton, P., Siegfried, W. R. 1979. Displays of the jackass penguin. *Ostrich* 50, 139–167.

- 624 Ehrlich, P. R., Dobkin, D. S., Wheye, D. 1988. The birders handbook. Simon and Schuster  
625 Incorporated, New York.
- 626 Faith, D. P. 1991. Cladistic tests for monophyly and nonmonophyly. *Systematic Zoology* 40,  
627 366-375.
- 628 Faith, D. P., Cranston, P. S. 1991. Could a cladogram this short have arisen by chance alone?  
629 On permutation tests for cladistic structure. *Cladistics* 7, 1-28.
- 630 Felsenstein, J. 1983. Parsimony in systematics: biological and statistical issues. *Ann. Rev.*  
631 *Ecol. Syst.* 14, 313–333.
- 632 Fleming, C. A. 1979. The geological history of New Zealand and its life. University of  
633 Auckland and Oxford University Press, Auckland.
- 634 Fordyce, R. E., Jones, C. M. 1990. Penguin history and new fossil material from New  
635 Zealand. In: Davis, L. S., Darby, J. T. (Eds.), *Penguin Biology*. Academic Press, San  
636 Diego, pp. 419–446.
- 637 Giannini, N. P., Bertelli, S. 2004. Phylogeny of extant penguins based on integumentary and  
638 breeding characters. *Auk* 121, 422–434.
- 639 Gittleman, J. L., Kot, M. 1990. Adaptation: Statistics and a null model for estimating  
640 phylogenetic effects. *Syst. Zool.* 39, 227–241.
- 641 Gray, R. D. 1989. Oppositions in panbiogeography: Can the conflicts between selection,  
642 constraint, ecology, and history be resolved? *N. Z. J. Zool.* 16, 787–806.
- 643 Gray, R. D. 1992. Death of the gene: Developmental systems strike back. In: Griffiths, P. E.  
644 (Ed.) *Trees of life*. Kluwer, Dordrecht, pp. 165-209.
- 645 Harper, D. G. C. 1991. Communication. In: Krebs, J. R., Davies, N. B. (Eds.), *Behavioural*  
646 *Ecology: an Evolutionary Approach*. Blackwell, Oxford, pp. 374–397.
- 647 Harvey, P. H., Pagel, M. D. 1991. *The comparative method in evolutionary biology*. Oxford  
648 University Press, Oxford.

- 649 Henry, C. S., Wells, M. L. M., Simon, C. M. 1999. Convergent evolution of courtship songs  
650 among cryptic species of the carnea group of green lacewings (Neuroptera:  
651 Chrysopidae: Chrysoperla). *Evolution* 53, 1165-1179.
- 652 Hillis, D. M., Huelsenbeck, J. P. 1992. Signal, noise, and reliability in molecular phylogenetic  
653 analyses. *Journal of Heredity* 83, 189-195.
- 654 Hinde, R. A., Tinbergen, N. 1965. The comparative study of species-specific behavior. In:  
655 T.E., M. (Ed.) *Readings in animal behavior*. Holt, Rinehart and Winston, New York,  
656 pp. 58–70.
- 657 Japyassu, H. F., de A. Machado, F. 2010. Coding behavioural data for cladistic analysis: using  
658 dynamic homology without parsimony. *Cladistics* 26, 625–642.
- 659 Jouventin, P. 1982. Visual and vocal signals in penguins, their evolution and adaptive  
660 characters. Verlag Paul Parey, Berlin and Hamburg.
- 661 Kennedy, M., Spencer, H. G., Gray, R. D. 1996. Hop, step and gape: Do the social displays of  
662 the Pelecaniformes reflect phylogeny? *Animal Behaviour* 51, 273-291.
- 663 Kishino, H., Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the  
664 evolutionary tree topologies from DNA sequence data, and the branching order in  
665 Hominoidea. *Journal of Molecular Evolution* 29, 170-179.
- 666 Kodric-Brown, A., Brown, J. H. 1984. Truth in advertising: the kinds of traits favored by  
667 sexual selection. *Am. Nat.* 124, 309–323.
- 668 Ksepka, D. T., Clarke, J. A. 2010. The basal penguin (Aves: Sphenisciformes) *Perudyptes*  
669 *devriesi* and a phylogenetic evaluation of the penguin fossil record. *Bulletin of the*  
670 *American Museum of Natural History* 337, 1–77.
- 671 Landis, C. A., Campbell, H. J., Begg, J. G., Mildenhall, D. C., Paterson, A. M., Trewick, S. A.  
672 2008. The Waipounamu erosion surface: questioning the antiquity of the New Zealand  
673 land surface and terrestrial fauna and flora. *Geological Magazine* 145, 173–197.

- 674 Langtimm, C. A., Dewsbury, D. A. 1991. Phylogeny and evolution of rodent copulatory  
675 behaviour. *Animal Behaviour* 41, 217-225.
- 676 Lehrman, D. S. 1953. A critique of Konrad Lorenz's theory of instinctive behaviour. *Quarterly*  
677 *Review of Biology* 28, 337-363.
- 678 Livezey, B. C. 1989. Morphometric patterns in recent and fossil penguins (Aves,  
679 Sphenisciformes). *J. Zool.* 219, 269–307.
- 680 Lorenz, K. Z. 1935. Der kumpan in der umwelt des vogels. *J. Orn.* 83.
- 681 Lorenz, K. Z. 1941. Vergleichende bewegungsstudien an anatinen. *J. Orn.* 89, 194-294.
- 682 Lorenz, K. Z. 1958. The evolution of behavior. *Sci. Am.* 199, 67–68.
- 683 Maddison, D. R., Maddison, W. P. 2005. *MacClade 4: Analysis of phylogeny and character*  
684 *evolution.* Sinauer Associates, Sunderland, Massachusetts.
- 685 Maddison, W. P., Maddison, D. R. 2011. *Mesquite: A modular system for evolutionary*  
686 *analysis.*
- 687 Marchant, S., Higgins, P. J. 1990. *Handbook of Australian, New Zealand and Antarctic birds,*  
688 *Vol.1: Ratites to ducks.* Oxford University Press, Oxford.
- 689 Martins, E. P. 1999. Estimation of ancestral states of continuous characters: A computer  
690 simulation study. *Syst. Biol.* 48, 642–650.
- 691 Martins, E. P. 2004. COMPARE, version 4.6b. Computer programs for the statistical analysis  
692 of comparative data. Department of Biology, Indiana University, Bloomington IN, pp.  
693 Distributed by the author at <http://compare.bio.indiana.edu/>.
- 694 Martins, E. P., Hansen, T. F. 1997. Phylogenies and the comparative method: A general  
695 approach to incorporating phylogenetic information into the analysis of interspecific  
696 data. *Am. Nat.* 149, 646–667.
- 697 Mattern, M. Y., McLennan, D. A. 2004. Total evidence phylogeny of Gasterostiedae:  
698 Combining molecular, morphological and behavioural data. *Cladistics* 20, 14–22.

- 699 McArdle, B., Rodrigo, A. G. 1994. Estimating the ancestral states of a continuous-valued  
700 character using squared-change parsimony: An analytical solution. *Syst. Biol.* 43, 573–  
701 578.
- 702 McKittrick, M. C. 1991. Phylogenetic analysis of avian hindlimb musculature. University of  
703 Michigan, Ann Arbor.
- 704 McKittrick, M. C. 1992. Phylogenetic analysis of avian parental care. *Auk* 109, 828-846.
- 705 McKittrick, M. C. 1994. On homology and the ontological relationship of parts. *Syst. Biol.* 43,  
706 1–10.
- 707 McLennan, D. A. 1991. Integrating phylogeny and experimental ethology: From pattern to  
708 process. *Evolution* 45, 1773-1789.
- 709 McLennan, D. A., Brooks, D. R., McPhail, J. D. 1988. The benefits of communication  
710 between comparative ethology and phylogenetic systematics: a case study using  
711 gasterosteid fishes. *Canadian Journal of Zoology* 66, 2177-2190.
- 712 McLennan, D. A., Mattern, M. Y. 2001. The phylogeny of the Gasterosteidae: Combining  
713 behavioral and morphological data sets. *Cladistics* 17, 11-27.
- 714 Miyamoto, M. M., Fitch, W. M. 1995. Testing species phylogenies and phylogenetic methods  
715 with congruence. *Systematic Biology* 44, 64-76.
- 716 Nixon, K. C., Carpenter, J. M. 2012. On homology. *Cladistics* 28, 160–169.
- 717 Noll, F. B. 2002. Behavioral phylogeny of corbiculate Apidae (Hymenoptera: Apinae), with  
718 special reference to social behavior. *Cladistics* 18, 137–153.
- 719 Nugent, G. 1982. Sexing Black-backed Gulls from external measurements. *Notornis* 29, 37-  
720 40.
- 721 O'Hara, R. J. 1989. An estimate of the phylogeny of the living penguins (Aves: Spheniscidae).  
722 *Am. Zool.* 29, A11.
- 723 Omland, K. E. 1999. The assumptions and challenges of ancestral state reconstructions. *Syst.*  
724 *Biol.* 48, 604–611.

- 725 Packer, L. 1991. The evolution of social behaviour and nest architecture in sweat bees of the  
726 subgenus *Evyllaesus* (Hymenoptera: Halictidae): a phylogenetic approach. *Behavioural*  
727 *Ecology and Sociobiology* 29, 153-160.
- 728 Page, R. D. M. 1990. Component analysis: A valiant failure? *Cladistics* 6, 119-136.
- 729 Page, R. D. M. 1992. COMPONENT user manual (Release 2.0). Natural History Museum,  
730 London.
- 731 Paterson, A. M., Gray, R. D., Wallis, G. P. 1993. Parasites, petrels and penguins: Does louse  
732 phylogeny reflect seabird phylogeny? *International Journal of Parasitology* 23, 515-  
733 526.
- 734 Paterson, A. M., Wallis, G. P., Gray, R. D. 1995. Penguins, petrels and parsimony: Does  
735 cladistic analysis of behaviour reflect seabird phylogeny? *Evolution* 49, 974-989.
- 736 Penny, D., Foulds, L. R., Hendy, M. D. 1982. Testing the theory of evolution by comparing  
737 phylogenetic trees constructed from five different protein sequences. *Nature* 297, 197-  
738 200.
- 739 Price, J. J., Lanyon, S. M. 2002. Reconstructing the evolution of complex bird song in the  
740 oropendolas. *Evolution* 56, 1514–1529.
- 741 Proctor, H. C. 1992. Sensory exploitation and the evolution of male mating behaviour: A  
742 cladistic test using water mites (Acari: Parasitengona). *Animal Behaviour* 44, 745-  
743 752.
- 744 Prum, R. O. 1990. Phylogenetic analysis of the evolution of display behaviour in the  
745 neotropical manakins (Aves: Pipridae). *Ethology* 84, 202–231.
- 746 Robertson, C. J. R. 1985. Reader's Digest complete book of New Zealand birds. Read  
747 Methuen, Sydney.
- 748 Robillard, T., Höbel, G., Gerhardt, H. C. 2006. Evolution of advertisement signals in North  
749 American hylid frogs: vocalizations as end-products of calling behavior. *Cladistics* 22,  
750 533–545.

- 751 Saunders, D. 1971. Seabirds. Hamlyn Press, London.
- 752 Schluter, D., Price, T., Mooers, A. O., Ludwig, D. 1997. Likelihood of ancestor states in  
753 adaptive radiation. *Evolution* 51, 1699–1711.
- 754 Scholes, E. 2008. Evolution of the courtship phenotype in the birds of paradise genus *Parotia*  
755 (Aves: Paradisaedae): homology, phylogeny, and modularity. *Biological Journal of the*  
756 *Linnean Society* 94, 491–504.
- 757 Schreiweis, D. O. 1972. A comparative study of the appendicular musculature of the Order  
758 Sphenisciformes. Washington State University., Pullman.
- 759 Shimodaira, H., Hasegawa, M. 1999. Multiple comparisons of log-likelihoods with  
760 applications to phylogenetic inference. *Molecular Biology and Evolution* 16, 1114–  
761 1116.
- 762 Sibley, C. G., Ahlquist, J. E. 1990. Phylogeny and classification of birds: A study in  
763 molecular evolution. Yale University Press, New Haven.
- 764 Sillen-Tullberg, B., Møller, A. P. 1993. The relationship between concealed ovulation and  
765 mating systems in anthropoid primates: A phylogenetic analysis. *American Naturalist*  
766 141, 1-25.
- 767 Simpson, G. G. 1975. Fossil penguins. In: Stonehouse, B. (Ed.) *The Biology of Penguins*.  
768 MacMillan Press, London, pp. 19–41.
- 769 Simpson, G. G. 1976. *Penguins: Past and present, here and there*. Yale University Press, New  
770 Haven.
- 771 Slack, K. E., Jones, C. M., Ando, T., Harrison, G. L. A., Fordyce, R. E., Arnason, U., Penny,  
772 D. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution.  
773 *Molecular Biology and Evolution* 23, 1144–1155.
- 774 Slikas, B. 1998. Recognizing and testing homology of courtship displays in storks (Aves:  
775 Ciconiiformes: Ciconiidae). *Evolution* 52, 884-893.
- 776 Soper, M. F. 1976. *New Zealand birds*. Whitcoulls Limited, Christchurch.



777 Stevens, G. R. 1989. The nature and timing of biotic links between New Zealand and  
778 Antarctica in Mesozoic and early Cenozoic times. In: Crame, J. A. (Ed.) *Origins and*  
779 *Evolution of the Antarctic Biota*. Geological Society Special Publication, pp. 141–166.

780 Stonehouse, B. 1975. *Penguin Biology*. MacMillan Press, London.

781 Stuart, A. E., Currie, D. C. 2001. Using caddisfly (Trichoptera) case-building behaviour in  
782 higher level phylogeny reconstruction. *Canadian Journal of Zoology* 79, 1842-1854.

783 Stuart, A. E., Currie, D. C. 2002. Behaviour is not reliably inferred from end-product structure  
784 in Caddisflies. *Ethology* 108, 837–856.

785 Stuart, A. E., Hunter, F. F. 1998. End-products of behaviour versus behavioural characters: A  
786 phylogenetic investigation of pupal cocoon construction and form in some North  
787 American black flies (Diptera: Simuliidae). *Systematic Entomology* 23, 387-398.

788 Stuart, A. E., Hunter, F. F., Currie, D. C. 2002. Using behavioural characters in phylogeny  
789 reconstruction. *Ethology, Ecology & Evolution* 14, 129–139.

790 Swofford, D. L. 2002. PAUP\*. *Phylogenetic Analysis Using Parsimony (\*and Other*  
791 *Methods)*. Sinauer Associates, Sunderland, Massachusetts.

792 Swofford, D. L., Maddison, W. P. 1987. Reconstructing ancestral character states under  
793 Wagner parsimony. *Math. Biosci.* 87, 199–229.

794 Templeton, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site  
795 maps with particular reference to the evolution of humans and the apes. *Evolution* 37,  
796 221–244.

797 Tinbergen, N. 1959. Comparative studies of the behaviour of gulls (Laridae): A progress  
798 report. *Behaviour* 15, 1-70.

799 Trewick, S. A., Paterson, A. M., Campbell, H. J. 2007. Hello New Zealand. *Journal of*  
800 *Biogeography* 34, 1–6.

801 Waas, J. R. 1991. Intraspecific variation in social repertoires: Evidence from cave- and  
802 burrow-dwelling Little Blue Penguins. *Behavior* 115, 63–126.

- 803 Wagner, G. P. 1989. The origin of morphological characters and the biological basis of  
804 homology. *Evolution* 43, 1157–1171.
- 805 Warham, J. 1974. The Fiordland Crested Penguin, *Eudyptes pachyrhynchus*. *Ibis* 116, 1–27.
- 806 Watson, G. E. 1975. *Birds of the Antarctic and Subantarctic*. William Byrd Press, Richmond.
- 807 Weimerskirch, H., Stahl, J. C., Jouventin, P. 1992. The breeding biology and population  
808 dynamics of King Penguins *Aptenodytes patagonicus* on the Crozet Islands. *Ibis* 134,  
809 107–117.
- 810 Williams, A. J. 1981a. The clutch size of macaroni and rockhopper penguins. *Emu* 81, 87–90.
- 811 Williams, A. J. 1981b. The laying interval and incubation period of rockhopper and macaroni  
812 penguins. *Ostrich* 52, 226–229.
- 813 Winkler, D. W. 2000. The phylogenetic approach to avian life histories: An important  
814 complement to within-population studies. *Condor* 102, 52–59.
- 815 Winkler, D. W., Sheldon, F. W. 1993. Evolution of nest construction in swallows  
816 (Hirundinidae): A molecular phylogenetic perspective. *Proceedings of the National*  
817 *Academy of Sciences, USA* 90, 5705-5707.
- 818 Zink, R. M., Avise, J. C. 1990. Patterns of mitochondrial DNA and allozyme evolution in the  
819 avian genus *Ammodramus*. *Systematic Zoology* 39, 148-161.
- 820 Zyskowski, K., Prum, R. O. 1999. Phylogenetic analysis of the nest architecture of neotropical  
821 ovenbirds (Furnariidae). *Auk* 116, 891-911.

822

823

824 Table 1

825 Matrix showing behavioral and life history character data for all taxa

826			1	2	3	4	5	6
827			123456789012345678901234567890123456789012345678901234567890123456789012345678					
828	<i>Larus dominicanus</i>	Southern black-backed gull	000000?00??0?000000000000001?000000?00000000000?000?00?000000000??0?					
829	<i>Aptenodytes forsteri</i>	Emperor penguin	021010101?0200130000101000110301104111001011122111101111211230101210					
830	<i>Aptenodytes patagonicus</i>	King penguin	021010111?0100130201101000110311003111001011020111001111111120101210					
831	<i>Pygoscelis antarctica</i>	Chinstrap penguin	13101?00000001?1?11111110?0121?1112000??000011??141?11112112010001?1					
832	<i>Pygoscelis adeliae</i>	Adelie penguin	031010000100011101111211000?210111010010110011?00201121120122?1001?1					
833	<i>Pygoscelis papua</i>	Gentoo penguin	03111000010000?20??1????001?240111100012000001?0020?11112?321?100110					
834	<i>Spheniscus demersus</i>	Black-footed penguin	121110000?20102112?11211100?221111?0001100211??1?40?1111204340?00?11					
835	<i>Spheniscus magellanicus</i>	Magellanic penguin	121010?0??2110?0????0???01?141111?100??02202??1041???11202200100111					
836	<i>Eudyptula minor</i>	Little blue penguin	121210100101001111112100010?121111000000002021?00401111120321001?211					
837	<i>Megadyptes antipodes</i>	Yellow-eyed penguin	12111100?0000021111132000?103110110100110000211114111101201210110011					
838	<i>Eudyptes pachyrhynchus</i>	Fiordland penguin	11101000??11002101113111100131111100000112002111031?1211203210012120					
839	<i>Eudyptes robustus</i>	Snares penguin	131010?0?11100211?11311110113111112?0001120121?10311121120200?002120					
840	<i>Eudyptes sclateri</i>	Erect-crested penguin	131010?0?1?1002111113111??0?3111112?0001120100110310111120200?00?11?					

841	<i>Eudyptes chrysocome</i>	Rockhopper penguin	13101000?10?002?111113?111111341111200011120021?1031?111120200?002110
842	<i>Eudyptes chrysolophus</i>	Macaroni penguin	13101001010?0?2?1?1131?110013511??2?001?1000111?0310111120200?102110
843	<i>Eudyptes schlegeli</i>	Royal penguin	131?1001?10?0?211111311?10013111112?0011100111?1031?11?1202000002111

---

844 Missing data indicated by "?". Characters and character states are identified in Appendix 1.

845

846 Table 2

847 Results of Kishino-Hasegawa tests comparing other estimates of phylogeny of the penguins (e.g., see Fig. 1) to those derived from the BLH data

848	Tree	Length	Length diff	s.d.(diff)	<i>t</i>	<i>P</i> (one tailed)	
849	<i>16 taxon comparison</i>						
850	BLH-dollo	194	4	4.00000	1.0000	0.1605	NS
851	Baker et al.	200	10	5.74586	1.7404	0.0432	
852	Giannini and Bertelli	206	16	6.06433	2.6384	0.0052	
853	Jouventin	198	8	6.29617	1.2706	0.1042	NS
854	Livezey (E)	199	9	5.11801	1.7585	0.0416	
855	Livezey (S)	205	15	5.84757	2.5652	0.0063	
856	<i>13 taxon comparison</i>						
857	Schreiweis ( <i>S.d.</i> )	184	9	4.25143	2.1169	0.0190	
858	Schreiweis ( <i>S.m.</i> )	182	9	4.91571	1.8309	0.0358	
859	<i>7 taxon comparison</i>						
860	McKittrick ( <i>S.d.</i> )	142	11	3.93036	2.7987	0.0034	
861	McKittrick ( <i>S.m.</i> )	133	8	4.16094	1.9226	0.0294	

862	OHara ( <i>S.d.</i> )	136	5	5.00000	1.0000	0.1605	NS
863	OHara ( <i>S.m.</i> )	129	4	4.00000	1.0000	0.1604	NS
864	<i>6 taxon comparison</i>						
865	Sibley and Ahlquist ( <i>S.d.</i> )	110	(best)				NS
866	Sibley and Ahlquist ( <i>S.m.</i> )	108	(best)				NS

---

867 Length = length of tree, Diff = difference from shortest tree, s.d.(diff) = standard deviation of difference,  $t$  = Kishino-Hasegawa test score,  $P$   
868 (one-tailed) = probability that the tree topology is significantly different from the shortest tree, Livezey (E) = external morphology, Livezey (S)  
869 = skeletal morphology. For the 16 taxon comparisons all taxa were included (the taxa represented in the previous estimates of phylogeny at only  
870 a generic level had no more than two species, and were represented as monophyletic). For the comparisons with less than 16 taxa the genera  
871 were collapsed to single exemplars. Because *Spheniscus* is not monophyletic in the BLH trees two comparisons were made. The first of these  
872 comparisons (labelled *S.d.*) used *S. demersus* as the exemplar for *Spheniscus*, whereas the second comparison (labelled *S.m.*) used *S.*  
873 *magellanicus* as the exemplar for *Spheniscus* in the BLH trees.  
874

875 Table 3

876 Results of Kishino-Hasegawa tests comparing other estimates of phylogeny of the penguins (e.g., see Fig. 1) to that derived from the BLH data  
877 with dollo coded characters

878	Tree	Length	Length diff	s.d.(diff)	<i>t</i>	<i>P</i> (one tailed)	
879	<i>16 taxon comparison</i>						
880	BLH-1	198	2	5.32524	0.3756	0.3542	NS
881	BLH-2	199	3	5.94879	0.5043	0.3079	NS
882	Baker et al.	211	15	7.78345	1.9272	0.0291	
883	Giannini and Bertelli	220	24	8.63868	2.7782	0.0036	
884	Jouventin	210	14	8.12955	1.7221	0.0449	
885	Livezey (E)	206	10	6.41360	1.5592	0.0619	NS
886	Livezey (S)	214	18	7.35050	2.4488	0.0085	
887	<i>13 taxon comparison</i>						
888	Schreiweis ( <i>S.d.</i> )	194	16	6.39029	2.5038	0.0074	
889	Schreiweis ( <i>S.m.</i> )	191	13	6.08767	2.1355	0.0182	
890	<i>7 taxon comparison</i>						

891	McKittrick ( <i>S.d.</i> )	146	13	4.78025	2.7195	0.0042	
892	McKittrick ( <i>S.m.</i> )	139	12	6.36220	1.8861	0.0318	
893	OHara ( <i>S.d.</i> )	141	8	6.13286	1.3044	0.0983	NS
894	OHara ( <i>S.m.</i> )	134	7	5.35738	1.3066	0.0979	NS
895	<i>6 taxon comparison</i>						
896	Sibley and Ahlquist ( <i>S.d.</i> )	112	(best)				NS
897	Sibley and Ahlquist ( <i>S.m.</i> )	110	1	3.89987	0.2564	0.3992	NS

---

898 Details as for Table 3.

899



900 Table 4

901 Results of Shimodaira-Hasegawa tests comparing estimates of phylogeny of the penguins (e.g., see Fig. 1) to that derived from the sequence data  
902 (Baker et al., 2006)

903	Tree	-ln L	Diff -ln L	<i>P</i>
904	<i>16 taxon comparison</i>			
905	BLH-1	16164.48336	560.67018	<0.001
906	BLH-2	16183.14848	579.33530	<0.001
907	BLH-dollo	16246.38993	642.57675	<0.001
908	Giannini and Bertelli	15754.26199	150.44881	<0.001
909	Jouventin	15696.39594	92.58276	0.047
910	Livezey (E)	15728.82132	125.00814	0.002
911	Livezey (S)	15847.73005	243.91687	<0.001
912	<i>14 taxon comparison</i>			
913	Schreiweis	15496.07268	241.91060	<0.001
914	<i>7 taxon comparison</i>			
915	McKittrick	13052.44033	86.51622	<0.001

916	OHara	13038.41343	72.48932	<0.001	
917	<i>6 taxon comparison</i>				
918	Sibley and Ahlquist	12489.93162	0.00000	0.517	NS
919	<hr/> Shimodaira-Hasegawa test using RELL bootstrap (one-tailed test) with 1000 bootstrap replicates.				
920					

921 Table 5

922 Individual CIs and RIs of behavioral display characters mapped (using ACCTTRAN) onto the  
 923 molecular tree (Baker et al., 2006)

				(i) gains = losses		(ii) dollo	
				minimum		minimum	
Character	CI	RI	gains	losses	gains	losses	
14 gaping	0.50	0.00	2	0	1	1	
15 shoulder hunching	0.67	0.67	1	0	1	0	
16 trumpeting	0.75	0.50	2	0	1	1	
17 bills interlocked	0.50	0.80	1	1	1	1	
18 stare	0.67	0.50	2	0	1	1	
19 charge	1.00	1.00	1	0	1	0	
20 bite nape fight	0.50	0.00	2	0	1	1	
21 hunched submissive	1.00	1.00	1	0	1	0	
22 face away	0.40	0.40	1	1	1	1	
23 shivering	0.33	0.00	1	2	1	2	
24 squeal	0.33	0.50	3	0	1	2	
25 bill hiding	0.50	0.80	2	0	1	2	
30 ecstatic	0.71	0.50	1	0	1	0	
31 allopreening	0.50	0.67	2	0	1	2	
32 mutual bowing	0.50	0.00	1	1	1	1	
33 beating wings	0.50	0.00	1	1	1	1	
34 treading	1.00	1.00	1	0	1	0	
35 head shake and whine	0.67	0.50	3	1	1	4	
37 head circling	1.00	1.00	1	0	1	0	

946	38	ear rubbing	1.00	1.00	1	0	1	0
947	39	stone carrying	0.25	0.50	1	3	1	3
948	40	quivering	0.67	0.75	2	1	1	2

---

949 The inferred number of times each character arose or was lost (ignoring multiple states in  
950 multi-state characters) on the molecular tree was calculated assuming (i) gains were as likely  
951 as losses (i.e., the minimal number of gains and losses combined), or (ii) dollo, where each  
952 character arose only once (i.e., the minimal number of losses). In some instances more than  
953 one combination of gains and losses could be inferred (i.e. give the same minimal number of  
954 combined gains and losses). Thus, the two scenarios, (i) and (ii), sometimes give the same  
955 minimal score.

956

957 Fig. 1. Published hypotheses of penguin relationships among their genera and among species  
958 within the crested penguins, *Eudyptes* and *Pygoscelis* (i.e., trimmed to include the same set of  
959 taxa as the BLH dataset). The studies and their data types are (a) Baker et al. (2006) molecular  
960 tree, (b) Giannini and Bertelli (2004) integumentary and breeding and Bertelli and Giannini  
961 (2005) general morphology, (c) Jouventin (1982) behavior, (d) Livezey (1989) external  
962 morphology, (e) Livezey (1989) skeletal morphology, (f) McKittrick (1991) myology and  
963 morphology, (g) O'Hara (1989) morphology, (h) Schreiweis (1972) myology, and (i) Sibley  
964 and Ahlquist (1990) DNA-DNA hybridization.

965

966 Fig. 2. Two most parsimonious trees generated from unordered BLH data for 15 penguin  
967 species and one outgroup. The bootstrap values (>50%) are shown.

968

969 Fig. 3. The single most parsimonious tree generated from BLH data with 15 out of the 68  
970 characters dollo coded for 15 penguin species and one outgroup. The bootstrap values  
971 (>50%) are shown.

972

973 Fig. 4. Frequency histogram of the retention index for each of the BLH characters when  
974 mapped onto the molecular phylogeny (autapomorphies excluded). The grey proportion of  
975 each bar indicates the display characters, whereas the black region of each bar indicates the  
976 non-display characters.

977

978 Fig. 5. A majority rule consensus tree of the 268 shortest trees derived from BLH display data.  
979 The bootstrap values (>50%) are shown. The percentage of 268 shortest trees that contain  
980 that component are shown in italics.

981

982 Fig. 6. A majority rule consensus tree of the 297 shortest trees derived from BLH non-display  
983 data. The bootstrap values (>50%) are shown. The percentage of 297 shortest trees that  
984 contain that component are shown in italics.

985

986 Fig. 7. Three BLH characters (a) treading, (b) stare, (and c) bill hide mapped onto a simplified  
987 penguin phylogeny. Two possible scenarios are presented for each character: independent  
988 gains as likely as losses (circles = independent gains), and independent gains less likely than  
989 losses (black squares = gain of character; open squares = loss of character).

990

991

992 **Appendix 1: Behavioral and Life History Characters used in the Analysis**

993 Most of the information was derived from Marchant and Higgins (1990) except *Larus*  
994 *dominicanis* (L.d.) (Paterson et al., 1995), *S. mendiculus* (Stonehouse, 1975), and *S. demersus*  
995 (Eggleton and Siegfried, 1979). Information on display characters in all penguins from  
996 Eggleton and Siegfied (1979) and *Eudyptes pachyrhynchus* and *Eudyptula minor* from  
997 Warham (1974) and Waas (1991) respectively. Additional information on character states was  
998 obtained from \*del Hoyo et al. (1992), †Costa (1991), ‡Weimerskirch et al. (1992), and  
999 ♂Williams (1981a; 1981b). Where appropriate taxa are labelled by their genus and species  
1000 initials (e.g., *S. demersus* as *S.d.*).

1001

- 1002 1. Chicks fed each day during first weeks of nestling stage: (0) yes, (1) no.
- 1003 2. Predominant food (% numbers and mass) in diet: (0) crustaceans, (1) cephalopods, (2)  
1004 fish, (3) molluscs or scavenges.
- 1005 3. Main method of food capture: (0) browsing, (1) pursuit diving.
- 1006 4. Time of return from foraging: (0) any, (1) dusk, (2) after dark.
- 1007 5. Land directly at nest-site: (0) yes, (1) no.
- 1008 6. No vocalizations at sea while foraging: (0) no vocalisations, (1) vocalisations.
- 1009 7. Average weight of chick's meal / average weight of adult: (0) <0.1, (1) >0.1 [†*S.d.*,  
1010 *Eu.m.*].
- 1011 8. Adults forage inshore during breeding season: (0) yes, (1) no.
- 1012 9. Time at sea foraging: (0) 24 hours or less, (1) greater than 24 hours [†all].
- 1013 10. Method of regurgitation: (0) complete, (1) incomplete.
- 1014 11. Foraging: (0) in flocks, (1) solo, (2) cooperative foraging [\*all].
- 1015 12. Territory defended: (0) nest/burrow, (1) area around nest/burrow (2) no territory.
- 1016 13. Predawn chorus: (0) no, (1) yes.

- 1017 14. Gaping and open yell: (0) not present, (1) neck stretched more or less horizontal, gape,  
1018 carpal joints raised.
- 1019 15. Shoulder hunching: (0) not present, (1) body horizontal and neck stretched to varying  
1020 extent, bill horizontal or slightly up, opened or closed, carpal joints raised, call, (2) neck  
1021 lowered, head and bill directed forward, shoulders raised, body leaning forward, wings  
1022 held slightly forward.
- 1023 16. Trumpeting: (0) bill open, directed down and forward, moves towards intruder,  
1024 repetitive call not present, (1) call, bird leans forward, raises wings, steps towards  
1025 intruder, (2) neck stretched in direction of intruder, bill opening and closing, call, (3) not  
1026 present.
- 1027 17. Bills interlocked: (0) not present, (1) birds lock bills together and wrestle.
- 1028 18. Stare: (0) not present, (1) bird in upright posture, turns body and bill towards opponent,  
1029 feathers sometimes erected, (2) bird faces opponent, wings raised, bill straight out.
- 1030 19. Charge: (0) not present, (1) defending bird moves towards opponent with short run,  
1031 upright head forward, crown erect, wings raised.
- 1032 20. Bite nape fight: (0) not present, (1) bill strikes or nips feathers and skin of opponents  
1033 head and neck, swatting of opponent with wing.
- 1034 21. Hunched submissive: (0) not present, (1) body stretched up, neck elongated, wings held  
1035 out from body, feathers sleeked, (2) body and head held low, wings held close to sides,  
1036 steps, (3) body and head held low, wings extended, feathers sleeked, steps.
- 1037 22. Face away: (0) not present, (1) bird looks obliquely at owners of other territories, wings  
1038 forward, bill upwards and to one side, (2) body and neck held upright, wings forward,  
1039 bill hung down parallel to neck, head may turn to one side.
- 1040 23. Shivering: (0) not present, (1) wings and sometimes head vibrate rapidly in agonistic  
1041 situations.
- 1042 24. Squeal: (0) not present, (1) bird utters high pitched squeal in response to sudden danger.



- 1043 25. Bill hiding: (0) not present, (1) female crouches over nest and lowers bill until hidden  
1044 under body.
- 1045 26. Nest /burrow-site chosen by: (0) male, (1) either.
- 1046 27. Average length of incubating shifts, male : female: (0) equal, (1) male longer shift.
- 1047 28. Males present during laying: (0) no, (1) yes [‡A.p.].
- 1048 29. Method by which nest/burrow constructed: (0) plant and earth material patted onto nest  
1049 wall, (1) burrow dug using bill and feet, (2) adult (generally male) lying in nest scrape  
1050 kicks out of bowl, shifting material to the rim, (3) adult (generally female) squats with  
1051 wings extended and resting on ground, pressing back with feet and rotates.
- 1052 30. Ecstatic: (0) not present, (1) bird stands on toes, wings held stiffly forward, head and  
1053 open bill held vertical, loud trilling call, (2) bird stands on toes, wings held stiffly  
1054 forward, head and open bill held vertical, loud trilling call, head swung in arcs, (3) bird  
1055 stands on toes, head and bill held vertical, followed by head waving, (4) bird stands on  
1056 toes, flippers held down, head and open bill held vertical, loud trilling call, (5) bird  
1057 stands on toes, flippers held down, head and open bill held vertical, loud trilling call,  
1058 head swung in arcs.
- 1059 31. Allopreening: (0) not present, (1) neck, throat and head of mate preened.
- 1060 32. Mutual bowing: (0) not present, (1) members of pair direct open bills downwards,  
1061 usually into nest-bowls, and call.
- 1062 33. Beating wings: (0) not present, (1) male wings vibrated gently on females sides prior to  
1063 copulation.
- 1064 34. Treading: (0) not present, (1) male treads on back of female prior to copulation.
- 1065 35. Head shake and whine: (0) not present, (1) head rapidly waved from side to side with  
1066 bill closed and pointing at other bird, stands on toes with erect fanned tail, whining call,  
1067 wings stretched, (2) rub heads together, calling, (3) birds face each other and head

- 1068 waved side to side, punctuated by short calls. (4) birds face each other slowing raising  
1069 while contracting neck muscles.
- 1070 36. Copulation outside nest/burrow: (0) no, (1) yes.
- 1071 37. Head circling: (0) not present, (1) head thrown back, and then moved forward in lateral,  
1072 half circular motions, while uttering groans.
- 1073 38. Ear rubbing: (0) not present, (1) birds rub ear against shoulder.
- 1074 39. Stone carrying: (0) not present, (1) stones, grass, or earth collected and placed around  
1075 nest (generally male).
- 1076 40. Quivering: (0) not present, (1) nesting bird vibrates bill in very small arcs as it bows  
1077 over nest-bowl to deposit nest material, (2) bird opens bill and hisses as nest material is  
1078 deposited.
- 1079 41. Most adults remain near breeding site during year: (0) yes, (1) no.
- 1080 42. Type of migration used by species during life cycle: (0) coastal or no migration, (1) pack  
1081 ice region, (2) north of Antarctic convergence.
- 1082 43. Nester or burrower (nesting in burrows): (0) nest, (1) neither, (2) either.
- 1083 44. Predominant burrow/nest-site terrain: (0) cliff, (1) flat.
- 1084 45. Vegetation present at nest/burrow site: (0) no vegetation, (1) tussock or grassland, (2)  
1085 forest.
- 1086 46. Pair-bond duration: (0) several breeding seasons, (1) lifelong, (2) one breeding season.
- 1087 47. Sex ratio in adult life: (0) equal, (1) more males, (2) more females.
- 1088 48. Chicks sometimes abandoned to die by parents: (0) yes, (1) no.
- 1089 49. Nest-site fidelity: (0) yes, (1) no [*P.p.*, *E.c.*].
- 1090 50. Clutch size: (0) more than two, survival approximately equal for each egg/chick, (1)  
1091 one, (2) two, survival of first egg/chick better, (3) two, survival of second egg/chick  
1092 better, (4) two, survival approximately equal for each egg/chick.
- 1093 51. Replacement laying of eggs in same season as failure of first nest: (0) yes, (1) no.

- 1094 52. Sexual differences in voice: (0) yes, (1) no.
- 1095 53. Nestling state of development: (0) altricial, (1) semi-altricial.
- 1096 54. Nestling behavior: 0 semi-nidicolous, (1) nidicolous, (2) nudifugous.
- 1097 55. Average fledging weight/average adult weight: (0) >1, (1) <1 [ $\dagger S.d$ ,  $\dagger E.c.$ ].
- 1098 56. Moults halts foraging: (0) no, (1) yes.
- 1099 57. General foraging habitat: (0) marine, (1) pelagic, (2) intertidal.
- 1100 58. Adults leave young before fledging: (0) no, (1) yes.
- 1101 59. Season that breeding commences: (0) spring, (1) summer, (2) autumn, (3) winter (4) no  
1102 set time.
- 1103 60. Month of main moult: (0) September, (1) November, December, (2) February, (3) no set  
1104 time.
- 1105 61. Season of nest construction: (0) winter, (1) spring, (2) summer, (3) autumn, (4) no set  
1106 time.
- 1107 62. Ratio of age first breeding/first return to breeding colonies: (0) ~1, (1) 2+.
- 1108 63. Ratio of nest preparation period/incubation length: (0) >0.66, (1) <0.66 [ $\text{all}$ ].
- 1109 64. Ratio of nestling period/incubation length: (0) approx equal, (1) incubation period  
1110 shorter than nestling.
- 1111 65. Laying interval between eggs: (0) less than four days, (1) no laying interval, (2) four  
1112 days or more [ $\text{all}$ ].
- 1113 66. Age of creching: (0) no creching, (1) less than 28 days, (2) 28 days or more [ $\text{all}$ ].
- 1114 67. Length of premoult forage: (0) no premoult forage, (1) 3-5 weeks, (2) >5 weeks.
- 1115 68. Chicks fledged/pair: (0) 0-0.9 [ $\text{all}$ ], (1) 1.0-1.9.
- 1116