

1 **Allopatric montane wren-babblers exhibit similar song notes but divergent vocal
2 sequences**

3

4 M. Abhinava Jagan¹, Ananda Shikhara Bhat¹ and Anand Krishnan^{2, #}

5

6 1- Department of Biology, Indian Institute of Science Education and Research (IISER) Pune,
7 Pashan Road, Pune 411008, India.

8 2- Department of Biological Sciences, Indian Institute of Science Education and Research
9 (IISER) Bhopal, Bhauri 462066, Madhya Pradesh, India

10 #- Email: anandk@iiserb.ac.in

11

12 **Running title:** Evolution of vocal sequences

13 **Keywords:** Syntactic structure; vocal sequences; geographic barriers; montane birds; wren-
14 babblers

15

16 **Abstract**

17

18 The songs of passerine birds consist of notes temporally arranged into vocal sequences
19 following syntactic structures, and function both in courtship and territorial defense. Geographic
20 barriers are important drivers of avian speciation, and also influence the divergence of song.
21 However, there is relatively little quantitative study of the relationship between geographic
22 barriers and the syntactic structure of vocal sequences. Here, we investigate interspecific
23 divergence in song notes and syntax within a genus of allopatric montane Asian wren-babblers
24 (*Spelaeornis*). Employing a robust quantitative analytical framework and song recordings from
25 publicly accessible databases, we find that *Spelaeornis* appear to have undergone
26 diversification in song syntax without divergence in note parameters. Broadly, we find three
27 different syntactic structures across the eight species in the genus, each occurring in a different
28 geographic region in Asia, with two species apparently exhibiting intermediate syntax. Species
29 within the genus appear to possess similar song notes, but subgroups confined to different
30 geographic regions (eg: hills south of the Brahmaputra river) arrange these notes according to
31 different syntactic rules to construct songs. Our computational framework to examine signal
32 structure and diversification across multiple scales of signal organization has potential
33 implications for our understanding of speciation, signal evolution and more broadly in fields such
34 as linguistic diversification.

35

36

37

38

39

40

41

42

43

44

45

46 Introduction

47

48 Allopatric speciation, the process underlying the formation of new species by geographic
49 isolation, has long been a subject of research interest (Price 2007). Geographic barriers such as
50 rivers, mountains, oceans or valleys can isolate populations of a species, thereby preventing
51 gene flow. This reproductive isolation, over time, results in accumulation of genetic differences,
52 such that the two species remain isolated from each other on secondary contact (Irwin et al.
53 2001). The initial reproductive isolation may be reinforced by pre-mating isolation mechanisms
54 such as temporal separation of breeding periods (Hillis 1981; Marshall and Cooley 2000),
55 divergent habitats (ecological isolation) (Egan and Funk 2009), morphological differences
56 (mechanical isolation) (Richmond et al. 2011) and behavioral differences (behavioral isolation)
57 (Marler and Slabbekoorn 2004; Uy et al. 2018). Of these, the last is particularly relevant to
58 organisms that learn behavioral traits, because these learned traits may diverge relatively
59 rapidly over evolutionary time either for adaptive reasons or through neutral evolutionary
60 processes (Lachlan and Servedio 2004; Yeh and Servedio 2015). Understanding allopatric
61 speciation thus requires a comprehensive understanding of how behaviors and sensory signals
62 vary among close relatives and across geographic barriers.

63

64 Perhaps the best-studied learned signaling traits in this regard are the complex, culturally
65 transmitted songs of passerine birds. This learned behavior exhibits complex structures that,
66 among other communicative functions, support species recognition and minimize interspecific
67 hybridization (Marler and Slabbekoorn 2004; Bradbury and Vehrencamp 2011). Each song
68 consists of individual vocal units - notes or syllables, which are ordered in complex temporal
69 arrangements to form vocal sequences. This temporal arrangement is referred to as the syntax
70 of the song (Marler and Peters 1988; Kershenbaum et al. 2012; Fishbein et al. 2020). Syntax
71 may follow a number of rules, including repetitive songs, repetitions of specific motifs, or highly
72 complex songs consisting of a number of note types. Both the acoustic structure of individual
73 notes and the syntactic structure of vocal sequences play an important role in mate attraction,
74 species recognition, and communication of context in diverse vertebrate taxa (Peters et al.
75 1980; Searcy et al. 1981; Searcy and Marler 1981; Balaban 1988; Marler and Peters 1988; Isler
76 et al. 1998; Nowicki et al. 2001; Wanker and Fischer 2001; Charrier and Sturdy 2005; Dahlin
77 and Wright 2012; Kershenbaum et al. 2012; Briefer et al. 2013; Kershenbaum et al. 2016;
78 Suzuki et al. 2018; Ciaburri and Williams 2019; Engesser and Townsend 2019; Leroux et al.
79 2021; Bhat et al. 2022). Thus, it is important to study acoustic signals at multiple hierarchical
80 scales of organization to understand their evolution and function.

81

82 A number of studies have shown that geographic barriers exert potent influences on the
83 structure of bird song (Marler and Tamura 1962; Baker 2006; Baker et al. 2006; Podos and
84 Warren 2007; Irwin et al. 2008; Kirschenbaum et al. 2009; Grant and Grant 2010; Lachlan et al. 2013;
85 Lachlan et al. 2016). In white-crowned sparrows, geographically separated populations exhibit
86 different dialects (Marler and Tamura 1962), and the same pattern is observed in insular
87 populations of birds separated from the mainland (Baker et al. 2006; Lachlan et al. 2013). Song
88 structure could evolve by simple addition and deletion of new notes to the repertoire, through

89 changes in the syntactic rules which dictate how songs are constructed (i.e. increased repetition
90 of the same notes, use of certain motifs, etc.), or through a combination of both these processes
91 (Fig 1A). Quantitative studies that examine both notes and syntactic structure of song in
92 allopatric bird species are therefore needed to better understand the role of note change versus
93 sequence change in the evolution of song. Such an approach would enable us to examine how
94 syntax changes across geographic barriers in comparison to divergence in note structure. For
95 example, do geographically proximate species exhibit similar syntactic rules but different notes,
96 or different syntactic structures with or without note change? Given the parallels between bird
97 song and human language (Hunt 1923; Doupe and Kuhl 1999; Salwiczek and Wickler 2004;
98 Berwick et al. 2011; Miyagawa et al. 2013; Collier et al. 2014; Henry et al. 2015; Sainburg et al.
99 2019), such research also helps understand how geography potentially influences human
100 linguistic evolution.

101

102 *Spelaeornis* is a genus of montane understory birds from Asia containing eight species with
103 non-overlapping ranges (del Hoyo et al. 2014). All species of this genus are endemic to
104 montane forest from the Himalayas to South-East Asia (see Methods, Fig 1B) (IUCN 2019).
105 They thus present a good system to comparatively examine how the geographic barriers
106 between hill ranges have shaped the evolution of vocal sequences. Here, we quantitatively
107 examine the syntactic structure of vocal sequences as well as the parameters of song notes in
108 this genus. Specifically, we first ask whether individual song notes vary between species.
109 Secondly, we quantitatively analyze the syntactic structure of *Spelaeornis* songs to address
110 whether different allopatric species exhibit distinct vocal sequences, arranged according to
111 different rules. We also hypothesized that that geographically proximate species possess
112 similarities in their syntactic structure and vocal sequence content. Our study thus examines
113 how geographic barriers influence diversification in complex learned signals across multiple
114 organizational scales.

115

116 Materials and Methods

117

118 Study species

119 The genus *Spelaeornis* consists of eight species of allopatric montane wren-babblers- *S.*
120 *caudatus* (*S. ca*) endemic to the Eastern Himalayas, *S. badeigularis* (*S. ba*) of the Mishmi Hills
121 in India, *S. troglodytoides* (*S. tr*), that replaces other *Spelaeornis* species at higher elevations, *S.*
122 *chocolatinus* (*S. ch*), *S. oatesi* (*S. oa*), and *S. longicaudatus* (*S. lo*) of hill ranges south of the
123 Brahmaputra river, *S. kinneari* (*S. ki*) and *S. reptatus* (*S. re*) from Southeast Asia (Fig 1B)
124 (Rasmussen and Anderton 2005; Collar 2006; King and Donahue 2006; del Hoyo et al. 2014).
125 These allopatric species are endemic to restricted ranges, replacing each other in wet montane
126 forests. They therefore provide us with a good model system to examine the effect of
127 geographic barriers on interspecific divergence in acoustic signals.

128

129 Recordings

130 We sourced song recordings of the genus *Spelaeornis* from multiple online song databases
131 including Macaulay library (<https://www.macaulaylibrary.org/>), Xeno-Canto (<https://xenocanto.org/>), and AVoCet (<https://avocet.integrativebiology.natsci.msu.edu/>) (see Supplementary

132

133 Data 1 for a full list of recordings, including recordist and location). This dataset was curated to
134 include recordings that were longer than 10 seconds in duration with relatively low background
135 noise, and also to ensure that none of the recordings were duplicates from different databases.
136 For purposes of sample size, we considered a ‘recording’ as being all audio database files
137 recorded by the same recordist on the same day, and also the total number of songs measured
138 per species and the number of notes. Our total sample for analysis was thus as follows: *S. ca*:
139 1864 notes, 244 songs analyzed from 28 recordings; *S. ba*: 1358 notes, 253 songs from 14
140 recordings; *S. tr*: 1574 notes, 213 songs from 19 recordings; *S. ch*: 759 notes, 135 songs from
141 12 recordings; *S. re*: 622 notes, 50 songs from 4 recordings; *S. ki*: 904 notes, 62 songs from 8
142 recordings; *S. oa*: 1561 notes, 219 songs from 14 recordings; *S. lo*: 940 notes, 156 songs from
143 11 recordings.

144

145 *Comparison of note parameters*

146 Using Raven Pro Version 1.6 (Cornell Laboratory of Ornithology, Ithaca, NY, USA), we first
147 digitized individual notes of each species and calculated 10 parameters that described the
148 properties of each note: note duration, 90% bandwidth (difference between the frequencies
149 containing 90% of energy in the power spectrum), peak frequency, minimum and maximum
150 frequencies of the peak frequency contour, start and end frequencies of the peak frequency
151 contour (these combined giving parameters of note shape), center frequency (frequency of 50%
152 energy in the power spectrum), average entropy, and the time of maximum energy relative to
153 the start of the note (peak time relative). We then performed a Principal Components Analysis
154 (PCA) on the correlation matrix of these parameters, to examine interspecific note overlap of
155 *Spelaeornis* in signal space. To quantify this overlap, we trained a Linear Discriminant classifier
156 on the note parameters described above, using the Classification Learner app in MATLAB
157 (Mathworks Inc., Natick, MA, USA) with 10-fold cross-validation, and assessed the accuracy
158 with which notes were correctly assigned to the different species. We additionally performed a
159 statistical randomization test of whether species overlapped more or less than expected by
160 chance. For this, we randomly reshuffled data points across species, maintaining the sample
161 size for each individual species (Chek et al. 2003; Luther 2009; Schmidt et al. 2013; Krishnan
162 2019; Chitnis et al. 2020). For each of 1000 such randomized datasets, we calculated the
163 average interspecific Euclidean distance (for the first three principal components) in signal
164 space. Finally, we compared the observed average interspecific distance to the distribution
165 obtained from the randomized datasets by computing a Z-score. If the song notes of
166 *Spelaeornis* overlapped in signal space, the observed interspecific distance in the signal space
167 should be less than that expected by chance alone, thus resulting in a negative Z-score.

168

169 *Classifying note types and note groups*

170 Before analyses of song syntactic structure, we constructed sequences illustrating the temporal
171 structure of each song. This required us to first classify song note types for each species based
172 on differences in spectral structure and the parameters we measured (Fig 2A). We verified
173 these note type classifications in two ways: first, by measuring classification accuracy using a
174 linear discriminant classifier as detailed above, and secondly, by cross-verifying the note
175 classifications using a second observer in a double-blind paradigm. In a small number of cases,
176 the classifier incorrectly assigned visually dissimilar notes to the same group, and we manually

177 defined the notes as subclasses of the same note type in these cases. However, this was rare
178 (three instances), and the combination of methods generally enabled us to reliably classify note
179 types. We used these note types to calculate measures of song complexity (see below).
180 However, note type classifications did not permit interspecific comparisons of syntax as the note
181 types and diversity were not statistically comparable across species (subsequent analyses
182 required that all species be compared within the same state space). We therefore also grouped
183 notes into “note groups”, which were consistent across species, based on their spectral shape
184 (Fig 2A). This enabled direct interspecific comparison of song syntax, as notes with minor
185 variations in spectral structure or frequency were pooled into the same note group. As before,
186 we verified note group classifications in a double-blind paradigm (i.e., a second author also
187 scored note groups). Across species, we thus hierarchically grouped notes into nine note
188 groups, abbreviated as letters in the relevant figures and tables. These were: Ascending (a),
189 Descending (b), Ascending-Descending (c), Descending-Ascending (d), Ascending-
190 Descending-Ascending (e), Descending-Ascending-Descending (f), Notes with constant
191 frequency (k), Complex notes with more than two inflection points and a total duration of less
192 than 200 ms (g), Long complex notes with more than two inflection points and a total duration of
193 more than 200 ms (l). To ensure that we had adequately sampled all the note types/note
194 groups, we computed note accumulation curves for each species by randomizing the order of
195 digitized notes for all songs.
196

197 Syntactic analyses

198 To examine and compare vocal sequences and syntactic structures, we employed a series of
199 metrics calculated using both note types and note groups. Using the note type classifications,
200 we first calculated a metric of song complexity- the number of different note types in a song
201 divided by the total number of notes in the same song. For each song (sample sizes above) for
202 each species, we computed and compared the average values of song complexity for each
203 recording (the unit “recording” defined above) using a Kruskal-Wallis test with a post-hoc Dunn
204 test. Because further syntactic comparisons relied on species having similar sample spaces, we
205 used note groups as a unit in subsequent analyses for direct interspecific comparison (although
206 we also repeated these analyses using note types to ensure that the type of classification did
207 not influence the patterns observed). To examine song syntax in *Spelaeornis* we constructed
208 note group sequences for each song, for each species. We first modeled *Spelaeornis* song
209 structure as a first order Markov chain and constructed a transition probability matrix (M) for
210 each species, where the $i-j^{\text{th}}$ entry of the matrix, N_{ij} , represents the probability of transition from
211 note group i to note group j . To further validate our analysis, we employed a co-occurrence
212 metric which is free of Markovian assumptions (Bhat et al. 2022). This analysis is based on the
213 premise that not all vocal sequences in animals may be Markovian (Kershenbaum et al. 2014),
214 and it is thus important to cross-verify results without making underlying assumptions about the
215 processes involved. Here, for a given set of note group sequences, we define ${}^dC_{ij}$ as the
216 probability that the note group j occurs within $d-1$ notes of note group i . We computed ${}^dC_{ij}$ for
217 song sequences for each species. Next, we constructed artificial sequences where notes were
218 randomly distributed following a stationary distribution. In this stationary distribution, the
219 probability of a note occurring in a sequence was given by the proportion of occurrence of that
220 note in the species’ song. We generated 50 such artificial sequences to compute a robust

221 estimate of expected co-occurrence ${}^dE_{ij}$. Following this, we calculated ${}^dR_{ij}$, defined as the ratio of
222 ${}^dC_{ij}$ to ${}^dE_{ij}$ for each species. ${}^dR_{ij}$ is a measure of whether note group j occurs within d-1 notes of
223 note group i more or less often than expected by chance alone. ${}^dR_{ij} > 1$ suggests note group j
224 occurs within d-1 notes of note group i more often than expected by chance alone, and ${}^dR_{ij} < 1$
225 suggests that j occurs within d-1 notes of i less often than expected by chance alone. For the
226 analysis reported in the Results we used a d-value of 4, but also repeated our analysis for d-
227 values of 2 and 6 to ensure that our choice of d did not influence the observed patterns;
228 previous studies also indicate that the choice of d does not change co-occurrence patterns in
229 vocal sequences (Bhat et al. 2022). We chose to report results for a d-value of 4 because, on
230 average, songs of *Spelaeornis* consist of 4-8 notes and a distance of 4 thus captures
231 biologically meaningful note co-occurrence patterns.
232

233 To statistically compare transition probability matrices between species, it is necessary that all
234 the Markov chains being compared share the same state space. Therefore, we added a
235 pseudocount of 0.0001 to each entry of the 9x9 transition matrix of each species (thus ensuring
236 that all pairwise transition probabilities were non-zero, and the state space of every species
237 comprised the nine possible note groups defined earlier; for visualizations of transition
238 probabilities we used the original matrices). We then computed the frequency of occurrence for
239 each transition ij in each species' songs and added this to the pseudocount. This gave us a 9 x
240 9 transition probability matrix for each species. Next, we employed a homogeneity test that
241 computes a minimum discrimination information statistic (mdis) (Bhat et al. 2022). This statistic
242 tests for differences in the distributions of transition probabilities between samples (in our case
243 species) based on the Kullback-Leibler divergence (Kullback et al. 1962). The test statistic is
244 asymptotically distributed as a χ^2 distribution with $S(S - 1)(r - 1)$ degrees of freedom where S
245 is the number of states of the Markov chain (in our case, the number of note groups, i.e., 9) and
246 r is the number of different samples to be compared (in our case, number of species). We
247 computed this statistic for a comparison of all 8 species, as well as pairwise between any two
248 species.
249

250 Finally, to test for sequence similarity between species across geographic regions, we
251 computed the pairwise Levenshtein distances between the vocal sequences of each species
252 pair. We then computed a median Levenshtein distance for each species pair and constructed a
253 distance matrix D, where D_{ab} represents the median Levenshtein distance between the song
254 sequences of species a and b. Using this matrix, we constructed a distance dendrogram using
255 the unweighted pair group method with arithmetic mean (UPGMA). Species closer together on
256 this dendrogram possessed similar vocal sequence structures.
257

258 **Results**

259

260 *Spelaeornis* do not exhibit interspecific note parameter divergence

261 The first three principal components of 10 note parameters explained more than 77% of
262 variation in the notes of *Spelaeornis* (Table S1). PC1 loaded moderately positively on all
263 frequency parameters and loaded negatively on note duration and peak time. PC2 loaded

264 positively on note duration and negatively on minimum peak frequency. PC3 loaded positively
265 on peak time and exhibited weak negative loading on average entropy.

266

267 Broadly, *S. ca* and *S. tr* occupy opposite ends of the three-dimensional signal space
268 encompassing all the note parameters of *Spelaeornis*. All other species of *Spelaeornis* heavily
269 overlap with each other in between these extremes (Fig 2B). This suggests that the song notes
270 of *Spelaeornis* have not diverged in allopatry, and that their songs have thus retained similar
271 building blocks across their range. This assertion was also supported by a randomization test
272 (see Methods). We found that the average interspecific distance in note parameter space of
273 *Spelaeornis* is much lower (Z -score = -40.248, $p < 0.001$) than expected by chance, as
274 calculated from 1000 randomized datasets (See Methods). Linear Discriminant Analysis further
275 supports an overlap of notes in signal space. The LDA classifier correctly assigned notes to
276 different species with an accuracy of only 55.8% (Fig S1). The notes of *S. ca* and *S. tr*,
277 occupying opposite edges in *Spelaeornis'* signal space, were correctly classified at rates of
278 81.8% and 89.0% respectively. However, these were the only two species with high
279 classification rates: the classification accuracies for all other species were less than 60%
280 (Range: 11.6% for *S. re* – 55.2% for *S. oa*) (Fig S1). Taken together, these statistical and
281 quantitative analyses suggest that allopatric *Spelaeornis* do not exhibit interspecific divergence
282 in song note parameters. Instead, they exhibit more overlap in signal space than expected by
283 chance, consistent with different species using similar notes.

284

285 *Note types and note groups in Spelaeornis*

286 We identified multiple note types in each species, and verified these classifications using both
287 linear discriminant classifiers and independent cross-verification by a second author. In total, the
288 number of note types identified were as follows: *S. ca*: 6 note types (1 note type with 2
289 subclasses, see Methods); *S. ba*: 8 note types; *S. tr*: 21 note types (1 note type with 2
290 subclasses); *S. ch*: 8 note types; *S. re*: 9 note types; *S. ki*: 6 note types; *S. oa*: 9 note types; *S.*
291 *lo*: 13 note types (1 note type with 2 subclasses) (Fig S2). Classification accuracies for note
292 types ranged from 71.5% (*S. ca*) to 85.6% (*S. ba*). Additionally, we cross-verified our note type
293 classification through a second observer in a double-blind paradigm. We found an average
294 inter-observer agreement of 76.33% for classification of note types across species.

295

296 Because different species of *Spelaeornis* possess similar note parameters, as evinced by heavy
297 overlap in signal space, we grouped notes into note groups for interspecific comparison based
298 on spectral shape (see Methods). We identified the following note groups across species (letters
299 indicate abbreviations for note groups as per the key provided in the Methods):- *S. ca*:
300 a,b,c,d,e,f,g,k; *S. ba*: a,b,c,d,e,f,k; *S. tr*: a,b,c,d,e,f,g,k,l; *S. ch*: a,b,c,d,f; *S. re*: a,b,c,d; *S. ki*:
301 a,b,c,d,f,k; *S. oa*: a,b,c,d,e,f,g; *S. lo*: a,b,c,d,e,f,g,k. The average interobserver agreement in a
302 double-blind note group classification paradigm was 77.75%. It is important to note here that
303 many note groups are shared across species, consistent with our interpretation of the
304 interspecific overlap in signal space. The I note group, however, was unique to *S. tr* and was not
305 observed in other species. Accumulation curves for all species for both note types and note
306 groups were characterized by early saturation, providing evidence that our dataset reliably
307 sampled the note repertoires in *Spelaeornis* (Fig S3, S4).

308

309 *Spelaeornis species differ in song complexity*

310 The song complexity of *Spelaeornis* exhibited significant interspecific differences (Kruskal-Wallis
311 test, $H=57.05$, $df=7$, $p\text{-value}= 5.83e-10$). Qualitatively, we observed two groups, one consisting
312 of species possessing lower song complexity (fewer note types as a proportion of song length),
313 and the other comprising species with higher song complexity (higher number of note types as a
314 proportion of song length) (Fig 2C). We performed post-hoc pairwise comparisons using Dunn's
315 Test, and 10 out of 28 species pairs exhibited statistically significant differences in song
316 complexity (Table S2). These differences are represented as compact letter displays (a-d) in Fig
317 2C. *S. ki* differed significantly from *S. oa*, *S. ch*, *S. lo* and *S. ba*, all species with higher song
318 complexity (a), but not from other species. *S. ca*, *S. tr* and *S. re* differed significantly from *S. ba*
319 and *S. lo*, but not from *S. oa*, *S. ch* and *S. ki* (b). *S. ch* and *S. oa* differed significantly only from
320 *S. ki* (c). *S. ba* and *S. lo* significantly differed from *S. ca*, *S. tr*, *S. re* and *S. ki* (d). Thus, in spite
321 of using similar building blocks (i.e. notes) in their songs, species of *Spelaeornis* differ in the
322 complexity of these songs. This suggests that the temporal sequence or syntax of the songs
323 varies between species, a possibility we next investigated using computational analyses of
324 vocal sequence structure.

325

326 *Syntactic divergence in the songs of Spelaeornis*

327 We employed a combination of first-order Markov chain analysis, and a note co-occurrence
328 analysis that does not assume Markovianity (see Methods) to examine the syntactic structure of
329 *Spelaeornis* songs. As hinted at by the results of the song complexity analysis, we found that
330 different species fell into three broad syntactic groups. *S. ki* and *S. re*, both occurring in
331 Southeast Asia, exhibited songs with relatively sparse inter-note transitions, with notes instead
332 exhibiting a high tendency to repeat themselves. This was also borne out by the note co-
333 occurrence analysis, where high probabilities of co-occurrence (${}^dR_{ij}$) were primarily along the
334 diagonal of the matrix, indicating that notes tended to co-occur with themselves more than other
335 notes (Fig 3). We defined this sequence structure as Repetitive syntax. On the other hand, the
336 songs of *S. ca* from the Eastern Himalayas consisted of 2 or 3 notes alternating with each other.
337 This is represented by the reciprocating arrows between different nodes in Fig 3, as well as the
338 note co-occurrence matrix. Here, higher probabilities occurred both along the diagonal and with
339 1-2 other notes, indicating that each note type co-occurred with itself and one or two other
340 notes. We defined this second syntactic type as Alternating syntax. Finally, we also observed
341 songs with Complex syntax, primarily exhibited by *S. oa*, *S. lo*, and *S. ch* which occur in the hills
342 south of the Brahmaputra River. In these species' songs, each note group exhibited transitions
343 to many different note groups, manifested as many arrows originating from each node in Fig 4.
344 In the co-occurrence matrices, probabilities were generally low, and did not show any pattern of
345 higher values either on- or off-diagonal, indicating that each note could generally co-occur with
346 many other note types.

347

348 Two of the eight species (*S. ba* and *S. tr*) in the genus exhibited evidence of intermediate
349 syntactic types, employing a combination of the syntactic rules outlined above to order notes. *S.*
350 *ba* songs suggested an intermediate between Repetitive and Complex Syntax. As depicted in
351 Fig 4, some notes have high probabilities of transition to many different note groups in addition

352 to having a relatively high probability of self-transitions (i.e., repetitions). This is also observed in
353 the co-occurrence matrix, where generally low values of ${}^dR_{ij}$ are coupled with higher values
354 along the diagonal for certain notes. An examination of the song sequences (Supplementary
355 File 1) suggested that these two types could occur within a single song, with multiple note
356 groups in a single song, some of them repeating (unlike the other species with complex syntax).
357 On the other hand, *S. tr* appears to use a combination of Alternating and Repetitive syntax. Both
358 the transition probability matrix and the note co-occurrence matrix confirm this, with some note
359 groups tending to repeat and others alternating between each other. The song sequences
360 indicate that unlike *S. ba*, *S. tr* uses either alternating or repetitive syntax in a song, and not a
361 combination of both (Supplementary File 2). This indicates the presence of different song types
362 with differing syntax within this species, although we did not have enough data to determine if
363 these followed a geographic pattern. The patterns described above are broadly consistent
364 across different values of d (Fig S5, S6), in keeping with the findings of other studies that have
365 used the same analysis (Bhat et al. 2022). Similar patterns were also observed when
366 considering note types instead of note groups (Fig S7), which, coupled with our finding of
367 overlap in note parameter space, suggests that species of *Spelaeornis* use similar notes, but
368 arrange them according to different syntactic rules. Our note group classification thus
369 represents a valid means of interspecific comparison.
370

371 Using the aforementioned homogeneity test for Markov chains, we found that the transition
372 probability matrices for different species of *Spelaeornis* differ significantly from each other ($\chi^2=$
373 8061.23, df=504, p-value=0). Pairwise species comparisons found that transition probability
374 matrices differ for all species pairs (Fig S8). Further, we also found that species exhibiting
375 similar syntactic structure generally tend to have a smaller value of the test statistic when
376 compared to species with different syntactic rules (Fig S8). For example, *S. ki* and *S. re* were
377 closest to each other (smaller values of the test statistic), whereas *S. ch*, *S. oa* and *S. lo* were
378 closer to each other than to other species. *S. ca* was relatively far apart from all other species
379 (higher values of the test statistic). *S. ba* and *S. tr* were closest to *S. re*, and *S. tr* was also close
380 to *S. ki* (all four species use at least some repetitive syntax), but neither *S. ba* nor *S. tr* were
381 particularly close to any of the other species.
382

383 Syntactic structure of vocal sequences follows geographic barriers

384 Our analyses suggested that species of *Spelaeornis* existing geographically closer to each other
385 possess similar syntactic structure in their songs. Although the proportions of different note
386 groups in the sequences analyzed differed across geographically proximate species (Fig 5A),
387 the syntactic rules employed to construct songs were conserved within each sub-region.
388 However, across sub-regions separated by geographic barriers, syntactic structures changed
389 altogether (Fig 3, 4). We thus predicted that the vocal sequences of geographically proximate
390 species should group together, i.e. exhibit greater similarity than those from different geographic
391 regions. To further quantitatively examine sequence similarity across species, we calculated
392 sequence similarity using the median interspecific Levenshtein Distance and constructed a
393 distance dendrogram, using the UPGMA method (see Methods). We observed that species
394 geographically closer to each other are also closer to each other on the distance dendrogram

395 (Fig 5B), supporting the idea that geographic barriers drive syntactic diversification in the
396 absence of note divergence. To illustrate, *S. ki* and *S. re*, found in South-East Asia and
397 employing repetitive syntax grouped together in the dendrogram (Fig 5B). Additionally, *S. oa*, *S.*
398 *ch*, and *S. lo*, found south of the Brahmaputra River in North-Eastern India and employing
399 complex syntax, formed a second group in the dendrogram (Fig 5B), close to *S. ba* which
400 occupies the junction of the hills north and south of the region (and exhibits an intermediate
401 syntax with elements of complex syntax, see above). Taken together, our results suggest that in
402 the absence of note divergence, acoustic signal diversification in *Spelaeornis* involves changes
403 in the syntactic rules employed to order sequences across geographic barriers. Thus, vocal
404 sequences of species that are geographically proximate tend to structurally resemble each
405 other, whereas species from different geographic regions exhibit distinct vocal sequence
406 structures.

407

408 Discussion

409

410 In summary, we uncover evidence that song notes in the genus *Spelaeornis* remain more or
411 less conserved across species (Fig 2B, S1), but are arranged into songs according to different
412 syntactic rules (Fig 3, 4) based on the presence of geographic barriers. We found three broad
413 syntactic groups (with two species exhibiting intermediate syntax), occupying the Eastern
414 Himalayas, the hills south of the Brahmaputra river, and Southeast Asia respectively (Fig 3, 4).
415 Within each of these regions, geographically proximate species exhibited similar syntactic
416 structure in their songs (supported by the Kullback homogeneity test, and by the UPGMA
417 distance dendrogram) (Fig 5B, S8), with only minor differences in the proportions of different
418 note groups within their vocal sequences (Fig 5A). Our analysis of note accumulation suggested
419 that our sampling of the note group repertoires for each species is adequate (Fig S3), and
420 therefore we hypothesize that our results represent genuine interspecific changes in syntactic
421 structure of vocal sequences across geographic barriers. Because a variety of methods suggest
422 that we have adequately sampled the note group repertoire in this genus, we believe that
423 additional recordings will not significantly alter the pattern we observed, particularly that of
424 groups across biogeographic regions defined by vocal sequence structure.

425

426 Previous studies on interspecific song divergence have largely focused on differences in note
427 parameters, or on broad, summarized metrics of song structure such as trill rate or song
428 duration (Slabbekoorn and Smith 2002a; Slabbekoorn and Smith 2002b; Haavie et al. 2004;
429 Podos and Warren 2007; Irwin et al. 2008; Grant and Grant 2010; Tobias et al. 2010; Podos et
430 al. 2013; Wilkins et al. 2013). Here, employing a combination of tools (Kershenbaum et al.
431 2012), including a robust computational analysis of note co-occurrence originally developed to
432 study syntactic structure in anuran vocalizations (Bhat et al. 2022), we demonstrate that
433 *Spelaeornis* represents an example of syntactic diversification. This implies that the rules by
434 which notes are temporally arranged into songs have diversified across geographic barriers
435 without significant change in the underlying notes themselves. Our statistically robust analytical
436 framework enables us to comparatively examine temporal structure in diverse animal signals
437 and thus investigate signal evolution at multiple levels, from individual notes to syntax.

438

439 Contrary to our initial predictions, we did not find any interspecific divergence in the note
440 parameters in *Spelaeornis*. Instead, the overlap in signal space as well as a statistical
441 randomization test suggest that *Spelaeornis* species use similar song notes. In light of our
442 findings that species geographically closer to each other exhibit similar syntax (Fig 5B), and also
443 share note groups (Fig 2B), future research should comparatively examine the role of notes
444 versus syntax in species recognition. This will be particularly interesting for the case of *S.*
445 *badeigularis*, which occupies a junction point between different biogeographic regions, and also
446 exhibits songs with elements of both repetitive and complex syntax. Finally, the high-altitude *S.*
447 *troglodytoides* exhibits distinct song types with alternating and repetitive syntax, and more data
448 is required to understand whether this variation also corresponds to geographic barriers at high
449 altitudes. Distinct song types with different syntactic structures may also occur as rarer songs in
450 other species, particularly those such as *S. badeigularis* that occupy junctions between
451 geographic regions (King and Donahue 2006). However, the addition of rare song types is
452 unlikely to significantly alter the salient pattern we observe, that of syntactic change across
453 geographic regions without change in the underlying song notes. Therefore, the syntactic types
454 we describe here likely represent the dominant forms within each geographic region.
455 *Spelaeornis* as a genus is very poorly studied, with some species going undetected for decades
456 before being rediscovered (King and Donahue 2006). In light of this, online song databases
457 provide us with valuable data which, coupled with a rigorous analytical framework, enables us to
458 effectively examine questions about behavior and signal evolution. Quantitative analysis (Fig
459 S3, S4) suggests that we have adequately sampled the note group and note type repertoire of
460 all species within the genus (including for *S. reptatus* for which we had the smallest sample
461 size), further supporting this assertion. With so little known about the different hill ranges in the
462 Himalayas and Southeast Asia, it is possible that undescribed species in this genus remain to
463 be found, and their songs may prove valuable in understanding the steps by which syntactic
464 diversification of bird song has occurred.

465

466 The evolution and function of bird song exhibits many parallels to human languages (Doupe and
467 Kuhl 1999; Salwiczek and Wickler 2004; Berwick et al. 2011; Miyagawa et al. 2013; Collier et al.
468 2014; Sainburg et al. 2019), and geography shapes the structure and syntax of human
469 languages as well (Henry et al. 2015; Dunn 2019; Huisman et al. 2019; Urban 2021). It is,
470 therefore, also interesting to note that diversification of human languages in the region follows a
471 similar geographic pattern (Sagart et al. 2019), with an Indo-Chinese trail of language
472 diversification north of the Brahmaputra River and the Burman language groups diversifying
473 south of the Brahmaputra River. The parallels in the evolution of diverse culturally transmitted
474 traits point to the importance of studying syntax and sequences in birdsong to gain broader
475 evolutionary insights.

476

477 Acoustic signals in general are shaped by various factors such as evolutionary history,
478 geographic separation, and sexual selection (Marler and Tamura 1962; McCracken and
479 Sheldon 1997; Baker 2006; Baker et al. 2006; Podos and Warren 2007; Kirschel et al. 2009;
480 Kershenbaum et al. 2012; Garland et al. 2013; Lachlan et al. 2013; Seddon et al. 2013; Wilkins
481 et al. 2013; Lachlan et al. 2016; Arato and Fitch 2021). By presenting evidence that a) despite
482 geographical isolation, *Spelaeornis* appear to possess similar song notes across species, and

483 b) syntactic structure of vocal sequences in *Spelaeornis* has diverged across geographical
484 barriers, we shed light on the diverse paths through which signals evolve in allopatry
485 (Kershenbaum et al. 2012). Because, as discussed above, bird song exhibits similar geographic
486 patterns to those observed for human languages, we suggest that the computational approach
487 employed here can have far-reaching implications in our understanding of speciation, signal
488 evolution and linguistic diversification. The Himalayas and South-east Asia are regions with
489 hyperdiverse ecological communities (Srinivasan et al. 2014; Munjee and Athreya 2021;
490 Ashokan et al. 2022). Studies such as ours emphasize the importance of studying biodiversity,
491 natural history and behavior of taxa in these regions to understand the biogeography of
492 communication strategies and signal evolution.

493

494 **Acknowledgments and Funding**

495 We are grateful to the many recordists (listed in the Supplementary), whose recordings from
496 public databases contributed to this project, as well as to Ramana Athreya, Umesh Srinivasan,
497 Vaibhav Chhaya, Aurnab Ghose and members of the Krishnan lab for useful feedback and
498 discussions. Finally, we thank Seshadri KS for help with maps. AK is funded by an INSPIRE
499 Faculty Award from the Department of Science and Technology, Government of India, and an
500 Early Career Research Grant (ECR/2017/001527) from the Science and Engineering Research
501 Board (SERB), Government of India. MAJ and ASB are the recipients of the KVPY Fellowship
502 from the Government of India.

503

504 **References**

- 505 Arato J, Fitch WT. 2021. Phylogenetic signal in the vocalizations of vocal learning and vocal
506 non-learning birds. *Philos Trans R Soc B Biol Sci.* 376(1836). doi:10.1098/rstb.2020.0241.
507 Ashokan A, Xavier A, Suksathan P, Ardiyani M, Leong-Škorničková J, Newman M, Kress WJ,
508 Gowda V. 2022. Himalayan orogeny and monsoon intensification explain species diversification
509 in an endemic ginger (*Hedychium*: Zingiberaceae) from the Indo-Malayan Realm. *Mol
510 Phylogenet Evol.* 170(February). doi:10.1016/j.ympev.2022.107440.
511 Baker MC. 2006. Differentiation of mating vocalizations in birds: Acoustic features in mainland
512 and island populations and evidence of habitat-dependent selection on songs. *Ethology*.
513 112(8):757–771. doi:10.1111/j.1439-0310.2006.01212.x.
514 Baker MC, Baker MSA, Tilghman LM. 2006. Differing effects of isolation on evolution of bird
515 songs: Examples from an island-mainland comparison of three species. *Biol J Linn Soc.*
516 89(2):331–342. doi:10.1111/j.1095-8312.2006.00677.x.
517 Balaban E. 1988. Bird song syntax: Learned intraspecific variation is meaningful. *Proc Natl
518 Acad Sci U S A.* 85(10):3657–3660. doi:10.1073/pnas.85.10.3657.
519 Berwick RC, Okanya K, Beckers GJL, Bolhuis JJ. 2011. Songs to syntax: The linguistics of
520 birdsong. *Trends Cogn Sci.* 15(3):113–121. doi:10.1016/j.tics.2011.01.002.
521 Bhat AS, Sane VA, Seshadri KS, Krishnan A. 2022. Behavioural context shapes vocal
522 sequences in two anuran species with different repertoire sizes. *Anim Behav.* 184(xxxx):111–
523 129. doi:10.1016/j.anbehav.2021.12.004.
524 Bradbury JW, Vehrencamp SL. 2011. Principles of Animal Communication. 2nd ed. Sinauer
525 Associates.
526 Briefer EF, Rybak F, Aubin T. 2013. Does true syntax or simple auditory object support the role
527 of skylark song dialect? *Anim Behav.* 86(6):1131–1137. doi:10.1016/j.anbehav.2013.09.019.
528 Charrier I, Sturdy CB. 2005. Call-based species recognition in black-capped chickadees. *Behav
529 Processes.* 70(3):271–281. doi:10.1016/j.beproc.2005.07.007.

- 530 Chek AA, Bogart JP, Lougheed SC. 2003. Mating signal partitioning in multi-species
531 assemblages: A null model test using frogs. *Ecol Lett.* 6:235–247. doi:10.1046/j.1461-
532 0248.2003.00420.x.
533 Chitnis SS, Rajan S, Krishnan A. 2020. Sympatric wren-warblers partition acoustic signal space
534 and song perch height. *Behav Ecol.* 31:559–567. doi:10.1093/beheco/artz216.
535 Ciaburri I, Williams H. 2019. Context-dependent variation of house finch song syntax. *Anim*
536 *Behav.* 147:33–42. doi:10.1016/j.anbehav.2018.11.001.
537 Collar NJ. 2006. A partial revision of the Asian babblers (Timaliidae). *Forktail.* 22:85–112.
538 Collier K, Bickel B, van Schaik CP, Manser MB, Townsend SW. 2014. Language evolution:
539 Syntax before phonology? *Proc R Soc B Biol Sci.* 281(1788). doi:10.1098/rspb.2014.0263.
540 Dahlin CR, Wright TF. 2012. Does syntax contribute to the function of duets in a parrot,
541 *Amazona auropalliata?* *Anim Cogn.* 15(4):647–656. doi:10.1007/s10071-012-0493-y.
542 Doupe AJ, Kuhl PK. 1999. Birdsong and human speech: Common themes and mechanisms.
543 *Annu Rev Neurosci.* 22(February):567–631. doi:10.1146/annurev.neuro.22.1.567.
544 Dunn J. 2019. Global Syntactic Variation in Seven Languages: Toward a Computational
545 Dialectology. *Front Artif Intell.* 2(August):1–22. doi:10.3389/frai.2019.00015.
546 Egan SP, Funk DJ. 2009. Ecologically dependent postmating isolation between sympatric host
547 forms of *Neochlamisus bebbiana* leaf beetles. *Proc Natl Acad Sci U S A.* 106(46):19426–
548 19431. doi:10.1073/pnas.0909424106.
549 Engesser S, Townsend SW. 2019. Combinatoriality in the vocal systems of nonhuman animals.
550 *Wiley Interdiscip Rev Cogn Sci.* 10(4):1–10. doi:10.1002/wcs.1493.
551 Fishbein AR, Idsardi WJ, Ball GF, Dooling RJ. 2020. Sound sequences in birdsong: How much
552 do birds really care? *Philos Trans R Soc B Biol Sci.* 375(1789). doi:10.1098/rstb.2019.0044.
553 Garland EC, Noad MJ, Goldizen AW, Lilley MS, Rekdahl ML, Garrigue C, Constantine R,
554 Daeschler Hauser N, Michael Poole M, Robbins J. 2013. Quantifying humpback whale song
555 sequences to understand the dynamics of song exchange at the ocean basin scale. *J Acoust*
556 *Soc Am.* 133(1):560–569. doi:10.1121/1.4770232.
557 Grant BR, Grant PR. 2010. Songs of Darwin's finches diverge when a new species enters the
558 community. *Proc Natl Acad Sci.* 107(47):20156–20163.
559 Haavie J, Borge T, Bures S, Garamszegi LZ, Lampe HM, Moreno J, Qvarnström A, Török J,
560 Sætre GP. 2004. Flycatcher song in allopatry and sympatry - Convergence, divergence and
561 reinforcement. *J Evol Biol.* 17(2):227–237. doi:10.1111/j.1420-9101.2003.00682.x.
562 Henry L, Barbu S, Lemasson A, Hausberger M. 2015. Dialects in Animals: Evidence,
563 Development and Potential Functions. *Anim Behav Cogn.* 2(2):132–155.
564 doi:10.12966/abc.05.03.2015.
565 Hillis DM. 1981. Premating Isolating Mechanisms among Three Species of the *Rana pipiens*
566 Complex in Texas and Southern Oklahoma. *Copeia.* 1981(2):312. doi:10.2307/1444220.
567 del Hoyo J, Collar NJ, Christie DA, Elliott A, Fishpool LDC. 2014. HBW and BirdLife
568 International Illustrated Checklist of the Birds of the World. Lynx Edicions BirdLife International.
569 Huisman JLA, Majid A, Van Hout R. 2019. The geographical configuration of a language area
570 influences linguistic diversity. *PLoS One.* 14(6):1–18. doi:10.1371/journal.pone.0217363.
571 Hunt R. 1923. The Phonetics of Bird-Sound. *Condor.* 25(6):202–208.
572 Irwin DE, Irwin JH, Price TD. 2001. Ring species as bridges between microevolution and
573 speciation. *Genetica.* 112–113:223–243. doi:10.1023/A:1013319217703.
574 Irwin DE, Thimgan MP, Irwin JH. 2008. Call divergence is correlated with geographic and
575 genetic distance in greenish warblers (*Phylloscopus trochiloides*): A strong role for stochasticity
576 in signal evolution? *J Evol Biol.* 21(2):435–448. doi:10.1111/j.1420-9101.2007.01499.x.
577 Isler ML, Isler PR, Whitney BM. 1998. Use of vocalizations to establish species limits in antbirds
578 (Passeriformes: Thamnophilidae). *Auk.* 115(3):577–590. doi:10.2307/4089407.
579 IUCN. 2019. The IUCN Red List of Threatened Species.

- 580 Kershenbaum A, Blumstein DT, Roch MA, Akçay Ç, Backus G, Bee MA, Bohn K, Cao Y, Carter
581 G, Cäsar C, et al. 2016. Acoustic sequences in non-human animals: A tutorial review and
582 prospectus. *Biol Rev.* 91(1):13–52. doi:10.1111/brv.12160.
- 583 Kershenbaum A, Bowles AE, Freeberg TM, Jin DZ, Lameira AR, Bohn K. 2014. Animal vocal
584 sequences: Not the Markov chains we thought they were. *Proc R Soc B Biol Sci.* 281(1792).
585 doi:10.1098/rspb.2014.1370.
- 586 Kershenbaum A, Ilany A, Blaustein L, Geffen E. 2012. Syntactic structure and geographical
587 dialects in the songs of male rock hyraxes. *Proc R Soc B Biol Sci.* 279(1740):2974–2981.
588 doi:10.1098/rspb.2012.0322.
- 589 King B, Donahue JP. 2006. The rediscovery and song of the Rusty-throated Wren Babbler
590 *Spelaeornis badeigularis*. *Forktail.* 22:113–115.
- 591 Kirschel ANG, Blumstein DT, Smith TB. 2009. Character displacement of song and morphology
592 in African tinkerbirds. *Proc Natl Acad Sci U S A.* 106(20):8256–8261.
- 593 Krishnan A. 2019. Acoustic community structure and seasonal turnover in tropical South Asian
594 birds. *Behav Ecol.* 30(5):1364–1374. doi:10.1093/beheco/arz087.
- 595 Kullback S, Kupperman M, Ku HH. 1962. Tests for Contingency Tables and Markov Chains.
596 *Technometrics.* 4(4):573–608.
- 597 Lachlan RF, van Heijningen CAA, ter Haar SM, ten Cate C. 2016. Zebra finch song phonology
598 and syntactical structure across populations and continents-a computational comparison. *Front
599 Psychol.* 7:980. doi:10.3389/fpsyg.2016.00980.
- 600 Lachlan RF, Servedio MR. 2004. Song learning accelerates allopatric speciation. *Evolution (N
601 Y).* 58(9):2049–2063.
- 602 Lachlan RF, Verzijden MN, Bernard CS, Jonker PP, Koese B, Jaarsma S, Spoor W, Slater PJB,
603 Ten Cate C. 2013. The progressive loss of syntactical structure in bird song along an Island
604 colonization chain. *Curr Biol.* 23(19):1896–1901. doi:10.1016/j.cub.2013.07.057.
- 605 Leroux M, Bosshard AB, Chandia B, Manser A, Zuberbühler K, Townsend SW. 2021.
606 Chimpanzees combine pant hoots with food calls into larger structures. *Anim Behav.* 179:41–
607 50. doi:10.1016/j.anbehav.2021.06.026.
- 608 Luther D. 2009. The influence of the acoustic community on songs of birds in a neotropical rain
609 forest. *Behav Ecol.* 20(4):864–871. doi:10.1093/beheco/arp074.
- 610 Marler P, Peters S. 1988. The role of song phonology and syntax in vocal learning preferences
611 in the Song Sparrow, *Melospiza melodia*. *Ethology.* 77:125–149.
- 612 Marler P, Slabbekoorn H. 2004. *Nature's Music: the Science of Birdsong*. London, UK: Elsevier
613 Academic Press.
- 614 Marler P, Tamura M. 1962. Song “Dialects” in Three Populations of White-Crowned Sparrows.
615 *Condor.* 64(5):368–377. doi:10.2307/1365545.
- 616 Marshall DC, Cooley JR. 2000. Reproductive character displacement and speciation in
617 periodical cicadas, with description of a new species, 13-year *Magicicada neotredecim*.
618 *Evolution (N Y).* 54(4):1313–1325. doi:10.1111/j.0014-3820.2000.tb00564.x.
- 619 McCracken KG, Sheldon FH. 1997. Avian vocalizations and phylogenetic signal. *Proc Natl Acad
620 Sci U S A.* 94(April):3833–3836.
- 621 Miyagawa S, Berwick R, Okanoya K. 2013. The Emergence of Hierarchical Structure in Human
622 Language. *Front Psychol.* 4. doi:10.3389/fpsyg.2013.00071.
- 623 Mungee M, Athreya R. 2021. Intraspecific trait variability and community assembly in
624 hawkmoths (Lepidoptera: Sphingidae) across an elevational gradient in the eastern Himalayas,
625 India. *Ecol Evol.* 11(6):2471–2487. doi:10.1002/ece3.7054.
- 626 Nowicki S, Searcy WA, Hughes M, Podos J. 2001. The evolution of bird song: Male and female
627 response to song innovation in swamp sparrows. *Anim Behav.* 62(6):1189–1195.
628 doi:10.1006/anbe.2001.1854.
- 629 Peters SS, Searcy WA, Marler P. 1980. Species song discrimination in choice experiments with

- 630 territorial male swamp and song sparrows. *Anim Behav.* 28(2):393–404. doi:10.1016/S0003-
631 3472(80)80048-0.
- 632 Podos J, Dybboe R, Ole Jensen M. 2013. Ecological speciation in Darwin's finches: Parsing the
633 effects of magic traits. *Curr Zool.* 59(1):8–19. doi:10.1093/czoolo/59.1.8.
- 634 Podos J, Warren PS. 2007. The Evolution of Geographic Variation in Birdsong. *Adv Study
635 Behav.* 37(07):403–458. doi:10.1016/S0065-3454(07)37009-5.
- 636 Price T. 2007. *Speciation in Birds.* CO, USA: Roberts and Co. Publishers.
- 637 Rasmussen PC, Anderton JC. 2005. *Birds of South Asia: The Ripley Guide.* Lynx Edicions,
638 Barcelona.
- 639 Richmond JQ, Jockusch EL, Latimer AM. 2011. Mechanical reproductive isolation facilitates
640 parallel speciation in western North American scincid lizards. *Am Nat.* 178(3):320–332.
641 doi:10.1086/661240.
- 642 Sagart L, Jacques G, Lai Y, Ryder RJ, Thouzeau V, Greenhill SJ, List JM. 2019. Dated
643 language phylogenies shed light on the ancestry of Sino-Tibetan. *Proc Natl Acad Sci U S A.*
644 116(21):10317–10322. doi:10.1073/pnas.1817972116.
- 645 Sainburg T, Theilman B, Thielk M, Gentner TQ. 2019. Parallels in the sequential organization of
646 birdsong and human speech. *Nat Commun.* 10(1):1–11. doi:10.1038/s41467-019-11605-y.
- 647 Salwiczek LH, Wickler W. 2004. Birdsong: An evolutionary parallel to human language.
648 *Semiotica.* 151(January):163–182. doi:10.1515/semi.2004.066.
- 649 Schmidt AKD, Römer H, Riede K. 2013. Spectral niche segregation and community
650 organization in a tropical cricket assemblage. *Behav Ecol.* 24(October):470–480.
651 doi:10.1093/beheco/ars187.
- 652 Searcy WA, Marler P. 1981. A Test for Responsiveness to Song Structure and Programming in
653 Female Sparrows. *Science* (80-). 213(4510):926–928.
- 654 Searcy WA, Marler P, Peters SS. 1981. Species song discrimination in adult female song and
655 swamp sparrows. *Anim Behav.* 29(4):997–1003. doi:10.1016/S0003-3472(81)80053-X.
- 656 Seddon N, Botero CA, Tobias JA, Dunn PO, Macgregor HEA, Rubenstein DR, Uy JAC, Weir JT,
657 Whittingham LA, Safran RJ. 2013. Sexual selection accelerates signal evolution during
658 speciation in birds. *Proc R Soc B Biol Sci.* 280(1766):20131065.
- 659 Slabbekoorn H, Smith TB. 2002a. Habitat-dependent song divergence in the little greenbul: An
660 analysis of environmental selection pressures on acoustic signals. *Evolution (N Y).* 56(9):1849–
661 1858. doi:10.1111/j.0014-3820.2002.tb00199.x.
- 662 Slabbekoorn H, Smith TB. 2002b. Bird song, ecology and speciation. *Philos Trans R Soc B Biol
663 Sci.* 357(1420):493–503. doi:10.1098/rstb.2001.1056.
- 664 Srinivasan U, Tammar K, Ramakrishnan U. 2014. Past climate and species ecology drive nested
665 species richness patterns along an east-west axis in the Himalaya. *Glob Ecol Biogeogr.*
666 23(1):52–60. doi:10.1111/geb.12082.
- 667 Suzuki TN, Wheatcroft D, Griesser M. 2018. Call combinations in birds and the evolution of
668 compositional syntax. *PLoS Biol.* 16(8):1–5. doi:10.1371/journal.pbio.2006532.
- 669 Tobias JA, Aben J, Brumfield RT, Derryberry EP, Halfwerk W, Slabbekoorn H, Seddon N. 2010.
670 Song divergence by sensory drive in amazonian birds. *Evolution (N Y).* 64:2820–2839.
- 671 Urban M. 2021. The geography and development of language isolates. *R Soc Open Sci.* 8(4).
672 doi:10.1098/rsos.202232.
- 673 Uy JAC, Irwin DE, Webster MS. 2018. Behavioral isolation and incipient speciation in birds.
674 *Annu Rev Ecol Evol Syst.* 49:1–24. doi:10.1146/annurev-ecolsys-110617-062646.
- 675 Wanker R, Fischer J. 2001. Intra- and interindividual variation in the contact calls of spectacled
676 parrotlets (*Forpus conspicillatus*). *Behaviour.* 138(6):709–726.
677 doi:10.1163/156853901752233361.
- 678 Wilkins MR, Seddon N, Safran RJ. 2013. Evolutionary divergence in acoustic signals: causes
679 and consequences. *Trends Ecol Evol.* 28(3):156–166. doi:10.1016/j.tree.2012.10.002.

680 Yeh DJ, Servedio MR. 2015. Reproductive isolation with a learned trait in a structured
681 population. *Evolution (N Y)*. 69(7):1938–1947. doi:10.1111/evo.12688.

682

683

684

685

686

687 **Figure Legends:**

688

689 **Fig 1: A.** Acoustic signal divergence may involve divergence in the frequency-time parameters
690 (acoustic structure) of individual notes (note divergence, left), and the rules followed to order
691 these individual notes into a song (syntactic divergence, right). **B.** Geographical ranges of the
692 genus *Spelaeornis*, and the abbreviated names we use throughout. *S.ca*: *S. caudatus*, *S. ba*: *S.*
693 *badeigularis*, *S. tr*: *S. troglodytoides*, *S. ch*: *S. chocolatinus*, *S. re*: *S. reptatus*, *S. ki*: *S. kineari*, *S.*
694 *oa*: *S. oatesi*, *S. lo*: *S. longicaudatus*. The species are allopatric, and *S. tr* is present at higher
695 altitudes than its congeners. The images of *Spelaeornis* were sourced with permission from the
696 Macaulay library (Cornell Laboratory of Ornithology, Ithaca, NY, USA) (ML101191261- Jon
697 Irvine, ML379201951- Anupam Nahardeka, ML379618361-Lakpa Tenzing Sherpa,
698 ML378245821- Pranjal J. Saikia, ML213328961- Xiwen Chen, ML269988921- Ngoc Sam
699 Thuong Dang, ML205760231- James Eaton, ML378358871- Firoz Hussain). Species ranges
700 were sourced from International Union for Conservation of Nature (IUCN 2019) and the figure
701 was created in QGIS (Quantum GIS Development Team 2013; <http://qgis.osgeo.org>).
702

703

704 **Fig 2: A.** Schematic representation of the framework followed to classify notes from different
705 species into note types and note groups (see Methods). Different note types of Species A and
706 Species B are depicted by alphabets (top), and different note groups in each species are
707 represented by different colors. Four examples of note groups present in Species A and
708 Species B are described in the bottom panel. The Y-axis represents the frequency of each note,
709 and the X-axis represents time. **B.** Three-dimensional signal space of *Spelaeornis* constructed
710 using the first three principal components (PC 1-3) of the correlation matrix of 10 measured note
711 parameters. Each point in this space represents a note, and the species are represented by
712 different colors. *Spelaeornis* exhibit overlap in note parameter space. **C.** Box plot depicting the
713 song complexity metric across species in *Spelaeornis*. Letters depicted represent groups
714 showing statistically significant differences after pairwise comparisons. a: differed significantly
715 from *S. oa*, *S. lo*, *S. ch* and *S. ba*; b: differed significantly from *S. ba* and *S. lo* but not from *S.*
716 *oa*, *S. ch*, *S. ki*; c: differed significantly only from *S. ki*; d: differed significantly from *S. ca*, *S. tr*,
717 *S. re* and *S. ki*. Further details are given in the Results and Table S2.
718

719

720 **Fig 3:** Transition probabilities obtained by modeling note group sequences of *Spelaeornis* as
721 first-order Markov chains (left side on all the panels). Different colors represent different note
722 groups (see key below the figure), and the size of the node is proportional to frequency of
723 occurrence in our dataset. The thickness of each arrow scales with increasing transition
probabilities between note groups. Ratio of observed to expected probability of co-occurrences
for note groups for a value of d=3 (right side on all the panels). Warmer or redder colors in the

724 *ij*-th square on the matrix depict that note group i occurs with note group j more often than by
725 chance. *S. ki* and *S. re* exhibit repetitive syntax, depicted by thicker arrows for self-transitions
726 and redder colors along the diagonal in the co-occurrence matrix. In *S. ca*, we find alternating
727 syntax, depicted by arrows going back and forth between different note groups and redder
728 colors both above and below the diagonal in the co-occurrence matrix. *S. tr* exhibits a
729 combination of alternating and repetitive syntax. Some note groups exhibit higher propensity for
730 self-transition in the transition probability figures and redder colors along the diagonal in the co-
731 occurrence matrix, whereas others exhibit alternation (redder colors for 1-2 other note groups in
732 the co-occurrence matrix).

733

734 **Fig 4:** Transition probabilities and the ratio of observed to expected probability of co-
735 occurrences for note groups for a value of d=3, as in Fig 3. *S. lo*, *S. oa* and *S. ch* exhibit
736 complex syntax, where each group can transition to many others with a low probability of self-
737 transition and a lack of discernible structure in the co-occurrence matrix. *S. ba*, on the other
738 hand, exhibits an intermediate between complex and repetitive syntax. Many of the above
739 features are present, but some note groups exhibit an increased propensity to self-transition,
740 resulting in some repetition within vocal sequences (see Supplementary).

741

742 **Fig 5: A.** Pie charts representing the proportions of different note groups across species in our
743 dataset. Note accumulation curves suggest that this dataset is an adequate sample of the note
744 group repertoire across species (see Fig S3). Species geographically closer to each other
745 exhibit similar syntactic structure (see alphabets), with changes across major geographic
746 barriers. Note in particular the intermediate syntax of *S. ba*, which occupies a junction between
747 two major geographic regions. **B.** UPGMA distance dendrogram constructed using median
748 Levenshtein distances between vocal sequences for each pair of species. Species with similar
749 vocal sequences cluster together in this distance dendrogram, and are also geographically
750 proximate, as shown in the map.

751

752 **Fig S1:** Interspecific confusion matrix representing correct classification rates of the note
753 parameters of different species in a linear discriminant classifier built using MATLAB, with the
754 10 note parameters described above as the input. Percentages in the blue boxes indicate
755 correct classification rates for each species, and the red boxes indicate misclassification rates.
756 We observe an overall classification accuracy of 55.8%, coupled with high misclassification
757 rates of LDA in assigning notes to species.

758

759 **Fig S2:** Intraspecific confusion matrices representing correct classification rates for individual
760 note types of each species using linear discriminant classifiers built in MATLAB.

761

762 **Fig S3:** The accumulation curves for note groups exhibit early saturation, suggesting that we
763 adequately sampled the note groups for each species.

764

765 **Fig S4:** The accumulation curves for note types exhibit early saturation, suggesting that we
766 adequately sampled the note types for each species.

767

768 **Fig S5:** Ratio of observed to expected probability of co-occurrences for note groups for a value
769 of d=2.

770

771 **Fig S6:** Ratio of observed to expected probability of co-occurrences for note groups for a value
772 of d=6.

773

774 **Fig S7:** Transition probabilities obtained by modeling note type sequences of *Spelaeornis* as
775 first-order Markov chains (left side on all the panels), as in Fig 3 and 4 (which depict the same
776 analysis for note groups). Ratio of observed to expected probability of co-occurrences for note
777 types for a value of d=3 (right side on all the panels), as in Fig 3 and 4. The syntactic patterns
778 described for note groups in *Spelaeornis* remain consistent when considering note types. Thus,
779 interspecific syntactic patterns observed are not influenced by our note group classification.

780

781 **Fig S8:** Heat map depicting the pairwise test statistic of the homogeneity test. Species
782 geographically closer to each other generally exhibit a smaller value of the test statistic,
783 suggesting they are more similar to each other than to other congeners.

784

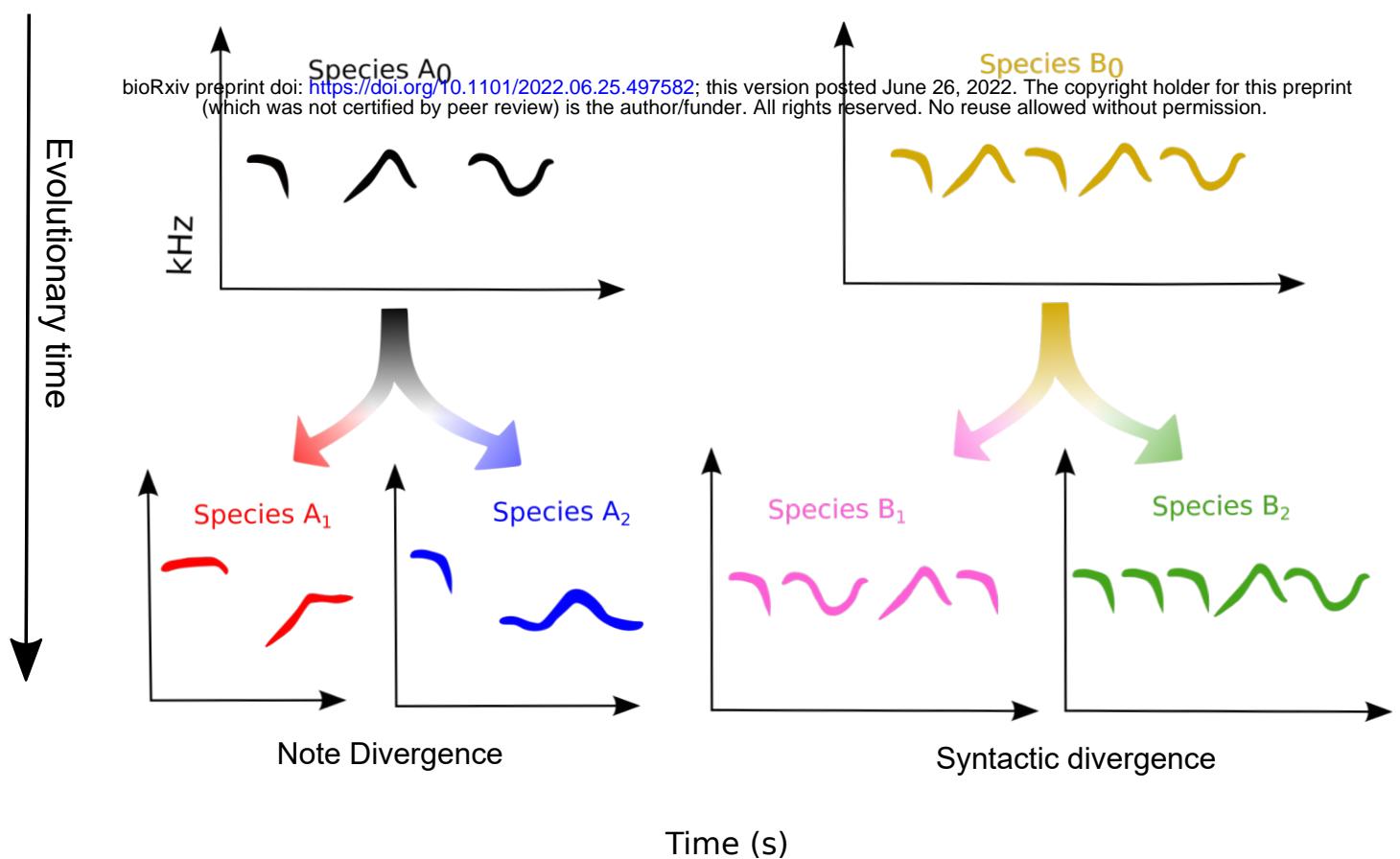
785 **Table S1:** Results of Principal Component analysis on the correlation matrix of the ten note
786 parameters measured from the songs of *Spelaeornis*. The values represent the loadings of
787 various parameters onto the Principal Components (PCs), as well as the proportion of explained
788 variance.

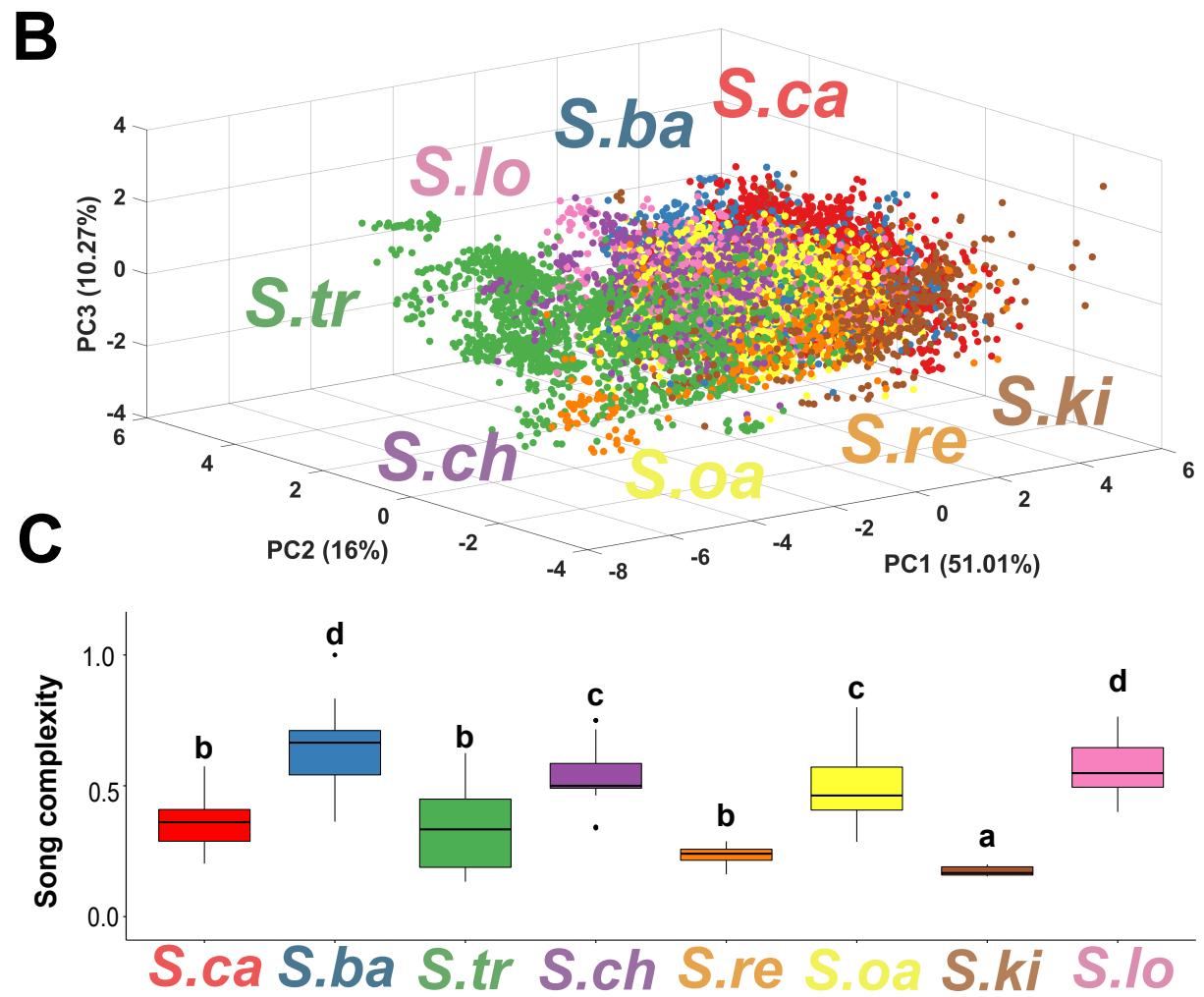
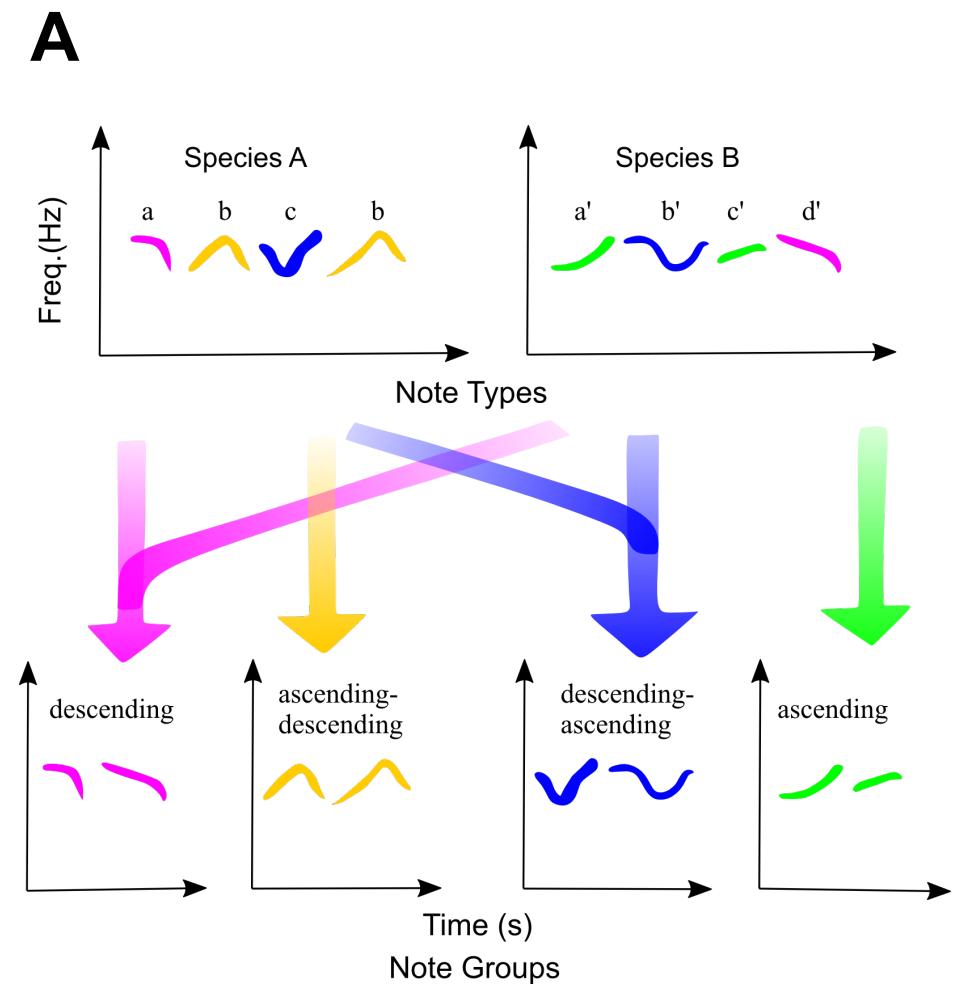
789

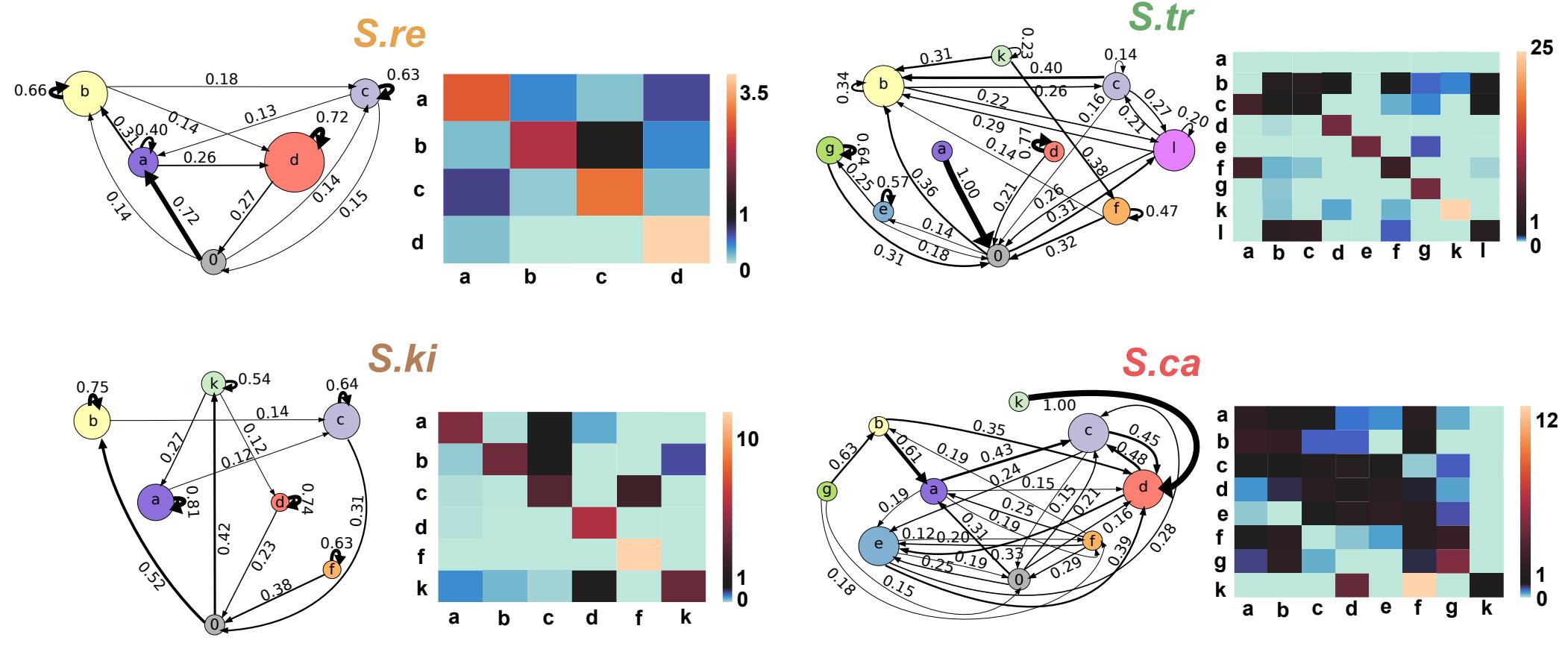
790 **Table S2:** Table depicting the results of post-hoc Dunn's tests for pairwise comparisons.
791 Species pairs showing statistically significant differences in song complexity are represented in
792 bold.

793

794

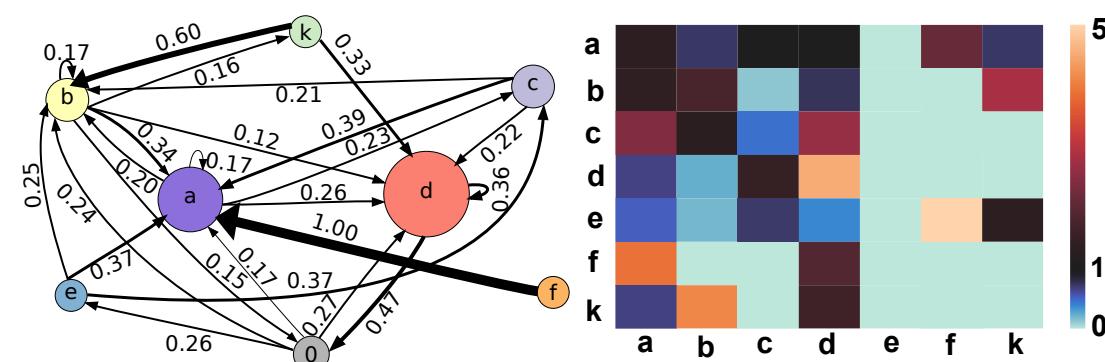
A**B**



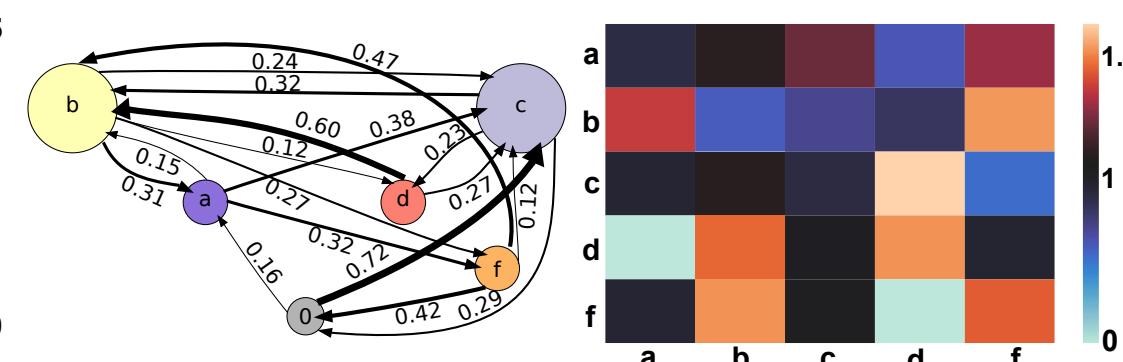


a- ascending; b- descending; c- ascending-descending; d- descending-ascending ;
e- ascending-descending ascending ; f- descending- ascending-descending ; g- complex ;
k- constant frequency ; l-long complex

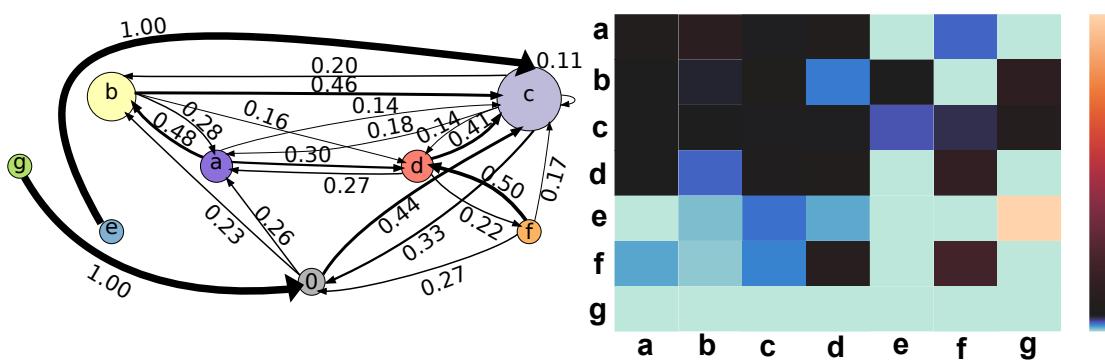
S.ba



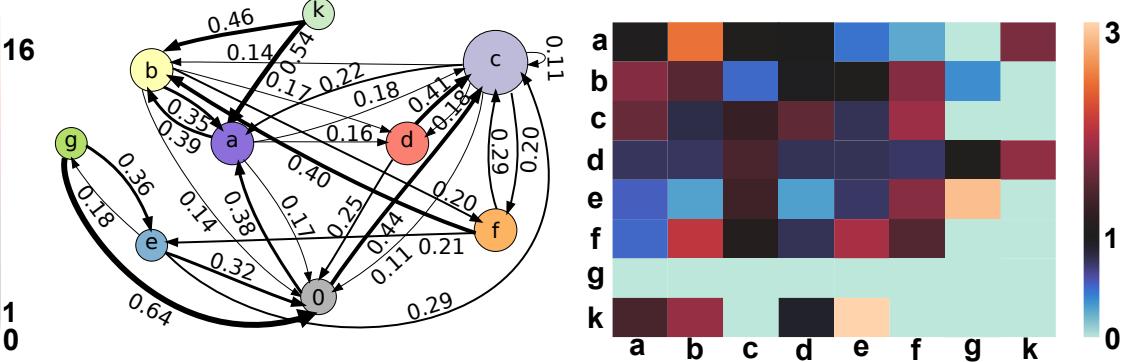
S.ch



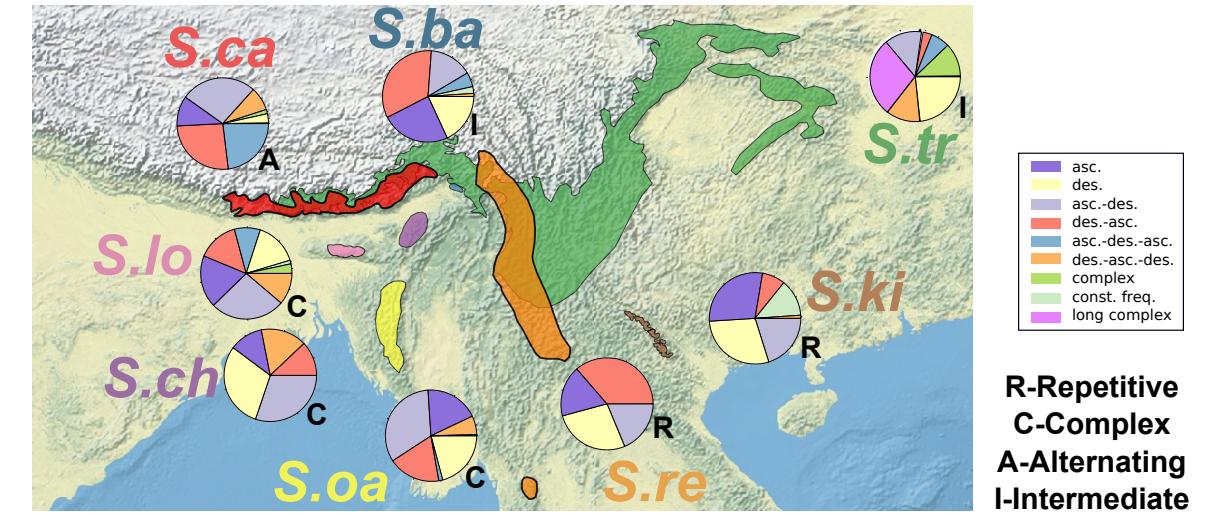
S.oa



S.lo



a- ascending; b-descending; c- ascending-descending; d-descending-ascending ;
e- ascending-descending ascending ; f- descending- ascending-descending ; g- complex ;
k- constant frequency ; l-long complex

A**B**