

# Isochrony as ancestral condition to call and song in a primate

Chiara De Gregorio<sup>1,2,†</sup>  | Marco Maiolini<sup>1,†,a</sup> | Teresa Raimondi<sup>1,3</sup> | Filippo Carugati<sup>1</sup> | Longondraza Miaretsoa<sup>4</sup> | Daria Valente<sup>1,5</sup> | Valeria Torti<sup>1</sup> | Cristina Giacomini<sup>1,§</sup> | Andrea Ravnani<sup>3,6,7,§</sup>  | Marco Gamba<sup>1,§</sup>

<sup>1</sup>Department of Life Sciences and Systems Biology, University of Torino, Turin, Italy

<sup>2</sup>Department of Psychology, University of Warwick, Coventry, UK

<sup>3</sup>Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy

<sup>4</sup>Groupe d'étude et de recherche sur les primates de Madagascar (GERP), Antananarivo, Madagascar

<sup>5</sup>Parco Natura Viva Garda Zoological Park (PNV), Verona, Italy

<sup>6</sup>Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

<sup>7</sup>Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus, Aalborg, Denmark

## Correspondence

Daria Valente, Department of Life Sciences and Systems Biology, University of Torino, Via Accademia Albertina 13, 10123 Turin, Italy. Email: [daria.valente@unito.it](mailto:daria.valente@unito.it)

Andrea Ravnani, Department of Human Neurosciences, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185 Rome, Italy. Email: [andrea.ravnani@uniroma1.it](mailto:andrea.ravnani@uniroma1.it)

<sup>a</sup>Present address: Behavioural Biology, Institute of Biology Leiden, Leiden University, Leiden, The Netherlands.

## Funding information

Max-Planck-Gesellschaft, Grant/Award Number: -; Danish National Research Foundation, Grant/Award Number: DNRF117; H2020 European Research Council, Grant/Award Numbers: TOHR, 101041885

## Abstract

Animal songs differ from calls in function and structure, and have comparative and translational value, showing similarities to human music. Rhythm in music is often distributed in quantized classes of intervals known as rhythmic categories. These classes have been found in the songs of a few nonhuman species but never in their calls. Are rhythmic categories song-specific, as in human music, or can they transcend the song-call boundary? We analyze the vocal displays of one of the few mammals producing both songs and call sequences: *Indri indri*. We test whether rhythmic categories (a) are conserved across songs produced in different contexts, (b) exist in call sequences, and (c) differ between songs and call sequences. We show that rhythmic categories occur across vocal displays. Vocalization type and function modulate deployment of categories. We find isochrony (1:1 ratio, like the rhythm of a ticking clock) in all song types, but only advertisement songs show three rhythmic categories (1:1, 1:2, 2:1 ratios). Like songs, some call types are also isochronous. Isochrony is the backbone of most indri vocalizations, unlike human speech, where it is rare. In indri, isochrony underlies both songs and hierarchy-less call sequences and might be ancestral to both.

## KEYWORDS

alarm call, categorical rhythm, hierarchical, singing primates, song

<sup>†</sup> Co-first authorship.

<sup>§</sup> Co-last authorship.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Authors. *Annals of the New York Academy of Sciences* published by Wiley Periodicals LLC on behalf of The New York Academy of Sciences.

## INTRODUCTION

Scholars across disciplines have long debated the extent of speech-music parallels in our species. This debate has partly focused on exaptations or selective pressures that may have led to the origin of human music and speech.<sup>1</sup> Both music and speech have rhythm, although potentially in different ways.<sup>2</sup> Rhythmicity across human displays prompts comparisons with behavior in species other than our own. This comparative approach is crucial to understanding how widespread and qualitatively similar musical abilities are among animals.<sup>3</sup>

Given its rarity and parallels with human musical behaviors,<sup>4</sup> the singing behavior of nonhuman primates is often hypothesized to be homologous to human song.<sup>5</sup> Most primate acoustic communication relies on the use of calls, such as contact or alarm calls; however, a few species also produce other types of vocal sequences, usually termed songs. Songs are a series of notes of different types, are characterized by a frequency variation, and are uttered following a hierarchical structure<sup>6</sup> as notes are combined into phrases and phrases are combined to form songs. Nevertheless, even if lacking the complexity and hierarchical structure of songs, primate calls can be organized in sequences with a variable degree of stereotypy and therefore might possess a rhythmic pattern that can be extracted and compared to that of songs. Geissmann<sup>5</sup> hypothesized that loud calls in apes and human music might derive from ancestral forms of loud callings, such as antipredator alarms.

Functionally, primate songs and human music are not referential, while calls are. Structurally, different human musical features show similarities to primate songs, such as notes being emitted in an organized pattern with small intervals and short phrases, with discrete pitches, melodic contours, and a rhythmic pattern of small integer ratios.<sup>5,7,8</sup> We focus on these particular ratios because auditory rhythms characterized by small integer ratios between following intervals are common in the world's musical repertoire.<sup>7,9,10</sup> Therefore, testing for comparative rhythmicity among songs and call sequences in the same primate species can provide insights into their joint/parallel evolution.

Do call and song, in the few mammals that produce both like us, have the same rhythmic building blocks? *Indri indri* is an ideal model species to address this question; this singing lemur is the only known mammal sharing multiple rhythmic categories with human music, and whose songs possess a strong rhythmic component.<sup>8,11</sup> Moreover, the indri vocal repertoire is not limited to singing behavior; indris also produce different rhythmic calls that serve to communicate between members of the monogamous family group.<sup>12,13</sup> While songs are produced during vocal interactions among individuals, calls are not interactive: this difference may affect the degree of hierarchy in temporal structure.<sup>14</sup> Songs may last several minutes and are composed of phrases that can last up to 10 s,<sup>15</sup> which in turn are composed of notes lasting usually around 1.5 seconds.<sup>16</sup> Adjacent phrases are separated by silent gaps that are longer than those occurring between notes of the same phrase. Conversely, indri calls are composed of units that last less than 1 s and are emitted as a single sequence. Therefore, it is likely that songs

and calls would differ in terms of rhythmic structure. Indri songs and calls are roughly analogous to their songbird counterparts. While the singing and calling behavior of birds has been extensively researched, mammalian song research is still fairly untapped. For the first time, we have a mammalian model suited to the structural comparison between songs and call sequences.

Indri songs feature two distinct rhythmic categories shared with human music,<sup>17</sup> showing multiple intervals that have an integer-ratio relationship.<sup>8</sup> Still, past work on indris only analyzed one of three possible song types and no calls. Here, we test for conserved rhythmicity: (a) among three song types (i.e., singing displays emitted in different contexts, such as during daily territorial advertisement when members belonging to the same family are out of eyesight and during encounters with different family groups as active territorial defense), (b) among three types of call sequences, and (c) between song and call sequences. In other words, we test for rhythmic similarities and differences between songs and/or calls in a rare, nonhuman primate species that produces both. We determine rhythmic similarity by testing whether songs and calls contain partly overlapping rhythmic categories or have no rhythmic categories in common.<sup>8</sup>

Our hypothesis is that the number of rhythm categories corresponding to small integer ratios changes across vocal displays. In particular, we predict that (a) songs would show a consistent number of rhythmic categories and that (b) the number of categories in songs differs from the categories in calls. Since Darwin,<sup>18</sup> animal songs have been hypothesized as precursors to human vocal displays.<sup>19</sup> Human speech is rarely isochronous,<sup>20</sup> while rhythmic categories with small integer ratios seem characteristic of musical displays.<sup>7,17,21</sup> Because of the presence of hierarchical structures (e.g., notes organized in phrases<sup>22</sup>), we predict that songs would show multiple rhythmic categories. Conversely, alarm calls show no hierarchical structures<sup>12,13</sup> (see findings on other species producing sequences of calls<sup>23</sup>), so we predict fewer rhythmic categories.

## MATERIALS AND METHODS

### Study species

Indris are strepsirrhine primates inhabiting the eastern rainforest of Madagascar. They live in family groups<sup>24</sup> that actively defend their territory through powerful long-distance vocal emissions, uttered by two or more individuals simultaneously.<sup>25–27</sup> These lemurs show long-term pair bonds with a genetically monogamous mating system.<sup>24</sup> Through songs, indris exchange information about group location and composition.<sup>28,29</sup> This species shows female dominance,<sup>30</sup> and pair formation is likely mediated by singing behavior.<sup>31</sup>

The indri vocal repertoire includes songs and calls that are functionally diverse acoustic displays<sup>13,32</sup> composed of a series of elements separated by silent gaps. This repertoire makes indri vocal production an excellent testbench for the rhythmic aspects of animal communication, as the co-occurrence of songs and calls is rare in mammalian

species. Songs differ from other vocalizations based on their duration, complexity, and note organization.<sup>6,12,33</sup> Indris produce at least three types of calls: (1) honks (or “clacsons”, Figure 1A) are rhythmic, rapid alerting calls given in a sequence of short, loud notes, produced when potential terrestrial predators are present<sup>12</sup>; (2) alarm roars (Figure 1B) are a sequence of harsh vocalizations emitted in the presence of aerial predators<sup>12</sup>; and (3) song roars (Figure 1C) are similar to alarm roars but are emitted in sequence before the starting of a song,<sup>34</sup> with a supposed attention-gathering function.

Songs are vocal displays composed of notes. Each note can be either part of a phrase, known as a descending phrase, where two to six note clusters follow a descending frequency pattern,<sup>15</sup> or produced as isolated units in single or long notes.<sup>31</sup> Long notes are scarcely modulated elements and generally emitted at the beginning of the song. They are longer than single notes and usually precede both phrases and isolated units (Figure 1). Indris emit songs in three different contexts: (1) during territorial encounters between different family groups (territorial songs, Figure 1D); (2) when individuals from the same social group are not in visual contact (cohesion songs, Figure 1E); or (3) as a daily territorial advertisement (advertisement songs, Figure 1F). These song types partly differ in their spectral features<sup>31</sup>; differences in their rhythmic parameters remain untested.

## Observations and recordings

We recorded songs and calls produced by 22 indri groups (51 individuals) living in five different rainforest patches in Madagascar: six groups in the Analamazaotra forest (18° 56' S, 48° 25' E), two groups in the Anjozorobe forest (18° 16' S, 47° 59' E), three groups in the Mitsinjo Station Forestière (18° 56' S, 48° 4' E), nine groups in the Maromizaha forest (18° 56' 49" S, 48° 27' 53" E), and two groups in Mantadia National Park (18° 28' S, 48° 28' E). We collected data in the field from 2005 to 2020 for a total of 70 months. We observed one group per week from 6:00 a.m. to 1:00 p.m., using natural marks to identify each individual. Our analyses focused on adult reproductive indris because temporal features of their songs may differ from those found in young or juvenile individuals.<sup>22,35</sup> We recorded songs and calls using solid-state recorders (SoundDevices 702, Olympus S100 and LS05, and Tascam DR-100, DR-40, and DR-05) equipped with Sennheiser (ME 66 and ME 67) or AKG (CK 98) shotgun microphones. All recordings were sampled at 44.1 kHz with a 16-bit amplitude resolution. Recordists pointed the microphone at the focal animals and attributed each vocalization to the signaler via focal animal sampling.<sup>36</sup> The distance between individuals and the microphone ranged from 2 to 20 meters.

## Acoustic analyses

From indris' choruses, we obtained a total of 820 individual song contributions. Seven hundred and ten contributions belonged to advertisement songs, 51 to territorial songs, and 59 to cohesion songs.<sup>16</sup> We also collected 469 introductory song roar sequences (harsh emissions

hypothesized to have an attention-gathering function). Moreover, we recorded 260 sequences of honks and 26 sequences of alarm roars. Using Praat 5.3.46,<sup>37</sup> we edited and saved the recorded portion as single mono audio files. Using field notes and video recordings, we then identified, annotated, and saved the onsets and offsets of each emission for each individual as a Praat TextGrid. Finally, we identified units and silences for both songs and calls via visual inspection of the spectrograms.

For each individual song contribution and call, we labeled the silent portions, differentiating silences depending on whether they occurred between two units of the same phrase versus two different phrases, or between two different isolated units. For each of the above units and intervals, we extracted the onset timing and duration using Praat. We imported these into R (R Core Team 2017; version 3.4.3) and calculated the inter-onset intervals (hereafter  $t_k$ ). A sequence of  $n$  notes will produce  $n-1$  inter-onset intervals (i.e.,  $t_1, \dots, t_{n-1}$ ).

We then calculated rhythmic ratios following the methodology of Roeske,<sup>17</sup> dividing each  $t_k$  by its duration plus the duration of the following interval:  $r_k = t_k / (t_k + t_{k+1})$ . Further analyses were performed both on the  $t_k$  values and their ratios (hereafter  $r_k$ ). Values of  $t_k$  and  $r_k$  in songs and calls are shown in Figure 1.

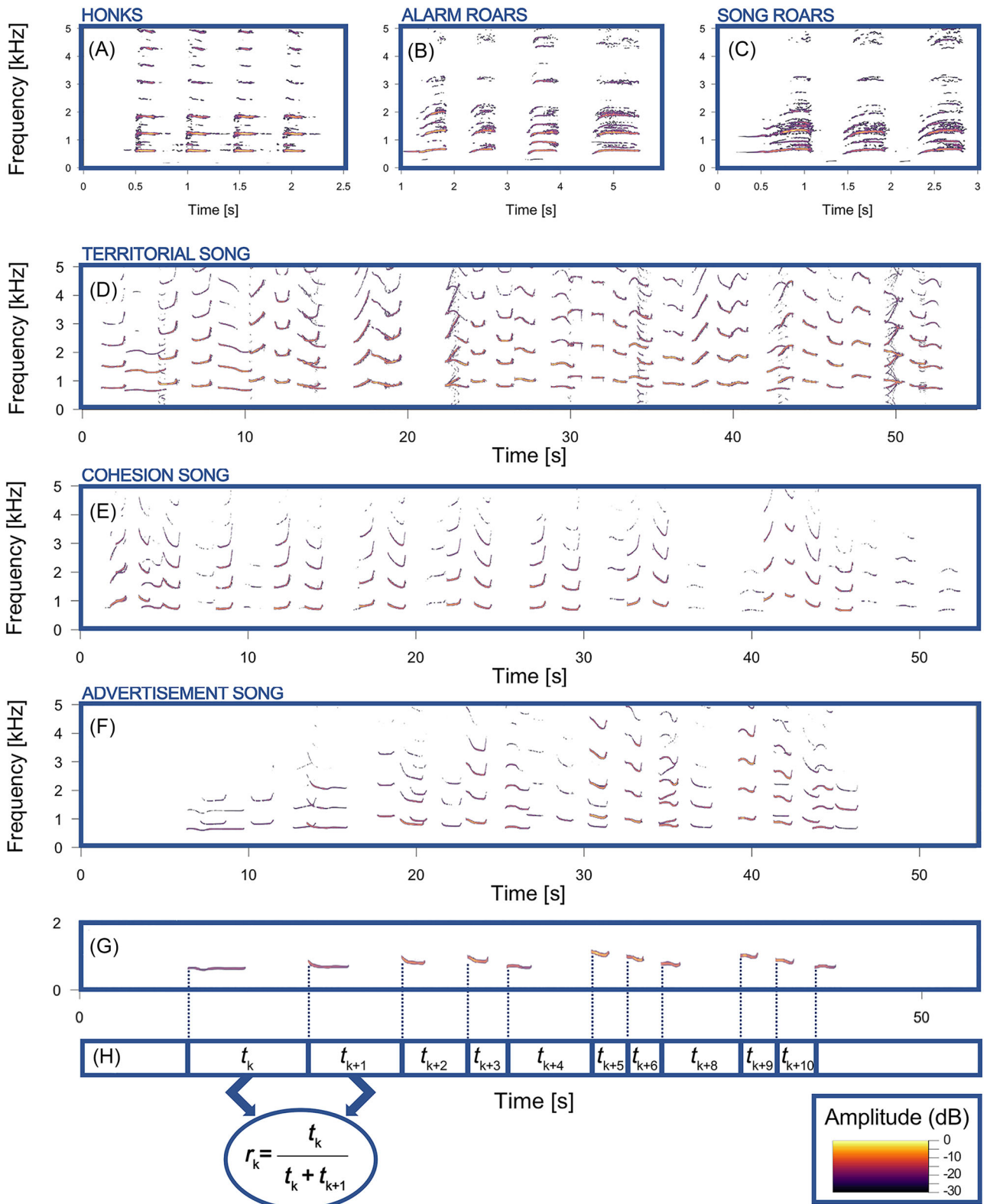
## Statistical analysis

### $t_k$ duration

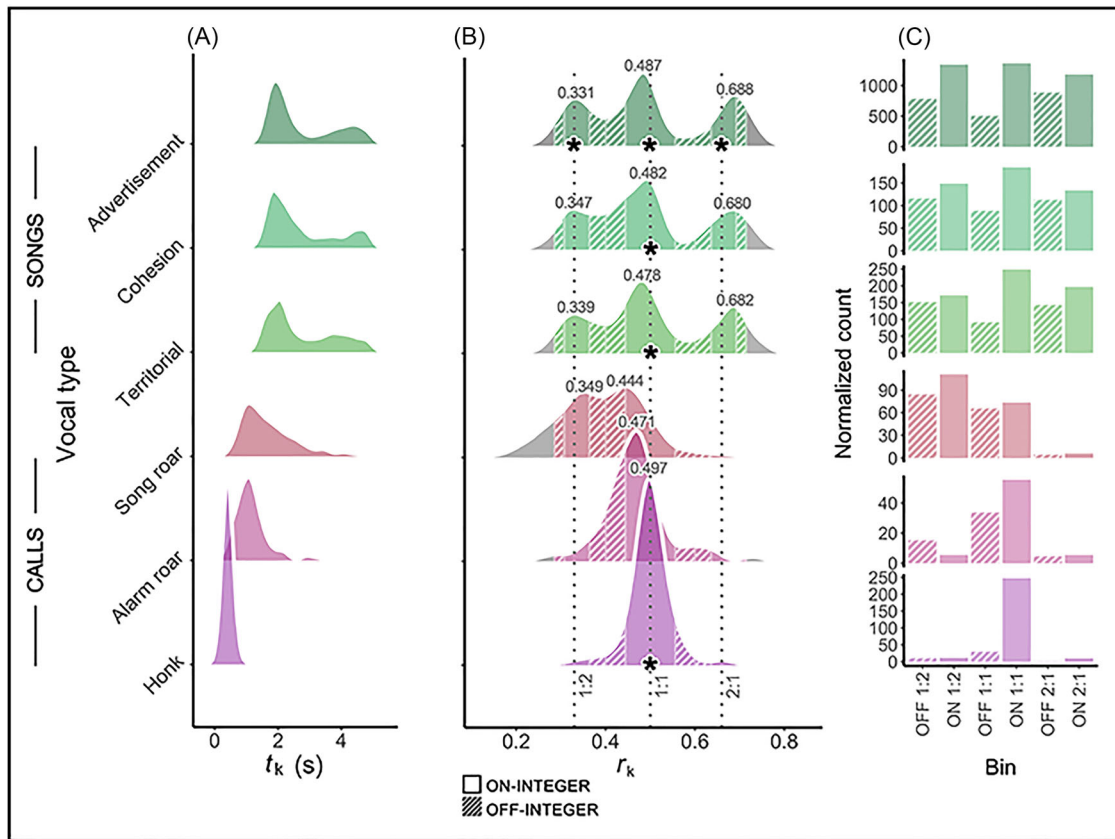
We tested whether the type of vocal sequence (songs versus calls) statistically predicted both the  $t_k$  duration and variability using two linear mixed models (LMM, *lmer* function of *lme4* package<sup>38</sup>). We calculated the individual means of the coefficient of variation (CV), and before fitting the models,  $t_k$  duration and  $t_k$  CV were log-transformed (base e) because they were not normally distributed. We entered the values of  $t_k$  and the  $t_k$  CV as response variables, and the type of vocal sequence (advertisement song, territorial song, cohesion song, honks, alarm roars, and song roars) as fixed factors. We included the individual's ID and the ID of the vocal emission from which we extracted the  $t_k$  as nested random factors. We used Tukey's test<sup>39</sup> to perform pairwise comparisons for all the levels of the fixed factor (multiple comparison package *multcomp* in R). We verified the assumptions of normality and homogeneity of residuals by visually inspecting the qqplot and the residuals' distribution (using a function written by R. Mundry). To test for the significance of the full models,<sup>40</sup> we compared them against the null models containing only the random factors, with a likelihood ratio test (ANOVA with argument test *Chisq*<sup>41</sup>). We report the results of the ANOVA, estimates, standard error (S.E.),  $z$ , and  $p$ -values for the two Tukey's tests (Tables S1 and S2).

### Rhythmic categories: $r_k$ distributions and their peaks

Following the methodology introduced by Roeske and colleagues,<sup>17</sup> the off-integer ratio ranges were centered around 1:3.5 (or 0.285),



**FIGURE 1** Representative spectrograms of the vocal signals analyzed. (A) Honks. (B) Alarm roars. (C) Song roars. (D) Territorial song. (E) Cohesion song. (F) Advertisement song. In all song types, two individuals are singing together. (G) Schematic representation of onsets (dotted blue lines) from an individual contribution to the song. (H) Pairs of onsets define inter-onset intervals ( $t_k$ ), marked by solid blue lines. Rhythmic ratios are calculated by dividing each interval by itself plus the next interval:  $r_k = t_k / (t_k + t_{k+1})$ .



**FIGURE 2**  $t_k$  and  $r_k$  distributions for songs and calls. (A) Probability density functions of  $t_k$  for each vocal type. (B) Probability density functions of rhythm ratios ( $r_k$ ) for each vocal type. Local maxima for  $r_k$  distribution are: 0.331, 0.487, and 0.688 for advertisement songs; 0.347, 0.482, and 0.680 for cohesion songs; 0.339, 0.478, and 0.682 for territorial songs; 0.444 and 0.349 for song roars; 0.471 for alarm roars; and 0.497 for honks. (C) Barplot showing the average normalized  $r_k$  occurrence for on-integer (solid bars) and off-integer (striped bars) ratio ranges for songs and calls. \* $p < 0.05$ ; statistically significant matching between the empirical distribution and a small integer ratio rhythmic category.

1:2.5 (or 0.400), 1-1:2.5 (or 0.600), and 1-1:3.5 (or 0.710); the boundaries of all on- and off-integer ratio ranges were 1:3.25 (or 0.307), 1:2.75 (or 0.363), 1:2.25 (or 0.444), 1-1:2.25 (or 0.555), 1-1:2.75 (or 0.637), and 1-1:3.25 (or 0.693). We counted all occurrences of ratio values that fell into each on- and off-integer ratio range for each individual for each song type (advertisement, cohesion, and territorial) and each call type (song roars, honks, and alarm roars), and we normalized these counts according to the size of their range on the x-axis. We used the Shapiro–Wilk test to assess whether our count data followed a normal distribution. Since our data were not normally distributed, we used three (paired) Wilcoxon signed-rank tests for each type of song and call, to compare on- and off-integer ratio ranges. Finally, we calculated effect sizes for each Wilcoxon signed-rank test.

## RESULTS

### $t_k$ duration

The average duration of  $t_k$  was  $2.670 \pm 1.070$  s for songs and  $0.502 \pm 0.319$  s for calls (Figure 2A). All three song types showed longer  $t_k$  than those found in the calls (Table S1). There was a significant difference in

$t_k$  duration between different calls but not between song types. Within call types, honks had shorter  $t_k$  than song roars ( $p < 0.001$ ) and alarm roars ( $p < 0.001$ ); song roars had longer  $t_k$  than alarm roars ( $p < 0.001$ ). Similarly, the CV did not differ between pairs of song types; in contrast, within call sequences, both roars and honks differed significantly from each other ( $p < 0.001$  for each comparison, Table S2). In particular, honks were the least variable in terms of  $t_k$  duration and showed lower CV values than both roar types, while song roars had higher CV values than honks and alarm roars. Finally, alarm roars showed an intermediate variability, having a higher CV than honks but a lower CV than song roars.

### Rhythmic categories: Ratio distributions and their peaks

#### Songs

The dependent two-group Wilcoxon tests between on-integer and off-integer ratio ranges (Figure 2B,C) confirmed previous findings<sup>8</sup> that indris produce advertisement songs characterized by two rhythmic categories: isochrony, corresponding to a 1:1 ratio ( $p < 0.001$ ,  $V = 4$ ;

effect size = 0.865), and a 1:2 ratio ( $p < 0.001$ ,  $V = 168$ ; effect size = 0.650). In addition, the inclusion in the sample of the long notes emitted at the beginning of the song also revealed the presence of a 2:1 rhythmic ratio ( $p < 0.001$ ,  $V = 264$ ; effect size = 0.485). In other words, the three peaks in Figure 2B result from indris-producing ratios falling within small on-integer ratio neighborhoods rather than within off-integer neighborhoods. The other two song types only had one significant, 1:1 peak: the Wilcoxon tests between on- and off-integer ratio ranges for cohesion and territorial songs showed that both these song types featured 1:1 ratios ( $p = 0.003$ ,  $V = 18$ , effect size = 0.679, and  $p = 0.001$ ,  $V = 9$ , effect size = 0.763, respectively). However, neither the 1:2 ratio ( $p = 0.485$ ,  $V = 54$ , effect size = 0.130, and  $p = 0.410$ ,  $V = 45$ , effect size = 0.213, respectively) nor the 2:1 ratio ( $p = 0.365$ ,  $V = 50$ , effect size = 0.250, and  $p = 0.147$ ,  $V = 34$ , effect size = 0.356, respectively) were significant for cohesion and territorial songs. Table S3 reports  $V$  value,  $p$ -value, and effect size for each comparison between on- and off-integer ratio ranges.

### Calls: Honks, alarm roars, and song roars

The dependent two-group Wilcoxon tests between on- and off-integer ratio ranges (Figure 2B,C) showed that honks were strongly isochronous, exhibiting a 1:1 ratio ( $p < 0.001$ ,  $V = 0$ ). Note that a  $V$  value of 0 for the isochronous 1:1 category means that each of the 51 indris produced more on-integer than off-integer ratios, not only statistically but also numerically at an individual level. As clearly shown in Figure 2B, the other two rhythmic categories were absent (1:2 ratio,  $p = 1.000$ ,  $V = 11$ ; 2:1 ratio,  $p = 1.000$ ,  $V = 0$ ). For both roar types (uttered at the beginning of the song or to signal an aerial predator), the Wilcoxon tests did not highlight the presence of any rhythmic categories; there was only a tendency toward the 1:2 ratio for song roars (song roars,  $p = 0.053$ ,  $V = 31$ ; alarm roars,  $p = 0.201$ ,  $V = 9$ ). Both roars were neither isochronous (1:1 ratio for song roars,  $p = 0.143$ ,  $V = 445$ ; alarm roars,  $p = 0.107$ ,  $V = 18$ ), nor characterized by a 2:1 ratio (song roars,  $p = 0.229$ ,  $V = 28$ ; alarm roars,  $p = 0.577$ ,  $V = 3$ ). Table S3 reports  $V$  values,  $p$ -values, and the effect size for every comparison between on- and off-integer ratio ranges.

## DISCUSSION

We investigated how rhythmic categories vary across indri vocalizations. In particular, we tested for similarities and differences in rhythmic structure between songs and call sequences. We also tested for rhythmic similarities and differences within functionally closer vocal signals (i.e., within song types and within call types). Our data showed rhythmic differences between songs and calls, within song types, and within call types. Overall, songs had a slower tempo than calls (see also Ref. 42). Song inter-onset intervals did not change depending on song type, while call inter-onset intervals changed depending on call type. Moreover, calls showed more variable intervals across types, while songs did not. Songs and calls also differed in their number of rhythmic categories, as songs showed from one to three rhythmic cate-

gories corresponding to small integer ratios, while calls showed zero to one.

As predicted, rhythmic categories are selectively deployed across vocal display types. However, our hypothesis was only partly supported: how these categories emerged in the vocal repertoire did not precisely match our predictions. We found that sequences of roars produced before the song or as aerial predator alarm calls did not show any categorical rhythm at small-integer ratios. Songs uttered in any of three contexts showed a significant peak corresponding to isochrony, as did terrestrial predator alarm calls. One particular song type (i.e., advertisement songs) showed all three tested rhythmic categories.

Our results partly support our hypothesis that the number of rhythm categories corresponding to small integer ratios is consistent within a vocal display type. In particular, not all songs show a consistent number of rhythmic categories. Songs' hierarchical structure may deliver the three categorical peaks visible across all song types (Figure 2B); crucially, as shown by our statistical results, these peaks do not always match small integer ratios. Advertisement, territorial, and cohesion songs all contain phrases<sup>16</sup> but do not share the same significant categorical rhythms. Thus, our results indicate that hierarchical structure in songs is not sufficient to explain the matching of different rhythmic categories. Although isochrony (1:1) is widespread across the three song types, only the advertisement song showed significant peaks at 1:2 and 2:1. Thus, we did not find support for the prediction that songs would show a consistent number of vocal categories corresponding to song-specific integer ratios. Different vocal displays in indris may selectively recruit a subset of categorical rhythms at small integer ratios.

Our results also did not support our second prediction, namely that the overall number of vocal categories appearing in the songs differs from the categories of calls. Territorial songs, cohesion songs, and honks all showed isochrony, while roar sequences did not show any statistical peak corresponding to small integer ratios. Roars at the initial position of songs tended toward 1:2. The  $t_k$  had similar bimodal distributions across the three song types. Conversely, all call types differed in inter-onset interval duration and variability; this difference may derive from the absence of an underlying hierarchical structure shared across call sequences, which instead may exist across songs.

### Three categorical rhythms in the songs of a primate

Our data showed three rhythmic categories in advertisement songs; this makes *Indri indri* and *Homo sapiens* the species to show the most rhythmic categories, even more than songbirds.<sup>17,43,44</sup> This result only superficially contradicts the findings of earlier work that found two, rather than three, rhythmic categories in indri.<sup>8</sup> Adding long notes emitted at the song's beginning to the previous advertisement song dataset<sup>8</sup> explains all three rhythmic categories. This result suggests that the first interval between the two long notes' onsets tends to be twice the second one, confirming the strongly regular nature of this particular singing display. More generally, it suggests that comparative rhythm research should consider how intro and outro elements of vocalizations may affect inference on the existence of categories.

Why do we find different number of rhythmic categories across song types? While rhythmic categories may be a building block of hierarchical structure in songs, these categories need not match precise small integer ratios. In fact, we hypothesize that the interactivity of a vocal display may mediate whether hierarchical structure is built on small integer ratio categories. Indris show different levels of interaction across song types,<sup>16</sup> and interaction modulates rhythmic regularity in another singing primate, the lar gibbon.<sup>44</sup> Mechanistically, a precise matching of 1:2 and 2:1 rhythmic categories would increase predictability and enable higher note overlap between group members.<sup>16</sup> Higher overlap would deliver amplitude summation, hence louder sounds. Functionally, higher overlap in territorial songs could have been positively selected to sound more intimidating to neighboring groups.<sup>45,46</sup> To corroborate this hypothesis, we know that cohesion songs show low overlap; these songs also contain silent intervals lasting several seconds, slow tempo, and solo song sections.<sup>34</sup> In brief, rhythmic differences across vocal displays may be modulated by synchrony and interactivity among group members.

Plots of rhythmic categories may look qualitatively similar across song types, but only advertisement songs show significant peaks at 1:2, 1:1, and 2:1. Why do advertisement songs emerge as the most rhythmic among the six vocalization types? Cohesion songs facilitate cohesion within a family group. Territorial songs serve to maintain possession of a boundary. Advertisement songs announce territorial occupancy and mediate inter-group spacing.<sup>25</sup> Advertisement songs are routinely produced each morning,<sup>47</sup> while cohesion and territorial songs are rarely produced overall.<sup>16,25</sup> From the advertisement song, more than from other types of songs, conspecifics might derive important information such as sex and age,<sup>27,35</sup> the formation of a new pair, or the decision to challenge a neighboring group in search of an extension of the territorial boundary.<sup>48</sup> Considering all of this, it would make sense if advertisement songs were the song type under the strongest selective pressures. As with rock hyraxes,<sup>43</sup> greater regularity, not just greater isochrony, could be a key signal selected by potential partners. Although direct empirical testing may be difficult because of indri sociobiology, multicategorical rhythmicity in advertisement songs but not in cohesion or territorial ones may be a sexually selected trait.

The context of emission may shape song rhythmicity. Advertisement songs are produced while individuals are physically close and in visual contact; they are also thought to be used by nonreproductive individuals to convey information on sex and age, hence mediating mate finding at a distance.<sup>22,49</sup> Cohesion songs, where individuals are hundreds of meters away in a dense rainforest environment, serve to locate the position of the missing partners.<sup>16</sup> Finally, territorial songs feature individuals from neighboring groups singing all together; indris acoustically overlap when engaging in vocal battles that may last hours. Based on these ecological factors, isochrony may be the only rhythmic feature that indris can maintain (and attend to) in territorial and cohesion songs. An alternative hypothesis is that the statistical differences in rhythm categories across songs may be of degree and not kind; in other words, different functions may map to similar rhythmic structures whose similarity we failed to detect here. Future work should empirically tackle these alternative hypotheses.

## Rhythmic categories in calls

Isochrony in modern humans is peculiar because it is omnipresent in music but only seldom occurs in speech.<sup>50</sup> In other species, isochrony has not been studied systematically.<sup>50</sup> Previous studies have shown that social interactions, such as duetting behavior, play an important role in modulating the isochronous structure of primates' song.<sup>44</sup> We show in the present study that isochrony is present not only during coordinated singing behavior but also in the spontaneous vocal emissions triggered by the presence of a terrestrial predator. In other words, while humans show isochrony in music but rarely in speech, indri show isochrony in both songs and calls—loose functional equivalents of human music and speech.

Finding isochrony in both indri songs and call sequences partly dovetails with Geissmann's<sup>11,51</sup> hypothesis, whereby loud calls in apes and human music might have derived from ancestral forms of loud callings, such as antipredator calls. Furthermore, the presence of rhythmic categories in alarm calls confirms that the temporal patterning of sounds shows cross-domain similarities, from speech to music and animal vocalizations.<sup>2</sup> Calls show zero to one rhythmic category, while songs show one to three. Song–call rhythms in indri seem to sit in a graded spectrum, with isochrony as a leitmotif. Following Geissmann's reasoning,<sup>11</sup> one hypothesis could entail that alarm call rhythmicity is ancestral to other vocal behavior rhythmicity. An alternative hypothesis is that isochrony may be an ancestral trait of both calls and songs, remaining agnostic on which one came first.

Is isochrony in indri songs fairly plastic or mostly physiologically determined? Finding isochrony in honks, but not in antipredator roars, may provide an answer. Honks and roars are elicited by different predators. Honks show almost perfect isochrony, with a peak at 0.497. However, the alarm roar ratio distribution is shifted toward values lower than 1:1 (0.471), suggesting a gradually decelerating tempo (also known as *ritardando*). Roars provide an important baseline: indris are able to produce nonisochronous rhythms, which in turn suggests that isochrony is not a default, automatic behavior. Given that alarm roars are three times longer than honks,<sup>12</sup> we hypothesize that indris emit honks at an optimal rate of respiration, while alarm roars might be subjected to respiratory constraints that lead to a delay in the emission of subsequent sounds. This would translate into a series of progressively longer inter-onset intervals. Because honks show the shortest and least variable inter-onset intervals, their isochrony may also be the byproduct of a physiological constraint that imposes a lower limit in the duration of those intervals; in lar gibbons, a higher call rate increases isochrony.<sup>44</sup> Crucially, however, both indris and lar gibbons produce isochrony at different tempi, suggesting that high levels of rhythmic regularity occur at different call rates.<sup>44</sup> Finally, we also show that roar elements, despite sounding quite similar, have different rhythmicity depending on whether they precede a song or are emitted as a sequence of calls. When preceding a song, roars have longer inter-onset intervals and a bimodal distribution tending toward a 1:2 ratio. Conversely, alarm roars show a unimodal distribution, as do honks. All this suggests that isochrony in indris is fairly plastic and not fully physiologically determined.

## CONCLUSIONS

Mechanistic work on singing animals is, with a few notable exceptions (e.g., songbirds, singing mice), almost impossible. Indris are critically endangered and cannot be bred in captivity, so directly tapping the mechanisms behind their singing is unfeasible.<sup>52,53</sup> This limits our understanding of primates' song production and perception. Still, fine-grained behavioral studies can highlight factors regulating the rhythmic organization of singing behavior. On the basis of indirect evidence, we hypothesize that rhythmic categories, besides having been mostly studied in singing displays in different animals<sup>17,43,51,54</sup> owing to their particular link to human music, have not evolved for song purposes only, at least in indris. Still, the difference in the number of different rhythmic categories between songs and other vocalizations is striking. We found that isochrony was common in all song types, but not in all call types, leading us to hypothesize that it could be ancestral, at least to a subset of the vocal repertoire.

Our results may inform debates on the origins of music and speech. Specific components of human musicality, with the present study focusing on multiple rhythm categories, are shared with nonhuman primate vocal repertoires. We suggest that, even if rhythm categories are absent in spoken language, they might have played a role in our ancestors' communication.

## AUTHOR CONTRIBUTIONS

C.D.G. and M.G. designed the study. C.D.G., D.V., M.M., T.R., F.C., L.M., and V.T. collected the data. C.D.G., T.R., and D.V. performed the analyses. T.R. and D.V. made the figures. C.D.G. wrote the original draft. All authors reviewed and edited the manuscript.

## ACKNOWLEDGMENTS

A.R. is funded by the European Union (ERC, TOHR, 101041885). Center for Music in the Brain is funded by the Danish National Research Foundation (DNRF117). The Comparative Bioacoustics Group was funded by Max Planck Group Leader funding to A.R.

Open access funding enabled and organized by Projekt DEAL.

## COMPETING INTERESTS

The authors declare no competing interests.

## ORCID

Chiara De Gregorio  <https://orcid.org/0000-0001-7017-6181>

Andrea Ravnani  <https://orcid.org/0000-0002-1058-0024>

## PEER REVIEW

The peer review history for this article is available at: <https://publons.com/publon/10.1111/nyas.15151>

## REFERENCES

- Haiduk, F., & Fitch, W. T. (2022). Understanding design features of music and language: The choric/dialogic distinction. *Frontiers in Psychology*, 13, 1126. <https://doi.org/10.3389/fpsyg.2022.786899>
- Patel, A. D. (2003). Rhythm in language and music. *Annals of the New York Academy of Sciences*, 999(1), 140–143. <https://doi.org/10.1196/ANNALS.1284.015>
- Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1664), 20140088. <https://doi.org/10.1098/RSTB.2014.0088>
- Ravnani, A., Bowling, D., & Fitch, W. T. (2014). Chorusing, synchrony and the evolutionary functions of rhythm. *Frontiers in Psychology*, 5, 1118. <https://doi.org/10.3389/fpsyg.2014.01118>
- Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 103–123).
- De Gregorio, C., Carugati, F., Valente, D., Raimondi, T., Torti, V., Miarsetsoa, L., Gamba, M., & Giacoma, C. (2022). Notes on a tree: Reframing the relevance of primate choruses, duets, and solo songs. *Ethology Ecology & Evolution*, 34(3), 205–219. <https://doi.org/10.1080/03949370.2021.2015451>
- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences of the United States of America*, 112(29), 8987–8992. <https://doi.org/10.1073/PNAS.1414495112>
- De Gregorio, C., Valente, D., Raimondi, T., Torti, V., Miarsetsoa, L., Friard, O., Giacoma, C., Ravnani, A., Gamba, M., & Ravnani, A. (2021). Categorical rhythms in a singing primate. *Current Biology*, 31(20), R1379–R1380. <https://doi.org/10.1016/j.cub.2021.09.032>
- Essens, P. J., & Povel, D. J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics*, 37(1), 1–7.
- Toussaint, G. T. (2013). The pairwise variability index as a measure of rhythm complexity. *Analytical Approaches to World Music*, 2(2), 1–42.
- De Gregorio, C., Zanolli, A., Valente, D., Torti, V., Bonadonna, G., Randrianarison, R. M., Giacoma, C., & Gamba, M. (2019). Female indris determine the rhythmic structure of the song and sustain a higher cost when the chorus size increases. *Current Zoology*, 65(1), 89–97.
- Maretti, G., Sorrentino, V., Finomana, A., Gamba, M., & Giacoma, C. (2010). Not just a pretty song: An overview of the vocal repertoire of *Indri indri*. *Journal of Anthropological Sciences*, 88, 151–165. <http://europepmc.org/abstract/MED/20834055>
- Valente, D., De Gregorio, C., Torti, V., Miarsetsoa, L., Friard, O., Randrianarison, R. M., Giacoma, C., & Gamba, M. (2019). Finding meanings in low dimensional structures: Stochastic neighbor embedding applied to the analysis of *Indri indri* vocal repertoire. *Animals*, 9(5), 243. <https://doi.org/10.3390/ani9050243>
- Kello, C. T., Bella, S. D., Médé, B., & Balasubramaniam, R. (2017). Hierarchical temporal structure in music, speech and animal vocalizations: Jazz is like a conversation, humpbacks sing like hermit thrushes. *Journal of the Royal Society Interface*, 14(135), 20170231. <https://doi.org/10.1098/RSIF.2017.0231>
- Valente, D., De Gregorio, C., Favaro, L., Friard, O., Miarsetsoa, L., Raimondi, T., Ratsimbazafy, J., Torti, V., Zanolli, A., Giacoma, C., & Gamba, M. (2021). Linguistic laws of brevity: Conformity in *Indri indri*. *Animal Cognition*, 24, 897–906. <https://doi.org/10.1007/s10071-021-01495-3>
- Torti, V., Gamba, M., Rabemananjara, Z. H., & Giacoma, C. (2013). The songs of the indris (Mammalia: Primates: Indridae): Contextual variation in the long-distance calls of a lemur. *Italian Journal of Zoology*, 80(4), 596–607. <https://doi.org/10.1080/11250003.2013.845261>
- Roeske, T. C., Tchernichovski, O., Poeppel, D., & Jacoby, N. (2020). Categorical rhythms are shared between songbirds and humans. *Current Biology*, 30(18), 3544–3555.e6. <https://doi.org/10.1016/j.cub.2020.06.072>
- Darwin, C. (1871). *The descent of man and selection in relation to sex* (Vol. 1). John Murray.
- Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biological Reviews*, 77(1), 57–76. <https://doi.org/10.1017/S1464793101005826>



20. Bowling, D. L., Herbst, C. T., & Fitch, W. T. (2013). Social origins of rhythm? Synchrony and temporal regularity in human vocalization. *PLoS ONE*, 8(11), e80402.
21. Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, 24(1), 99–104. <https://doi.org/10.1525/MP.2006.24.1.99>
22. Gamba, M., Torti, V., Estienne, V., Randrianarison, R. M., Valente, D., Rovara, P., Bonadonna, G., Friard, O., & Giacoma, C. (2016). The Indris have got rhythm! Timing and pitch variation of a primate song examined between sexes and age classes. *Frontiers in Neuroscience*, 10(JUN), 249. <https://doi.org/10.3389/fnins.2016.00249>
23. Burchardt, L. S., Norton, P., Behr, O., Scharff, C., & Knörnschild, M. (2019). General isochronous rhythm in echolocation calls and social vocalizations of the bat *Saccopteryx bilineata*. *Royal Society Open Science*, 6(1), 181076. <https://doi.org/10.1098/rsos.181076>
24. Bonadonna, G., Torti, V., De Gregorio, C., Valente, D., Randrianarison, R. M., Pozzi, L., Gamba, M., & Giacoma, C. (2019). Evidence of genetic monogamy in the lemur Indri (*Indri indri*). *American Journal of Primatology*, 81(6), e22993. <https://doi.org/10.1002/ajp.22993>
25. Bonadonna, G., Zaccagno, M., Torti, V., Valente, D., De Gregorio, C., Randrianarison, R. M., Tan, C., Gamba, M., & Giacoma, C. (2020). Intra- and intergroup spatial dynamics of a pair-living singing primate, *Indri indri*: A multiannual study of three Indri groups in Maromizaha forest, Madagascar. *International Journal of Primatology*, 41, 224–245. <https://doi.org/10.1007/s10764-019-00127-5>
26. Zanoli, A., De Gregorio, C., Valente, D., Torti, V., Bonadonna, G., Randrianarison, R. M., Giacoma, C., & Gamba, M. (2020). Sexually dimorphic phrase organization in the song of the indris (*Indri indri*). *American Journal of Primatology*, 82(6), e23132. <https://doi.org/10.1002/AJP.23132>
27. De Gregorio, C., Zanoli, A., Carugati, F., Raimondi, T., Valente, D., Torti, V., Miarretsoa, L., Rajaonson, A., Gamba, M., & Giacoma, C. (2022). Parent-offspring turn-taking dynamics influence parents' song structure and elaboration in a singing primate. *Frontiers in Ecology and Evolution*, 10, 906322. <https://doi.org/10.3389/fevo.2022.906322>
28. Torti, V., Valente, D., De Gregorio, C., Comazzi, C., Miarretsoa, L., Ratsimbazafy, J., Giacoma, C., & Gamba, M. (2018). Call and be counted! Can we reliably estimate the number of callers in the indri's (*Indri indri*) song? *PLoS ONE*, 13(8), e0201664.
29. Zanoli, A., Raimondi, T., De Gregorio, C., Valente, D., Carugati, F., Torti, V., Friard, O., Longondraza, M., Giacoma, C., & Gamba, M. (2023). "The song remains the same": Not really! Vocal flexibility in the song of the indris. *Animal Cognition*, 26, 2009–2021.
30. Pollock, J. I. (1979). Female dominance in *Indri indri*. *Folia Primatologica*, 31(1–2), 143–164.
31. Giacoma, C., Sorrentino, V., Rabarivola, C., & Gamba, M. (2010). Sex differences in the song of *Indri indri*. *International Journal of Primatology*, 31(4), 539–551. <https://doi.org/10.1007/s10764-010-9412-8>
32. Valente, D., Miarretsoa, L., Anania, A., Costa, F., Mascaro, A., Raimondi, T., De Gregorio, C., Torti, V., Friard, O., Ratsimbazafy, J., Giacoma, C., & Gamba, M. (2022). Comparative analysis of the vocal repertoires of the indri (*Indri indri*) and the diademmed sifaka (*Propithecus diadema*). *International Journal of Primatology*, 43(4), 733–751. <https://doi.org/10.1007/s10764-022-00287-x>
33. Spector, D. A. (1994). Definition in biology: The case of "bird song". *Journal of Theoretical Biology*, 168(4), 373–381.
34. Pollock, J. I. (1986). The song of the Indris (*Indri indri*; Primates: Lemuroidea): Natural history, form, and function. *International Journal of Primatology*, 7(3), 225–264. <https://doi.org/10.1007/BF02736391>
35. De Gregorio, C., Carugati, F., Estienne, V., Valente, D., Raimondi, T., Torti, V., Miarretsoa, L., Ratsimbazafy, J., Gamba, M., & Giacoma, C. (2021). Born to sing! Song development in a singing primate. *Current Zoology*, 67(6), 597–608. <https://doi.org/10.1093/cz/zoab018>
36. Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, 49(3–4), 227–266. <https://doi.org/10.1163/156853974x00534>
37. Boersma, P., & Weenink, D. (2016). PRAAT: Doing phonetics by computer (5.0. 21).
38. Bates, D., Maechler, M., & Bolker, B. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
39. Bretz, F., Hothorn, T., & Westfall, P. (2011). Multiple comparisons using R.
40. Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, 65(1), 47–55. <https://doi.org/10.1007/s00265-010-1038-5>
41. Dobson, A., & Barnett, L. A. (2002). *An introduction to generalized linear models*. CRC Press.
42. Ozaki, Y., Tierney, A., Pfordresher, P. Q., McBride, J. M., Benetos, E., Proutskova, P., Chiba, G., Liu, F., Jacoby, N., Purdy, S., Opondo, P., Fitch, T., Hegde, S., Rocamora, M., Thorne, R., Nweke, F. E., Sadaphal, D., Sadaphal, P., Hadavi, S., ... Savage, P. E. (2024). Globally, songs and instrumental melodies are slower and higher and use more stable pitches than speech: A Registered Report. *Science Advances*, 10(20), eadm9797.
43. Demartsev, V., Haddas-Sasson, M., Ilany, A., Koren, L., Geffen, E., & Koren, L. (2022). Male rock hyraxes that maintain an isochronous song rhythm achieve higher reproductive success. *Journal of Animal Ecology*, 92(8), 1520–1531. <https://doi.org/10.1111/1365-2656.13801>
44. Raimondi, T., Di Panfilo, G., Pasquali, M., Zarantonello, M., Favaro, L., Savini, T., Gamba, M., & Ravignani, A. (2023). Isochrony and rhythmic interaction in ape duetting. *Proceedings of the Royal Society B: Biological Sciences*, 290(1990), 20222244. <https://doi.org/10.1098/RSPB.2022.2244>
45. Briseño-Jaramillo, M., Berthet, M., Estrada, A., Biquand, V., & Lemasson, A. (2021). Socially mediated overlap in vocal interactions between free-ranging black howler monkeys. *American Journal of Primatology*, 83(8), e23297. <https://doi.org/10.1002/AJP.23297>
46. Dalziell, A. H., Maisey, A. C., Magrath, R. D., Welbergen, J. A., Dalziell, A. H., Maisey, A. C., Magrath, R. D., & Welbergen, J. A. (2021). Male lyrebirds create a complex acoustic illusion of a mobbing flock during courtship and copulation. *Current Biology*, 31, 1970–1976. <https://doi.org/10.1016/j.cub.2021.02.003>
47. Ravaglia, D., Ferrario, V., De Gregorio, C., Carugati, F., Raimondi, T., Cristiano, W., Torti, V., Hardenberg, A. V., Ratsimbazafy, J., Valente, D., Giacoma, C., & Gamba, M. (2023). There you are! Automated detection of Indris' songs on features extracted from passive acoustic recordings. *Animals*, 13(2), 241. <https://doi.org/10.3390/ANI13020241>
48. Spezie, G., Torti, V., Bonadonna, G., De Gregorio, C., Valente, D., Giacoma, C., & Gamba, M. (2022). Evidence for acoustic discrimination in lemurs: A playback study on wild indris (*Indri indri*). *Current Zoology*, 69, 41–49. <https://doi.org/10.1093/CZ/ZOAC009>
49. Torti, V., Bonadonna, G., De Gregorio, C., Valente, D., Randrianarison, R. M., Friard, O., Pozzi, L., Gamba, M., & Giacoma, C. (2017). An intra-population analysis of the indris' song dissimilarity in the light of genetic distance. *Scientific Reports*, 7(1), 10140. <https://doi.org/10.1038/s41598-017-10656-9>
50. Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, 8, 280885. <https://doi.org/10.3389/fpsyg.2017.01820>
51. De Gregorio, C., Raimondi, T., Bevilacqua, V., Pertosa, C., Valente, D., Carugati, F., Bandoli, F., Favaro, L., Lefaux, B., Ravignani, A., & Gamba, M. (2023). Isochronous singing in 3 crested gibbon species (*Nomascus* spp.). *Current Zoology*, zoad029. <https://doi.org/10.1093/cz/zoad029>
52. Andriaholinirina, N., Baden, A., Blanco, M., Chikhi, L., Cooke, A., Davies, N., Dolch, R., Donati, G., Ganzhorn, J., Golden, C., Groeneveld, L.

- F., Irwin, M., Johnson, S., Kappeler, P., King, T., Lewis, R., Louis, E. E., Markolf, M., Mass, V., & Mittermeier, A. (2020). *Indri indri*. The IUCN Red List of Threatened Species. Version 2020–3. <https://www.iucnredlist.org>
53. Britt, A., Randriamandratonirina, N. J., Glasscock, K. D., & Iambana, B. R. (2003). Diet and feeding behaviour of *Indri indri* in a low-altitude rain forest. *Folia Primatologica*, 73(5), 225–239. <https://doi.org/10.1159/000067455>
54. Anichini, M., de Reus, K., Hersh, T. A., Valente, D., Salazar-Casals, A., Berry, C., Keller, P. E., & Ravignani, A. (2023). Measuring rhythms of vocal interactions: A proof of principle in harbour seal pups. *Philosophical Transactions of the Royal Society B*, 378(1875), 20210477. <https://doi.org/10.1098/RSTB.2021.0477>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** De Gregorio, C., Maiolini, M., Raimondi, T., Carugati, F., Miaretsoa, L., Valente, D., Torti, V., Giacomini, C., Ravignani, A., & Gamba, M. (2024). Isochrony as ancestral condition to call and song in a primate. *Ann NY Acad Sci*, 1537, 41–50. <https://doi.org/10.1111/nyas.15151>