

<https://doi.org/10.1038/s42003-024-06253-y>

Vocal convergence and social proximity shape the calls of the most basal Passeriformes, New Zealand Wrens

Check for updates

Ines G. Moran^{1,2}✉, Yen Yi Loo^{1,2}, Stilianos Louca³, Nick B. A. Young⁴, Annabel Whibley¹, Sarah J. Withers¹, Priscila M. Salloum⁵, Michelle L. Hall^{6,7,8}, Margaret C. Stanley^{1,2} & Kristal E. Cain^{1,2}

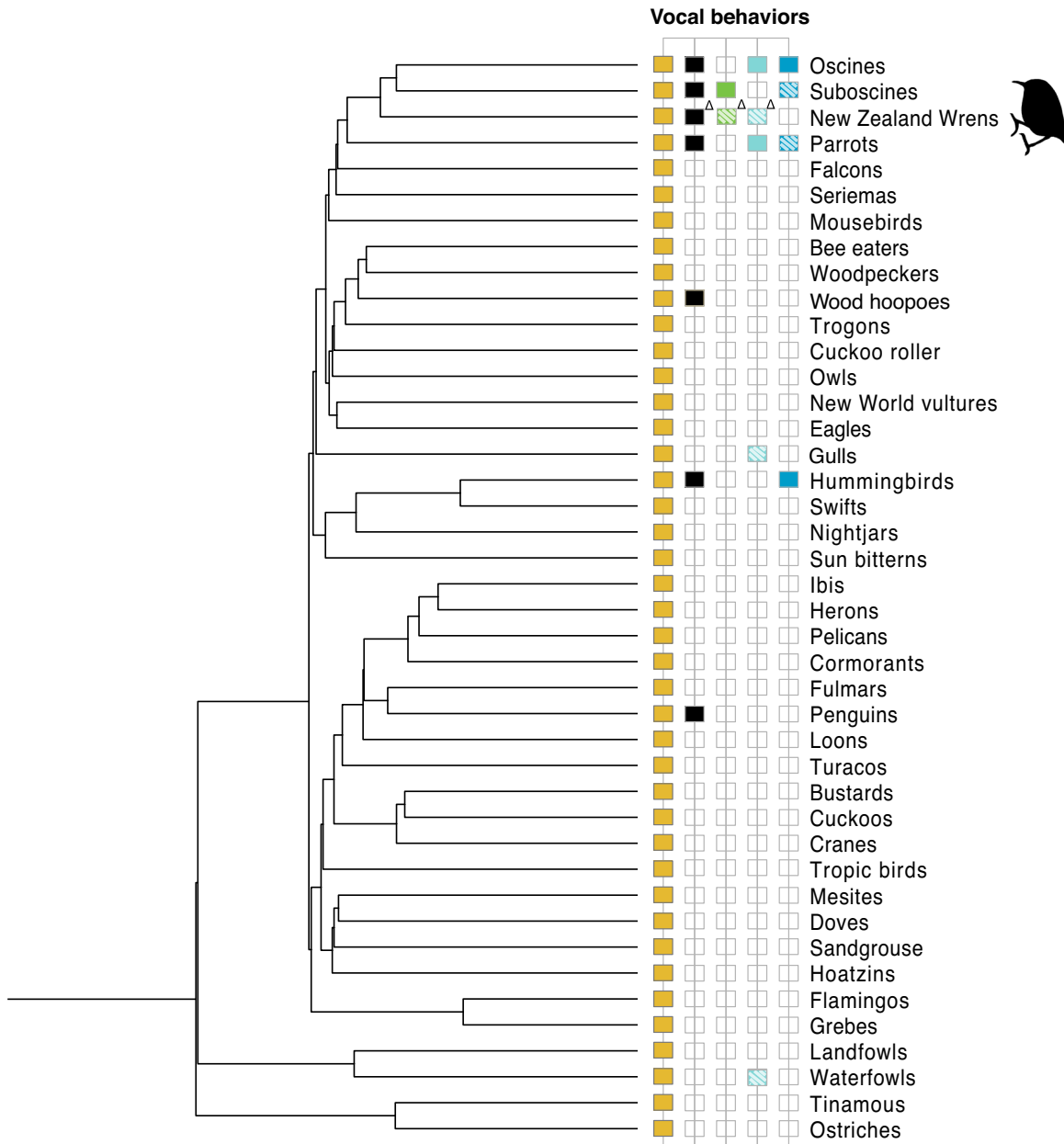
Despite extensive research on avian vocal learning, we still lack a general understanding of how and when this ability evolved in birds. As the closest living relatives of the earliest Passeriformes, the New Zealand wrens (*Acanthisitti*) hold a key phylogenetic position for furthering our understanding of the evolution of vocal learning because they share a common ancestor with two vocal learners: oscines and parrots. However, the vocal learning abilities of New Zealand wrens remain unexplored. Here, we test for the presence of prerequisite behaviors for vocal learning in one of the two extant species of New Zealand wrens, the rifleman (*Acanthisitta chloris*). We detect the presence of unique individual vocal signatures and show how these signatures are shaped by social proximity, as demonstrated by group vocal signatures and strong acoustic similarities among distantly related individuals in close social proximity. Further, we reveal that rifleman calls share similar phenotypic variance ratios to those previously reported in the learned vocalizations of the zebra finch, *Taeniopygia guttata*. Together these findings provide strong evidence that riflemen vocally converge, and though the mechanism still remains to be determined, they may also suggest that this vocal convergence is the result of rudimentary vocal learning abilities.

Most vocal animals communicate with innate vocalizations, but a few taxa are capable of vocal production learning – a behavior that provides animals with the learning ability to copy, match, or imitate sound¹. Species that vocally learn include a wide range of distantly related taxa such as cetaceans², pinnipeds³, elephants⁴, bats⁵, humans, hummingbirds⁶, parrots⁷, songbirds⁸, and although more research is needed, it appears they also include a few suboscines (e.g. bellbirds, *Procnias* spp⁹), African naked mole-rats (*Heterocephalus glaber*)¹⁰, musk ducks (*Biziura lobata*)¹¹ and black-headed gulls (*Larus ridibundus*), among others¹². This paraphyly of vocal learners has led to many hypotheses about the evolution of vocal learning, along with a relatively new hypothesis, which suggests that vocal production learning exists along a continuum consisting of modules^{1,13} – as opposed to a binary dichotomy between vocal learners and vocal non-learners (absence vs. presence). In this hypothesis, vocal production learning is made up of distinct, yet connected behavioral modules (e.g. vocal convergence, vocal

matching, mimicry, and song sharing) – resulting in varying levels of vocal learning complexity (i.e., absent, limited/rudimentary, advanced)^{1,14–17}.

Birds are an excellent group to explore this hypothesis due to their diverse vocal production learning abilities. While advanced vocal learning is well established in parrots (Psittaciformes)⁷, hummingbirds (Trochiliformes)⁶, and oscine songbirds (Passeriformes)⁸, the picture is less clear for suboscines (Passeriformes) and the New Zealand wrens (Passeriformes and sister sub-order to oscines and suboscines, Fig. 1). Suboscines have traditionally been classified as vocal non-learners^{18,19}, but some species have been reported as vocal learners⁹, or as limited learners with a rudimentary neural circuitry related to vocal learning in oscine songbirds^{14,20}. As for New Zealand wrens, their vocal learning abilities have never been directly tested, and have been assumed to be nonexistent based on their simple syrinx morphology (i.e., lacking the intrinsic muscles present in vocal learners^{21,22}), and their basic and short call structure (i.e. lacking

¹School of Biological Sciences, University of Auckland, Auckland 1142, Aotearoa New Zealand. ²Centre for Biodiversity and Biosecurity, University of Auckland, Auckland 1142, Aotearoa New Zealand. ³Department of Biology, University of Oregon, Eugene 97403-1210 OR, USA. ⁴Centre for eResearch, University of Auckland, Auckland 1142, Aotearoa New Zealand. ⁵Department of Zoology, University of Otago, Dunedin 9016, Aotearoa New Zealand. ⁶School of BioSciences, University of Melbourne, Melbourne, VIC 3010, Australia. ⁷Bush Heritage Australia, Melbourne, VIC 3000, Australia. ⁸School of Biological Sciences, University of Western Australia, Perth, WA 6009, Australia. ✉e-mail: inesmoran.publications@gmail.com



Vocal behaviors found in some bird species

- Production of innate calls
- Vocal convergence (i.e. convergence of calls or songs towards common group vocal features)
- Predisposition for rudimentary/limited vocal learning
- Vocal production learning – call learning (e.g. learning of functional, short vocalizations such as contact calls)
- Vocal production learning – song learning (e.g. learning of territorial and courtship vocalizations)
- Vocal behaviors related to vocal production learning have not yet been systematically assessed/or studied
- The presence of the vocal behavior is supported by some evidence in some species, but more research is needed
- Δ Vocal behavior found in this current study

complex and broadcast songs^{23–25}). However, recent revisions of the avian phylogeny show that the New Zealand wrens share a close common ancestor with vocal learning parrots and oscines, and with suboscines^{26–30} (Fig. 1). According to the continuum/module hypotheses^{1,15}, this opens the possibility for New Zealand wrens to have rudimentary learning abilities. Investigating New Zealand wrens’ vocal behaviour and plasticity, and

determining where this species fits into the rudimentary/vocal learning continuum hypotheses is hence key to resolve gaps in the evolution of vocal production learning in Passeriformes.

In wild and at-risk animal populations, such as New Zealand wrens, well-established approaches that enable the detection of vocal production learning abilities (e.g., cross-fostering, social isolation, deafening)^{31–33} are

Fig. 1 | Schematic phylogenetic tree with associated vocal behaviors of New Zealand wrens and other birds. Avian vocal behaviors are categorized based on neurobiological and/or behavioral studies or anecdotes (see below) and are represented with colored rectangles. Behaviors, such as vocal production learning for calls and songs, are separated due to their distinct neurological and functional basis, which may have evolved separately^{131,132}. Diagonally barred rectangles indicate that, although there is some evidence for vocal production learning in this taxon, further testing is needed. When no data is available on the presence or absence of vocal production learning in a clade, vocal behaviors are represented with empty rectangles. Triangles highlight novel and previously unknown behaviors found in this current study (i.e. in the rifleman). Innate vocalizations (labeled in yellow) are present in all bird clades suggesting that they were likely present in the common ancestor to all birds. Seven bird clades vocally converge toward common group vocal

signatures (labeled in black) and include oscines^{41,133,134}, suboscines¹⁰¹, New Zealand wrens (based on current study), parrots^{44,76,135}, wood-hoopoes⁷⁶, hummingbirds⁸² and penguins⁹². Two bird clades have possible rudimentary/limited vocal learning predispositions (labeled in green), the suboscines – based on neurological evidence^{14,20} and New Zealand Wrens – based on the quantitative genetic model comparisons from this study. Call learning (labeled in light blue) has been demonstrated in oscines^{42,51–53}, parrots⁵⁶ and although more research is needed, it is also present in one species of gulls^{12,136} and ducks^{11,12} and may be present in New Zealand Wrens (this study). While song learning (labeled in cyan blue) has clearly been demonstrated in oscines^{8,137} and hummingbirds^{6,138}, other forms of song learning have also been found in suboscines⁹ and parrots^{139–141}. This simplified subtree was derived from Jetz et al.¹⁴².

often unfeasible. Alternative innovative methods are needed, such as the detection of behavioral modules and behavioral predispositions unique to vocal production learning. An example of such a behavioral module is vocal convergence – a form of vocal modification controlled and maintained over time among socially close conspecifics. To achieve vocal convergence, individuals change their unique vocal signatures toward common group vocal features, resulting in group vocal signatures^{34–39}. Vocal convergence is a particularly important behavior in the investigation of vocal production learning because it may have been a precursor of more complex forms of vocal production learning (e.g., vocal matching or mimicry)^{17,36,39,40}.

Another powerful tool for detecting vocal production learning is quantitative genetics as it facilitates the investigation of the role of genetics and social environment in shaping vocal behavior. In vocal non-learners, kin are expected to sound more similar due to their shared genetics⁴¹, but in vocal learners, individuals often copy sounds from distantly related social partners, eroding that similarity^{34,37,42–44}. Thus, if distantly related individuals sound more similar than their close kin, this could suggest that a form of vocal imitation occurs. Furthermore, according to the phenotypic plasticity continuum⁴⁵, it is possible to distinguish vocal learners from vocal non-learners, by partitioning genetics from social environment and by determining how these latter factors contribute to the phenotypic plasticity and variation of vocalizations⁴⁶. Accordingly, vocal learners are expected to have a phenotypic vocal plasticity strongly associated with social environment, while vocal non-learners are expected to show limited voluntary vocal control and display minimal phenotypic vocal plasticity associated with social environment (i.e., indicative of a stronger genetic basis of vocalizations)^{45,47}.

By using alternative and integrative approaches, we aim to determine whether the rifleman (titipounamu, *Acanthisitta chloris*), one of the only two extant species of New Zealand Wrens, has predispositions for vocal production learning (e.g., rudimentary vocal learning abilities). Among riflemen's large call repertoire⁴⁸, feeding calls are a good candidate for this investigation. They are produced in a cooperative breeding social context by both parents, kin and unrelated helpers at nests^{49,50}. Vocal learning (if present) is most likely to have evolved in such a social context, in contrast to non-interactive solitary contexts. Furthermore, rifleman feeding calls are social contact calls, which are often learned in avian vocal learners, such as in parrot and zebra finches' contact calls and the flight calls of the *Carduelinae* subfamily^{42,51–54}.

Here, we search for predispositions for vocal production learning in the rifleman in two ways: (1) by investigating the presence of individual and group vocal signatures and determining how genetic relatedness and social proximity influence the acoustic similarity and feeding call features of distantly related individuals living in close proximity; (2) by disentangling the genetically driven phenotypic variances of rifleman call features from their socially driven vocal counterparts using quantitative genetics, and by comparing those ratios to a vocal learner, the zebra finch (*Taeniopygia guttata*)⁵⁴.

Results

Individual and group vocal signatures in riflemen

Vocal production learning can be revealed when animals copy the unique and distinctive vocalizations of another individual, known as individual

vocal signatures^{10,37}. Most animals that communicate with sounds, including vocal non-learners, are distinguishable thanks to their unique vocal signatures^{55–60}. But, vocal learners can go one step further and learn to copy other's unique vocal signature^{37,40}. In riflemen, only weak signs of vocal individuality have been found to date in their feeding calls⁶¹, and it remains unknown whether they learn to copy conspecifics' individual vocal signatures.

Thanks to novel high-quality recording techniques, we detected the presence of strong individual vocal signatures in rifleman feeding calls and found that adults provisioning at the same nest were more similar to one another than random individuals. Riflemen produced visually distinctive individual vocal signatures with a strong stereotypy (i.e. structural consistency between vocalizations within an individual, Fig. 2A; $n = 6839$ calls; $n = 13$ individuals; *isoMDS Kruskal stress* = 0.27; *iterations* = 200; *k-dimensional* = 2), which were more similar within individuals than between individuals (PERMANOVA of cross-correlations: $F = 289.9$; $P < 0.01$ and Mantel $q = 0.24$; $P = 0.001$; number of permutations: 10,000; Fig. 2A, B). In addition, rifleman social partners (i.e. parents and helpers provisioning the same nest) were more similar to one another than to individuals from other nests (*isoMDS Kruskal stress* = 0.27; *iterations* = 200; *k-dimensional* = 2; PERMANOVA of cross-correlations: $F = 541.7$; $P = 0.03$ and Mantel $q = 0.23$; $P = 0.001$; number of permutations=10,000; Fig. 2C), revealing the presence of group vocal signatures in riflemen.

Machine learning algorithms are capable of recognizing and distinguishing individual and group vocal signatures^{10,62,63}. Thus, to further confirm the presence of individual and group vocal signatures in riflemen, we trained a Random Forest machine learning algorithm⁶⁴ with the above datasets to examine classification accuracies at the individual and nest (group) level (Fig. 2D). The Random Forest Classifier accurately classified calls to the right individual with 82.95% accuracy (95% *CI* = 0.80, 0.85; $\kappa = 0.80$; $P < 2.2e-16$; number of times cases are 'out-of-bag'; computing OOB estimate of error rate: 20%, *number of trees* = 500; Fig. 2D), and to the correct nest with 85.96% accuracy (95% *CI* = 0.84, 0.88; $P < 2.2e-16$; $\kappa = 0.84$; computing OOB estimate of error rate: 17.5%, *number of trees* = 500; Fig. 2D). These results outperform previous vocal identification classification results for riflemen that used discriminant analysis (82.95% in this study vs. 26% Khwaja et al.⁶¹). Our high-quality audio recordings (with little signal to noise ratio) of rifleman feeding calls likely facilitated the distinction between individuals and groups. This result confirmed that both individual and nest vocal signatures are present, distinctive, and identifiable in riflemen.

High acoustic similarity is not explained by genetic similarity

We hypothesized that if distantly related riflemen sound more similar than their close kin, this could suggest that a form of vocal imitation occurs. By examining the relationship between pairwise acoustic similarity (using mean spectrographic cross-correlations⁶⁵) and genetic relatedness in a wild population of riflemen (using 32,948 Single Nucleotide Polymorphisms or SNPs generated with Genotyping-By-Sequencing^{66–68}; Fig. 3A), we found that the correlation between genetic similarity (i.e., relatedness) and acoustic similarity of feeding calls was low (Spearman's correlation $\rho = 0.0028$,

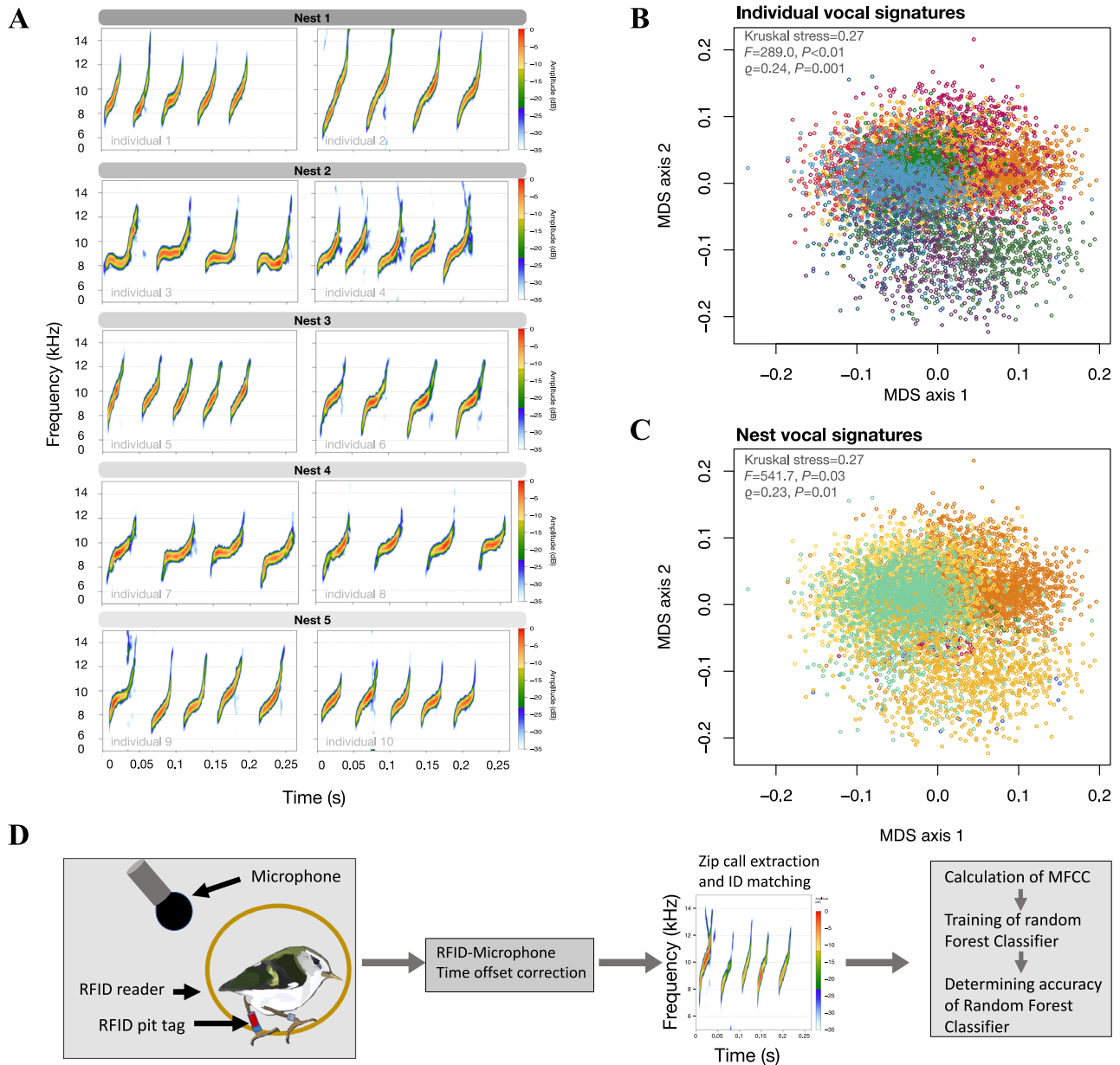


Fig. 2 | Individual and group vocal signatures in the feeding calls of rifleman. **A** Examples of rifleman feeding call spectrograms concatenated together from different time events showing unique individual vocal signatures. Concatenated calls in each frame come from a single individual. The gray colored bar above each frame groups individuals provisioning the same nest. **B**, **C** The multidimensional scaling is based on call dissimilarities of 6839 rifleman feeding calls and shows clusters of (**B**)

distinct individual vocal signatures (each color represents calls from a single individual) and (**C**) group vocal signatures at the nest level (each color represents calls from individuals feeding chicks at the same nest; $n = 6$ nests, $n = 13$ individuals). **D** Training of a Random Forest Classifier resulted in the accurate classification of rifleman feeding calls to the correct individual with 82.95% accuracy and the correct nest with 85.96% accuracy.

$P = 0.94$, $n = 49$ individuals with genetic data and with a maximum of 50 randomly selected feeding calls; $n_{pairs} = 1176$ Fig. 3B.a). This indicates that genetic similarity is a poor predictor of acoustic similarity in rifleman. Furthermore, we found no correlation between genetic relatedness and mean difference in a comprehensive range of specific acoustic parameters, including call frequency and call duration (Tables S1-S2), consistent with the idea that factors other than genetic relatedness influence call similarity and vocal feature differences in rifleman.

Socially-close, but distantly related rifleman sound similar

Vocal learners often imitate individuals that are in close proximity, regardless of their relatedness^{8,69,70}. Following a similar approach to the above section, and by examining the correlations between pairwise acoustic similarity and social proximity (i.e., low mean geographic

distance between individuals based on nest attendance indicates high social proximity), we found that acoustic similarity was positively correlated with social proximity in all birds (Spearman’s correlation, $\rho = 0.20$, $P = 0.0011$, $N_{perm} = 10,000$, $N_{individuals} = 70$, $n_{pairs} = 2415$ – note that the number of individuals $n = 70$ differs from $n = 49$ from above – see method). In other words, social proximity appears to play a crucial role in shaping rifleman call similarity.

Rifleman are facultative cooperative breeders that live within close proximity with relatives and helpers in kin-based neighborhoods^{50,71}. Our results confirmed the presence of kin-based neighborhoods in rifleman and showed that relatives were socially closer than distantly related pairs of individuals (Mantel statistics Spearman’s correlation: $\rho = 0.15$, $P = 0.0005$, $N_{perm} = 10,000$, $N_{individuals} = 49$ birds, $n_{pairs} = 1176$; Fig. S1.a). However, this meant we could not exclude the possibility that genetic relatedness

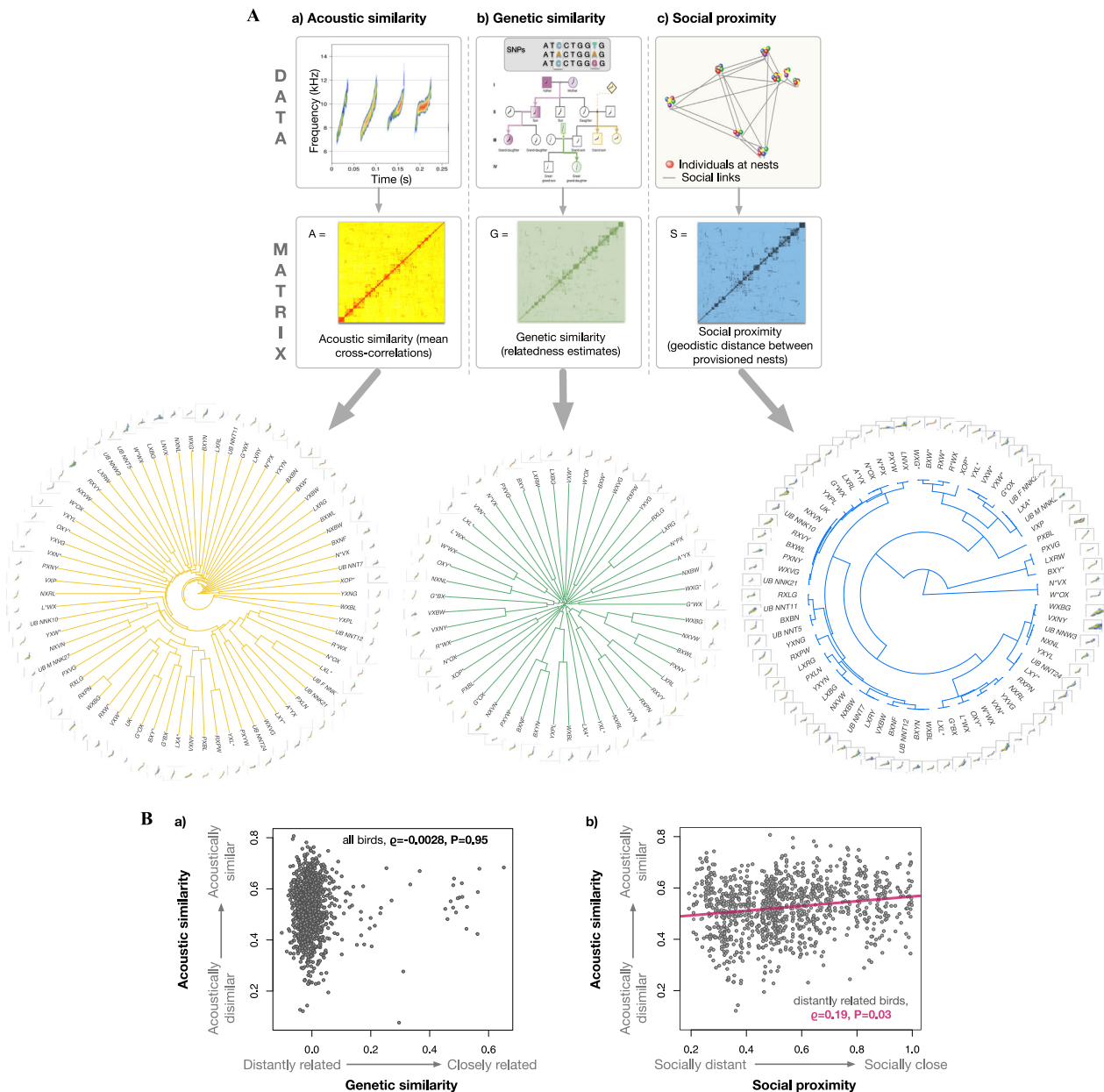


Fig. 3 | Methodology and relationship between acoustic, genetic and social similarity in rifleman feeding calls. **A** Methodology used to obtain acoustic, genetic and social similarity multiple-matrices and hierarchical-clustering phylogenetic trees of rifleman calls. **a** The acoustic similarity matrix and the hierarchical-clustering phylogenetic tree were derived from spectrographic cross-correlations of rifleman feeding calls (i.e. zip calls; 1110 sound clips from 70 individuals across 29 nests); **b** The genetic relatedness matrix was derived from SNPs data from 186 rifleman and the hierarchical-clustering phylogenetic tree was based on genetic relatedness and acoustic data from 49 individuals; **c** The social proximity distance matrix and the hierarchical-clustering phylogenetic tree were derived from geodesic distance proximity (i.e., based on nest locations) of social partners provisioning the

same nests ($n = 70$ individuals across 29 nests). The hierarchical-clustering phylogenetic trees have at their tips, bird identity with one representative feeding call spectrogram. The diagrams for “DATA” and “MATRIX” are for methodology illustration purposes only. **B (a)** Relationship between acoustic similarity (based on mean spectrographic cross-correlation of rifleman feeding calls) and genetic similarity (based on SNP relatedness estimates) (1176 bird pairs across 49 individuals). Each circle represents one bird pair. **b** Relationship between pairwise acoustic similarity (mean spectrographic cross-correlation of rifleman feeding calls) and social proximity (based on geodesic distances between nests visited by individuals) among distantly related pairs of rifleman (1149 bird pairs across 49 individuals). Each circle represents one bird pair.

was driving the high acoustic similarity among socially close individuals. To control for this, we repeated the above analysis (i.e. acoustic similarity vs social proximity), but excluded pairs of close genetic relatives from our pairwise comparisons: $G \geq 0.2$. The positive correlation between pairwise social proximity and acoustic similarity persisted in distantly related individuals (Spearman’s correlation: $\rho = 0.19$, $P = 0.032$, $N_{perm} = 10,000$, $N_{individuals} = 49$, $n_{pairs} = 1149$; Fig. 3B.b). A similar pattern was also detected among closely related individuals (excluding distantly related individuals), but although the relationship was stronger than in distantly

related individuals, it was not statistically significant ($\rho = 0.41$, $P = 0.98$, $N_{perm} = 1262$, $N_{individuals} = 29$, $n_{pairs} = 27$; Fig. S1.b). Overall, this provides additional evidence that a form of vocal convergence or possibly imitation is present in the rifleman.

To understand which aspects of rifleman vocalizations were adjusted in response to social proximity, we further examined the relationship between acoustic parameters of feeding calls and social proximity among distantly related individuals (Fig. 4A, Table S3). Acoustic parameter-specific Mantel tests revealed statistically significant correlations for 7 out

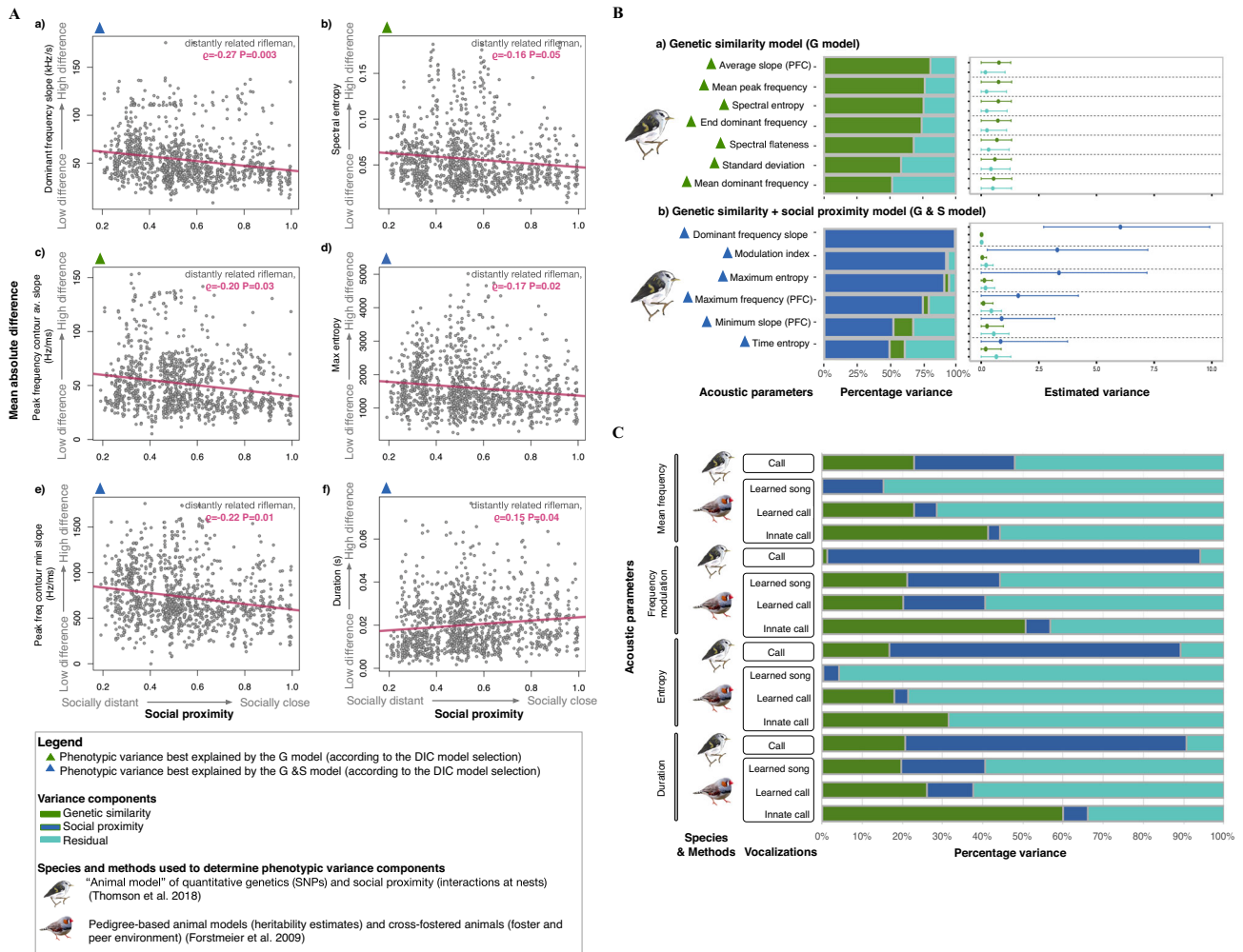


Fig. 4 | Influence of social proximity on rifleman feeding calls and comparison of phenotypic vocal variances with zebra finches. A Significant correlations between the mean absolute difference of acoustic parameters and social proximity among distantly related pairs of riflemen (1149 bird pairs across 49 birds). **B** Examples of phenotypic variances of selected acoustic parameters that were best explained by: a) the genetic similarity model (G model; $n = 39$ individuals), and b) the genetic similarity and social proximity model (G & S model; $n = 39$ individuals). The complete set of comparisons for each model can be found in Figs. S4–S6. The social

proximity animal model (S model) is not represented because none of the acoustic parameter variances could be explained by this model (Fig. S5). **C** Comparison between the phenotypic vocal variances of rifleman feeding calls, zebra finch females' innate calls, and zebra finch males' learned calls and songs. Two different methods and models were used to determine phenotypic variances in riflemen and zebra finches^{46,54}. For best comparison, the social components (i.e., peer and foster effects) in zebra finches were combined and the maternal effects were not represented.

of 37 acoustic parameters (at a two-sided significance threshold of 0.05) (Table S3). These parameters were related to frequency, such as frequency slope of feeding calls (i.e., the change in frequency through time), duration, entropy and inflections (Table S3). These correlations ranged from $\rho = -0.27$ ($P = 0.003$) to $\rho = 0.20$ ($P = 0.01$), with dominant frequency slope (dfslope) and minimum frequency contour slope (PFC) having the strongest correlations, and frequency median having the weakest correlations (Table S3). After accounting for multiple comparisons and correlations between acoustic parameters, the chance of obtaining a significant correlation between at least 7 acoustic parameters (out of 37 parameters) and social proximity was 0.03. In other words, it is very unlikely that one would see this many significant correlations just by chance. This suggests that some parameters are biologically relevant to rifleman social interactions. It is also worth noting that, despite some correlations being significant, the values of ρ were generally low, suggesting that the correlations observed may result from subtle call feature modulations. Overall, these results show that social proximity plays a role in shaping some aspects of rifleman calls, and that individuals provisioning for the same nests share higher acoustic similarity, regardless of their genetic relatedness.

The phenotypic variance of the rifleman calls resembles those of a vocal learner

Vocal learners are expected to have higher vocal plasticity and phenotypic vocal variances than vocal non-learners^{45,54}. If riflemen vocally learn, their vocal plasticity and phenotypic vocal variance should be similar to those of vocal learners. To test this hypothesis, we built three multiple-matrix “animal models” (i.e., Genetic similarity model; Social proximity model; and Genetic similarity & Social proximity model)⁴⁶ and used model selection with Deviance Information Criterion (DIC; Table S5) to determine which of these three models best predict the proportion of phenotypic variance components for each acoustic parameter⁴⁶ (Fig. 4A, B).

Based on the DIC model selection, the genetic similarity model (G model) best explained the phenotypic variance of 16 out of 37 feeding calls' acoustic parameters (Fig. 4B, Fig. S4 and Table S5). For example, the phenotypic variance of the average slope of the peak frequency contour was best explained by this model and had the strongest genetic influence (i.e., largest genetic proportion of variance and smallest residual variance). In addition, the credible interval ranged from < 0.0001 to 1.3 (Fig. 4B.a; Fig. S4). These results indicate that some acoustic parameters have a stronger genetic basis with a minimal influence by social environment.

The social proximity model (S model), on its own, did not explain the phenotypic variances of any acoustic parameters (Fig. S5; DIC Table S5). In this model, the credible interval ranged from 0 to 11.3 (Fig. S5). This result aligns with our expectations that some acoustic parameters have a strong genetic component (i.e., morphology and auditory innate vocal templates⁷²), and that the social environment alone cannot contribute to the entirety of the phenotypic variance of calls.

However, the combined genetic and social model (G&S model) best explained the phenotypic variance of 21 out of 37 acoustic parameters (Fig. 4B.b; Fig. S6), according to the DIC model selection (Table S5). In this model, social variance had the largest credible intervals compared to genetic variance (range < 0.0001 to 9.9; Fig. S6). The social variance of the dominant frequency slope and the peak frequency contours of the average slope of calls had non-overlapping credible intervals which diverged away from zero, supportive of the effect of the social proximity on these parameters. This result indicates that the phenotypic variance of rifleman call features is influenced by both genetic similarity and social proximity, and that some acoustic parameters are more influenced by social proximity than genetic similarity, as would be expected in vocal learners. However, as will be discussed below, non-genetic variation may also be caused by other factors, such as shared environmental or social conditions.

Discussion

By combining diverse quantitative approaches, we demonstrate that New Zealand wrens have unique individual vocal signatures that converge toward common vocal features among distantly related individuals that share high social proximity. This result reveals the presence of vocal convergence in rifleman. Further, we show that the phenotypic variance ratios of rifleman calls are most similar to vocal learners. These results align with the vocal learning continuum/module hypotheses, and suggest that rifleman may possess rudimentary vocal learning abilities.

Detecting whether and how individuals match each other's unique vocal features (i.e. vocal signatures^{73–75}) is one of the first stages in the investigation of vocal production learning^{56,76–78}. In our study, each individual had unique and distinctive vocal signatures with high stereotypy (i.e., structural consistency between vocalizations within an individual), possibly due to individual differences in the morphology of their vocal tract (e.g. bill, syrinx size). Individual vocal signatures in social species, such as the rifleman, are beneficial because they may help rifleman parents distinguish among the multiple helpers that visit nests simultaneously⁴⁹. This is true especially in the context of “pay-to-stay” situations commonly found in cooperative breeding systems, in which helpers support parents in order to be accepted and tolerated within the group territory⁷⁹. These unique individual vocal signatures may also help rifleman distinguish their breeding partner from others and may benefit parent-offspring interactions by helping nestlings and fledglings recognize and locate their parents. Although the goal of our study was not to investigate call differences within individual call repertoire, it is interesting to note the small variations within an individual's call repertoire. This is relevant to our search as it may be indicative of syringeal or bill control.

In the context of vocal learning, animals that copy each other's unique individual vocal signatures provide a strong indication for vocal production learning. Animals that imitate or converge toward others' individual vocal signatures result in group vocal signatures or “vocal passwords” or accents^{80,81}. This may help animals gain considerable social advantages, such as stronger bonding and group membership^{35,56,70,82,83}. In species that learn their vocalizations, such as in many birds^{44,82,84}, pinnipeds³, cetaceans^{37,85,86}, bats⁸⁰ and humans, group vocal signatures are obtained either through a process called vocal matching – an exact copy of individual vocal signatures from the same group⁸⁷, or vocal convergence – a form of vocal imitation in which individuals slowly modify their call features toward common group vocal features (e.g., when pairing with a partner or assimilating to a new group)^{70,82,83}. Vocal matching was not demonstrated in this study, but we detected vocal convergence in socially and geographically close rifleman that were provisioning for the same nest, but were not closely related.

This suggests that individuals have some degree of vocal control with their vocalizations. This result contrasts with the traditional assumptions that rifleman calls have relatively little vocal plasticity and genetically encoded vocalizations.

According to Janik et al.⁸⁸, such vocal convergence toward a common group vocal signature is a strong indicator for vocal learning ability. But recent work has also shown that vocal convergence is present in animals traditionally labeled as vocal non-learners, making it ambiguous whether vocal convergence is the result of vocal production learning^{36,76,89–92}. For example, goitred gazelles (*Gazella subgutturosa*)⁹³, domestic goats (*Capra hircus*)⁹⁴, pygmy marmosets (*Cebuella pygmaea*)^{95,96} and orangutans (*Pongo sp.*)^{97,98} show high levels of vocal plasticity and modify their vocalizations toward the individual signatures of their partners or group members. Similarly, young agile gibbons (*Hylobates agilis agilis*) produce strong innate vocal templates from birth, but modulate and refine their vocalizations during ontogeny to match their mothers' calls and songs⁹⁹. Whether vocal convergence is learned in these animals, traditionally known as vocal non-learners, is not well understood or thoroughly tested. One possible hypothesis is that vocal convergence is expressed along a learning continuum with non-learned, rudimentary and learned forms. For example, in some species, vocal convergence may be attributable to reward-based operant conditioning¹⁷, which may not require cortical involvement of vocal control and may only need the lower brain region (for example, midbrain thalamic pathway) for neuromuscular modulation of simple call matching. Thus, while convergence may be the result of simple learning in some species, it may have an entirely different mechanism in other species. Further research is needed to test this hypothesis and to investigate the mechanisms underlying vocal convergence in multiple and diverse species.

Importantly, unlike vocal non-learners that vocally converge as an immediate response to a conspecific call (e.g. frequency matching in frogs¹⁰⁰), rifleman vocal convergence was maintained even in the absence of other conspecifics at the nests. This indicates that the observed vocal convergence in rifleman is not a simple one-time matching or a short-term frequency shift response. This type of sustained vocal convergence has previously been observed in vocal learners such as parrots^{34,44}, oscines⁴², and hummingbirds⁸², as well as in species traditionally termed vocal non-learners^{76,92}, including suboscines¹⁰¹ (Fig. 1), but never in New Zealand wrens. This may suggest that the vocal convergence detected in rifleman is more similar to that of vocal learners than vocal non-learners.

Our quantitative genetic analyses further helped clarify whether learning could cause the observed vocal convergence in the rifleman. Similar to vocal learners – which generally have higher phenotypic plasticity than vocal non-learners due to their strong social influence (see phenotypic plasticity continuum by Mesoudi et al.⁴⁵) – the rifleman had call phenotypic variances that were best explained by the combined genetic and social model, and had high proportions of social variance. Vocal production learning is one potential mechanism that could explain these results; however, some interpretative caution is needed because factors such as shared physiological and motivational parameters¹⁰² may have also affected call phenotypic variance. However, in our study, environmental factors, such as shared food resources or habitat differences, can be excluded because the birds shared the same general habitat. Furthermore, it is important to highlight that innate vocal behaviors produced in the total absence of vocal learning can also show strong levels of vocal plasticity under the influence of social environment (although considerably less than in vocal learners)¹⁰³. For example, white-lipped frogs (*Leptosactylus albilabris*) are vocal non-learners that perform complex, plastic, yet instinctive vocal behaviors that match the frequency of conspecific calls¹⁰⁰. But, unlike white-lipped frogs' phenotypic vocal variance, which is predicted to be best explained by genetics (as would be expected in vocal non-learners)¹⁰⁴, rifleman call phenotypic ratios were more similar to that of vocal learners (in addition to being variable and sustained over a long time period). This may indicate that their vocal plasticity is closer to that of vocal learners (with their ability to both instantaneously match a sound as well as maintain a learned sound overtime).

Comparing the phenotypic vocal variances between rifleman and zebra finches^{46,54} (Fig. 4C) helped further illustrate how rifleman's vocal plasticity fitted in the phenotypic plasticity continuum⁴⁵. Although differing comparative methods and models were used for this comparison, the relative genetic versus social ratios of our models were similar to that of zebra finches' learned calls and songs (i.e., produced by males; Fig. 4C). Rifleman had an even larger social versus genetic variance ratio than zebra finches in some parameters. This further indicates that the mechanism underlying the observed variance in the rifleman calls may be closer to that of vocal learners. Conducting an extensive cross-species comparative analysis of phenotypic vocal variance components would be a valuable future comparison and contribution. Ideally, such future studies should use a single and standardized comparative method, such as using MCMCgmm of phenotypic variances between species⁴⁶. This would further situate rifleman phenotypic variances relative to other vocal non-learners and vocal learners along a phenotypic plasticity continuum (Mesoudi et al.⁴⁵).

In conclusion, this study offers new research avenues and methodologies to investigate potential predispositions for vocal learning abilities in presumed vocal non-learners. It reveals the presence of vocal convergence and possibly a rudimentary form of vocal learning in the rifleman. This raises important questions about the evolution of vocal convergence, which might have been present in the shared ancestor of Psittacopasserans. Vocal convergence and any predispositions for vocal production learning are behaviors that are easily overlooked without extensive analyses, so sophisticated and in-depth future studies are needed to explore their possible existence and their origins in other animal groups. In particular, exploring the vocal phenotypic variances of traditional vocal non-learners would provide an important contrast to our findings. It would also provide a framework for others trying to interpret phenotypic variances in the context of the vocal learning continuum hypothesis. Further, our study highlights the importance of revisiting the definition of vocal convergence in the context of vocal learning. Overall, this work offers new insights into the vocal behavior and learning predispositions of a presumed vocal non-learner, the rifleman – a species which is key to understanding the evolutionary origin of vocal production learning in Passeriformes.

Methods

Study field site and population monitoring

We monitored a wild population of rifleman at the Boundary Stream Mainland Island reserve, New Zealand (39°06'15.8"S, 176°48'06.1"E) during two austral breeding seasons, from September to February 2018–2020. The timing of rifleman breeding is asynchronous²⁵, which enabled us to monitor nests simultaneously and continuously throughout the breeding seasons. For monitoring and identification purposes, we caught individuals using mist-nets and speakers with conspecific lure calls, we then sexed and aged adults and fledglings based on their sexual dimorphism in size and coloration¹⁰⁵, and we banded individuals with unique color band combinations. One of the leg bands was equipped with a Passive Integrated Transponder (2.3 mm EM4102 PIT tag; Eccel Technology Embedded RFID) that could be read by Radio-frequency identification receivers (RFID; engineered by the University of Auckland) placed at nests (Fig. 2D). RFID receivers logged the identity of individuals entering or leaving nests during the 2019–2020 breeding season. After locating nests, we monitored nest activity and recorded vocal behavior of individuals provisioning nests (see recording methods below). Due to the inaccessibility of most natural nests (i.e., tight spherical nests in tree cavities), we could not band nestlings which limited information about the relatedness among individuals. To mitigate this, we banded fledgling groups when siblings still clustered together outside their nests and were fed by banded adults.

We have complied with all relevant ethical regulations for animal use. This research was approved and facilitated by mana whenua from the Maungaharuru region and the Department of Conservation, New Zealand (Department of Conservation permit obtained in 2018, act No FAU 55391) and approved by the University of Auckland Animal Ethics Committee

(Approval no. 001866). Bird capture and banding activities were conducted under the New Zealand National Bird Banding Scheme.

Study species

Rifleman are socially monogamous cavity-nesters that build their nests at various heights (i.e., on the ground under leaf litter or in tree cavities) in the native forests of New Zealand²⁵. Nest building can last a few days (i.e., 3–6 days, personal observation) and is followed by an incubation period which can last 20 days²⁵. Once hatched, nestlings remain in the nest for around 24 days²⁵. Rifleman are facultative cooperative breeders (i.e., helpers in addition to parents feed nestlings at some nests), so nestlings receive frequent visits from both parents and helpers which are often genetically related to the parents^{49,50,71,106}. The frequency of visits at the nests increases as nestlings grow, with a feeding rate ranging from 4 to 20 times an hour (i.e., every 3 to 12 min) in the later nestling stages²⁵. Each time adults visit nestlings, they produce feeding calls (also known as zip calls; e.g. Fig. 2A) prior to feeding them which seems to function as contact calls between parents and offspring²⁵. These feeding calls are distinguishable from other call types due to their "S" shape and higher frequency (~6–14 kHz; Fig. 2A)⁴⁸. Rifleman feeding calls are a good candidate to test the presence of vocal learning abilities because they are used in social contexts (between parents, helpers and nestlings, and between partners) in which vocal learning is most likely to have evolved and be detected if present. The rifleman does not produce vocalizations that classify as songs in the traditional sense; indeed, it is a non-territorial species and rifleman pairs do not seem to produce courtship songs to attract mates. Instead, the rifleman has a large repertoire of calls⁴⁸, which has been assumed to be innate based on its simple vocal features, nonterritorial or non-courtship vocalizations, and a syrinx that lacks complex intrinsic muscles found in bird vocal learners^{21,22,24,25}.

Sound recording

We collected rifleman feeding calls (zip calls; e.g. Fig. 2A) at nests using a combination of methods which we implemented to accommodate the diversity of nest heights found in our study population. First, we collected focal recordings of the feeding calls at nest (i.e. one observer with binoculars, a digital Zoom H6 recorder and a shotgun Sennheiser ME62 K6 microphone -20,000 Hz frequency) of banded individuals (i.e. identified by their unique color band combo). During focal monitoring sessions, we combined focal recordings with visual monitoring and video recordings to match the identity of individuals visiting nests to their corresponding feeding calls. We also recorded the vocal behavior of individuals at nests using passive Bioacoustic Automated Recorders (BAR; Frontier Labs version 1.4; WAV format with a sampling frequency of 44,100 Hz and 32 bits sampling depth; breeding season 2018/2020; Fig. 2D). Each BAR recorder was connected via a long cable to a small omnidirectional microphone placed close to nest entrances (range: 0.1 m–1 m). Each BAR recorder was placed further away from the focal nest (10–15 meters) to minimize nest disturbance when changing batteries and SD cards. BARs were programmed to record daily from 1 hour before sunrise to 2 h after sunset. In addition to the BAR recorders, we also deployed trail cameras (Bushnell Trophy Cameras 24MP and E2 12MP) and PIT tag readers (2.3 mm EM4102 PIT tag; Eccel Technology Embedded RFID; breeding season 2019–2020) to facilitate the synchronization of individual identity and feeding calls. The trail cameras pointed toward nest entrances (i.e., placed between ~30 cm and ~1 m) to capture photos and videos of the leg color bands combinations, and PIT tag readers were placed around nest entrances to log PIT tag numbers.

Sound processing and annotations

Time offset correction and synchronization. Time drift occurred in some of the BAR recordings due to SD card write-speed differences, which resulted in a mismatch between the timestamps of the BAR recorder and RFID (Fig. 2D). We corrected for time offset using known sound timestamps produced in the BAR recordings for which we had an exact RFID timestamp. For both the time offset correction and the

following synchronization step, we used the following custom pipeline (available at https://uoa-eresearch.github.io/bird_recognition/time_sync_check.nb.html). The synchronization between the timestamps of PIT tags and nest recordings enabled us to associate calls to the correct individuals. We extracted feeding calls (zip calls; e.g.: Fig. 2A) based on PIT tag readings of individuals entering nests within 5 s of a PIT tag time-stamp reading. Samples with two or more individuals detected within these 5 seconds were removed to avoid errors in identity attribution. We then manually annotated each feeding call within these 5-second filtered time windows using Raven Pro v1.6.1¹⁰⁷.

Sound libraries. The above sound collection method yielded a large and high-quality RFID/BAR call library with 6839 calls from 13 individuals (12 parents and one nest with one helper; $mean = 526.1 \pm sd = 456.4$ calls per individual, $min = 12$ calls, $max = 1378$ calls) across 6 nests ($mean = 977.0 \pm sd = 1,012.8$ calls, $min = 12$ calls, $max = 2457$ calls). This library offered a large number of calls per individual and high-quality sound clips with no background or overlapping sounds which is ideal to detect vocal signatures and train our machine learning algorithm (Fig. 2).

We also manually synchronized the BAR or focal recordings to the timestamps of our visual observations, trail cameras, and video recording timestamps using Raven Pro v1.6.1¹⁰⁷. This combination of recording techniques (BAR/focal/trail camera/video) enabled us to build a second extensive feeding call library with 4207 calls from 70 birds across 29 nests (see details below under “Creation of acoustic, genetic relatedness, and social matrices”). We used this call library and the above RFID/BAR call library for the remaining analysis of our study (Figs. 3–4).

Spectrograms and call feature measurements. All spectrograms of feeding calls were plotted with *Seewave* v.2.2.0¹⁰⁸ (settings set to $wl=250$, $ovlp=50$) and *warbleR* v.1.1.27⁶⁵ (Figs. 2A, 3A respectively). We measured feeding call features using three vocal measurement tools (*Seewave* v.2.2.0¹⁰⁸, *warbleR* v.1.1.27⁶⁵, and *RavenPro Raven Pro* v.1.6.1¹⁰⁷), and extracted 37 acoustic parameters (Table S1).

Detection of vocal signatures

To assess the presence of vocal signatures at the individual and nest level in rifleman, we compared feeding calls within and between individuals attending the same nest. We performed spectrographic cross-correlations between the feeding calls of each individual, which is a method that slides spectrograms over one another to obtain a similarity score between calls (i.e., similarity matrix, based on sound dissimilarities – one minus cross-correlations rescaled to be between 0 and 1). We used the R function *xcorr* from *warbleR* v.1.1.26⁶⁵ to perform the pairwise cross-correlations (Fig. 2B, C). All pairwise combinations of calls resulted in $6,839 \times (6,839 - 1)/2$ distinct cross-correlations of calls for 13 individuals ($mean = 526.1 \pm sd = 456.4$ calls per individual; $min = 12$ calls, $max = 1378$ calls) across 6 nests ($mean = 977.0 \pm sd = 1012.8$ calls; $min = 12$ calls, $max = 2457$ calls).

We then compared the pairwise cross-correlations using Kruskal’s non-metric multidimensional scaling with *MASS* (isoMDS) v.7.3.51.6¹⁰⁹ (Fig. 2B, C), and analyzed the goodness of fit with isoMDS Kruskal’s stress¹⁰⁹. We then used a Mantel test based on Pearson’s correlation from *vegan* v.2.5.6¹¹⁰ to determine whether vocal signature was significant at the individual or nest level. To do this, we created a binary matrix representing feeding call membership in which 0 was assigned to feeding calls belonging to the same individual and 1 was assigned to feeding calls belonging to different individuals⁸². The cross-correlation dissimilarity matrix (1-correlation similarity matrix) was then compared with the membership matrix using a Mantel test, and it used group membership matrices as predictors in Mantel correlations against acoustic dissimilarity matrices. We also used PERMANOVA¹¹¹ (adonis from *vegan* v.2.5.6¹¹⁰) as an alternative method to detect individuality in the feeding calls of riflemen. For nest signatures, we controlled for individual vocal signatures (i.e., by only permuting individuals, rather than clips between individuals) because individual vocal

signatures could otherwise cause an apparent nest vocal signature due to the small number of individuals per nest.

Random forest classifier

We created and trained a Random Forest Classifier using calls collected from our BAR/RFID library (Fig. 2A) to test if the classification of feeding calls could be accurately achieved at the individual and group level (Fig. 2D). We normalized feeding calls at -1 dB using *ffmpeg* v1.20.1¹¹² to ensure that amplitude differences between recording clips would not cause classification inaccuracies. We compiled the training library by loading .wav files of feeding calls and then calculated Mel-frequency Cepstral Coefficients (MFCC) of the feeding calls using *tuneR* (melfcc) v1.3.2¹¹³. MFCC have been extensively used for human voice recognition¹¹⁴ and focus on dynamic features of the vocal tract, and have been successfully applied to animal vocalization recognition¹¹⁵. We then set up the duration of each time frame (i.e., used as a “sample”) at 0.25 s. A longer duration meant that more call clips were included in one “sample”, thus making the classification easier; however, it also meant that from each particular .wav file, fewer “samples” could be extracted. Consequently, we set hoptime to be the same as wintime, so that successive time frames could not overlap, making the samples truly independent. Next, we randomized the order of time frames (e.g., in case bird clips had not been concatenated randomly). We then trained a Random Forest Classifier, which used multiple learning algorithms (ensemble learning) to obtain high predictive performance for classification purposes, using the function *randomForest* from *randomForest* v.4.6.14⁶⁴. We then estimated the accuracy of the classifier with the samples from each group (individual and nest).

Blood sample collection and DNA extraction

We collected blood samples (10–35 μ l) during banding sessions using brachial venipuncture with a BD PrecisionGlide sterile needle (26 G 1/2–0.45 mm x 13 mm) and a 70 μ l capillary tube (Microhematocrit tubes). Blood samples were preserved in 95% ethanol and temporarily stored at 4 °C in the field before being transferred for long-term storage at -20 °C. In addition to collecting blood samples from Boundary Stream Mainland Island (i.e., focal population), we also collected blood samples from Mohi Bush (39°51'25.34"S, 176°54'7.49"E), a geographically close but genetically distant population to increase our sample size to 186 individuals to provide robust genetic relatedness estimates¹⁶. We extracted DNA from individual blood samples using a Qiagen DNeasy Blood and Tissue kit and followed the protocol for total DNA purification from nucleated blood with elution in 80 μ l AE buffer and without RNase treatment (Spin-column Protocol; DNeasy Blood & Tissue Handbook version 07/2006). DNA was quantified using the Qubit Broad Range DNA assay (Qubit 2.0 Fluorometer), and further quality checks were performed by spectrophotometry (Implen NanoPhotometer N60) and by gel electrophoresis. DNA aliquots were dried (3 h. at 30 °C) and sealed for shipping to AgResearch, Mosgiel, Aotearoa, New Zealand, where a Genotyping-By-Sequencing (GBS) protocol was carried out.

GBS library construction and SNPs detection

High-throughput Genotyping-By-Sequencing (GBS) was used to generate Single Nucleotide Polymorphisms (SNPs). The GBS library was prepared following the method outlined in Elshire et al.⁶⁶ with modification as in Dodds et al.⁶⁷ (*KGD* v0.8.4) for 186 individuals (total of 192 wells with 6 negative control samples) using *PstI-MspI* double-digest restriction enzymes. The Library underwent size selection with Pippin Prep (SAGE Science, Beverly, Massachusetts, USA) to select fragments in the size range of 193–318 bp (genomic sequence plus 123 bp of adapters). Single-end sequencing (1x101bp) was performed on an Illumina HiSeq 2500 utilizing v4 chemistry. Raw fastq files were quality checked using a custom QC pipeline (available at <https://github.com/AgResearch/DECONVQC> and *FastQC* v.0.10.1, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, Andrew¹¹⁷).

Demultiplexing, SNP detection, and filtering were undertaken with the reference-free SNP detection pipeline called *UNEAK*¹¹⁸, implemented in *Tassel3* v.3.0.174¹¹⁹. We filtered SNPs to avoid the inclusion of erroneous SNPs in downstream analyses. The following parameters were used: *s* (maximum number of barcoded reads per lane) set to 400 M, *t* (merge taxa option) set to “no”, *m* (maximum tag number in the merged TagCount file) set to 600 M, *x* (Maximum tag number in TagCount file for each taxa) set to 100 M, *c* (minimum count of a tag that must be present to be output) set to 30, *mnMAF* (minimum minor allele frequency) set to 0.03, and *mnC* (minimum call rate) set to 0.1 (10%). All other parameters (e.g., Error tolerance rate – ETR; Maximum minor allele frequency – *mxMAF*; maximum call rate – *mxC*) were left at the default option (i.e., ETR = 0.03; *mxMAF* = 0.5; *mxC* = 1). We excluded SNPs with a significant ($p < 0.05$) deviation from Hardy-Weinberg equilibrium (Fig. S2). We identified 32,948 SNPs (*mean co-call rate for sample pairs* = 0.55; *min co-call rate for sample pairs* = 0.23; *proportion of missing genotypes* = 0.36; *call rate* = 0.64) and the average sequencing read depth for called SNPs was 7.6.

Creation of acoustic, genetic relatedness, and social matrices

To further investigate the mechanisms resulting in nest vocal signatures, we created three matrices (i.e., acoustic matrix, genetic relatedness, and social proximity matrices; Fig. 3A) for downstream analyses to disentangle the genetic and social factors influencing rifleman call structure (i.e., similarity). In vocal learners, social factors influence acoustic structure and similarity between individuals, especially in individuals that are not closely related, so these analyses can help detect predispositions for vocal imitation or learning abilities^{45,120}.

First, we determined acoustic similarities between individuals by creating an acoustic matrix for all individuals for which we had collected call clips (70 birds across 29 nests; Fig. 3A.a). Some individuals had considerably fewer feeding calls than others due to rare nest visitations (e.g., helpers) or limited available recordings, which resulted in an imbalanced dataset. Thus, to balance out our dataset of calls per individual, we randomly selected a maximum of 50 feeding calls per individual. This resulted in a total of 1,110 filtered sound clips (*mean* = 14.4 calls per individual, *sd* = 14.3 calls, *min* = 1, *max* = 50). We then calculated the mean of pairwise spectrographic cross-correlations between rifleman feeding calls using the function *xcorr* from *warblerR* v.1.1.27⁶⁵. The *warblerR* settings for *xcorr* were set as follows: window length (*wl*) = 300 and overlap (*ovlp*) = 90. All pairwise combinations of calls resulted in 1110 × 1110 distinct cross-correlations of calls and 70 × 70 distinct mean cross-correlations for individuals.

Next, following the KGD v0.8.4 pipeline (G5 method, as in Dodds et al.^{67,68}), we determined the genetic relatedness between individuals using the 32,948 filtered SNPs (obtained from the above method) to generate a genetic relatedness matrix (GRM) and to measure genetic similarity (or genetic distance) between individuals ($n = 186$ individuals, Fig. 3A.b). Genetic relatedness between individuals was visualized using a relatedness heatmap (Fig. S2). It is important to note that genetic relatedness was determined here, not the heritability of any vocal trait.

Finally, to create a social proximity matrix that reflected social closeness between individuals, each individual for which we had acoustic data ($n = 70$ birds across $n = 29$ nests) was given a location (GPS points) based on the location of the nest they provisioned (Fig. 3A.c). Because several adults attended the same nest, we “blurred/jittered” the GPS points to a few centimeters away from the commonly shared GPS nest point to avoid having all individuals provisioning for a same nest with a social distance of 0. This “blurring/jittering” was essential to retain a true social proximity matrix that could become invertible (i.e., positive definite), as recommended by Thomson et al.⁴⁶. We used the package *castor* v.1.7.2¹²¹ and the function *all_pairwise_geodesic_angles* to calculate the distance between two sets of individual locations/nests coordinates. This function returns a 2D matrix of size $N1 \times N2$ (in this case, 70×70). If one individual was found at several nests, we calculated the average distance for that individual between all provisioned nests. We used this distance as a proxy for a shared social environment (i.e., social interactions) to generate a social proximity matrix (Fig. 3A.c).

Hierarchical-clustering trees with spectrograms

We visualized call similarity between individuals by generating hierarchical-clustering trees with spectrograms and individual identity at their tips (Fig. 3A), based on acoustic, genetic and social proximity distances between individuals (based on GPS locations of individuals provisioning at nests) using the package *ape* v.5.6-2¹²² and *phylo_spectro* from *warblerR* v.1.1.27⁶⁵. One randomly selected feeding call from each individual call repertoire was represented as a spectrogram at the tips of the trees’ branches. The *warblerR* settings for the *phylo_spectro* function were set as follows: *wl* = 300, *ovlp* = 90, *wl.freq* = 512.

Correlations between acoustic, genetic, and social similarity

To determine whether and how the social environment and genetics contribute to shaping rifleman nest vocal signatures and whether this could reveal any predispositions for vocal learning (e.g. high acoustic similarity among distantly related pairs of individuals), we examined the relationship between the acoustic (i.e., mean spectrographic cross-correlation of calls per bird), genetic (i.e., SNP relatedness estimates) and social similarity (geodesic distances between nests provisioned by individuals) (Fig. 3; following similar reasoning as in Lemasson et al.¹²⁰).

To assess the correlations between these matrices, we used Mantel tests (Spearman’s correlation with permutation null model; two-sided significance) with the function *cor* from the R package *stats* v4.0.2¹²³ (Figs. 3B, 4A; Tables S2, S3). Note that when comparing pairwise distances, the individual pairs cannot be considered independent samples because different pairs may correspond to shared individuals, hence the statistical significance of any given correlation is generally estimated using permutation tests that account for this data structure^{124,125}.

First, we used a Mantel test (Spearman’s correlation with permutation null model; two-sided significance) to examine the relationship (i.e., correlation) between genetic similarity and acoustic similarity (1176 bird pairs across 49 birds; Fig. 3B.a). 49 individuals (out of 186 individuals for which we had genetic data) had both genetic relatedness and acoustic data. The mean spectrographic cross-correlation for these 49 individuals was based on 929 sound clips (*mean* = 19.0 calls per bird, *sd* = 15.4). We expected that if acoustic and genetic similarity were strongly correlated, this would indicate that genetics has likely a strong influence on acoustic similarity. In contrast a weak or negative correlation would indicate that other factors shape acoustic similarity. For this analysis, we excluded self-related acoustic similarity from our analysis because the “within-individual” vocal signatures would have otherwise biased the correlations between the two entries.

Next, we used a similar Mantel test as above to examine the relationship between acoustic similarity and social proximity and the relationship between social similarity and genetic similarity (Fig. S1a). This enabled us to determine whether individuals living in close proximity sounded similar and whether they were kin (i.e., kin-neighborhoods). After confirming the influence of social proximity on acoustic similarity in all birds (Spearman’s correlation, $q = 0.20$, $P = 0.0011$, $N_{perm} = 10,000$, $N_{individuals} = 70$, $n_{pairs} = 2415$) and detecting kin-neighborhood effect in rifleman (Spearman’s correlation: $q = 0.15$, $P = 0.0005$, $N_{perm} = 10,000$, $N_{individuals} = 49$ birds, $n_{pairs} = 1176$, Fig. S1a), we controlled for genetic relatedness, by re-examining the relationship between acoustic similarity and social proximity, but this time we restricted our analysis to distantly related pairs of individuals by excluding genetically close pairs of individuals (Fig. 3B.b). To exclude genetically close pairs of individuals we set a maximum relatedness threshold to 0.2 (i.e., excluding siblings, parents, uncles, and aunts from the genetic matrix and the correlation calculations). Among those 1176 bird pairs for which genetic relatedness was known, 1149 bird pairs had a genetic similarity below 0.2 (i.e., “distantly related” or “unrelated”), and the remaining 27 pairs had a genetic similarity above 0.2 (i.e., they are “closely related”). The 1149 distantly related pairs covered all 49 birds, in other words for every bird in our dataset there exists another bird that is unrelated to it. The 27 closely related pairs cover 29 individuals, in other words there are 29 birds in our dataset for which there exists another closely related bird.

Next, to determine which specific acoustic parameters were driving the significant positive correlation between acoustic similarity and social proximity among distantly related pairs of individuals (Fig. 3B.b), we investigated the relationship between the mean absolute difference in 37 acoustic parameters between any two sound clips of two individuals and social proximity in distantly related pairs of birds (1149 across 49 individuals, Fig. 4A and Table S3). The *warbleRv.1.1.27* function *specan*⁶⁵ and *Raven Pro v.1.6.1*¹⁰⁷ were used to measure the 37 acoustic parameters of rifleman feeding calls (see list of the acoustic parameters and their descriptions can be found in Table S1). The *warbleR* settings for the *specan* function were set as follows: *wl* = 300, *ovlp* = 90, *wl.freq* = 512. We then used acoustic-parameter-specific Mantel tests (two-sided significance threshold of 0.05) to examine correlations between mean absolute acoustic difference and genetic similarity (Table S2), and mean absolute acoustic difference and social proximity for each acoustic parameter (Fig. 4A; Table S3). Acoustic-parameter-specific Mantel tests revealed statistically significant correlations with social proximity for 7 out of 37 vocal parameters (at a two-sided significance threshold of 0.05; Table S3). Mutual pairwise spearman correlations showed these latter parameters were independent of each other (Table S4). To account for multiple comparisons and examine how probable it would be to erroneously obtain at least this many significant correlations by chance (i.e., under the null hypothesis of no correlations with social proximity), we used an adjusted permutation test. Our test fully accounted for correlations between acoustic parameters. Specifically, rows and columns of the social distance matrix were synchronously permuted (to break any association with acoustic parameters, as in a regular Mantel test), and subsequently all 37 acoustic-parameter-specific Mantel tests were repeated to re-compute the number of significant correlations. This permutation step was repeated 1000 times, to obtain the fraction of permutations that yielded at least 7 significant correlations. This fraction is an estimate for the probability of erroneously seeing a significant correlation between at least 7 acoustic parameters and social proximity, under the null hypothesis. We found that this probability was 0.03.

Multiple-matrix animal models: Estimation of the phenotypic vocal variance of rifleman feeding calls

Next, we built multiple-matrix animal models to partition the phenotypic variance of the rifleman feeding calls for each acoustic parameter (Fig. 4B, C). Models with strong social influence (i.e., acoustic parameters with high social variance components) are more likely to indicate the presence of social learning possibly via a vocal production learning mechanism⁴⁵. We followed methods from Thomson et al.⁴⁶, which uses Markov Chain Bayesian generalized linear mixed effect models (MCMCglmm) – an approach that fits an animal model into a Bayesian framework (Fig. 4B, C) to estimate acoustic traits' genetic, social, and residual variance components for each acoustic parameter. We used *MCMCglmm v.2.34*¹²⁶ and built generalized linear mixed effect models (GLMM) with the previously generated genetic, and social matrices (Fig. 3A). For all three models, acoustic traits were continuous traits added as fixed effects.

The MCMC ran for $n = 1,000,000$ iterations with thinning interval ($n = 100$), burn-in period ($n = 100,000$) for each acoustic parameter. We used $n = 1,000,000$ iterations to obtain an effective size between 1000 and 10,000 as recommended by Hadfield et al.^{126,127} and de Villemereuil¹²⁸, de Villemereuil et al.¹²⁹. The number of sound clips per individual was set to a minimum of 5 and maximum of 50 clips per individual. The first model (G model) determined genetic similarity and residual variances ($n = 39$ individuals; *mean* = 23.3 calls per bird; *sd* = 14.2; $n = 911$ sound clips), and genetic-relatedness was added as a random effect (Fig. 4B.a; Fig. S4). The second model (S model) determined the social and residual variance ($n = 49$ individuals, *mean* = 21.7 calls per bird, *sd* = 13.8, $n = 1066$ sound clips, *min* = 5 and *max* = 50), and social proximity was added as a random effect (Fig. S5). The final model combined genetic, social, and residual variances (G & S model; fixed = trait values ~ 1, random = ~ genetic relatedness + social proximity; $n = 39$ individuals; *mean* = 23.3 calls per bird; *sd* = 14.2; $n = 911$ sound clips, Fig. 4B.b; Fig. S6).

For each model, we plotted the trace of MCMC chains for each acoustic parameter and accessed the curves of traces and posterior density based on their relatively symmetry and unimodality (e.g., Fig. S3). We then reported the estimated percentage and the variances for each model by extracting the post-distribution means, credible intervals, and effective sample sizes for each acoustic parameter (Fig. 4B; Figs. S4–S6). The effective sample sizes ranged from 915.4 to 2883.6 for the G model, 3213.4 and 9,000.0 for S model, and 966.7 and 8,335.8 for the G & S model which satisfied the recommendations of $1000 < \text{effective sample size} < 10,000$ to run our models^{126–129}. Non-overlapping credible intervals indicated a strong separation between variance components (genetic, social, and residual), and a credible interval diverging away from zero best supported the effect of the variance components on call parameters.

Finally, we used DIC for model selection (i.e., based on the smallest DIC values) to determine which model best predicted the proportion of variance components for each acoustic parameter of rifleman feeding calls (Fig. 4A; Table S5). We conducted the above analyses with *R* (v4.2.0; 2022-04-22)¹²³.

Comparisons of rifleman phenotypic vocal variances with a known vocal learner

We compared the phenotypic vocal variance profiles of rifleman to a known vocal learner, the zebra finches to further situate rifleman phenotypic vocal variance along a phenotypic plasticity continuum⁴⁵ and assess whether vocal learning may be a potential mechanism underlying rifleman phenotypic vocal variance (Fig. 4C). Different methods were used to estimate social proximity and acoustic measurements in riflemen and zebra finches, thus we only conducted qualitative rather than quantitative comparisons to assess differences between the phenotypic vocal variance profiles of rifleman and zebra finches. We extracted the phenotypic vocal variance profiles of zebra finches from a reference study by Forstmeier et al.⁵⁴ (Tables A6–A8), which relied on pedigree-based animal models to determine the phenotypic variance and heritability of female zebra finch innate calls and male. The social variance in zebra finches was based on cross-fostering methods (“Foster parent” and “Peer” variances), while for riflemen it was based on geodesic distances between individuals provisioning at the same nest. For zebra finches, we combined the “Foster parent” and “Peer” variances⁵⁴, and we excluded the maternal effects on zebra finches' vocalizations because maternal effects were not accounted for in the rifleman models. We then selected a subset of variance components of rifleman feeding calls (i.e., four acoustic parameters: duration, entropy, frequency modulation, and mean frequency) and compared them against those of zebra finches (Fig. 4C). We then compared the relative social and genetic ratios between the phenotypic variances of rifleman feeding calls (best explained by our genetic and social proximity model) and non-learned zebra finch female calls, learned male calls, and learned male songs⁵⁴ (Fig. 4C).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study is available in Figshare under a license CC BY 4.0: Moran et al.¹³⁰. *Dataset. Vocal learning in a vocal non-learner? Social proximity and vocal convergence shape the calls of the most basal Passeriformes, New Zealand Wrens. 2024. figshare. Dataset. https://doi.org/10.17608/k6.auckland.25549466.*

Code availability

The R code which performed the study analyses is available in Figshare under a license CC BY 4.0: Moran et al.¹³⁰. *Dataset. Vocal learning in a vocal non-learner? Social proximity and vocal convergence shape the calls of the most basal Passeriformes, New Zealand Wrens. 2024. figshare. Dataset. https://doi.org/10.17608/k6.auckland.25549466.*

Received: 7 September 2022; Accepted: 26 April 2024;
Published online: 15 May 2024

References

- Jarvis, E. D. Evolution of vocal learning and spoken language. *Science* **366**, 50–54 (2019).
- Abramson, J. Z. et al. Imitation of novel conspecific and human speech sounds in the killer whale (*Orcinus orca*). *Proc. Biol. Sci.* **285**, 20172171 (2018).
- Stansbury, A. L. & Janik, V. M. Formant modification through vocal production learning in gray seals. *Curr. Biol.* **29**, 2244–2249.e4 (2019).
- Stoeger, A. S. et al. An Asian elephant imitates human speech. *Curr. Biol.* **22**, 2144–2148 (2012).
- Vernes, S. C. & Wilkinson, G. S. Behaviour, biology and evolution of vocal learning in bats. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190061 (2019).
- Jarvis, E. D. et al. Behaviourally driven gene expression reveals song nuclei in humming bird brain. *Nature* **406**, 628–632 (2000).
- Chakraborty, M. et al. Core and shell song systems unique to the parrot brain. *PLoS One* **10**, e0118496 (2015).
- Mennill, D. J. et al. Wild birds learn songs from experimental vocal tutors. *Curr. Biol.* **28**, (2018).
- Kroodsma, D. E. et al. Behavioral evidence for song learning in the Suboscine Bellbirds (*Procnias* spp.; Cotingidae). *Wilson J. Ornithol.* **125**, 1–14 (2013).
- Barker, A. J. et al. Cultural transmission of vocal dialect in the naked mole-rat. *Science* **371**, 503–507 (2021).
- Ten Cate, C. & Fullagar, P. J. Vocal imitations and production learning by Australian musk ducks (*Biziura lobata*). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200243 (2021).
- Ten Cate, C. Re-evaluating vocal production learning in non-oscine birds. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200249 (2021).
- Petkov, C. & Jarvis, E. Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* **4**, 1–24 (2012).
- Liu, W.-C., Wada, K., Jarvis, E. D. & Nottebohm, F. Rudimentary substrates for vocal learning in a suboscine. *Nat. Commun.* **4**, 2082 (2013).
- Wirthlin, M. et al. A Modular approach to vocal learning: Disentangling the diversity of a complex behavioral trait. *Neuron* **104**, 87–99 (2019).
- Wright, T. F. & Derryberry, E. P. Defining the multidimensional phenotype: New opportunities to integrate the behavioral ecology and behavioral neuroscience of vocal learning. *Neurosci. Biobehav. Rev.* **125**, 328–338 (2021).
- Vernes, S. C. et al. The multi-dimensional nature of vocal learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200236 (2021).
- Kroodsma, D. E. & Konishi, M. A suboscine bird (Eastern phoebe, *Sayornis phoebe*) develops normal song without auditory feedback. *Animal Behav.* **42**, 477–487 (1991).
- Touchton, J. M., Seddon, N. & Tobias, J. A. Captive rearing experiments confirm song development without learning in a Tracheophone suboscine bird. *PLoS One* **9**, 95746 (2014).
- de Lima, J. L. R. et al. A putative RA-like region in the brain of the scale-backed antbird, *Willisornis poecilinotus* (Furnariidae, Suboscines, Passeriformes, Thamnophilidae). *Genet. Mol. Biol.* **38**, 249–254 (2015).
- Ames, P. L. *The morphology of the syrinx in passerine birds*. vol. 37 (Peabody Museum of Natural History, Yale University New Haven, CT, 1971).
- Moran, I. G. The evolutionary roots of vocal learning: Exploring vocal learning abilities in vocal non-learners in birds. (University of Auckland, 2021).
- McLean, J. C. Bush birds of New Zealand. Part 1. *Emu* **11**, 1–17 (1911).
- Stidolph, R. H. D. Classified summarised notes. *Notornis* **3** (1950).
- Sherley, G. H. The breeding system of the South Island rifleman (*Acanthisitta chloris*) at Kowhai Bush, Kaikoura, New Zealand. (University of Canterbury. Zoology., 1985). <https://doi.org/10.26021/6852>.
- Jarvis, E. D. & Kaas, J. H. The evolution of vocal learning systems in birds and humans. in *Evolution of Nervous Systems* 213–227 (Academic Press, 2007).
- Hackett, S. J. et al. A phylogenomic study of birds reveals their evolutionary history. *Science* **320**, 1763–1768 (2008).
- Suh, A. et al. Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds. *Nat. Commun.* **2**, 443 (2011).
- Jarvis, E. D. et al. A phylogeny of modern birds. *Science* **346**, 1126–1138 (2014).
- Zhang, G. et al. Comparative genomic data of the Avian Phylogenomics Project. *Gigascience* **3**, 1–8 (2014).
- Heaton, J. T. & Brauth, S. E. Effects of deafening on the development of nestling and juvenile vocalizations in budgerigars (*Melopsittacus undulatus*). *J. Comp. Psychol.* **113**, 314 (1999).
- Chaiken, M. L. & Böhrner, J. Song learning after isolation in the open-ended learner the European starling: Dissociation of imitation and syntactic development. *Condor* **109**, 968–976 (2007).
- Favaro, L. et al. Evidence suggests vocal production learning in a cross-fostered Risso's dolphin (*Grampus griseus*). *Anim. Cogn.* **19**, 847–853 (2016).
- Hile, A. G., Plummer, T. K. & Striedter, G. F. Male vocal imitation produces call convergence during pair bonding in budgerigars, *Melopsittacus undulatus*. *Animal Behav.* **59**, 1209–1218 (2000).
- Wanker, R., Sugama, Y. & Prinage, S. Vocal labelling of family members in spectacled parrotlets, *Forpus conspicillatus*. *Animal Behav.* **70**, 111–118 (2005).
- Tyack, P. L. Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *J. Comp. Psychol.* **122**, 319–331 (2008).
- King, S. L., Sayigh, L. S., Wells, R. S., Fellner, W. & Janik, V. M. Vocal copying of individually distinctive signature whistles in bottlenose dolphins. *Proc. Royal Soc. B: Biol. Sci.* **280**, 20130053 (2013).
- Pardo, J. Measuring phonetic convergence in speech production. *Front. Psychol.* **4**, 559 (2013).
- Tyack, P. L. A taxonomy for vocal learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20180406 (2019).
- Vernes, S. C., Janik, V. M., Fitch, W. T. & Slater, P. J. B. Vocal learning in animals and humans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200234 (2021).
- McDonald, P. G. & Wright, J. Bell miner provisioning calls are more similar among relatives and are used by helpers at the nest to bias their effort towards kin. *Proc. Royal Soc. B: Biol. Sci.* **278**, 3403–3411 (2011).
- Keenan, P. C. & Benkman, C. W. Call imitation and call modification in red crossbills. *Condor* **110**, 93–101 (2008).
- Baptista, L. F. & Schuchmann, K.-L. Song learning in the Anna hummingbird (*Calypte anna*). *Ethology* **84**, 15–26 (1990).
- Scarl, J. C. & Bradbury, J. W. Rapid vocal convergence in an Australian cockatoo, the galah (*Eolophus roseicapillus*). *Anim. Behav.* **77**, 1019–1026 (2009).
- Mesoudi, A., Chang, L., Dall, S. R. X. & Thornton, A. The evolution of individual and cultural variation in social learning. *Trends Ecol. Evol.* **31**, 215–225 (2016).
- Thomson, C. E., Winney, I. S., Salles, O. C. & Pujol, B. A guide to using a multiple-matrix animal model to disentangle genetic and nongenetic causes of phenotypic variance. *PLoS One* **13**, e0197720 (2018).
- Nieder, A. & Mooney, R. The neurobiology of innate, volitional and learned vocalizations in mammals and birds. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190054 (2019).

48. Loo, Y. Y. et al. Structure and function of the vocal repertoire of the Rifleman (*Acanthisitta chloris*), a member of the earliest diverging passerine suborder, Acanthisitti. *Journal of Field Ornithology* **94**, 11 (2023).
49. Sherley, G. H. Co-operative breeding in Riflemen (*Acanthisitta chloris*) benefits to parents, offspring and helpers. *Behaviour* **112**, 1–22 (1990).
50. Preston, S. A. J. Routes to cooperation in the rifleman *Acanthisitta chloris*. (University of Sheffield, 2012).
51. Mundinger, P. C. Vocal imitation and individual recognition of Finch calls. *Science* **168**, 480 (1970).
52. Mundinger, P. C. Call Learning in the Carduelinae: Ethological and Systematic Considerations. *Syst. Biol.* **28**, 270–283 (1979).
53. Walløe, S., Thomsen, H., Balsby, T. J. & Dabelsteen, T. Differences in short-term vocal learning in parrots, a comparative study. *Behaviour* **152**, 1433–1461 (2015).
54. Forstmeier, W., Burger, C., Temnow, K. & Derégnaucourt, S. The genetic basis of Zebra Finch vocalizations. *Evolution* **63**, 2114–2130 (2009).
55. Robisson, P., Aubin, T. & Bremond, J.-C. Individuality in the Voice of the Emperor Penguin *Aptenodytes forsteri*: Adaptation to a Noisy Environment. *Ethology* **94**, 279–290 (2010).
56. Berg, K. S., Delgado, S., Cortopassi, K. A., Beissinger, S. R. & Bradbury, J. W. Vertical transmission of learned signatures in a wild parrot. *Proc. Biol. Sci.* **279**, 585–591 (2012).
57. Kremers, D., Lemasson, A., Almunia, J. & Wanker, R. Vocal sharing and individual acoustic distinctiveness within a group of captive orcas (*Orcinus orca*). *J. Comp. Psychol.* **126**, 433–445 (2012).
58. Janik, V. M. & Sayigh, L. S. Communication in bottlenose dolphins: 50 years of signature whistle research. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **199**, 479–489 (2013).
59. King, S. L. & Janik, V. M. Bottlenose dolphins can use learned vocal labels to address each other. *Proceedings of the National Academy of Sciences* **110**, 13216 (2013).
60. Mumm, C. A. S., Urrutia, M. C. & Knörnschild, M. Vocal individuality in cohesion calls of giant otters, *Pteronura brasiliensis*. *Anim. Behav.* **88**, 243–252 (2014).
61. Khwaja, N., Briskie, J. V. & Hatchwell, B. J. Individuality, kin similarity and experimental playback of contact calls in cooperatively breeding rifleman. *N. Z. J. Zool.* **46**, 334–347 (2019).
62. Valletta, J. J., Torney, C., Kings, M., Thornton, A. & Madden, J. Applications of machine learning in animal behaviour studies. *Anim. Behav.* **124**, 203–220 (2017).
63. Zhang, K. et al. Comparing context-dependent call sequences employing machine learning methods: An indication of syntactic structure of greater horseshoe bats. *J. Exp. Biol.* **222**, jeb214072 (2019).
64. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).
65. Araya-Salas, M. & Smith-Vidaurre, G. warbleR: an r package to streamline analysis of animal acoustic signals. *Methods Ecol. Evol.* **8**, 184–191 (2017).
66. Elshire, R. J. et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**, e19379 (2011).
67. Dodds, K. G. et al. Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genomics* **16**, 1047 (2015).
68. Dodds, K. G. et al. Exclusion and genomic relatedness methods for assignment of parentage using genotyping-by-sequencing data. *G3: Genes, Genomes, Genetics* **9**, 3239–3247 (2019).
69. Pardo, J. S. On phonetic convergence during conversational interaction. *J. Acoust. Soc. Am.* **119**, 2382–2393 (2006).
70. McDonald, P. G., Rollins, L. A. & Godfrey, S. The relative importance of spatial proximity, kin selection and potential ‘greenbeard’ signals on provisioning behaviour among helpers in a cooperative bird. *Behav. Ecol. Sociobiol.* **70**, 133–143 (2016).
71. Preston, S. A. J., Briskie, J. V., Burke, T. & Hatchwell, B. J. Genetic analysis reveals diverse kin-directed routes to helping in the rifleman *Acanthisitta chloris*. *Mol. Ecol.* **22**, 5027–5039 (2013).
72. Soha, J. A. The auditory template hypothesis: a review and comparative perspective. *Anim. Behav.* **124**, 247–254 (2017).
73. Smolker, R. A., Mann, J. & Smuts, B. B. Use of signature whistles during separations and reunions by wild bottlenose dolphin mothers and infants. *Behav. Ecol. Sociobiol.* **33**, 393–402 (1993).
74. Sousa-Lima, R. S., Paglia, A. P. & Da Fonseca, G. A. B. Signature information and individual recognition in the isolation calls of Amazonian Manatees, *Trichechus inunguis* (Mammalia: Sirenia). *Anim. Behav.* **63**, 301–310 (2002).
75. Feng, A. S. et al. Diversity of the vocal signals of concave-eared torrent frogs (*Odorrana tormota*): Evidence for individual signatures. *Ethology* **115**, 1015–1028 (2009).
76. Radford, A. N. Group-specific vocal signatures and neighbour-stranger discrimination in the cooperatively breeding green woodhoopoe. *Anim. Behav.* **70**, 1227–1234 (2005).
77. Chuang, M.-F., Kam, Y.-C. & Bee, M. A. Territorial olive frogs display lower aggression towards neighbours than strangers based on individual vocal signatures. *Anim. Behav.* **123**, 217–228 (2017).
78. Elie, J. E. & Theunissen, F. E. Zebra Finches identify individuals using vocal signatures unique to each call type. *Nat. Commun.* **9**, 4026 (2018).
79. Mulder, R. A. & Langmore, N. E. Dominant males punish helpers for temporary defection in superb fairy-wrens. *Anim. Behav.* **45**, 830–833 (1993).
80. Knörnschild, M., Nagy, M., Metz, M., Mayer, F. & von Helversen, O. Learned vocal group signatures in the polygynous bat *Saccopteryx bilineata*. *Anim. Behav.* **84**, 761–769 (2012).
81. Colombelli-Négrel, D. et al. Embryonic learning of vocal passwords in superb fairy-wrens reveals intruder cuckoo nestlings. *Curr. Biol.* **22**, 2155–2160 (2012).
82. Araya-Salas, M. et al. Social group signatures in hummingbird displays provide evidence of co-occurrence of vocal and visual learning. *Proc. Royal Soc. B: Biol. Sci.* **286**, 20190666 (2019).
83. Fischer, J., Wegdell, F., Trede, F., Dal Pesco, F. & Hammerschmidt, K. Vocal convergence in a multi-level primate society: Insights into the evolution of vocal learning. *Proc. Royal Soc. B: Biol. Sci.* **287**, 20202531 (2020).
84. Sharp, S. P. & Hatchwell, B. J. Development of family specific contact calls in the Long-tailed Tit *Aegithalos caudatus*. *Ibis* **148**, 649–656 (2006).
85. Miller, P. J. O., Shapiro, A. D., Tyack, P. L. & Solow, A. R. Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca*. *Anim. Behav.* **67**, 1099–1107 (2004).
86. Nakahara, F. & Miyazaki, N. Vocal exchanges of signature whistles in bottlenose dolphins (*Tursiops truncatus*). *J. Ethol.* **29**, 309–320 (2011).
87. King, S. L. & McGregor, P. K. Vocal matching: the what, the why and the how. *Biol. Lett.* **12**, 20160666 (2016).
88. Janik, V. M. & Slater, P. J. B. The different roles of social learning in vocal communication. *Anim. Behav.* **60**, 1–11 (2000).
89. Crockford, C., Herbinger, I., Vigilant, L. & Boesch, C. Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology* **110**, 221–243 (2004).
90. Lemasson, A. & Hausberger, M. Patterns of vocal sharing and social dynamics in a captive group of Campbell’s monkeys (*Cercopithecus campbelli campbelli*). *J. Comp. Psychol.* **118**, 347–359 (2004).
91. Zaccaroni, M. et al. Group specific vocal signature in free-ranging wolf packs. *Ethol. Ecol. Evol.* **24**, 322–331 (2012).
92. Luigi, B. et al. Vocal accommodation in penguins (*Spheniscus demersus*) as a result of social environment. *Proc. Royal Soc. B: Biol. Sci.* **289**, 20220626 (2022).
93. Volodin, I. A., Volodina, E. V., Lapshina, E. N., Efremova, K. O. & Soldatova, N. V. Vocal group signatures in the goitred gazelle *Gazella subgutturosa*. *Anim. Cogn.* **17**, 349–357 (2014).
94. Briefer, E. F. & McElligott, A. G. Social effects on vocal ontogeny in an ungulate, the goat, *Capra hircus*. *Anim. Behav.* **83**, 991–1000 (2012).

95. Snowdon, C. T. & Elowson, A. M. Pygmy marmosets modify call structure when paired. *Ethology* **105**, 893–908 (1999).
96. Zhao, L., Rad, B. B. & Wang, X. Long-lasting vocal plasticity in adult marmoset monkeys. *Proc. Royal Soc. B: Biol. Sci.* **286**, 20190817 (2019).
97. Wich, S. A. et al. A case of spontaneous acquisition of a human sound by an orangutan. *Primates* **50**, 56–64 (2009).
98. Lameira, A. R. et al. Sociality predicts orangutan vocal phenotype. *Nat. Ecol. Evol.* **6**, 644–652 (2022).
99. Koda, H., Lemasson, A., Oyakawa, C., Pamungkas, J. & Masataka, N. Possible role of mother-daughter vocal interactions on the development of species-specific song in gibbons. *PLoS One* **8**, e71432 (2013).
100. Lopez, P. T., Narins, P. M., Lewis, E. R. & Moore, S. W. Acoustically induced call modification in the white-lipped frog, *Leptodactylus albilabris*. *Anim. Behav.* **36**, 1295–1308 (1988).
101. Trainer, J. M., McDonald, D. B. & Learn, W. A. The development of coordinated singing in cooperatively displaying long-tailed manakins. *Behav. Ecol.* **13**, 65–69 (2002).
102. Briefer, E. F., Tettamanti, F. & McElligott, A. G. Emotions in goats: mapping physiological, behavioural and vocal profiles. *Anim. Behav.* **99**, 131–143 (2015).
103. Wei, D., Talwar, V. & Lin, D. Neural circuits of social behaviors: Innate yet flexible. *Neuron* **109**, 1600–1620 (2021).
104. Parris, K. M., Velik-Lord, M. & North, J. M. A. Frogs call at a higher pitch in traffic noise. *Ecol. Soc.* **14**, (2009).
105. Hunt, G. R. & Mclean, I. G. The ecomorphology of sexual dimorphism in the New Zealand Rifleman *Acanthisitta chloris*. *Emu* **93**, 71–78 (1993).
106. Preston, S. A. J., Briskie, J. V. & Hatchwell, B. J. Adult helpers increase the recruitment of closely related offspring in the cooperatively breeding rifleman. *Behav. Ecol.* **27**, 1617–1626 (2016).
107. Cornell Lab of Ornithology, R. P. Raven Pro: Interactive sound analysis software (Version 1.6.1) [Computer software] Center for conservation bioacoustics. *Ithaca, NY* The Cornell Lab of Ornithology (2019).
108. Sueur, J., Aubin, T. & Simonis, C. Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics* **18**, 213–226 (2008).
109. Venables, W. N. & Ripley, B. D. *Modern applied statistics with S*. (Springer, 2010).
110. Oksanen, J. et al. *vegan: Community Ecology Package*. (2019).
111. Anderson, M. J. & Walsh, D. C. I. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? *Ecol. Monogr.* **83**, 557–574 (2013).
112. Tomar, S. Converting video formats with FFmpeg. *Linux J* **2006**, 10 (2006).
113. Ligges, U., Krey, S., Mersmann, O. & Schnackenberg, S. tuneR: Analysis of Music and Speech. *Ecol. Appl.* **30**, e02140 (2018).
114. Muda, L., Begam, M. & Elamvazuthi, I. Voice recognition algorithms using Mel Frequency Cepstral Coefficient (MFCC) and dynamic time warping (DTW) techniques. *Journal of Computing* **2**, 138–143 (2010).
115. Spillmann, B., van Schaik, C. P., Setia, T. M. & Sadjadi, S. O. Who shall I say is calling? Validation of a caller recognition procedure in Bornean flanged male orangutan (*Pongo pygmaeus wurmbii*) long calls. *Bioacoustics* **26**, 109–120 (2017).
116. Wang, J. Estimating pairwise relatedness in a small sample of individuals. *Heredity* **119**, 302–313 (2017).
117. Andrew, S. *Fastqc: A quality control tool for high throughput sequence data*.
118. Lu, F. et al. Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genet* **9**, e1003215 (2013).
119. Bradbury, P. J. et al. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
120. Lemasson, A., Ouattara, K., Petit, E. J. & Zuberbühler, K. Social learning of vocal structure in a nonhuman primate? *BMC Evol. Biol.* **11**, 362 (2011).
121. Louca, S. & Doebeli, M. Efficient comparative phylogenetics on large trees. *Bioinformatics* **34**, 1053–1055 (2017).
122. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
123. RCoreTeam. *A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2016). <http://www.r-project.org>.
124. Dietz, E. J. Permutation Tests for Association Between Two Distance Matrices. *Syst. Biol.* **32**, 21–26 (1983).
125. Legendre, P. & Legendre, L. *Numerical Ecology*. (Elsevier, 1998).
126. Hadfield, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33**, 1–22 (2010).
127. Hadfield, J. D., Heap, E. A., Bayer, F., Mittell, E. A. & Crouch, N. M. A. Disentangling genetic and prenatal sources of familial resemblance across ontogeny in a wild passerine. *Evolution* **67**, 2701–2713 (2013).
128. de Villemereuil, P. Quantitative genetic methods depending on the nature of the phenotypic trait. *Annals of the New York Academy of Sciences. The Year in Evolutionary Biology* **1422**, 29–47 (2018).
129. de Villemereuil, P., Gimenez, O. & Doligez, B. Comparing parent-offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: a simulation study for Gaussian and binary traits. *Methods Ecol. Evol.* **4**, 260–275 (2013).
130. Moran, I. G. et al. Dataset. Vocal learning in a vocal non-learner? Social proximity and vocal convergence shape the calls of the most basal Passeriformes, New Zealand Wrens. *Figshare Dataset* <https://doi.org/10.17608/k6.auckland.25549466> (2024).
131. Fukushima, Y. & Aoki, K. The Role of the Dorsomedial Nucleus (DM) of Intercollicular Complex with regard to sexual difference of distance calls in Bengalese finches. *Zoolog. Sci.* **17**, 1231–1238 (2000).
132. Fukushima, Y. & Aoki, K. Neural function of the Mesencephalic Dorsomedial Nucleus (DM) on distance call production in Bengalese Finches. *Zoolog. Sci.* **19**, 393–402 (2002).
133. Nowicki, S. Vocal plasticity in captive black-capped chickadees: the acoustic basis and rate of call convergence. *Anim. Behav.* **37**, 64–73 (1989).
134. Luef, E. M., Maat, A. T. & Pika, S. Vocal similarity in long-distance and short-distance vocalizations in raven pairs (*Corvus corax*) in captivity. *Behav. Processes* **142**, 1–7 (2017).
135. Hile, A. G. & Striedter, G. F. Call convergence within groups of female budgerigars (*Melopsittacus undulatus*). *Ethology* **106**, 1105–1114 (2000).
136. Groothuis, T. The influence of social experience on the development and fixation of the form of displays in the black-headed gull. *Anim. Behav.* **43**, 1–14 (1992).
137. Williams, H. & Lachlan, R. F. Evidence for cumulative cultural evolution in bird song. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **377**, 20200322 (2022).
138. Araya-Salas, M. & Wright, T. Open-ended song learning in a hummingbird. *Biol. Lett.* **9**, 20130625 (2013).
139. Nespor, A. A., Lukazewicz, M. J., Dooling, R. J. & Ball, G. F. Testosterone induction of male-like vocalizations in female budgerigars (*Melopsittacus undulatus*). *Horm. Behav.* **30**, 162–169 (1996).
140. Bradbury, J. W. & Balsby, T. J. S. The functions of vocal learning in parrots. *Behav. Ecol. Sociobiol.* **70**, 293–312 (2016).
141. Madabhushi, A. J., Wewhare, N., Binwal, P. & Krishnan, A. Higher-order dialectic variation and syntactic convergence in the complex warble song of budgerigars. *J. Exp. Biol.* **226**, jeb245678 (2023).
142. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).

Acknowledgements

We would like to thank mana whenua of the Maungaharuru region for welcoming our research team on their land and the Boundary Stream

Mainland Island team from the Department of Conservation (DOC) for facilitating our research and for providing generous support to our research team from 2018 to 2021. We are grateful to T. and J. Wells, J. Leung, A. Menzies, M. Bidmead, F. Patterson, for their help during data collection and L. Zantis, and E. Carroll for their help in the lab. We are thankful to the AgResearch team, J. McEwan, T. Vanstijn, R. Brauning, K. Dodds, and C. Shannon, for their sequencing work. We thank the University of Auckland's engineering team for building the RFID units and the eResearch team for their help with the RFID data processing and time offset correction. We would also like to thank the editors Richard Holland and Luke R. Grinham and anonymous referees who provided detailed comments on previous versions of the manuscript. Finally, we are grateful to our sources of funding that supported this research, The Royal Society of New Zealand's Marsden Fund [number MFP-UOA1707], the University of Auckland Doctoral Scholarship and Press Account, Birds New Zealand Research Funds, and the Centre for Biodiversity and Biosecurity.

Author contributions

All authors made substantial contributions. This work was designed by I.G. Moran and K.E. Cain with additional input from S.J. Withers and M.L. Hall. I.G. Moran wrote the manuscript. I.G. Moran conducted the data collection with input from K.E. Cain, Y.Y. Loo, S.J. Withers, and S. Louca. I.G. Moran conducted lab work with input from A. Whibley, and P.M. Salloum. I.G. Moran analyzed and interpreted the data with input from N.B.A. Young, S. Louca, A. Whibley, and K.E. Cain. S. Louca and N.B.A. Young provided statistical and programming input. K.E. Cain and M.C. Stanley supervised this research and edited the manuscript. Y.Y. Loo, N.B.A. Young, A. Whibley, P.M. Salloum, S.J. Withers, S. Louca, and M.L. Hall provided editorial input to the manuscript drafts.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-024-06253-y>.

Correspondence and requests for materials should be addressed to Ines G. Moran.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Richard Holland and Luke R. Grinham.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024