De novo establishment of wild-type song culture in the zebra finch

Olga Feher[1], **Haibin Wang**[2], **Sigal Saar**[1], **Partha P. Mitra**[2], and **Ofer Tchernichovski**[1]

[1] Department of Biology, City College, City University of New York, New York 10031, USA

[2] Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, US

Abstract

What sort of culture would evolve in an island colony of naive founders? This question cannot be studied experimentally in humans. We performed the analogous experiment using socially learned birdsong. Culture is typically viewed as consisting of traits inherited epigenetically, via social learning. However, cultural diversity has species-typical constraints[1], presumably of genetic origin. A celebrated, if contentious, example is whether a universal grammar constrains syntactic diversity in human languages[2]. Oscine songbirds exhibit song learning and provide biologically tractable models of culture: members of a species show individual variation in song[3] and geographically separated groups have local song dialects [4,5]. Different species exhibit distinct song cultures[6,7], suggestive of genetic constraints[8,9]. Absent such constraints, innovations and copying errors should cause unbounded variation over multiple generations or geographical distance, contrary to observations[9]. We asked if wild-type song culture might emerge over multiple generations in an isolated colony founded by isolates, and if so, how this might happen and what type of social environment is required[10]. Zebra finch isolates, unexposed to singing males during development, produce song with characteristics that differ from the wild-type song found in laboratory[11] or natural colonies. In tutoring lineages starting from isolate founders, we quantified alterations in song across tutoring generations in two social environments: tutor-pupil pairs in sound-isolated chambers and an isolated semi-natural colony. In both settings, juveniles imitated the isolate tutors, but changed certain characteristics of the songs. These alterations accumulated over learning generations. Consequently, songs evolved toward the wild-type in 3–4 generations. Thus, species-typical song culture can appear *de novo*. Our study has parallels with language change and evolution[12,13]. In analogy to models in quantitative genetics[14,15], we model song culture as a *multi-generational* phenotype, partly encoded genetically in an isolate founding population, influenced by environmental variables, and taking multiple generations to emerge.

Young male zebra finches develop individually distinct song by imitating adult males[16]. The adult wild-type (WT) song includes stereotyped syllables repeated in fixed order (song motifs, Fig. 1a) in both wild and domesticated zebra finch colonies. Birds deprived of song during vocal development, develop a less structured isolate (ISO) song with more noisy, broadband notes and high pitch upsweeps[11] (Fig. 1b). ISO syllables are often prolonged, monotonic or stuttered, and the songs appear to have an irregular rhythm. Despite these anomalies, young zebra finches readily imitate songs of adult isolates[17] even in the presence of WT adults[11].

We quantified the differences between WT and ISO songs over three time-scales. At the 10 ms time-scale, we used spectral frame features (e.g., frequency modulation; Supplementary 4a). Over the 10–100 ms time-scale, we used the correlation time of the spectral shape, termed *Duration of Acoustic State* (DAS, Supplementary 4b). At even longer (200–1000 ms) time-scales, we used measures of song rhythm (Supplementary 4d)[18]. Feature probability distributions across birds differed between ISO and WT (Fig. 1c–e). ISO songs had lower frequency modulation, longer durations of acoustic state, and less structured rhythms.

These distributions provide a high-dimensional song phenotype for each bird. We reduced the dimensionality by applying Principal Component Analysis (PCA) to the collection of feature distributions of all birds (WT & ISO), and retained the first two principal components (PCs) to obtain two-dimensional song phenotype values (Supplementary 4e). PCs at all three time-scales show separable clusters for ISO and WT songs along a continuum (Fig 2a–c). The mean values of the first PC were significantly different between ISO and WT at all time-scales of song structure ($p<0.001$, t-tests, $n_{WT}=52$ birds, $n_{iso}=17$ birds, FDR adjusted, Supplementary 5). We found that these differences are largely an outcome of tutoring deprivation and not of social isolation (Supplementary 3f).

To examine the imitation of isolate songs, we trained 13 juvenile birds (pupils) by isolate tutors one-to-one in a sound-isolated chamber. This allowed us to control genetic relatedness, and to minimize social effects, e.g., to eliminate feedback from female listeners. Four isolate tutors, with songs stable over the course of tutoring, were used 2–4 times to train unrelated pupils. We projected the feature distributions of the pupils on the PCs derived earlier from the WT/ISO data (Fig. 2a–c), and displayed vectors connecting each ISO tutor to his pupils (Fig. 2d–f). As shown, most of these vectors point in the direction of the WT cluster, indicating a shift toward WT features in pupils of ISO tutors. The mean values of the first PC for the first generation pupils differed significantly from both ISO and WT means for the spectral-frame features and for DAS ($p=0.018$-0.001, $n=13$), but not for rhythm. Feature distributions of most individual pupil songs were closer to WT songs than were their tutor's songs (12/13 at at least one time-scale, 10/13 at all time-scales, FDR significance=0.01, binomial test, $n=52$, supplementary 5d).

Although pupils typically imitated all of the tutor syllables[20] and did not invent new syllables (Supplementary 2), pupil songs deviated consistently from tutor songs. Fig. 2g presents an example where a long ISO syllable (red bar, mean duration=367ms, s.d.=29ms) was copied by a pupil, but was shortened by about 30% (mean=243ms, s.d.=7.6ms). Across all the syllables and all pupils, the durations of pupil syllables accurately matched those of

the corresponding ISO tutor syllables for syllables shorter than 230ms (Fig. 2h, r, 2=0.98, slope=0.97, n=20 syllables). Copies of longer ISO syllables, however, were shorter than the originals (r²=0.84, slope=0.56, n=11 syllables). Across birds, the ratio between the longest and shortest syllable within a bout was significantly smaller in pupils compared to their ISO tutors (p<0.01 n=13, Wilcoxon sign test, Supplementary 4c). Overall, the range where durations of ISO syllables were accurately copied is similar to the range of WT syllable durations (25–75 percentile range = 67–180ms, n=52 WT birds). In addition, pupils only copied the abundance (relative frequency) of syllables when it was within the WT range (up to about 30%). In cases where one syllable dominated the ISO song (Fig 2i), pupils decreased its abundance to 20–30% (Supplementary Fig. 5), thereby creating more structured song motifs.

Imitation of spectral features, as judged by the first PC of the feature distribution, was also biased: linear regression analysis of pupil versus tutor yielded a nonzero intercept and a slope slightly less than one (Fig 2j). The equality line, corresponding to faithful copying (pupil=tutor, dashed blue line), was rejected in favor of the alternative hypothesis represented by the linear fit shown in red (P<0.001, likelihood ratio test, n=13). Note that imitation that was inaccurate but unbiased would have only increased the spread around the equality line.

Because the songs of ISO-tutored birds differed significantly from both their respective ISO tutors and WT, we examined whether recursive tutoring would cause further progression toward WT over multiple generations. We used four of the first-generation pupils as tutors of a second generation of unrelated pupils, and continued recursively over 2–5 generations (Fig. 3a). Similarity to WT songs increased over 3–4 generations, as can be appreciated from the audio in Supplementary 1 and the three examples of multiple generations of recursive tutoring in Fig. 3b. In the first example, both ISO syllables become shorter in the songs of the first and second generation pupils (blue and red rectangles), but the second syllable is also differentiated into three distinct notes. The middle panel shows spectral and temporal differentiation of syllables, and omission by the 3rd generation pupil. In the right lineage, the duration of the final syllable (red rectangle) decreased over two generations and then stabilized. The spectral structure, however, continued to change in the 3rd and 4th generations.

To judge if the imitation of ISO songs progressed toward WT song over multiple generations, we displayed vectors in the PC space (as in Fig. 2d–f) with each tutoring lineage labeled by a different color (Fig. 3c–e). As shown, the multi-generational trajectories penetrate more deeply into the WT cluster (purple shading). Direct comparisons across first and later generation pupils reach significance only for DAS (p=0.02), but multi-generational comparisons suggest further progression toward WT for all song traits. For spectral frame features, we found that the first principal component of song features changes monotonically toward WT over generations. Its mean values for ISO, first generation, later generations, and WT songs were 1.3, 0.3, 0.03, −0.4 respectively. First PC values for later generation songs were significantly different from ISO song (p<0.005, t-test, n=8 for later generations) but not from WT songs (p=0.17). For DAS, first PC values also decreased monotonically with generations: 1.1, 0.3, 0.02, −0.3. Higher generation songs were significantly different

(p<0.01) from both WT and ISO, suggesting that WT approximation was not complete. For rhythm, first PC values also decreased monotonically with generations: 4.1, 2.2, 1.4, −2, and differences from WT and ISO were marginally significant (p=0.02, 0.056 respectively).

Although the one-to-one training provided a well defined learning environment, the multi-generational changes that would occur in a complex social setting may be more representative of natural evolutionary processes. Therefore, we established a semi-natural island colony (Supplementary 3d) starting with one of our isolate tutors and three unrelated females in a large sound chamber (Supplementary Fig. 1).

In this social situation, too, the isolate colony approached the WT cluster over a few generations (Fig. 4). To judge the transition toward WT clusters, we examined PC projections with the isolate tutor song marked as a red dot. Comparing the trajectory shown in Fig. 4e to that of Fig. 3b, right panel (originating from the same tutor), we see that the outcome in the colony is similar to that observed in one-to-one tutoring. Even though the outcome of the colony experiment can only be judged qualitatively, we find it remarkable that despite intense social interactions, female presence and mating competition, there were only mild differences between birds in the two conditions. In the colony, juveniles also imitated sibling syllables and female long calls, leading to more complex songs (Supplementary 1c). In contrast to one-to-one tutoring, the best progress toward WT song occurred in rhythm, perhaps because birds incorporated additional syllable types into their song motifs.

Our findings resemble the well-known case of deaf children in Managua, Nicaragua, spontaneously developing sign language[21], as well as linguistic phenomena such as creolization. Models of language change and evolution[12–14], which contain a developmental account of the language acquisition process, are germane to our study (Supplementary Model 3).

We further discuss our findings using a simple recursive model which motivated this study. PCs of feature distributions (Fig. 2) give us phenotypic measures of song. Consider the distribution of a quantitative phenotype P in the ISO population. Since some of the variation in ISO songs is heritable, we partition P into a genotypic and an environmental value $P = G + E$, assuming an additive model for genetic variance[22] $V_P = V_G + V_E$.

We consider an Isolated Lineages Model, in which the environmental component of the pupil phenotype $P(n+1)$ in the $n+1$'th generation is further divided into a portion $E_0(n+1)$ independent of the tutor, and a portion proportional to the tutor song phenotype $c_0 P(n)$. We therefore have the recursion $P(n+1) = G(n+1) + c_0 P(n) + E_0(n+1)$ [Eq. 1]. The partitioning of the phenotypic variance is analogous to the parental effects model in quantitative genetics[1,23]. In the one-to-one study, tutor and pupil genotypic values are approximately uncorrelated, and c_0 may be estimated by regressing the pupil against the tutor (cf. Fig. 2j, c_0 = 0.86, s.d = 0.15). The literature on cultural transmission[24,25] also contains models analogous to Eq. 1 and has similar implications. Half-sib or cross-fostering experimental designs[26] should be useful for separating the genetic[27] and learning-related components of song transmission in future studies[28].

Our one-to-one experimental design may be modeled using Eq. 1 by initializing $P(1)=G(1)+E(1)$ for the ISO generation. The recursion then causes the distribution of phenotypic values to exponentially relax to an asymptotic "WT" distribution, the relaxation being rapid if c_0 is close to 0. The largest changes occur in the first generation (consistent with our results). The case $c_0 =1$ corresponds to a simple random walk $V[P(n)]\sim n$, where the song phenotype would drift indefinitely (unbiased song copying with errors). The "copying bias" $(1-c_0)$ plays the role of a spring constant, confining the walker to a parabolic potential well. Notably, the WT variance in the model is a combination of the ISO variance and the learning parameter, emphasizing how ISO song and learning ability combine to produce WT song. Extensions of the model predict that both genetic relatedness between tutor and pupil and horizontal transmission alter the asymptotic "WT" distributions (Supplementary Model). Therefore we would expect our two designs to yield slightly different song cultures.

In a sense, the results of our study show that song culture is the result of an extended developmental process, a 'multi-generational' phenotype partly genetically encoded in a founding population and partly in environmental variables, but taking multiple generations to emerge. The functional significance of our findings remains open, i.e. whether WT females prefer the songs of multi-generation pupils to those of ISO tutors. Since our findings suggest that song culture is the result of an extended developmental process, it would be interesting to examine if changes in gene expression, neuronal reorganization or neurogenesis associated with song development show orderly multi-generational progression during the evolution of song culture.

METHODS SUMMARY

Animal care

All experiments were performed in accordance with guidelines of the National Institutes of Health and have been reviewed and approved by the IACUC of CCNY.

Experimental design

We used zebra finches (*Taenyopygia guttata*) from the CCNY breeding colony. Colony management and isolation procedures have been described previously[29]. Except for the colony experiment, all birds were kept either singly (isolates) or pair-wise (one-to-one tutored) in sound attenuation chambers (Supplementary 3e) from day 30 to 120 post-hatch. Wild-type songs (n=52) were obtained from birds raised in two well-established colonies. Isolates (n=17) were raised by their mothers from day 7–29 post-hatch and were kept in complete isolation from day 30 until day 120 or later. One-to-one tutored birds (n=13 and 8, for first and later generations, respectively), were randomly selected from 40 breeding pairs, and paired with one of 6 isolate tutors on day 30. For the colony setting, we made a sound isolation chamber from an old 20 cubic ft refrigerator (Supplementary Fig. 1). All birds in the colony (except for the 3 female founders) were the descendants of the founder male.

Data analysis

All the analysis was performed using *Matlab 7*, except for spectral feature calculations, which were done using *Sound Analysis Pro 2*. Isolate song syllables are often prolonged and monotonic. To quantify this notion, we estimated the time interval where acoustic features remain highly correlated and named this feature *duration of acoustic state* (Supplementary 4b). Rhythm spectrum18 was used to detect periodicity in song features at the syllabic and the song-motif levels (Supplementary 4d). We constructed song feature PCs by first computing cumulative frequency distributions (CDF) for each feature time-series (Supplementary Fig. 8). These CDFs were the input vectors for the Principal Component Analysis (Fig. 2a–c). Statistical tests are described Supplementary 5.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Marler P, Sherman V. Innate differences in singing behavior of sparrows reared in isolation from adult conspecific song. Animal Behaviour. 1985; 33:57–71.

2. Chomsky, N. Aspects of the Theory of Syntax. MIT Press; 1965.

3. Catchpole, CK.; Slater, PBJ. Bird Song: Biological Themes and Variations. Cambridge University Press; 2008.

4. Marler P, Tamura M. Culturally transmitted patterns of vocal behavior in sparrows. Science. 1964; 146:1483–6. [PubMed: 14208581]

5. Olofsson H, Servedio MR. Sympatry affects the evolution of genetic versus cultural determination of song. Behavioral Ecology. 2008; 19:596–604.

6. Soha JA, Marler P. A species-specific acoustic cue for selective song learning in the white-crowned sparrow. Animal Behaviour. 2000; 60:297–306. [PubMed: 11007638]

7. West, MJ.; King, AP. Issues in the Ecological Study of Learning. Johnston, TD.; Pietrewicz, AT., editors. Lawrence Erlbaum Associates; Hillsdale, NJ: 1985.

8. Gardner TJ, Naef F, Nottebohm F. Freedom and rules: the acquisition and reprogramming of a bird's learned song. Science. 2005; 308:1046–1049. [PubMed: 15890887]

9. Adret, P. In search of the song template. In: Zeigler, HP.; Marler, P., editors. Behavioral Neurobiology of Birdsong. Vol. 1016. Annals of the New York Academy of Sciences; 2004. p. 303-324.

10. Volman SF, Khanna H. Convergence of untutored song in group-reared zebra finches (*Taeniopygia guttata*). Journal of Comparative Pshychology. 1995; 109:211–221.

11. Williams H, Kilander K, Sotanski ML. Untutored song, reproductive success and song learning. Animal Behaviour. 1993; 45:695–705.

12. Niyogi, P. The computational nature of language learning and evolution. Cambridge, Mass: MIT Press; 2006.

13. Nowak MA, Komarova NL, Niyogi P. Computational and evolutionary aspects of language. Nature. 2002; 417:611–617. [PubMed: 12050656]

14. Cheverud JM. Evolution in a genetically heritable social environment. Proc Natl Acad Sci U S A. 2003; 100:4357–9. [PubMed: 12682282]

15. Dickerson G. Comparison of hog carcasses as influenced by heritable differences in rate and economy of gain. Iowa Agr Exper Sta Res Bull. 1947; 354:489–524.

16. Jones AE, Ten Cate C, Slater PJB. Early experience and plasticity of song in adult male zebra finches (*Taeniopygia guttata*). Journal of Comparative Phychology. 1996; 4:354–369.

17. Bohner J. Song learning in the zebra finch (*Taeniopygia guttata*): Selectivity in the choice of a tutor and accuracy of song copies. Animal Behaviour. 1983; 31:231–237.

18. Saar S, Mitra PP. A technique for characterizing the development of rhythms in bird song. PLoS ONE. 2008; 3:1461.

19. Tchernichovski O, Nottebohm F, Ho CE, Bijan P, Mitra PP. A procedure for an automated measurement of song similarity. Animal Behaviour. 2000; 59:1167–1176. [PubMed: 10877896]

20. Nelson DA, Marler P. Selection-based learning in bird song development. Proc Natl Acad Sci U S A. 1994; 91:10498–10501. [PubMed: 7937982]

21. Kegl, J. Language Emergence in a Language-Ready Brain: Acquisition Issues. In: Morgan, G.; Woll, B., editors. Language Acquisition in Signed Languages. Cambridge University Press; 2002. p. 207-254.

22. Falconer, DS.; Mackay, TCF. Introduction to quantitative genetics. 4. Harlow, England: Pearson, Prentice Hall; 1996.

23. Willham RL. The covariance between relatives for characters composed of components contributed by related individuals. Biometrics. 1963; 19:18.

24. Cavalli-Sforza, LL.; Feldman, MW. Cultural transmission and evolution: a quantitative approach. Princeton University Press; Princeton, N. J: 1981.

25. Boyd, R.; Richerson, PJ. Culture and the evolutionary process. University of Chicago Press: Chicago; 1985.

26. Boake, CRB. Quantitative genetic studies of behavioral evolution. University of Chicago Press; Chicago: 1994.

27. Mundinger PC. Behaviour-genetic analysis of canary song: inter-strain differences in sensory learning, and epigenetic rules. Animal Behaviour. 1995; 50:1491–1511.

28. Soha JA, Marler P. Cues for Early Discrimination of conspecific song in the white-crowned sparrow (*Zonotrichia leucophrys*). Ethology. 2001; 107:813–826.

29. Tchernichovski O, Lints T, Mitra PP, Nottebohm F. Vocal imitation in zebra finches is inversely related to model abundance. Proceedings of the National Academy of Sciences USA. 1999; 96:12901–12904.
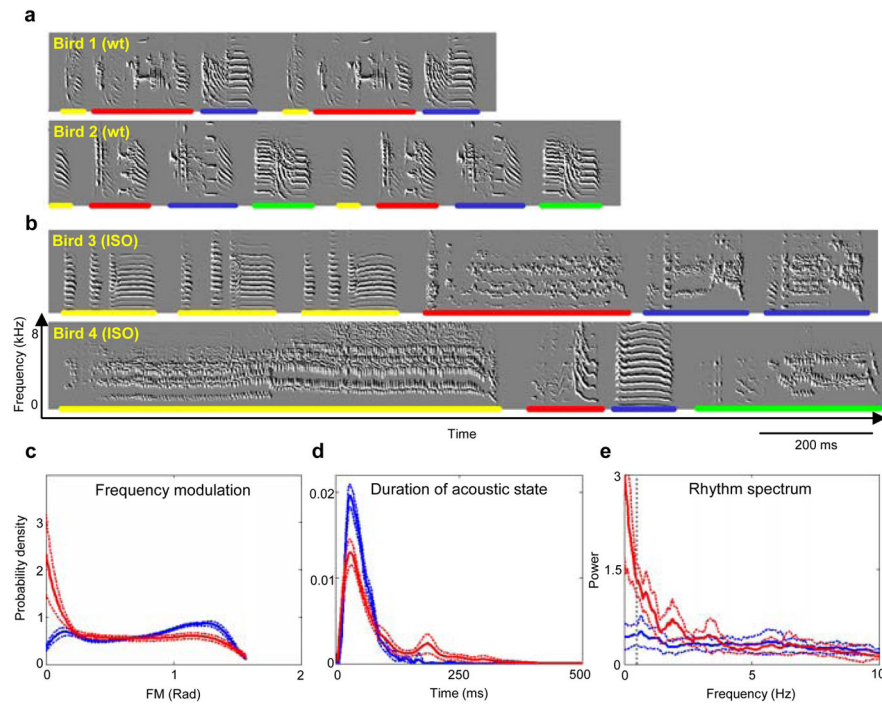
Figure 1. Wild-type songs versus isolate songs

a, Spectral derivatives19 of two WT song bouts. Different syllable types are underlined in different colors. Syllables show stereotypical organization into song motifs and rapid acoustic transitions within syllables. **b**, Isolate song bouts. Some syllables are extremely long (Bird 4, yellow) and others are stuttered (Bird 3, yellow and blue). **c**, Mean distribution histogram of *frequency modulation* in WT birds (blue, n=52) versus ISO birds (red, n=17). Dotted lines represent 95% confidence intervals. **d,** Histogram of *duration of acoustic state*, demonstrating longer durations in ISO. **e,** Spectra of rhythm frequencies showing less structured rhythm in ISO. The dotted gray line marks the minimum frequency that we used for further analysis (0.5 Hz).
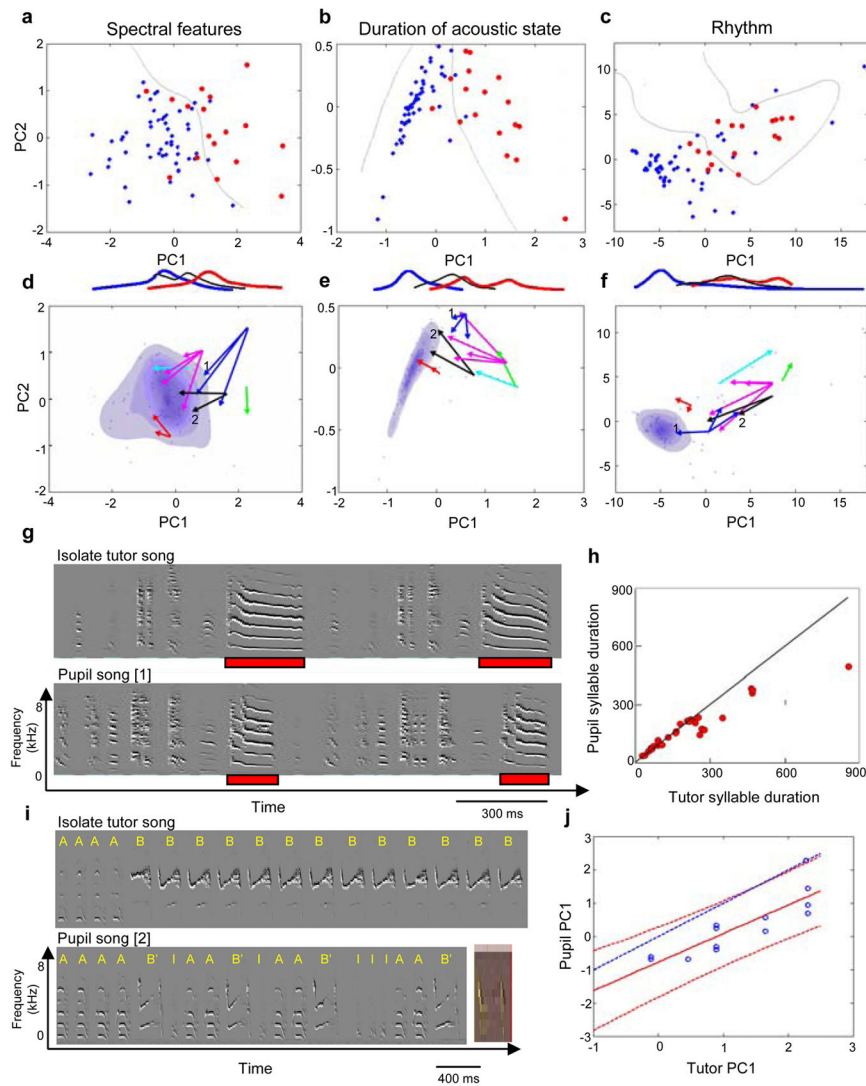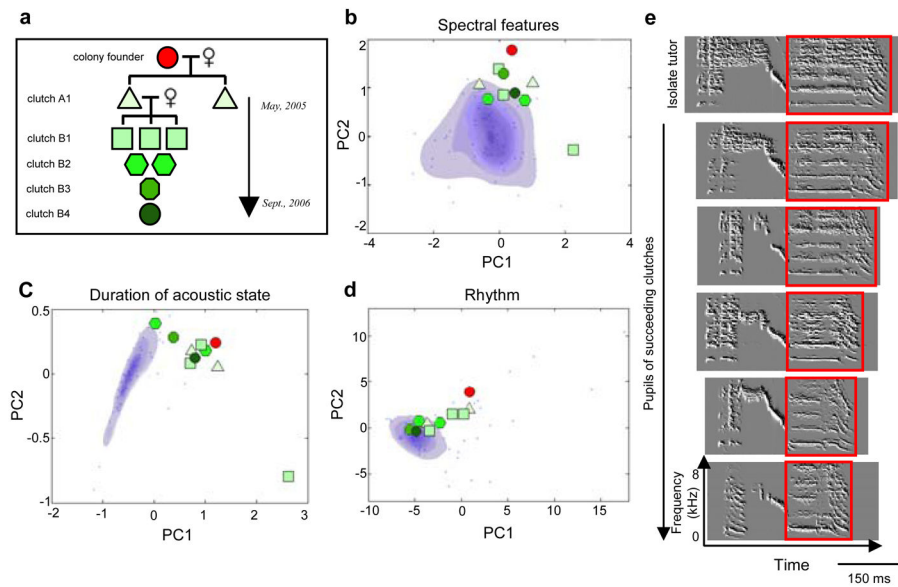
Figure 2. Progression toward WT song in pupils of isolates

First two PCs constructed from **a**, spectral features; **b**, DAS; **c**, rhythm frequencies. Dots represent individual WT (blue, n=52) and ISO (red, n=17) birds. Bayes classification lines are shown in gray. Histogram (bottom) of PC1 in first-generation (black, n=13) pupils falls between WT and ISO. **d**–**f**, Same data as in **ac**. Arrows originate at the tutors and point toward pupils. Different colors represent different tutors. Purple shading indicates center of WT cluster. Numerals indicate the arrows corresponding to the songs in **g** and **i. g**–**h**, Biased copying of syllable durations. **i**, Biased copying of syllable abundance and emergence of song motif. Shaded rectangle: overlay of syllable B and its imitation, B′. **j**, Correlation between first PCs of pupil versus tutor, indicating biased imitation. Dashed red line represents 95% confidence band, and the dashed blue line is the identity line.
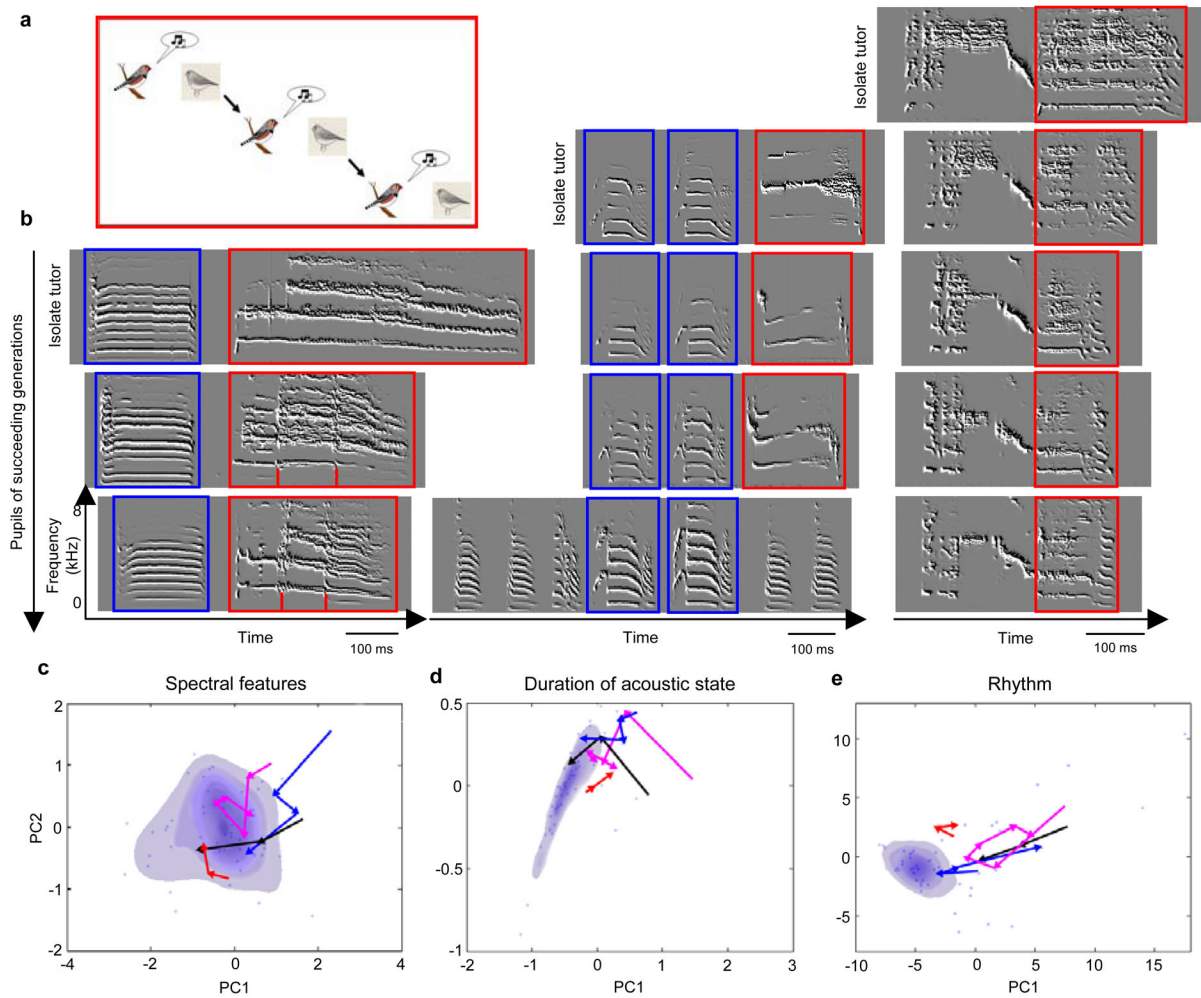
Figure 3. Multi-generational progression toward WT song
a, Schematic diagram of the experimental paradigm. Pupils become tutors when they reach adulthood (day 120–140). **b**, Three examples of the songs of isolate tutors and the succeeding generations of learners. Blue and red boxes show individual syllable types that are altered by pupils. Long, monotonic syllables become shorter and more differentiated (left and right panels). Rarely, syllables were omitted (middle panel) in later generations of learners **c**–**e**, PCA of song features, state duration and rhythm spectra. As in Fig. 2d–f, arrows originate at the tutors and point toward pupils. The progression toward the WT cloud (purple ovals) continues over generations.

Figure 4. Progression toward WT song in an isolated colony
a, Family relationships in the first 5 clutches based on behavioral observations. **b–d**, PCA of song features, state duration and rhythm (as in Fig. 2d–f). The colony founder is marked by red dot. Colors and symbols identify individuals in (a). Successive clutches approach the WT cloud (purple shading) in the song features, especially in rhythm frequencies. **e**, A long syllable that dominates the founder isolate song motif, and its imitations in successive clutches.