

LETTERS

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

Olga Fehér¹, Haibin Wang², Sigal Saar¹, Partha P. Mitra² & Ofer Tchernichovski¹

Culture is typically viewed as consisting of traits inherited epigenetically, through social learning. However, cultural diversity has species-typical constraints¹, presumably of genetic origin. A celebrated, if contentious, example is whether a universal grammar constrains syntactic diversity in human languages². Oscine songbirds exhibit song learning and provide biologically tractable models of culture: members of a species show individual variation in song³ and geographically separated groups have local song dialects^{4,5}. Different species exhibit distinct song cultures^{6,7}, suggestive of genetic constraints^{8,9}. Without such constraints, innovations and copying errors should cause unbounded variation over multiple generations or geographical distance, contrary to observations⁹. Here we report an experiment designed to determine whether 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 required¹⁰. Zebra finch isolates, unexposed to singing males during development, produce song with characteristics that differ from the wild-type song found in laboratory¹¹ 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 towards the wild-type in three to four generations. Thus, species-typical song culture can appear *de novo*. Our study has parallels with language change and evolution^{12–14}. In analogy to models in quantitative genetics^{15,16}, we model song culture as a multigenerational 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 males¹⁷. 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 upsweeps¹¹ (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 isolates¹⁸ even in the presence of WT adults¹¹.

We quantified the differences between WT and ISO songs on three timescales. At the 10-ms timescale, we used spectral-frame features (for example frequency modulation; Supplementary Information, section 4a). At the 10–100-ms timescale, we used the correlation time of the spectral shape, termed the duration of acoustic state (DAS; Supplementary Information, section 4b). At longer (200–1,000-ms) timescales, we used measures of song rhythm (Supplementary

Information, section 4d)¹⁹. Feature probability distributions across birds differed between ISO and WT songs (Fig. 1c–e). ISO songs had lower frequency modulation, longer DAS 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 plus ISO), and retained the first two principal components (PC1 and PC2) to obtain two-dimensional song phenotype values (Supplementary Information, section 4e). Principal components at all three timescales show separable clusters for ISO and WT songs along a continuum (Fig. 2a–c). The mean values of PC1 were significantly different between ISO and WT at all timescales of song structure ($P < 0.001$, t tests, $n_{WT} = 52$ birds, $n_{ISO} = 17$ birds, adjusted for false discovery rate; Supplementary Information, section 5). We found that these differences are largely an outcome of tutoring deprivation and not of social isolation (Supplementary Information, section 3f).

To examine the imitation of isolate songs, we trained 13 juvenile birds (pupils) one to one with isolate tutors in a sound-isolated chamber. This allowed us to control genetic relatedness, and to minimize

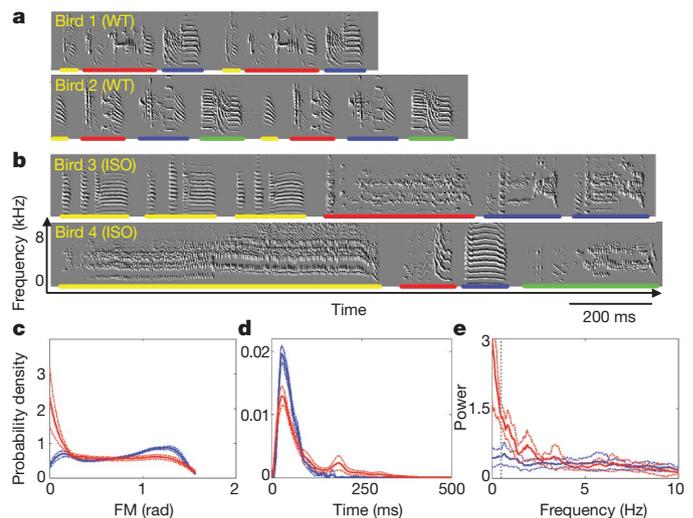


Figure 1 | Wild-type songs versus isolate songs. **a**, Spectral derivatives³⁰ of two WT song bouts. Different syllable types are underlined in different colours. Syllables show stereotypical organization into song motifs and rapid acoustic transitions within syllables. **b**, ISO song bouts. Some syllables are extremely long (bird 4, yellow) and others are stuttered (bird 3, yellow and blue). **c**, Mean distribution histogram of frequency modulation (FM) in WT birds (blue, $n = 52$) and ISO birds (red, $n = 17$). Dotted lines delineate the 95% confidence intervals. **d**, Histogram of DAS, showing longer durations in ISO. **e**, Spectra of rhythm frequencies, showing less structured rhythm in ISO. The dotted grey line marks the minimum frequency that we used for further analysis (0.5 Hz).

¹Department of Biology, City College, City University of New York, New York 10031, USA. ²Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.

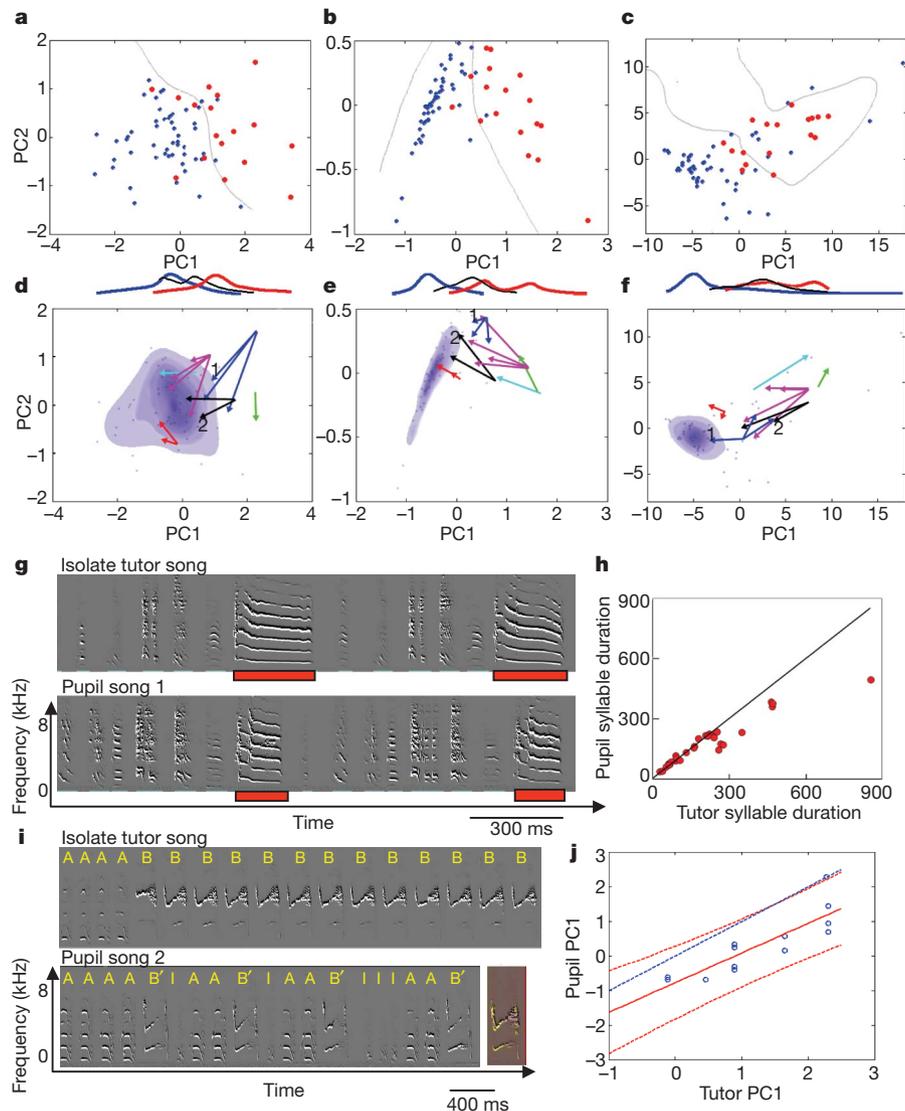


Figure 2 | Progression towards WT song in pupils of isolates. First two principal components constructed from spectral features (**a**), DAS (**b**) and rhythm frequencies (**c**). Dots represent individual WT (blue, $n = 52$) and ISO (red, $n = 17$) birds. Bayes classification lines are shown in grey. The histogram (bottom) of PC1 in first-generation pupils (black, $n = 13$) falls between those of WT and ISO. **d–f**, Same data as in **a–c**. Arrows originate at the tutors and point towards pupils. Different colours represent different

social effects, for example to eliminate feedback from female listeners. Four isolate tutors, with songs stable over the course of tutoring, were used two to four times to train unrelated pupils. We projected the feature distributions of the pupils onto the principal components 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 towards WT features in pupils of ISO tutors. The mean values of PC1 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 of 13 at at least one timescale, 10 of 13 at all timescales, false discovery rate significance of 0.01, binomial test, $n = 52$; Supplementary Information, section 5d).

Although pupils typically imitated all of the tutor syllables²⁰ and did not invent new syllables (Supplementary Information, section 2), pupil songs deviated consistently from tutor songs. In Fig. 2g, we show an example in which a long ISO syllable (red bar; mean duration of 367 ms, $\sigma = 29$ ms) was copied by a pupil, but was shortened by about

tutors. Purple shading indicates the centre of the WT cluster. Numerals indicate the arrows corresponding to the songs in **g** and **i**. **g, h**, Biased copying of syllable durations (milliseconds). **i**, Biased copying of syllable abundance and emergence of song motif. Shaded rectangle, overlay of syllable B and its imitation, B'. **j**, Correlation between PC1s of pupil and tutor, indicating biased imitation. Dashed red lines delineate the 95% confidence band, and the dashed blue line is the identity line.

30% (mean duration of 243 ms, $\sigma = 7.6$ ms). 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 230 ms (correlation coefficient $r^2 = 0.98$, slope of 0.97, $n = 20$ syllables; Fig. 2h). Copies of longer ISO syllables, however, were shorter than the originals ($r^2 = 0.84$, slope of 0.56, $n = 11$ syllables). Across birds, the ratio between the longest and shortest syllables within a bout was significantly smaller in pupils than in their ISO tutors ($P < 0.01$, $n = 13$, Wilcoxon sign test; Supplementary Information, section 4c). Overall, the range over which durations of ISO syllables were accurately copied is similar to the range of WT syllable durations (25–75 percentile range of 67–180 ms, $n = 52$ 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 in which one syllable dominated the ISO song (Fig. 2i), in pupils its abundance decreased to 20–30% (Supplementary Fig. 5), thereby creating more structured song motifs.

Imitation of spectral features, as judged by PC1 of the feature distribution, was also biased: linear regression analysis of pupil versus tutor yielded a non-zero intercept and a slope slightly less than one

(Fig. 2j). The equality line, corresponding to faithful copying (identical PC1s; dashed blue line), was rejected in favour of the alternative hypothesis represented by the linear fit shown in red ($P < 0.001$, likelihood ratio test, $n = 13$). We 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 towards 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 two to five generations (Fig. 3a). Similarity to WT songs increased over three to four generations, as can be appreciated from Supplementary Audio 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 third-generation pupil. In the rightmost 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 third and fourth generations.

To judge whether or not the imitation of ISO song progressed towards WT song over multiple generations, we displayed vectors in the principal-component space (as in Fig. 2d–f) with each tutoring lineage labelled using a different colour (Fig. 3c–e). As shown, the multigenerational 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 multigenerational comparisons suggest further progression towards WT for all song traits. For spectral-frame features, we found that PC1

of song features changes monotonically towards WT over generations. Its mean values for ISO, first generation, later generations and WT songs were 1.3, 0.3, 0.03 and -0.4 , respectively. PC1 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, PC1 values also decreased monotonically with increasing generation: 1.1, 0.3, 0.02, -0.3 . Later-generation songs were significantly different ($P < 0.01$) from both WT and ISO song, suggesting that WT approximation was not complete. For rhythm, PC1 values also decreased monotonically with increasing generation (4.1, 2.2, 1.4, -2), and differences from WT and ISO were marginally significant ($P = 0.02$ and 0.056, respectively).

Although the one-to-one training provided a well-defined learning environment, the multigenerational 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 Information, section 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 towards WT clusters, we examined principal-component projections, with the isolate tutor song marked as a red dot. Comparing the trajectory shown in Fig. 4e with that in the rightmost panel of Fig. 3b (originating from the same tutor), we see that the outcome in the colony is similar to that observed in one-to-one tutoring. Although the outcome of the colony experiment can only be judged qualitatively, we find it noteworthy 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 Information, section 1c). In

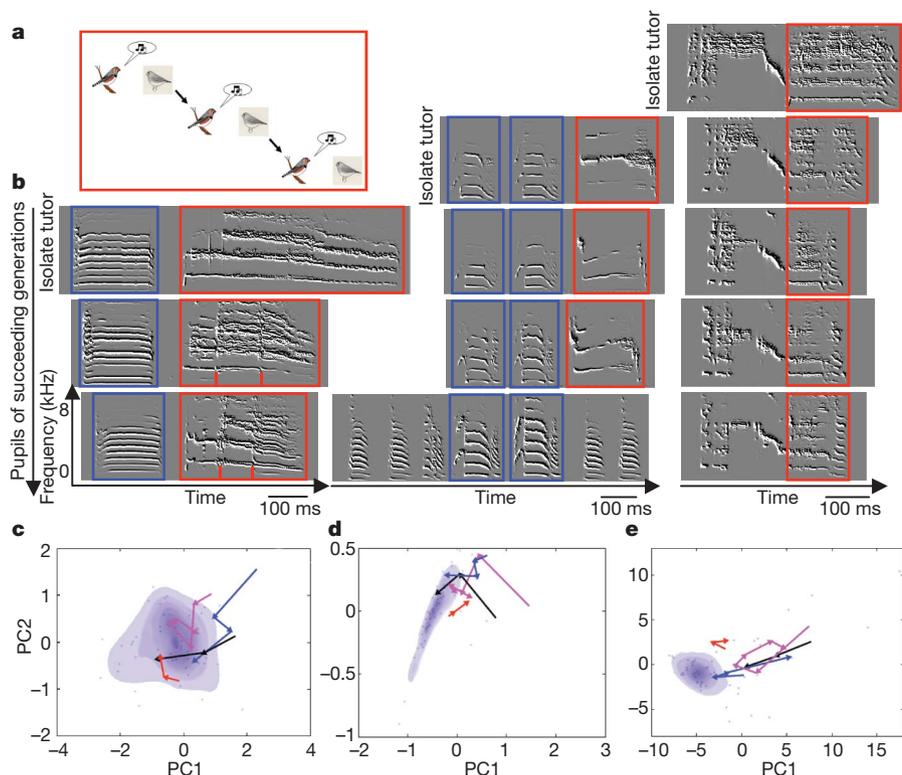


Figure 3 | Multigenerational progression towards WT song. **a**, Schematic diagram of the experimental set-up. Pupils become tutors when they reach adulthood (day 120–140). **b**, Three examples of the songs of isolate tutors and the succeeding generations of learners. Blue and red boxes show individual syllable types that are altered by pupils. Long, monotonic syllables become shorter and more differentiated (left- and right-hand panels).

Rarely, syllables were omitted (middle panel) in later generations of learners. **c–e**, PCA of spectral features (**c**), state duration (**d**) and rhythm spectra (**e**). As in Fig. 2d–f, arrows originate at the tutors and point towards pupils. The progression towards the WT cluster (purple ovals) continues over generations.

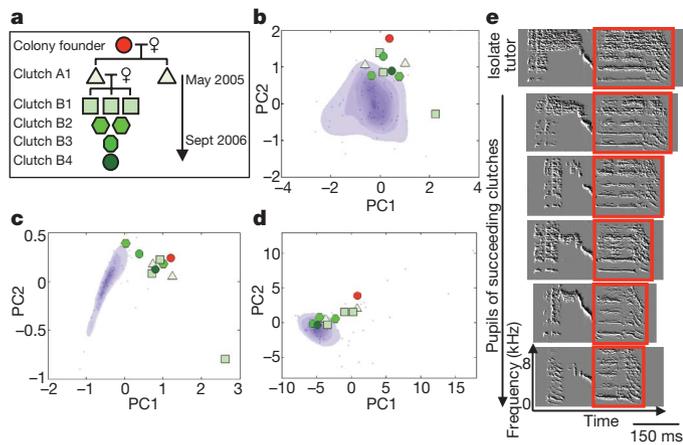


Figure 4 | Progression towards WT song in an isolated colony. **a**, Family relationships in the first five clutches, based on behavioural observations. **b–d**, PCA of spectral features (**b**), DAS (**c**) and rhythm (**d**) (as in Fig. 2d–f). The colony founder is marked with a red dot. Colours and symbols identify individuals in **a**. Successive clutches approach the WT cluster (purple shading) in spectral features, DAS, and most strongly in rhythm frequencies. **e**, A long syllable that dominates the founder isolate song motif, and its imitations in successive clutches.

contrast to one-to-one tutoring, the farthest progress towards 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 language²¹, as well as linguistic phenomena such as creolization. Models of language change and evolution^{12,13,15}, which contain a developmental account of the language acquisition process, are germane to our study (Supplementary Material, model details, section 3).

We further discuss our findings using a simple recursive model that motivated this study. Principal components of feature distributions (Fig. 2) give us phenotypic measures of song. Consider the distribution of a quantitative phenotype P in the ISO population. Because some of the variation in ISO song is heritable, we partition P into genotypic and environmental values, as $P = G + E$, assuming an additive model for genetic variance²², $V_P = V_G + V_E$.

We consider an isolated-lineages model, in which the environmental component of the pupil phenotype in the $(n + 1)$ th generation, $E(n + 1)$, is further divided into a portion, $E_0(n + 1)$, independent of the tutor, and a portion, $c_0 P(n)$, proportional to the tutor song phenotype. We therefore have the recursion

$$P(n + 1) = G(n + 1) + c_0 P(n) + E_0(n + 1) \quad (1)$$

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 (see Fig. 2j, where $c_0 = 0.86$ and $\sigma = 0.15$). The literature on cultural transmission^{24,25} also contains models analogous to equation (1) and has similar implications. Half-sib or cross-fostering experimental designs²⁶ should be useful for separating the genetic²⁷ and learning-related components of song transmission in future studies²⁸.

Our one-to-one experimental design may be modelled using equation (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 near 0. The largest changes occur in the first generation (consistent with our results). The case $c_0 = 1$ corresponds to a simple random walk ($V_{P(n)} \approx \sqrt{n}$), in which 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 Material, model details, sections 1 and 2). 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 ‘multigenerational’ 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—that is, their bearing on whether WT females prefer the songs of multigeneration pupils to those of ISO tutors—remains open. Because our findings suggest that song culture is the result of an extended developmental process, it would be interesting to examine whether changes in gene expression, neuronal reorganization or neurogenesis associated with song development show orderly multigenerational progression during the evolution of song culture.

METHODS SUMMARY

Animal care. All experiments were performed in accordance with guidelines of the US National Institutes of Health and have been reviewed and approved by the Institutional Animal Care and Use Committee of City College, City University of New York.

Experimental design. We used zebra finches (*Taeniopygia guttata*) from the City College breeding colony. Colony management and isolation procedures have been described previously²⁹. Except for the colony experiment, all birds were kept either singly (isolates) or in pairs (one-to-one tutored) in sound attenuation chambers (Supplementary Information, section 3e) between days 30 and 120 post hatch. WT songs ($n = 52$) were obtained from birds raised in two well-established colonies. Isolates ($n = 17$) were raised by their mothers from day 7 to day 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 six isolate tutors on day 30. For the colony setting, we made a sound isolation chamber from an inoperational 20-ft³ refrigerator (Supplementary Fig. 1). All birds in the colony (except for the three female founders) were the descendants of the founder male.

Data analysis. All the analysis was performed using MATLAB7, 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 over which acoustic features remain highly correlated and named this feature the DAS (Supplementary Information, section 4b). Rhythm spectrum¹⁹ was used to detect periodicity in song features at the syllabic and song-motif levels (Supplementary Information, section 4d). We constructed song-feature principal components by first computing cumulative frequency distributions for each feature time series (Supplementary Fig. 8). These cumulative frequency distributions were the input vectors for the PCA (Fig. 2a–c). Statistical tests are described Supplementary Information, section 5.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 October 2008; accepted 16 March 2009.

Published online 3 May 2009.

1. Marler, P. & Sherman, V. Innate differences in singing behavior of sparrows reared in isolation from adult conspecific song. *Anim. Behav.* **33**, 57–71 (1985).
2. Chomsky, N. *Aspects of the Theory of Syntax* (MIT Press, 1965).
3. Catchpole, C. K. & Slater, P. B. J. *Bird Song: Biological Themes and Variations* (Cambridge Univ. Press, 2008).
4. Marler, P. & Tamura, M. Culturally transmitted patterns of vocal behavior in sparrows. *Science* **146**, 1483–1486 (1964).
5. Olofsson, H. & Servedio, M. R. Sympatry affects the evolution of genetic versus cultural determination of song. *Behav. Ecol.* **19**, 596–604 (2008).
6. Soha, J. A. & Marler, P. A species-specific acoustic cue for selective song learning in the white-crowned sparrow. *Anim. Behav.* **60**, 297–306 (2000).
7. West, M. J. & King, A. P. in *Issues in the Ecological Study of Learning* (eds Johnston, T. D. & Pietrewicz, A. T.) 245–274 (Erlbaum, 1985).
8. Gardner, T. J., Naef, F. & Nottebohm, F. Freedom and rules: the acquisition and reprogramming of a bird’s learned song. *Science* **308**, 1046–1049 (2005).
9. Adret, P. In search of the song template. *Ann. NY Acad. Sci.* **1016**, 303–324 (2004).

10. Volman, S. F. & Khanna, H. Convergence of untutored song in group-reared zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* **109**, 211–221 (1995).
11. Williams, H., Kilander, K. & Sotanski, M. L. Untutored song, reproductive success and song learning. *Anim. Behav.* **45**, 695–705 (1993).
12. Niyogi, P. *The Computational Nature of Language Learning and Evolution* (MIT Press, 2006).
13. Nowak, M. A., Komarova, N. L. & Niyogi, P. Computational and evolutionary aspects of language. *Nature* **417**, 611–617 (2002).
14. Kirby, S., Cornish, H. & Smith, K. Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proc. Natl Acad. Sci. USA* **105**, 10681–10686 (2008).
15. Cheverud, J. M. Evolution in a genetically heritable social environment. *Proc. Natl Acad. Sci. USA* **100**, 4357–4359 (2003).
16. Dickerson, G. Comparison of hog carcasses as influenced by heritable differences in rate and economy of gain. *Iowa Agric. Exp. Sta. Res. Bull.* **354**, 489–524 (1947).
17. Jones, A. E., ten Cate, C. & Slater, P. J. B. Early experience and plasticity of song in adult male zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* **110**, 354–369 (1996).
18. Böhner, J. Song learning in the zebra finch (*Taeniopygia guttata*): selectivity in the choice of a tutor and accuracy of song copies. *Anim. Behav.* **31**, 231–237 (1983).
19. Saar, S. & Mitra, P. P. A technique for characterizing the development of rhythms in bird song. *PLoS ONE* **3**, e1461 (2008).
20. Nelson, D. A. & Marler, P. Selection-based learning in bird song development. *Proc. Natl Acad. Sci. USA* **91**, 10498–10501 (1994).
21. Kegl, J. in *Language Acquisition in Signed Languages* (eds Morgan, G. & Woll, B.) 207–254 (Cambridge Univ. Press, 2002).
22. Falconer, D. S. & Mackay, T. C. F. *Introduction to Quantitative Genetics* 4th edn (Prentice Hall, 1996).
23. Willham, R. L. The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics* **19**, 18–27 (1963).
24. Cavalli-Sforza, L. L. & Feldman, M. W. *Cultural Transmission and Evolution: A Quantitative Approach* (Princeton Univ. Press, 1981).
25. Boyd, R. & Richerson, P. J. *Culture and the Evolutionary Process* (University of Chicago Press, 1985).
26. Boake, C. R. B. *Quantitative Genetic Studies of Behavioral Evolution* 94–98 (University of Chicago Press, 1994).
27. Mundinger, P. C. Behaviour-genetic analysis of canary song: inter-strain differences in sensory learning, and epigenetic rules. *Anim. Behav.* **50**, 1491–1511 (1995).
28. Soha, J. A. & Marler, P. Cues for early discrimination of conspecific song in the white-crowned sparrow (*Zonotrichia leucophrys*). *Ethology* **107**, 813–826 (2001).
29. Tchernichovski, O., Lints, T., Mitra, P. P. & Nottebohm, F. Vocal imitation in zebra finches is inversely related to model abundance. *Proc. Natl Acad. Sci. USA* **96**, 12901–12904 (1999).
30. Tchernichovski, O., Nottebohm, F., Ho, C. E., Bijan, P. & Mitra, P. P. A procedure for an automated measurement of song similarity. *Anim. Behav.* **59**, 1167–1176 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Wallman and H. Williams for critical reading of the manuscript and consultation and F. Nottebohm, C. Harding and N. Leader for recordings of WT songs. The study was supported by US National Institutes of Health (NIH) grants to O.T. and P.P.M., by a NIH Research Centers in Minority Institutions grant to City College, City University of New York, and by the Crick-Clay Professorship to P.P.M.

Author Contributions The idea for the study originated with P.P.M., with important modifications by O.T. and O.F. The experiments were carried out by O.F. and O.T. The model was developed by P.P.M. with help from H.W. All authors participated in the data analysis, with major efforts by H.W. and O.F.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to O.F. (olcifeher@gmail.com).

METHODS

For a more thorough description of our methods and supplementary data, see Supplementary Information.

Animal care. All experiments were performed in accordance with guidelines of the US National Institutes of Health and have been reviewed and approved by the Institutional Animal Care and Use Committee of City College, City University of New York.

Experimental design. We used zebra finches (*Taeniopygia guttata*) from the City College breeding colony.

Isolation from song culture: All birds (except for the isolated colony birds) were raised by their parents in a dedicated cage until day 7 post hatch. The father was then removed and the cage was taken to a nursery area housing mothers (who do not sing) and chicks only. Zebra finches do not imitate songs heard before day 20 (ref. 31), and measurements obtained on day 10 post hatch show elevated auditory thresholds³². Therefore, possible exposure to the father WT songs before day 7 and later exposure to siblings is unlikely to have an effect (Supplementary Fig. 2). Except for the colony experiment, all birds were kept either singly (isolates) or in pairs (one-to-one tutored) in sound-attenuating chambers from day 30 to day 120 post hatch. Our custom-made sound chambers provide a sufficient level of box-to-box isolation for zebra finch songs (Supplementary Information, section 3e).

Experimental groups. WT songs ($n = 52$): To obtain a baseline for WT songs produced by domesticated zebra finches, we used recordings from birds raised in two well-established colonies, The Rockefeller University colony and Hunter College colony. Both colonies have existed for over 20 years.

Isolates ($n = 17$): Birds kept visually and acoustically separated from other birds during the sensitive period for song learning are called isolates. Birds were isolated from songs from day 7 to day 29 post hatch. They were kept in complete isolation from day 30 until day 120 post hatch or later. Six of these birds were used as tutors. Tutors' ages ranged between 140 and 1,571 days (median age at beginning of first tutoring, 316 days).

One-to-one tutoring ($n = 13$ and 8 for first and later generations, respectively): We randomly selected 13 hatchlings from 40 breeding pairs and paired them with one of six isolate tutors on day 30. The isolate tutor and his pupil were kept together for 90 days in a sound-isolated chamber. The isolate tutors were removed when the pupils were 120 days old. At this time, we recorded the song of the pupil and obtained an additional recording from the isolate tutor (in a separate box) to test whether his song remained unchanged. After confirming that the isolate tutor's song remained stable over the tutoring period, a new pupil was placed with the tutor and the new training began.

Tutoring lineages: For four of the six isolate tutors, we established a line of learners, where the first-generation pupils tutored another generation that, upon reaching adulthood (between days 120 and 140), tutored another generation, and so on. This allowed us to track the same song as it was passed down over a few generations.

Isolate colony setting: We constructed a large isolation chamber from an inoperational 20-ft³ refrigerator (Supplementary Fig. 1). We used one of our

isolate tutors to establish the 'island colony'. Three females and this isolate male were placed into the chamber and kept completely isolated, acoustically as well as socially, from other birds over a period of two years. Once a pair bond had been established, we followed (by audio and video recording) the evolution of this colony. All birds in the colony (except for the three female founders) were the descendants of the founder male. However, he fathered only one clutch, after which one of his sons from this first clutch paired up with another of the original female founders and produced all the successive clutches. On the basis of partial video observations, we suspect that the rest of the colony birds were all descendants of this pair. We allowed the colony to grow, with the occasional removal of female offspring, over five generations of learners. The colony founder was removed just before the hatching of the fifth-generation learner.

Data analysis. All the analysis presented below was performed using MATLAB 7, except for feature calculations, which were done using SOUND ANALYSIS PRO 2.

Spectral-frame (10-ms) features: We used SOUND ANALYSIS PRO 2 to calculate song features in 10-ms windows. We analysed 20 s of singing bouts for each bird. To construct the PCA shown in Fig. 2a–c, we used the three features that showed the best separation between ISO and WT in the cumulative frequency distributions (Supplementary Fig. 5): frequency modulation, amplitude modulation and goodness of pitch.

DAS: Isolate song syllables and notes are often prolonged and monotonic. To quantify this notion, we estimated correlation time, namely, the interval over which acoustic features remain highly correlated. This procedure was implemented in SOUND ANALYSIS PRO 2. See Supplementary Information, section 4b, for details.

Rhythm spectrum: Rhythm spectrum¹⁹ was used to detect periodicity (rhythm) in song features on the timescale of the song bout. Rhythm frequencies can capture patterns of repetitions at the syllabic level and at the song-motif level. See Supplementary Information, section 4d, for details.

Constructing the song features PCA: The song of each individual bird is described by a set of feature vectors spanning multiple timescales. For example, when analysing 20 s of singing, we obtained several spectral-frame features (pitch, frequency modulation and so on), with 20,000 time-series values for each feature. We first computed cumulative frequency distributions for each feature time series (Supplementary Fig. 8). The cumulative frequency distributions, which summarize the distribution of each feature in a song, are the input vectors of the PCA shown in Fig. 2a–c. Note that each red dot is a two-dimensional projection of cumulative frequency distributions (red lines) presented in Supplementary Fig. 8 for each bird (combining frequency modulation, amplitude modulation and goodness of pitch).

Statistical tests: See Supplementary Information, section 5.

31. Roper, A. & Zann, R. The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology* **112**, 458–470 (2006).
32. Amin, N., Doupe, A. & Theunissen, F. E. Development of selectivity for natural sounds in the songbird auditory forebrain. *J. Neurophysiol.* **97**, 3517–3531 (2007).